hexagon:2611.00577v1

AN INFINITE CASCADE FOR THE DEFOCUSING CUBIC NONLINEAR SCHRÖDINGER EQUATION ON THE SQUARE TORUS

XI CHEN

Abstract. We prove that the defocusing cubic nonlinear Schrödinger equation on the square torus has a global solution whose $H^{6/5}$ norm tends to infinity. This result resolves a conjecture posed by Bourgain. In a weaker Sobolev topology the same solution approaches a horizontal almost periodic motion. We choose a one-dimensional invariant torus whose spatial mean has a nonzero stationary component in the mass gauge. This component couples pairs of high-frequency waves with orthogonal wave vectors. We arrange finitely many such interactions so that energy passes to horizontal modes, then realize the transfer along an exact trajectory in the stable manifold of a larger horizontal torus. A comparison in two Sobolev norms allows successive transfers to preserve the growth already attained at earlier times.

1. Introduction

Consider the defocusing cubic Schrödinger equation

$$i\partial_t u + \Delta u = |u|^2 u, \qquad (t,x,y) \in \mathbb{R}\times\mathbb{T}^2, \qquad \mathbb{T}^2 = (\mathbb{R}/2\pi\mathbb{Z})^2. \tag{1.1}$$

With normalized Haar measure, write $\|u\|_{H^r}^2 = \sum_{k\in\mathbb{Z}^2}(1+|k|^2)^r|\widehat{u}(k)|^2$, where $\widehat{u}(k)$ is the coefficient of $e^{ik\cdot(x,y)}$. Conservation of the mass $\int|u|^2$ and of the energy $\int(|\nabla u|^2+\frac{1}{2}|u|^4)$ bounds the H^1 norm on each trajectory. For $s>1$, a mode of length $N\gg 1$ carrying kinetic energy e contributes $(1+N^2)^s e/N^2$ to the squared H^s norm. Thus a transfer of a small amount of energy to sufficiently high frequencies can give a large Sobolev norm while respecting both conservation laws. In his list of problems on Hamiltonian PDE, Bourgain asked whether the higher Sobolev norm of a solution to the periodic defocusing cubic equation can grow without bound [6, §5, p. 43]. Our main result answers this existence question on the square torus.

Theorem 1.1. *There is an initial datum $u_0 \in H^{6/5}(\mathbb{T}^2)$ such that the global solution $u \in C(\mathbb{R};H^{6/5}(\mathbb{T}^2))$ of (1.1) with $u(0)=u_0$ satisfies*

$$\limsup_{t\to+\infty}\|u(t)\|_{H^{6/5}} = +\infty.$$

The construction yields more information about this trajectory. Its initial datum may be chosen so that $\|u(t)\|_{H^{6/5}}$ tends to infinity while $u(t)$ approaches, in a weaker Sobolev norm, a solution depending on one spatial variable (Corollary 5.2). The underlying global flow comes from the periodic restriction estimates of Bourgain [3, 5]. The theory below the energy space extends through the almost conservation argument of De Silva, Pavlović, Staffilani and Tzirakis [12] to the sharp periodic Strichartz estimates and the mass-critical global theory of Herr and Kwak [35, 36]. A different line of estimates controls the size of high Sobolev norms over long times. Bourgain and Staffilani developed polynomial bounds and the associated multilinear estimates [4, 46, 47]. Catoire and Wang treated flat tori with general periods [9]. Modified energies on compact manifolds and improved space-time estimates on irrational three-dimensional tori give

2020 *Mathematics Subject Classification.* Primary 35Q55; Secondary 35B40, 37K45.

Key words and phrases. Defocusing nonlinear Schrödinger equation, growth of Sobolev norms, resonant interactions, stable manifolds.

further bounds adapted to the underlying geometry, the latter for nonlinear powers strictly between three and five [44, 13].

Energy transfer. Kuksin obtained growth estimates for Schrödinger equations in the small-dispersion regime [38]. For (1.1) itself, Colliander, Keel, Staffilani, Takaoka and Tao showed that, given $0 < \delta < 1 < K$ and $s > 1$, a smooth solution can start with $\|u(0)\|_{H^s} < \delta$ and reach $\|u(T)\|_{H^s} > K$ at a finite time [11]. Their construction organizes resonant lattice rectangles into generations and transfers amplitude across them. Guardia and Kaloshin obtained a polynomial time bound for a prescribed factor of amplification [30]. The geometry of resonant wave mixing is also explicit in multiphase geometric optics [7]. Starting from five modes, Carles and Faou identified successive times at which new, higher frequencies appear [8]. The finite-transfer construction has counterparts for the quintic and analytic gauge-invariant nonlinearities [34, 29]. Guardia treated convolution potentials on the square torus [27], and Diophantine approximation provides the required near-resonances on irrational rectangular tori [22, 21].

The passage from a resonant interaction to a particular PDE trajectory depends on the time scale over which one can follow it. Periodic exchanges between a few modes can already be seen in a cubic equation with a spatially modulated coefficient [26] and in the one-dimensional quintic equation [25]. At the other extreme, Hani obtained a single unbounded orbit for the resonant cubic equation on $\mathbb{T}^2$ as part of a general criterion for long-time instability [31]. Hani, Pausader, Tzvetkov and Visciglia identified the asymptotic resonant dynamics of small data on $\mathbb{R} \times \mathbb{T}^d$ for $1 \leq d \leq 4$ by modified scattering. Their work gives unbounded Sobolev orbits on these waveguides for $d \geq 2$ [32]. A related description holds under partial harmonic trapping [33]. A different limit gives the continuous resonant equation. Faou, Germain and Hani derived it in a weakly nonlinear large box [16], and Germain, Hani and Thomann described its invariant subspaces and its relation to resonant Hermite dynamics [20]. For the cubic Szegő flow, the Lax pair and spectral coordinates give another mechanism for unbounded Sobolev orbits [17, 19]. For a class of small initial data the Szegő flow also approximates the cubic half-wave evolution on the circle over long time intervals [18]. For the defocusing half-wave equation itself, Chen has constructed solutions with exponentially growing H^s norms for $1/2 < s < 3/2$ [10].

Invariant tori. The horizontal functions $u(x, y) = U(x)$ form an invariant subspace of (1.1). The defocusing cubic NLS on the circle admits global Birkhoff coordinates [24]. Gauge rotations and spatial translations preserve each spectral action [23, Proposition 3.1]. We compute their angle weights in the physical Fourier numbering in Section 2.1. The nonlinear part of the coordinate map gains one derivative [37]. Quantitative Sobolev estimates, estimates on the frequencies and tame analytic bounds near the origin appear in [42, 43, 40]. These results make it possible to distinguish a change of an action from a change of phase along a finite-gap torus. For Schrödinger equations, KAM theory gives families of quasiperiodic motions in one and several space dimensions [39, 14, 45], while normal forms control the drift of small solutions over long intervals [1, 2]. Close to a plane wave, Faou, Gauckler and Lubich proved long-time Sobolev stability under suitable nonresonance conditions [15]. Maspero and Procesi obtained long-time stability near a family of small finite-gap tori of (1.1) [41]. Guardia, Hani, Haus, Maspero and Procesi constructed orbits with prescribed finite growth near such tori [28]. For $0 < s < 1$ their orbit can start arbitrarily close to the torus in H^s and reach an arbitrarily large H^s norm. For $s > 1$ they obtain an arbitrarily large amplification factor with initial proximity measured in L^2.

The construction. We choose the horizontal torus to satisfy a particular resonance. After removal of the common mass rotation, its spatial mean has a nonzero temporal Fourier coefficient at frequency zero. The three active one-dimensional modes have indices $0, 1, 2$, and the angular character producing this coefficient is $(2, -2, 1)$. When new horizontal modes are added, we vary the actions to keep the old angular velocities fixed. This exact equality is needed at arbitrarily late times. Expanding the cubic nonlinearity about the horizontal motion now produces a

quadratic interaction between high modes. With one zero spatial frequency supplied by the background, its resonance condition is $p+q=r$, $p\cdot q=0$, for lattice vectors p,q,r. In Section 2.2 we arrange these triples in finite families whose last generation lies on the horizontal axis. Small auxiliary horizontal modes give the limiting equilibrium contracting directions in every interior generation. Following a suitable orbit backwards from that equilibrium gives a section where its weighted norm is smaller by an arbitrarily prescribed factor.

The reason to prescribe a horizontal endpoint becomes apparent when passing to the PDE. At a large integer dilation of the finite family, we construct an exact orbit which approaches an enlarged horizontal torus exponentially and which, on any prescribed early time interval, is close to the old horizontal motion. The linear estimate in Section 3 locates the desired directions inside the full stable space, which can contain additional resonant pairs with the same contraction rate. The nonlinear estimate in Section 4 uses different amplitude scales for the selected modes and the error. We prescribe the stable projection of the finite orbit and set the additional stable parameters to zero. A bounded time primitive of the oscillatory mean and the exponential decay at the endpoint then give a comparison on the entire positive half-line. This allows us to find the low-norm end of the transfer on the same exact PDE trajectory. In Section 5 we use the exponential tail to carry each earlier part of the orbit through the next enlargement. The successive initial data can then be chosen in nested closed balls of $H^{6/5}$, each of which preserves every previously obtained strict inequality at its observation time. Their intersection supplies the initial datum in Theorem 1.1.

Conventions and the flow. Write $\langle k\rangle=(1+|k|^2)^{1/2}$, $\|u\|_r=\|u\|_{H^r}$, and $\|z\|_{h^r}=(\sum_k\langle k\rangle^{2r}|z_k|^2)^{1/2}$. We use $\langle f\rangle_x$ for the spatial average in x. A horizontal field depends only on the first spatial coordinate. The subspace of $H^r(\mathbb{T}^2)$ with second Fourier index in $d\mathbb{Z}$, $d\in\mathbb{N}$, is denoted by H^r_d. Convolution preserves this sublattice. We equip this subspace with the ordinary H^r norm. Here $A\lesssim B$ means $A\le CB$ with the same conventions on constants as for $O(\cdot)$. A subscript on $O(\cdot)$ indicates dependence of its constant, or the space in which a vector is estimated. The notation $o_R(1)$ refers to the last integer dilation $R\to\infty$ at a fixed stage of the construction.

The two conserved quantities we use are the mass $M(u)=\int|u|^2$ and the energy $E(u)=\int(|\nabla u|^2+|u|^4/2)$. For completeness, the global flow and its continuity on every H^r, $r>1$, follow here from an elementary estimate. Local existence follows from the algebra property of H^r and the Duhamel formula. On the existence interval the energy bounds $\|u\|_1$ by a constant B, using conservation first for smooth approximants and then by local continuity. Cauchy–Schwarz on $|k|\le N$ and $|k|>N$ gives

$$\|\widehat{u}\|_{\ell^1}\le CB\sqrt{\log(2+N)}+C_rN^{1-r}\|u\|_r.$$

Choose N to be the smallest integer above $(2+\|u\|_r)^{1/(r-1)}$. Weighted convolution bounds $\||u|^2u\|_r$ by $C_r\|\widehat{u}\|_{\ell^1}^2\|u\|_r$. The integral inequality for $Y=2+\|u\|_r$ is therefore bounded by the scalar equation $Y'=C_{B,r}Y\log(e+Y)$, which has no finite-time blow-up. Iterating the local continuity estimate gives continuity of the solution map on every compact time interval in both directions.

It will be convenient to remove the common mass rotation. We write

$$iv_t=-\Delta v+(|v|^2-2M(v))v,\qquad u(t)=e^{-2iM(v)t}v(t).\tag{1.2}$$

All comparisons between different trajectories are made in this gauge. The scalar factor has modulus one on each trajectory.

Acknowledgments and use of generative artificial intelligence.

The author acknowledges the use of ChatGPT 6.2 Astra during the research process to explore mathematical ideas, discuss possible proof strategies, and assist with the drafting and editing of the manuscript.

2. Horizontal tori and finite resonant transfers

A horizontal solution evolves by the one-dimensional equation. We first choose its frequencies so that a nonzero component of its spatial mean is stationary in (1.2). We then construct the finite systems of transverse modes coupled by that component. The time dependence of the interaction comes from the torus and the gain in Sobolev norm from the placement of the transverse frequencies.

2.1. **A stationary component of the spatial mean.** Use normalized Haar measure in both periods. Put $L = 2\pi$, $\kappa_0 = \sqrt{2}/L$, and $u(t,x) = \kappa_0 w(t/L^2, x/L)$. The period-one equation for w has nonlinearity $2|w|^2 w$. Let Φ be the Birkhoff map of [37, Theorem 2.1]. Its first component has differential $(d_0\Phi)_1(w)_n = -\widehat{w}(-n)$. Let Ω denote the resulting Birkhoff map in physical variables, given by

$$z_j(u) = -\kappa_0 \Phi_1(w, \bar{w})_{-j}, \qquad I_j = |z_j|^2 = \kappa_0^2 J_{-j}, \qquad z_j = \sqrt{I_j}\, e^{-i\theta_j}.$$

Here $z = \Omega(u)$, J_n is the period-one action and θ_j is defined for $I_j > 0$. Scaling both amplitudes by κ_0 preserves canonicality for normalized Haar measure. The map Ω is a real-analytic isomorphism $H^m \to h^m$ for every integer $m \geq 0$, with physical Fourier differential at zero [37, Theorem 2.1]. The mass trace identity gives $M(u) = \sum_j I_j$ [42, Section 2]. Write $\nu_j = \partial_{I_j} E$ for the physical frequencies.

Spectral numbering. Number the spectral gaps by continuation from the free double eigenvalues $n\pi$, $n \in \mathbb{Z}$. For $k \in \mathbb{Z}$ and $c \in \mathbb{C} \setminus \{0\}$, set $v_k(y) = ce^{2\pi iky}$. Multiplication by $D_k = \operatorname{diag}(e^{\pi iky}, e^{-\pi iky})$ preserves the two-periodic boundary condition. In the convention of [23, (1.5)], the spectral operator on two-periodic pairs is

$$\mathcal{L}(v) = i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \partial_y + \begin{pmatrix} 0 & v \\ \bar{v} & 0 \end{pmatrix}.$$

Conjugation gives $D_k^{-1} \mathcal{L}(v_k) D_k = \mathcal{L}(c) - \pi k \mathrm{Id}$. The unique open gap with edges $\pm|c|$ for the constant potential is therefore shifted to the raw index $-k$.[1]

Covariance in the physical chart. For every $u \in L^2(\mathbb{T})$ and real τ, β,

$$z_j(u(\cdot + \tau)) = e^{ij\tau} z_j(u), \qquad z_j(e^{i\beta} u) = e^{i\beta} z_j(u). \tag{2.1}$$

Indeed translation conjugates the monodromy based at a different point, and phase rotation conjugates it by a constant diagonal matrix. Each preserves the discriminant and hence every spectral action [23, Proposition 3.1(i),(iv)]. Both are canonical transformations. Fix a finite support A with all its actions positive. Its Cartesian coordinate subspace is invariant, and on its positive-action part the symplectic form is the standard action–angle form. For a fixed group element, comparing the coefficients of $dI_i \wedge d\theta_j$ in its pullback gives $\partial_\theta \theta' = \mathrm{Id}$. The shift in each angle therefore depends only on I and the group element. The circle law makes each multiplier on z_j, $j \in A$, a character. Its integer weight is constant on the connected cone $(0,\infty)^A$ by continuity (equivalently, use homotopy invariance of the degree). Fix nonzero amplitudes in A and multiply all of them by $\epsilon \downarrow 0$. The differential at the origin conjugates each physical

[1] With this operator the sign of the index of the single open gap in [23, Lemma 3.5] is reversed. For $k = 1$, translation by $1/4$ in y equals rotation by $\pi/2$ on v_1. At its active raw index -1 the angle formulas in Propositions 3.2 and 3.4 would assign opposite increments. The argument below uses only the invariance of the spectral actions.

symmetry by the ordinary Fourier transform. Dividing by the original nonzero coordinate identifies the translation weight as j and the gauge weight as one. Inactive coordinates are zero, and density of finite sequences together with L^2 continuity proves (2.1). Thus our active labels $0, 1, 2$ correspond to raw labels $0, -1, -2$.

In the original NLS gauge $\dot\theta_j = \nu_j$. Constant solutions of mass b have phase frequency b. Linearizing at such a solution and diagonalizing the pair $j, -j$ gives relative frequencies $\omega_j = \sqrt{j^2(j^2+2b)}$.

We seek an angular Fourier coefficient of the mean which stays fixed along the mass-rotating flow. For $A = \{0, 1, 2\}$ the integer vector $\alpha = (2, -2, 1)$ has gauge charge $\sum \alpha_j = 1$ and spatial momentum $\sum j\alpha_j = 0$. These are the two selection rules for the spatial mean. The next calculation shows that this coefficient actually occurs and that the torus can be varied while keeping it stationary.

Lemma 2.1 (Three resonant actions)**.** *There are tori with three positive actions on $A = \{0, 1, 2\}$ such that, with $\alpha = (2, -2, 1)$, $\alpha \cdot \nu = 2M$. The angular Fourier coefficient of the spatial mean at α is nonzero. The compression to A of $\partial_I(\nu - 2M)$ is invertible at these tori.*

Proof. Choice of the mass. At the plane wave, the desired resonance reduces to $f(b) = \sqrt{16+8b} - 2\sqrt{1+2b} - b = 0$. The function decreases strictly on $b \geq 0$ and changes sign. Its root b_* is in $(1.3340704772439, 1.3340704772441)$. For example $4/\sqrt{16+8b} \leq 1$, so $f' < 0$. Use total mass M and I_1, I_2 as independent coordinates. Analyticity and the simple root give the required pin for small positive transverse actions.

The mean coefficient. We compute the coefficient of the first monomial with angular index α. Set $u = \sqrt{b}\, e^{-ibt}(1 + w)$. At signed frequency (n, σ) the pencil for $(w, \bar w)$ has matrix

$$L_n = \begin{pmatrix} n^2 + b & b \\ -b & -n^2 - b \end{pmatrix}, \qquad (\sigma - L_n)^{-1} = \frac{\sigma + L_n}{\sigma^2 - \omega_n^2}.$$

Define $h_n = (\omega_n - n^2 - b)/b$. Record a Fourier input as $(n, \sigma; X, Y)$ when it contributes $(X, Y)e^{inx - i\sigma t}$ to $(w, \bar w)$. To extract the desired cubic coefficient, use one labelled copy of $(2, \omega_2; 1, h_2)$ and two copies of $(-1, -\omega_1; h_1, 1)$. For each subset of two or three leaves, sum over its ordered splittings into nonempty sets, the pairs for these sets being $(x, y), (z, v)$. Add the quadratic source $b(xz + 2xv, -yv - 2yz)$. At the triple also add $b(x_1x_2y_3, -y_1y_2x_3)$ over the six orders. Apply the displayed resolvent at the summed spatial and temporal indices. For two leaves of index -1 the divisor denominator is $4\omega_1^2 - \omega_2^2 = -12$. For the mixed pair it is $(\omega_2 - \omega_1)^2 - \omega_1^2 = \omega_2 b_*$ by the pin. At the root $(n, \sigma) = (0, b_*)$ the denominator is b_*^2. To give the algebra explicitly, write $x = \sqrt{1+2b_*}$, let $+$ denote the leaf of index 2 and $-$ either distinguished leaf of index -1. Reducing the two components of each resolvent output by the polynomial relation below gives

	first component	second component
$+, -$	$(-37 + 3x + 9x^2 + x^3)/8$	$(24 + 5x - 8x^2 - x^3)/4$
$-, -$	$(-35 + 3x + 7x^2 + x^3)/12$	$1 - x$
$+, -, -$	$(7 + 13x - 7x^2 - x^3)/2$	$(33 + x - 9x^2 - x^3)/4.$

The pin gives $x^4 + 8x^3 - 2x^2 - 8x - 47 = 0$. Dividing the root coefficient by two for the duplicate leaves and reducing with this polynomial gives $(x+5)(x-2)/(x+1) < 0$, since $1 < x < 2$ by the bracket for b_*. The linear changes to normalized complex action amplitudes multiply it by nonzero scalars. To see that action-dependent frequency shifts do not enter this coefficient, complexify the embedding, fix I_0 and retain only z_2 and the independent variable $\bar z_1$, setting $\bar z_2 = z_1 = 0$. Then $I_1 = I_2 = 0$ on this slice. The frequencies are precisely their plane values there, and the coefficient of $z_2\bar z_1^2$ is fixed by the three invertible internal pencils just described.

The charge and momentum equations force the angular indices of the mean to be $(1+l,-2l,l)$. The calculated term is the first possible term for $l=1$. Each higher term includes an additional pair of small variables. This proves nonvanishing on a sufficiently small positive regular torus.

The frequency minor. Use (M,I_1,I_2) as coordinates and $(\nu_0-2M,\nu_1-\nu_0,\nu_2-\nu_0)$ as frequency variables. At the plane their Jacobian is

$$\begin{pmatrix}-1 & C^{\mathsf{T}}\\ C & H\end{pmatrix},\qquad C_j=\partial_b\omega_j=j^2/\omega_j,\qquad H_{pj}=\partial_{I_j}(\nu_p-\nu_0)\Big|_M.$$

Indeed, in these coordinates $\nu_0=\partial_M E$ and $\nu_p-\nu_0=\partial_{I_p}E|_M$. Equality of cross derivatives of the energy gives the first row. It remains to compute the two-action Hessian H.

Fix total mass b and use $u=e^{i\gamma}(\sqrt{b-Q}+v)$, where $\int v=0$, $Q=\int|v|^2$, and every integral uses normalized Haar measure. This is a chart near the constant solution since its spatial mean is nonzero. The identity $\operatorname{Im}\int\bar u\,\mathrm{d}u=b\,\mathrm{d}\gamma+\operatorname{Im}\int\bar v\,\mathrm{d}v$ shows that the v_k, $k\neq0$, are canonical on the reduced space. The homological equation on the reduced space therefore involves only nonzero modes. Writing $S=v+\bar v$, the terms of degrees two through four in the reduced Hamiltonian are

$$H_2=\int|v_x|^2+\tfrac{b}{2}\int S^2,\qquad H_3=\sqrt{b}\int S|v|^2,\qquad H_4=\tfrac12\Big(\int|v|^4-Q^2-Q\int S^2\Big).$$

For $k\neq0$ put

$$c_k=\left(\frac{(k^2+b)/\omega_k+1}{2}\right)^{1/2},\qquad d_k=-\frac{b}{2\omega_k c_k}.$$

The canonical change $v_k=c_k\zeta_k+d_k\bar\zeta_{-k}$ gives $H_2=\sum_{k\neq0}\omega_k|\zeta_k|^2$. Its linear coordinates agree with the reduced Birkhoff linear coordinates up to individual phases. Indeed translations distinguish k from $-k$. The two branches in a fixed translation character have opposite nonzero frequencies, so the positive branch at zero mass continues as the positive branch for all $b>0$. The identity $c_k^2-d_k^2=1$ fixes the symplectic normalization.

For a monomial m let $\Delta_m=\sum_{(s,k)\in m}s\omega_k$, where $s=1$ labels ζ_k and $s=-1$ its conjugate. No momentum-zero cubic monomial has $\Delta_m=0$. Such a resonance would require one of the three frequencies to equal the sum of the other two, while momentum forces the corresponding absolute wavenumber to be the sum or the absolute difference of the other two. The function $x\sqrt{x^2+2b}$ is strictly increasing and strictly superadditive for positive x. With $\{\zeta_j,\bar\zeta_k\}=-i\delta_{jk}$, put $\chi_m=i(H_3)_m/\Delta_m$, so $\{H_2,\chi\}=-H_3$. The transformed quartic polynomial is $H_4+\frac12\{H_3,\chi\}$. The cubic generator is unique among translation-invariant generators. Every subsequent degree-four generator adds $\{H_2,\chi_4\}$, whose coefficients on $|\zeta_i|^2|\zeta_j|^2$ vanish. Hence the coefficients of these monomials give the actual reduced action Hessian H, with a factor two on the diagonal, since differentiating the exact energy at fixed mass gives $\nu_j-\nu_0$.

The following coefficient formula gives the required sign bounds. Expand each H_j, $j=2,3,4$, in the commuting variables $Z_{+,k}=\zeta_k$ and $Z_{-,k}=\bar\zeta_k$. In the spatial integrals retain only products of momentum zero. Write $h_m=[m]H_3$ for the coefficient of a monomial m. For $w=|\zeta_p|^2|\zeta_j|^2$, the coefficient of the reduced quartic Hamiltonian is

$$[w]H_4+\frac12\sum_{m,n}h_mh_n\sum_{\substack{(e,k)\in m,\,(-e,k)\in n\\(m\setminus(e,k))(n\setminus(-e,k))=w}}\frac{e}{\Delta_n}.$$

Indices in the inner sum denote occurrences, so a repeated variable is counted with its multiplicity. Momentum bounds the modulus of the contracted index by four. One may consequently form the coefficients using only $0<|k|\le4$, with

$$v_k=c_kZ_{+,k}+d_kZ_{-,-k},\qquad(\bar v)_k=c_kZ_{-,-k}+d_kZ_{+,k}.$$

Multiply the resulting quartic coefficient by $1+\mathbf{1}_{p=j}$ to obtain H_{pj}. To display the sign margins, the following enclosures use the rational endpoints for b_* above. In the square root of a positive rational x, the endpoints $10^{-23}\lfloor\sqrt{\lfloor 10^{46}x\rfloor}\rfloor$ and this number plus 10^{-23} suffice. Products and quotients use the extreme endpoint values of the two factors. The finite sum gives

(p,j)	interval containing $H_{pj}+C_pC_j$
$(1,1)$	$(-1.363692,-1.363690)$
$(1,2)$	$(-0.000001,0.000001)$
$(2,2)$	$(-1.200067,-1.200065)$.

Thus $H+CC^{\mathsf{T}}$ is negative definite at the plane. Invertibility persists on the nearby positive tori constructed on the pinned frequency surface. □

Extension of the reference trajectory. Append zeros to α on every added label. For a finite support A let $\lambda=(\nu_j-2M)_{j\in A}$. By a Diophantine vector modulo α we mean that $\alpha\cdot\lambda=0$ and, for some $\gamma_0>0$ and $\tau>0$,

$$|n\cdot\lambda|\geq\gamma_0\langle[n]\rangle^{-\tau}\quad(0\neq[n]\in\mathbb{Z}^A/\mathbb{Z}\alpha).$$

Here $[n]$ denotes the class of n, and its norm is measured in any fixed integral basis of the quotient. The quotient is free since α is primitive. The adjective *old* below refers to the support and the angles of the particular reference being extended. At the next stage all the positive actions of the enlarged reference belong to its old support.

Lemma 2.2 (Extension without drift of the old angles)**.** *Let $A\supset\{0,1,2\}$ be a finite active support, with positive actions I_A, such that $\alpha\cdot\lambda=0$ and $\det(\partial_{I_j}(\nu_i-2M))_{i,j\in A}\neq0$. Assume the fixed old rotating vector is Diophantine on $\mathbb{Z}^A/\mathbb{Z}\alpha$. Fix a finite set of distinct nonzero horizontal integers P and an amplitude constant $d>0$. With $a=dR^{-6/5}$, one can activate $I_{RP}=a^2(d_P+o(1))^2$, $d_P>0$ fixed, preserving every old frequency ν_i-2M exactly and changing the old actions by $O_A(a^2)$. The enlarged minor stays invertible. The rotating frequency of the new label is $R^2P^2+O_A(R^{-2}+a^2)$ for fixed configuration. The enlarged rotating vector can be chosen Diophantine modulo α. The primary vector can be chosen with this property.*

Proof. By [43, Theorem 18.1 and Corollary 18.2(ii)], $F(I)=(\nu_n-n^2-2M)_n$ is analytic on an open complex ℓ^2 neighborhood of the old finite point and $DF=-\mathrm{Id}+K$ with K compact. To see the normalization, let $(PJ)_j=J_{-j}$ and write $\omega_n^*=\omega_n^{(1)}-(2\pi n)^2-4\sum_m J_m$ on summable actions, where $\omega_n^{(1)}$ are the period-one NLS frequencies. The scaling in Section 2.1 gives $I=\kappa_0^2PJ$, $\nu=L^{-2}P\omega^{(1)}$ and $F(I)=L^{-2}P\omega^*(J)$. Since $D\omega^*=-2\mathrm{Id}+$ compact, the coefficient of the identity in DF is -1. The complex ℓ^2 extension is that of the whole renormalized map.

A fixed compression on the old indices and the new compression have limiting diagonal blocks DF_{AA} and $-\mathrm{Id}$ and vanishing cross blocks, since for each escaping index n both Ke_n and K^*e_n tend to zero in ℓ^2. In particular the Jacobian with old frequencies held fixed is the Schur complement of the old block. For precision, Theorem 18.1 with $p=4$ gives a complex neighborhood containing the nonnegative ℓ^1 actions, so it contains each of the finite points considered here. In the proof of Corollary 18.2(ii) [43, p. 96], the derivative of the corrected map is bounded into ℓ^r, $4/3<r<2$. The compactness there follows from Pitt's theorem for bounded maps $\ell^2\to\ell^r$. The old implicit-function equation for $F_A(I_A,I_{\mathrm{new}})=F_A(I_A^{\mathrm{old}},0)$ is uniform for large R by bounded analytic derivatives on a smaller ℓ^2 ball. This gives the claimed old changes and the enlarged minor.

At a fixed smooth horizontal torus the inactive frequencies have $F_n(I)=O_A(|n|^{-2})$. For a direct proof take the high quasimode of Section 4 on ray $(n,0)$, with unit complex amplitude, base phase $e^{inx-i(n^2+2M)t}$ in the original gauge. Its linearized PDE defect is $O_A(n^{-2})$ in L^2,

uniformly in t. Apply the differential of the Birkhoff map along the old solution. In the inactive n coordinate the result y_n satisfies $\dot{y}_n = -i\nu_n y_n + O_A(n^{-2})$. Let $\mathcal{F}$ denote the linear Fourier transform. The smoothing estimate $\Omega - \mathcal{F} : H^1 \to h^2$ and its inverse counterpart are Theorem 1.1 and Corollary 1.1(i) of [37].[2] It follows that $y_n = e^{-i(n^2+2M)t}(1 + O_A(1/|n|))$ since the quasimode has H^1 norm $O(|n|)$. Put $\delta = \nu_n - (n^2 + 2M)$. If $\delta \neq 0$, take $T = \pi/|\delta|$. The two free phases are opposite at T, so Duhamel's formula for y_n yields $2 \leq C_A/|n| + C_A T/n^2$. This proves the asserted $O_A(n^{-2})$ estimate.

The $O(a^2)$ variation follows from the bounded derivative of F on the common ℓ^2 ball.

An integer relation involving only old entries is unchanged by this procedure. We therefore impose the Diophantine condition first on the three original frequencies. On the primary pin use the invertible minor and absolute continuity in the frequency variables to choose a vector Diophantine on $\mathbb{Z}^A/\mathbb{Z}\alpha$. At an extension the Schur complement of the old block is invertible. In an arbitrarily small relative open box of added actions, after fixing the old frequencies, every combination involving a new coefficient removes a slab of measure $O(\gamma\langle n\rangle^{-\tau}/\max|n_{\text{new}}|)$ in local frequency coordinates. Use one representative in integral quotient coordinates, sum with τ greater than the new dimension plus two, and let $\gamma \downarrow 0$. The old divisors are already controlled. There is no uniform lower bound required for the Diophantine constant of the *new* vector. □

For a fixed old reference its rotating spatial mean $m(t)$ has the form $c + \partial_t F(t)$, with F smooth and bounded uniformly over all starting old angles (with the corresponding c). Indeed the spatial mean of the finite smooth Birkhoff embedding is smooth on the angle torus. Divide its nonconstant temporal terms by their frequencies, using Lemma 2.2. Of the zero-frequency characters, gauge charge one forces the multiple of α to be α itself. Thus $c = c_\alpha e^{-i\alpha\cdot\theta(0)}$ with the chosen sign convention. The coefficient stays away from zero under all the extensions used here. The inverse Birkhoff map is uniformly Lipschitz into L^2 near the compact primary phase torus. Evaluation with all extra coordinates removed differs by at most $C\|z_{\text{extra}}\|_{\ell^2}$ uniformly over angles. Sum this tail budget, and the primary action adjustments, to less than a fixed sufficiently small number. At a new step the smooth old embedding with slightly adjusted actions is sampled along the exactly preserved *old* rotating velocities. The primitive just used for that embedding therefore has bounds depending on the preceding stage only. The averaged phase of c can be absorbed in a common rotation of all the marked amplitudes.

2.2. **A finite resonant system with a horizontal endpoint.** To see which interactions survive at high frequency, expand the cubic nonlinearity about a horizontal field A. The quadratic terms in a perturbation v are $\bar{A}v^2 + 2A|v|^2$. The spatial mean of A couples waves at p and q to $p+q$ with kinetic mismatch $|p|^2+|q|^2-|p+q|^2 = -2p\cdot q$. Its stationary coefficient thus selects orthogonal vectors. After a constant phase rotation and a change of time the resulting Hamiltonian is the sum, over unordered nonzero orthogonal $p, q \in \mathbb{Z}^2$, of

$$\overline{b_{p+q}}b_p b_q + b_{p+q}\overline{b_p b_q}.$$

Here $b_k \in \mathbb{C}$ is the amplitude at k and the convention is $i\dot{b}_k = \partial_{\bar{b}_k}\mathcal{H}$, where $\mathcal{H}$ denotes this sum. A finite family of modes is called closed if the forward rule $(p, q) \mapsto p + q$ and the reverse rule $(p + q, p) \mapsto q$ generate precisely its designated interactions whenever both input modes belong to the family. Zero modes are excluded throughout. Write $|b|^2_{s,\mathcal{G}} = \sum_{k\in\mathcal{G}}|k|^{2s}|b_k|^2$.

Proposition 2.3 (Finite transfer)**.** *Let $s > 1$ and $G > 1$. There is a closed finite family $\mathcal{G} \subset \mathbb{Z}^2 \setminus \{0\}$ and an exact trajectory of its quadratic Hamiltonian equation which converges*

[2] On pp. 4–5 of arXiv:1505.04455v1 and in the published version [37, pp. 7216–7217], use the identity initial matrix and the factor λ in the action differential, as in [24, Sections 1 and 13] and [42]. These conventions give the action-to-gap estimate used in the proof.

exponentially in positive time to an equilibrium b^+ supported on horizontal modes. At some time t_0,

$$|b^+|_{s,\mathcal{G}} > G|b(t_0)|_{s,\mathcal{G}}.$$

Every mode has nonzero longitudinal component. Given a finite spatial cutoff after choosing the desired transfer and the size of the auxiliary terminal amplitudes, the geometry can also satisfy the isolation conditions of Lemma 2.6.

The proof has three parts. We first place the modes so that distinct rectangles interact only at prescribed corners. We then construct a trajectory on this finite set. Finally we choose integer coordinates that isolate the linearized resonances of the full equation.

2.2.1. *Incidences between rectangles.* Fix two distinct opposite corners $C_i = (x_i, y_i)$, $i = 1, 2$, and a nonzero finite slope z. Let P_z, Q_z be the orthogonal projections onto $\mathbb{R}(1, z)$ and $\mathbb{R}(-z, 1)$. Set

$$\alpha_i = x_i + y_i z, \qquad \beta_i = x_i - y_i/z, \qquad p(c) = P_z(c, 0), \quad q(c) = Q_z(c, 0), \quad H(c, d) = p(c) + q(d).$$

In particular $p(\alpha_i) = P_z C_i$, $q(\beta_i) = Q_z C_i$, and $H(\alpha_i, \beta_i) = C_i$. The two other main corners $H(\alpha_1, \beta_2)$ and $H(\alpha_2, \beta_1)$ will be in the preceding generation. For an interior rectangle retain all eight side modes and all sixteen corners for $c, d \in \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$. The extra side opposite $p(c)$ is $q(c)$, and their sum $H(c, c) = (c, 0)$ is the horizontal terminal pump for this pair. When both specified corners C_i are horizontal, use just its two main rows and two main columns. Their diagonal corners already serve as horizontal pumps.

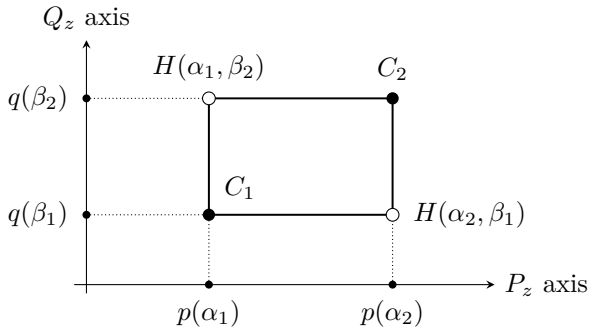


FIGURE 1. One step in the backward construction, drawn in the two projection axes. Each corner is the sum of its two orthogonal axis vectors. The filled corners are prescribed. The open corners become the prescribed pair in other rectangles at the next step.

A mode will be called private to a rectangle if it occurs in no other rectangle of the family.

Lemma 2.4 (Closure of the rectangular family)**.** *For every depth $J \geq 1$ one can construct $J + 1$ main generations of 2^J modes and join successive generations by rectangles as above. The last generation is horizontal. A main corner of each intermediate generation occurs in exactly one rectangle on each side of it. All side modes and all other corners are private to one rectangle. The only orthogonal triads created by any two distinct retained modes are the side-pair triads of the individual rectangles. These properties hold on a nonempty real algebraic open set of the parameters, dense in every ordinary open box away from the coordinate denominators.*

Proof. Begin with 2^J horizontal terminal coordinates as independent parameters. Partition them into singletons. At a backward step, merge two equal-sized blocks of 2^l terminal labels, match their 2^l current main corners bijectively, and assign an independent slope to each matched

pair. The two new off-diagonal corners for each pair make the next list of 2^{l+1} corners in the merged block. At $l = 0$ use the smaller rectangles. Later steps use the sixteen-corner rectangles. The two prescribed inputs of any interior rectangle vary independently with two-dimensional open images. Indeed they come from disjoint ancestral blocks. In each block the preceding slope makes the selected corner traverse the circle with diameter the preceding two prescribed corners. Common rescaling in that ancestral block supplies a radial variation independent of the circle tangent at a generic point. The same argument starts with independent one-dimensional horizontal terminal parameters.

All forbidden pairs are polynomial conditions after clearing the slope denominators. First consider relations among the modes of a single rectangle. It suffices to exhibit one choice of parameters at which every unintended incidence fails. For an interior rectangle take $C_1 = (7, 11)$, $C_2 = (19, -5)$ and $z = 3/5$. After a common dilation in orthogonal axes its modes are exactly

$$S = (\{0\} \cup 5\Gamma) \times (\{0\} \cup (-3\Gamma)) \setminus \{(0, 0)\}, \qquad \Gamma = \{102, 120, -85, 205\}.$$

Indeed the four parameters α_i, β_i are $68/5, 16, -34/3, 82/3$. For the final smaller rectangle take $C_1 = (7, 0)$, $C_2 = (19, 0)$ and the same slope, giving the same formula with $\Gamma = \{7, 19\}$. The following counts hold for each mode $v = (d, e)$ in each of these two finite tests, with $m = |\Gamma|$.

v	$\#\{w \in S \setminus \{v\} : v \cdot w = 0\}$	$\#\{w \in S \setminus \{v\} : v \cdot w = \lvert v\rvert^2\}$
$de = 0$	m	m
$de \neq 0$	0	0

For reproducibility the two entries are obtained by looping $D \in \{0\} \cup 5\Gamma$, $E \in \{0\} \cup (-3\Gamma)$, omitting $(0, 0)$, (d, e), and adding respectively the indicators of $dD+eE = 0$ and $dD+eE = d^2+e^2$. The axis counts are all supplied by the designated triads. These counts and the open-image property exclude all within-rectangle exceptional conditions as polynomial identities.

Next freeze the previous rectangles, with no zero or vertical mode, and vary just the fresh slope. For any previous mode O a persistent undesired pairing with a moving new mode would place its whole trajectory in a line perpendicular to O or on the circle with diameter $[0, O]$. The elementary expressions

$$p(c) = \frac{(c, zc)}{1 + z^2}, \qquad q(c) = \frac{(z^2c, -zc)}{1 + z^2}$$

classify all trajectories. The sides $p(\alpha_i), q(\beta_i)$ traverse the circle with diameter $[0, C_i]$. Its only possible circle coincidence is $O = C_i$, giving the designated reverse interaction. For unequal i, j the trajectory $H(\alpha_i, \beta_j)$ traverses the circle with diameter $[C_i, C_j]$. This circle misses zero for generic $C_i \cdot C_j \neq 0$. The equal-index cases are the already specified C_i. This covers all moving modes of a final smaller rectangle. For an interior rectangle take also $y_1y_2 \neq 0$. The sides $p(\beta_i)$ diverge horizontally as $z \to 0$, and $q(\alpha_i)$ do so as $|z| \to \infty$. The corners $H(\alpha_i, \alpha_j)$ diverge horizontally at infinity, $H(\beta_i, \beta_j)$ at zero, and $H(\beta_i, \alpha_j)$ at either end. Consequently no such trajectory is on a bounded circle. Its only possible containing line is horizontal, which cannot be perpendicular to the nonvertical O. These descriptions exclude accidental coincidences with frozen modes and show that no moving mode is identically vertical.

Order the rectangles within a layer arbitrarily. Other rectangles already constructed in the same layer are independent of the current slope. The restrictions on the current prescribed pair hold on a nonempty algebraic open subset of its ancestral parameters by their independent open images. They may therefore be intersected with all previous nonidentity conditions. This proves the assertion at every finite depth. □

The family can have an arbitrarily large ratio of the terminal and initial main weights. Start near the limiting configuration with one terminal horizontal label of length one and all other terminal labels zero. At the mergers along the distinguished block's ancestry each nonzero

distinguished descendant is paired with a zero donor. Choose its projection axes at 45 degrees to the distinguished vector, so each output has length $2^{-1/2}$ times that input. At depth J all 2^J initial main corners then have length $2^{-J/2}$. Continuous perturbation in the real terminal labels and angles yields an open set with one terminal length at least $1/2$ and all initial main lengths at most $2\,2^{-J/2}$. The main projections are continuous even at horizontal or vertical limiting axes, so we may move to finite nonzero slopes before using any auxiliary completion formulas. Intersect with Lemma 2.4, and take a sufficiently small open neighborhood with compact closure in this set. All auxiliary sizes are then bounded. The squared ratio of main weights obeys

$$\frac{\sum_{k \text{ in last main generation}} |k|^{2s}}{\sum_{k \text{ in first main generation}} |k|^{2s}} \geq 2^{-4s} 2^{J(s-1)}. \tag{2.2}$$

All strict lower bounds needed below may be enforced with a margin in this neighborhood.

2.2.2. *A transfer orbit with prescribed endpoint.* Each rectangle is a matrix H of corner amplitudes with side columns p, q. The individual contributions of its unordered orthogonal triads are

$$i\dot{H} = pq^{\mathsf{T}}, \qquad i\dot{p} = H\bar{q}, \qquad i\dot{q} = H^{\mathsf{T}}\bar{p}.$$

At a shared main corner add the contributions of its two rectangles. Index main rows by α_1, α_2, main columns by β_1, β_2, auxiliary rows by β_1, β_2, and auxiliary columns by α_1, α_2. Write $u = (1, -1)^{\mathsf{T}}/\sqrt{2}$ and let X be the matrix interchanging its two coordinates. Use one real scalar b_j for each main generation $0 \leq j \leq J$. At each rectangle of layer j, $0 \leq j < J$, set

$$p = q = e^{i\pi/4}(g_j u, h_j u), \qquad H = \begin{pmatrix} b_{j+1}\mathrm{Id} + b_j X & D_j \\ D_j & K_j uu^{\mathsf{T}} \end{pmatrix}, \quad D_j = \delta \mathrm{Id} + (A_j - \delta) uu^{\mathsf{T}}.$$

Omit the last two coordinates and the auxiliary matrix blocks at layer $J-1$. Direct substitution shows that the ansatz is invariant and yields

$$\begin{aligned} \dot{b}_j &= (g_{j-1}^2 - g_j^2)/2, & \dot{g}_j &= (b_j - b_{j+1}) g_j - A_j h_j, \\ \dot{h}_j &= -A_j g_j - K_j h_j, & \dot{A}_j &= g_j h_j, \qquad \dot{K}_j = h_j^2. \end{aligned} \tag{2.3}$$

Here $g_{-1} = g_J = 0$ and $A_{J-1} = h_{J-1} = K_{J-1} = 0$. At every interior layer $A_j^2 + K_j^2 + h_j^2$ is conserved. The terminal equilibrium has $b_J = 1$, all other $b_j = 0$, and $A_j = \delta$, $h_j = K_j = g_j = 0$ in the interior. The nonzero entries of its two cross matrices $\delta\mathrm{Id}$ are exactly the horizontal pump modes $H(c, c)$. Its last main diagonal is horizontal as well.

Lemma 2.5 (Transfer to the horizontal equilibrium)**.** *For every J and $\varepsilon_0 > 0$, and all sufficiently small fixed $\delta > 0$, this equilibrium has a forward exponentially convergent trajectory on which at some time $|b_0 - 1| < \varepsilon_0$ and $\max_{j \geq 1} |b_j| + \max_j |g_j| < \varepsilon_0$. All auxiliary corner and side amplitudes on that trajectory are $O(\delta)$.*

Proof. The linearization at the equilibrium is the real symmetric matrix $\left(\begin{smallmatrix} 0 & -\delta \\ -\delta & 0 \end{smallmatrix}\right)$ on each interior (g_j, h_j), the scalar -1 on the last g, and zero on the remaining variables. Let $r = \delta^{3/2}$. Prescribe stable eigenprojection with g_j component r at every layer. In the supremum norm weighted by $e^{\delta t/2}$ on $t \geq 0$, the stable integral from zero to t and complementary integral from t to infinity both cost $O(1/\delta)$. The nonlinear part of (2.3) is quadratic with constants fixed by J, so the stable integral equation contracts on an $O_J(r)$ ball with Lipschitz constant $O_J(r/\delta)$. Its exact trajectory at $t = 0$ has the prescribed stable vector plus $O_J(r^2/\delta) = o(r)$ in all coordinates and converges exponentially. Conservation now gives $A_j^2 + K_j^2 + h_j^2 = \delta^2$ throughout its continuation.

Use reverse time $\tau = -t$. The real symmetric Jacobi matrix L has diagonal b_j and off-diagonal $x_j = g_j/2$. For the skew matrix $S_{j,j+1} = x_j = -S_{j+1,j}$,

$$L' = [S, L] + D, \qquad D_{j,j+1} = D_{j+1,j} = A_j h_j/2, \qquad \|D\|_{\mathrm{op}} \leq \delta^2.$$

Orthogonal conjugation for S implies that throughout $0 \le \tau \le T = 10(J+1)\log(1/r)$ there is one eigenvalue $\lambda = 1 + O(\kappa)$ and all other eigenvalues are $O(\kappa)$, where $\kappa = C_J(r + T\delta^2)$. These bounds also give continuation on that interval. Thus $L = \lambda v v^{\mathsf{T}} + O_{\mathrm{op}}(\kappa)$ with $|v| = 1$, every $b_j \ge -C\kappa$, and $|g_j| \le C$. The partial sums $B_m = \sum_{j \le m} b_j$ satisfy $B_m' = g_m^2/2$ and $B_J = 1$ exactly.

Choose fixed $\varepsilon > 0$ with $(J+1)\varepsilon < 1/10$ and $30(J+1)\varepsilon < 1/6$. Until B_m first reaches ε, we have $b_m \le 2\varepsilon$ and $b_{m+1} \ge -C\kappa$. Hence, while $\tau \le T$, integration of the g_m equation on the positive branch gives

$$g_m(\tau) \ge (r/2)e^{-3\varepsilon\tau} - \delta^2\tau \ge r^{7/6}/4$$

for all sufficiently small δ. The bound itself keeps the branch positive, and $r^{7/6} \gg \delta^2$.

Induct downwards in m starting from $B_J = 1$. Once $B_{m+1} \ge 1 - (J-m-1)\varepsilon$, while $B_m \le \varepsilon$ one has $b_{m+1} - b_m > 2/3$, so $g_m' \ge g_m/2$. The uniform upper bound on $|g_m|$ forces B_m to reach ε within $4\log(1/r)$. If that threshold has already been crossed this part takes no time. While $\varepsilon \le B_m \le 1 - (J-m)\varepsilon$, one has $b_{m+1} \ge \varepsilon$. For $i < m$ the tridiagonal entry $L_{i,m+1}$ vanishes, so $|v_i v_{m+1}| = O(\kappa)$ and $|b_i| = O_\varepsilon(\kappa)$. Consequently $b_m \ge \varepsilon/2$ and $(g_m/2)^2 = b_m b_{m+1} + O(\kappa) \ge \varepsilon^2/4$ for sufficiently small δ. The previous partial sum therefore reaches $1 - (J-m)\varepsilon$ in at most $3\varepsilon^{-2}$ more time. The J steps all fit inside T for sufficiently small δ. At $b_0 \ge 1 - J\varepsilon$, the other b_j are $O_J(\varepsilon + \kappa)$ and every g_j is $O_J(\sqrt{\varepsilon + \kappa})$ by the rank-one estimate. First choose ε for the requested tolerance, then decrease δ. The auxiliary assertion follows from the conserved quantities and the explicit blocks $D_j, K_j u u^{\mathsf{T}}$. □

Choose J in (2.2) to give a strict factor larger than G^2 with room to spare. On the fixed relatively compact neighborhood of geometries, continuity of the finite weighted norm allows one choice of ε_0, then of δ, to preserve the desired transfer. This allows the spectral cutoff below to be chosen after the nonzero auxiliary amplitude has been fixed.

2.2.3. *Isolation of the transverse resonances.* In the full variational equation a horizontal mode can also couple two transverse modes outside the finite family. The following choice of integer coordinates determines which of these couplings persist in the limit. Here H is a fixed cutoff on Fourier harmonics of the old horizontal field.

Lemma 2.6 (Isolation on a sequence of sublattices)**.** *Given an integer $H \ge 1$, the preceding configuration can be chosen integral and then dilated through arbitrarily large integers R, divisible by any one prescribed integer, with the following properties. Denote all distinct horizontal terminal and auxiliary pump indices in the undilated configuration by P. For every nonzero transverse row Rm, every $|j| \le H$, and every integer k, the anomalous kinetic divisor*

$$k^2 + (RP + j - k)^2 + 2R^2m^2 - R^2P^2$$

has modulus at least cR apart from the exact pairs $j = 0$, $k = Rx$, $x(P-x) = m^2$, $x \in \mathbb{Z}$. These pairs form disjoint two-mode pairs in each doubled collection of opposite rows. None of their longitudinal indices is zero or equals any horizontal pump index RQ. The Sobolev transfer margin chosen before H persists.

Proof. For a prime p let v_p denote its additive valuation and $\mathbb{Z}_p$ the ring of p-adic integers. Let M be twice the product of u^2+v^2 over coprime distinct positive integers u, v with $|u^2-v^2| \le 2H+2$. There are finitely many since $|u^2 - v^2| \ge 2\min(u,v) + 1$. At every prime $p \mid M$ take all terminal root parameters p-integral, and impose $v_p(z) = 1$ at every slope. These rational choices are dense in each real parameter interval by ordinary congruences on numerators and denominators. They can be made inside the preceding open set. When $x \in \mathbb{Z}_p$, $y \in p\mathbb{Z}_p$, both $x + yz$ and $x - y/z$ are integral. Conversely each of $p(c), q(c)$ has integral first and p-divisible second coordinate for integral c, since $1 + z^2$ is a unit. This propagates down the entire tree, including the mirrors.

Clear all prototype denominators by a common integer prime to M. Call the resulting distinct horizontal labels c_i.

Choose an odd prime $\pi \nmid M$ outside the finite list of nonzero $c_i - c_j$ divisors and slope projection denominators. Choose an arbitrarily large K with $M \mid K$ and $v_\pi(K) = 1$. Scale and translate every specified main corner by $C \mapsto KC + (1, 0)$. At each rectangle the induced transformation is, for every c, d including auxiliary labels,

$$\begin{aligned} \alpha_i &\mapsto K\alpha_i + 1, \qquad \beta_i \mapsto K\beta_i + 1, \\ p(c) &\mapsto Kp(c) + P_z(1, 0), \qquad q(c) \mapsto Kq(c) + Q_z(1, 0), \\ H(c, d) &\mapsto KH(c, d) + (1, 0). \end{aligned}$$

Thus all shared corners still agree even when adjacent rectangles have different axes. Every horizontal mirror pump is translated by the same amount. After normalization by K the deformation tends to zero, so the finite strict nonincidences and transfer margin survive. A final common denominator D for the translated projections is prime to $M\pi$. The new horizontal labels are $P_i = D(1 + Kc_i)$, prime to M. For distinct P, Q the quantity

$$Q(P - Q) = D^2(1 + Kc_Q)K(c_P - c_Q)$$

has odd π-adic valuation, hence cannot be a nonnegative square.

It suffices to take $m > 0$. Each P is odd, so $|P| \neq 2m$ and there is no double root. For $|P| < 2m$ the divisor is uniformly of size at least cR^2 for $|j| \leq H$, as seen by completing the square. For each of the finitely many $0 < 2m < |P|$, the rescaled simple roots at $j = 0$ are $x = (P \pm \sqrt{P^2 - 4m^2})/2$. For an irrational x the corresponding root in k is

$$Rx - \frac{(P - x)j}{2x - P} + O_{P,H}(R^{-1}).$$

Simultaneously keep the leading quantities in this formula a positive distance from the integers for all irrational roots and $|j| \leq H$. Indeed in the compact group closure of the cyclic vector made of these irrational coordinates modulo one, each coordinate projection has full circle image. Each forbidden translated coordinate fiber has Haar measure zero. The complement of sufficiently small neighborhoods of their finite union contains a nonempty open subset of the compact group, visited unboundedly often. The same argument applies on multiples of the prescribed integer, as multiplying an irrational by an integer preserves irrationality. Distance from both simple roots gives a lower bound cR for the divisor.

A rational root is integral, since the square root of the integer $P^2 - 4m^2$ is then integral with the parity of P. Write $|x| = gu^2$ and $|P - x| = gv^2$ for coprime positive u, v, using $x(P - x) = m^2$. Then $|P| = g(u^2 + v^2)$ and $|u^2 - v^2| > 2H + 2$ by coprimality with M. For $k = Rx + l$ near this root the coefficient of $2R$ in the divisor is $(2x - P)l + (P - x)j$. If it vanishes, then $(u^2 - v^2)l + v^2 j = 0$ up to an overall sign. Coprimality forces $u^2 - v^2$ to divide j, so $j = l = 0$. At bounded l every other coefficient is bounded away from zero. Outside fixed neighborhoods of the two k roots the same cR bound follows by factoring the quadratic. Finally, for fixed nonzero x, m, the parent must be $P = x + m^2/x$. This yields the disjointness. The alternative $x = Q$ for a horizontal pump would require $m^2 = Q(P - Q)$, already excluded (also trivially when $P = Q$). The assertion follows. □

The order of choices also permits an arbitrary small initial norm in a later amplitude-scaled comparison. If an initial tolerance $d > 0$ and a terminal threshold A are fixed first, take the strict norm gain larger than $2A/d$ before the sieve. After constructing the integer prototype let $W = |b(t_0)|_{s,\mathcal{G}} > 0$ and dilate its labels by R. Quadratic time and amplitude scaling by $a = d/(2WR^s)$ gives initial homogeneous norm $d/2$ and terminal norm greater than A. The constant aR^s may depend on this fixed configuration but is independent of the final dilation.

Equivalence with the inhomogeneous norm costs $1 + o(1)$ on these nonzero dilated labels. This proves Proposition 2.3 for the auxiliary Hamiltonian and records the quantitative freedom used in selecting the dilations.

2.3. **An unbounded orbit of the comparison equation.** For the autonomous quadratic system, disjoint families can be placed so that the restriction to their union is a direct sum. This gives a first application of the finite transfer. Give the nonzero lattice labels canonical complex amplitudes and use, with unit coefficient, the Hamiltonian sum of $\bar{b}_{p+q} b_p b_q$ and its conjugate over unordered nonzero orthogonal pairs. The stationary coefficient of the background has already been scaled to one.

Corollary 2.7. *For every $s > 1$ there is a single $b(0) \in h^s(\mathbb{Z}^2)$ for this quadratic equation whose solution tends in h^1 to a horizontal $b_\infty \notin h^s$. In particular its h^s norm tends to infinity.*

Proof. Inductively take disjoint finite networks from Section 2.2. In the generic choices for each new network also impose the finitely many tests against the previous labels $v = (v_x, v_y)$, all with $v_x \neq 0$. For a freely variable horizontal root $(x, 0)$ an unintended reverse triad would require $x v_x = |v|^2$ or $x(v_x - x) = 0$. For each new moving side or corner use the same line and circle tests as in the private-grid proof, with v now fixed. In particular choose every new direction nonperpendicular to every old direction. Subsequent dilation of the new block by N preserves this last condition. A coincidence or either reverse test involves only finitely many additional values of N, since $(Nk - v) \cdot v = 0$ or $(v - Nk) \cdot k = 0$ for one of the finitely many new k and old v. Thus the larger subtori required for this corollary can have arbitrary integer dilations, preserving closure with every previous block.

In block j require a gain greater than 2^{2j}. At the early section of the finite orbit let its homogeneous h^s norm before final dilation be $W_j > 0$. Put $\eta_j = 2^{-j}/(N_j^s W_j)$. Dilate the labels by N_j and take η_j times the orbit evaluated at $\eta_j t$. These are exact symmetries of the quadratic equation. The early homogeneous norm is 2^{-j} and its limiting value is greater than 2^j. By continuity there is an internal time neighborhood of the early section on which the norm is at most twice its initial value. Take N_j sufficiently large that $|t| \leq j$ maps into this neighborhood. Orthogonality and the summable squares then give convergence to a continuous h^s orbit on every compact time interval. The restricted convolutions are bounded in h^s, by ordinary absolute convolution for $s > 1$. Hence the componentwise solutions give the strong solution of this equation.

Finite blocks exist in both time directions because the positive definite quadratic integral below bounds every coordinate on their nonzero finite supports. Each block conserves $\sum_k |k|^2 |b_k|^2$ because $|p + q|^2 = |p|^2 + |q|^2$ on every triad. These values are summable over the blocks. The finite blocks each tend to their horizontal endpoint and their supports avoid zero, so dominated convergence gives h^1 convergence of the whole orbit. Orthogonality shows that the limit has infinite h^s norm. Taking finitely many endpoint coordinates at a time gives the assertion about the full limit of the norm. □

3. Linearized dynamics at a horizontal endpoint

3.1. **A single transfer in the full equation.** Fix $1 < q < s = 6/5$. A horizontal reference C is *admissible* if it is constructed from the three-action torus of Section 2.1 by a finite number of its frequency-preserving extensions. Its active frequency minor is invertible, its rotating frequencies are Diophantine modulo α, and its coefficient at α is bounded away from zero in the specified small neighborhood. We specify its starting angles at physical time zero. The following proposition formulates the transfer we shall iterate.

Proposition 3.1 (Finite landing step)**.** *Given an admissible reference C, an integer $d_0 \geq 1$, an initial tolerance $\varepsilon > 0$ and a terminal threshold $K > 0$, one can fix a finite configuration and a constant $L < \infty$, independently of the final dilation R. Along a sequence of arbitrarily large multiples R of d_0 there exist an admissible horizontal C_R and an exact solution V_R on H^s_R such that*

$$\sup_{t\geq 0} \|V_R(t) - C(t)\|_q = o_R(1), \quad \sup_{t\geq 0} \|V_R(t) - C(t)\|_s \leq L,$$
$$\limsup_{R\to\infty} \sup_{0\leq t\leq D} \|V_R(t) - C(t)\|_s < \varepsilon \quad \textit{for each fixed } D < \infty,$$
$$\inf_{t\geq 0} \|C_R(t)\|_s > 2K, \qquad \|V_R(t) - C_R(t)\|_s = O_R(e^{-\gamma_R t}).$$

The variational equation at C_R in H^s_R and H^q_R has compatible exponential dichotomies at a weight $\gamma_R > 0$, with a positive margin on both sides in the sense of Lemma 5.1. The last estimate is in the corresponding strong stable fibre. The same small tail budget can be imposed on the new coordinates in ℓ^2 at every finite extension.

At early times the orbit is compared with C. At large times the solution tends to C_R, whose norm already exceeds $2K$ at time zero. We shall build the trajectory from the stable directions at C_R and use the finite orbit from Section 2.2 to determine which stable data to prescribe.

3.2. **The transverse variational equation.** We recall the roles of the small parameters. The undilated integer family of Section 2.2 is fixed for this transfer. Its dilated lattice is $\mathbb{Z} \times R\mathbb{Z}$, and its finite designated modes are of the form Rp with $p \in \mathcal{G}$. Set

$$s = 6/5, \qquad a = dR^{-s}, \qquad h = R^{-3/4}, \qquad \sigma = at.$$

Here $d > 0$ is a constant fixed by the initial tolerance after the integer family is known. The coefficient of a designated mode is measured in units of a. The other high modes are measured in units of ah. We indicate the time variable explicitly when passing between physical time t and slow time σ. For $n = (k, Rj)$ put

$$\|w\|^2_{X_R} = \sum_{k,j} \langle n/R\rangle^4 |\widehat{w}(k, Rj)|^2, \qquad \|w\|_{Y_R} = \sum_{k,j} \langle n/R\rangle^2 |\widehat{w}(k, Rj)|.$$

The same symbols on a horizontal sequence mean that only $j = 0$ is retained. These are the scaled Hilbert and Wiener norms, respectively. Throughout this section $O(\varepsilon a)$ denotes a physical-time operator error. We choose the spatial cutoff before fixing the integer coordinates of the family.

Denote the nonzero horizontal parent indices in the undilated integer family by P, and their prescribed terminal amplitudes by $a\beta_{P,R}$, where $\beta_{P,R} \to \beta_P \neq 0$. In formulas below β_P may denote the actual coefficient at R. The number of parents and their limiting amplitudes are fixed before the choice of integer coordinates. Constants after that choice may depend on the undilated configuration. The estimates below hold on the entire time line.

Projection of the original vector field. Freeze the mass at the terminal value M_* by writing $u = e^{-2iM_* t}v$. Let P_0 average in the second space variable, $U = P_0 v$, $V = (1 - P_0)v$. The identity

$$|U + V|^2(U + V) = |U|^2U + 2|U|^2V + U^2\bar{V} + 2U|V|^2 + \bar{U}V^2 + |V|^2V$$

and $P_0V = P_0\bar{V} = 0$ give the exact equations

$$iU_t = -U_{xx} + (|U|^2 - 2M_*)U + 2UP_0|V|^2 + \bar{U}P_0(V^2) + P_0(|V|^2V), \tag{3.1}$$
$$iV_t = (-\Delta + 2|U|^2 - 2M_*)V + U^2\bar{V} + (1 - P_0)(2U|V|^2 + \bar{U}V^2 + |V|^2V). \tag{3.2}$$

At $V = 0$ both first-order off-diagonal blocks between horizontal and transverse variables vanish. In the transverse linearization the conjugate potential couples row Rm to $-Rm$. If A denotes the old part of U and a terminal parent is $a\beta_P e^{iRPx-i\lambda_P^* t}$, the leading new coefficient in U^2 is $2aA\beta_P e^{iRPx-i\lambda_P^* t}$. Writing b for a Fourier index of A, this sends $\bar{v}_{q,-Rm}$ to mode (k, Rm) with $k + q = RP + b$. The interaction-picture phase has derivative $k^2 + q^2 + 2R^2m^2 - \lambda_P^*$. For $b = 0$ and a rescaled longitudinal index $x = k/R$, its kinetic part is $2R^2[m^2 - x(P - x)]$. This recovers both the anomalous-pair condition below and its coefficient directly from the NLS. For the normal coefficient the real multiplier $4\,\mathrm{Re}(\bar{A}a\beta_P e^{iRPx-i\lambda_P^* t})$ instead couples modes in the same transverse row.

For later use, a leading perturbation $a\sum_p E_p B_p$ consists of waves at physical lattice vectors p, with spatial-temporal phase $E_p = e^{ip\cdot(x,y)-i(|p|^2+d_p)t}$ and a small frequency correction d_p specified in Section 4. At $p + q = r$ with $p \cdot q = 0$ and $p, q \neq 0$, the two inputs are distinct. The square $\bar{A}(\sum E_p B_p)^2$ gives the parent coefficient $2\bar{m}B_p B_q$, where $m = \langle A\rangle_x$. The term $2A|\sum E_p B_p|^2$ gives each daughter coefficient, for instance $2mB_r\bar{B}_q$ at p. The common factor a^2 is removed by $\sigma = at$ and amplitude division by a. For a horizontal output this calculation is applied to the transverse forcing in (3.1), with the horizontal Birkhoff differential as estimated in Section 4. For a transverse output the horizontal inputs also enter through the variation of the two linear potentials in (3.2). Purely horizontal evolution is integrated by the one-dimensional chart.

We use mass-rotating coordinates with the parameter M_* fixed at the mass of the horizontal terminal. Let $U_*(t)$ be the exact horizontal solution. In physical variables the horizontal profile has the expansion

$$U_* = A(t, x) + a\sum_P \beta_P e^{iRPx-i\lambda_P^* t} + E_A, \qquad \|E_A\|_{Y_R} \le C(a/R + a^2). \tag{3.3}$$

Here A is the old finite-gap embedding evaluated at its actual old actions and angles, Y_R is the Fourier ℓ^1 norm with weight $\langle k/R\rangle^2$, and fixed spatial and temporal derivatives of A are bounded uniformly. Initial horizontal phases can be included in β_P. Writing $M_A = \langle |A|^2\rangle_x$, we have

$$M_* - M_A = O(a^2), \qquad \lambda_P^* = (RP)^2 + O(R^{-2} + a^2).$$

Indeed, the old torus high-frequency quasimode gives $\nu_{RP}(I_A, 0) = (RP)^2 + 2M_A + O(R^{-2})$, and changing the high actions of total size $O(a^2)$ changes the rotating frequency by $O(a^2)$ using the relative frequency map. The expansion in (3.3) follows from the one-smoothing inverse Birkhoff derivative at the compact old core and its symplectic transpose. Along the high-coordinate segment the h^1 norm of the added coordinates is $O(aR) = o(1)$, also on a fixed complex angular strip. Interpolation gives the differential gain $h^{-1} \to L^2$ for the nonlinear part, hence an $O(a/R)$ remainder in L^2_x. Spatial translation covariance fixes the spatial frequency of an angular Fourier coefficient as the scalar product of its integer angular multi-index and the active mode indices. Exponential angular decay on a smaller strip therefore upgrades the same estimate to Y_R, with a configuration-dependent constant. Only the spatial norm of E_A enters the variational equation.

The mean $m(t) = \langle A(t)\rangle_x$ has the exact decomposition

$$m(t) = c_R + \partial_t F_R(t), \qquad |c_R| \ge c_0 > 0, \qquad \|F_R\|_{L^\infty_t} \le C_{\mathrm{old}}. \tag{3.4}$$

Indeed the old angular frequencies in the rotating gauge are fixed exactly under extension and are Diophantine on the integral quotient by $\mathbb{Z}\alpha$, as in Lemma 2.2. The entire old finite-torus mean is smooth in these angles, uniformly for the slightly retuned old actions. Divide its temporal oscillatory characters by their divisors and sum the absolutely convergent series. The same argument bounds any fixed finite number of old angular derivatives of this primitive. Of the characters $l\alpha$, the gauge charge one leaves only $l = 1$. Thus $c_R = \widehat{m}_\alpha(I_A(R))e^{-i\alpha\cdot\theta^0}$, and the lower bound follows from Section 2.1. For each fixed invariant launch phase, c_R converges to its

preceding old-torus value. None of these bounds uses the Diophantine constant of the enlarged torus.

Lemma 3.2 (Unweighted transverse splitting)**.** *Assume the global arithmetic sieve of Section 2.2, with its cutoff for the old spatial Fourier series chosen as below. In L^2 on the transverse rows $j \neq 0$, the variational equation at U_* admits an exponential trichotomy in slow time, uniformly along the selected sequence of R. The same holds in the scaled norm X_R. There are two real stable and two real unstable directions for each independent pair of complex daughter modes. The other transverse directions have growth in both time directions bounded by $Ce^{\epsilon|\sigma-\sigma'|}$, where ϵ can be chosen smaller than a fixed fraction of the least target rate. The constants and gaps may depend on the finite construction. The assertion includes all additional resonant daughter pairs allowed by the sieve.*

Proof. On rows $\pm Rj$ double the unknown as the particle entries $v_{k,Rj}$ and hole entries $\overline{v_{q,-Rj}}$ (one may use the positive j to avoid repeating the real degrees of freedom). The free generator in the equation $i\partial_t w = Lw$ is diagonal with entries $k^2 + R^2 j^2$ and $-(q^2 + R^2 j^2)$. From the physical variational equation

$$iv_t = (-\Delta + 2|U_*|^2 - 2M_*)v + U_*^2 \overline{v}$$

and (3.3), the old normal potential is the doubled self-adjoint multiplication by $\pm(2|A|^2 - 2M_*)$. The old anomalous potential is the doubled pair coming from $A^2\overline{v}$. At first order in a the other terms are the real normal multiplication by $2(A\overline{H} + \overline{A}H)$ and the anomalous $2AH\overline{v}$, where $H = a\sum_P \beta_P e^{iRPx - i\lambda_P^* t}$. All omitted terms have operator norm $O(a/R + a^2)$ in L^2 and in X_R.

Removal of nonresonant entries. We begin with the changes of variable that leave bounded remainders on the whole line. A normal spatial coefficient of frequency b translates q to $k = q + b$. An anomalous one has $k + q = b$. Remove the old anomalous entries by a matrix S_0 obtained by dividing their coefficients, with homological sign, by the free difference across opposite branches $\pm(k^2 + q^2 + 2R^2 j^2)$. The Schur test gives $\|S_0\| + \|\partial_t S_0\| = O(R^{-2})$. Choose a fixed old cutoff H_0 so that the spatial tail of A in the unscaled weighted Wiener norm of order two is below an arbitrarily prescribed ε. Uniformity follows from the fixed smooth old family. The cutoff is imposed before the integer construction. For the retained high-parent coefficients, keep the rapid factor $e^{\mp i\lambda_P^* t}$ in the correcting matrix entry. The anomalous divisor is then, up to overall sign,

$$k^2 + q^2 + 2R^2 j^2 - \lambda_P^*, \qquad k + q = RP + b, \quad |b| \leq H_0.$$

By the sieve it is at least cR in absolute value except for $b = 0$ and the exact pairs

$$k = Rx, \quad q = R(P - x), \quad x(P - x) = j^2. \tag{3.5}$$

Thus the removing matrix for the nonexceptional entries has norm $O_{H_0}(a/R)$. Differentiating the slow old spatial coefficient costs the same order, since $\partial_t A$ is bounded in the required smooth norms. One can retain all old temporal harmonics this way. At fixed (x, j) with $j \neq 0$ the parent in (3.5) is unique, $P = x + j^2/x$. There are finitely many pairs in the undilated configuration and the two members of a pair are distinct by the sieve. For each ordered positive row, the two possible daughter orderings, when distinct, are retained.

Normal parent links incident to an exceptional index have divisors of order R^2, since the only leading degeneracies of a normal parent shift would have $x = 0$ or $x = Q$ for a parent Q, both excluded in Section 2.2. Remove these links together with their adjoints. When proving the weighted estimate, one may also remove the normal links outside a finite range of longitudinal frequencies. Set $L_0 = 1 + \max|P|$ and retain only links for which both longitudinal labels have modulus at most $3L_0 R$. We call their compression the normal block. Any discarded link has, for fixed $|b| \leq H_0$ and sufficiently large R, a normal divisor of magnitude $\gtrsim R^2$. Indeed

$k = q \pm (RP + b)$ gives a leading difference $\pm 2RPq$ (or the same expression from the other end). Outside this range neither end is close to 0 or to the relevant RQ. Each removed normal or anomalous matrix has bounded row and column sums by the same elementary Schur estimates. All claims here about operator tails are in physical-time units.

To justify these changes on the domain of the free generator, combine the time derivative of the rapidly oscillating coefficient with its spatial commutator. For $T = 1 + S$, free part D and desired bounded retained potential V_{ret}, expand

$$(D + V)T - iT' - T(D + V_{\mathrm{ret}}) = V - V_{\mathrm{ret}} + [D, S] - iS' + VS - SV_{\mathrm{ret}}.$$

For each fixed R the derivative of the correcting matrix is bounded (its oscillating scalar factors have finite, possibly large frequencies). The homological identity then also bounds $[D, S]$ for each fixed R. The Fourier core and graph closure give preservation of the free domain by S, and the bounded commutator identity gives domain preservation by $(1 + S)^{-1}$. Only the combined fast terms, evaluated by the displayed divisors, need a uniform estimate. The last two products are bounded by $C\|S\|$. Then multiply by the bounded inverse of T. The result is the old self-adjoint normal part, the remaining Hermitian normal parent links, and only the bare anomalous couplings (3.5), up to a bounded error $O_{H_0}(R^{-2} + a/R + a^2) + O(\varepsilon a)$.

The scaled Hilbert norm. We now obtain the same assertions in X_R with the cutoff chosen before the arithmetic sizes $|P|$ are known. First use the equivalent norm obtained by the row operator

$$W_j = 1 + j^2 + f\left(\frac{-\partial_x^2 + 2|A|^2}{(L_0 R)^2}\right).$$

Choose $f \geq 0$ smooth, equal to zero on $[0, 16]$ and to the identity on $[36, \infty)$. On the hole component take the conjugate operator. Its free version w_j uses $-\partial_x^2$ instead. The norm $\|W_j v\|_{\ell_j^2 L_x^2}$ is equivalent to X_R uniformly in R, with an equivalence constant allowed to depend on L_0. Furthermore

$$\|W_j - w_j\|_{2\to 2} + \|\partial_t W_j\|_{2\to 2} \leq C(L_0 R)^{-2}.$$

For $f - \mathrm{Id}$ extend its restriction from $[0, \infty)$ as a compactly supported smooth function on $\mathbb{R}$. The Fourier representation and the Duhamel identity for unitary groups bound the difference under a bounded operator perturbation by $\int |t\widehat{f - \mathrm{Id}}(t)|dt$ times its norm. The same argument gives the derivative. The linear part is explicit. All operator norms in this display are on L_x^2. The ratio of the free weights across any parent shift $RP + b$ is bounded by $C\langle b\rangle^2$ with an absolute C, independent of L_0, since $|P|/L_0 \leq 1$. This proves the previous Schur estimates in this norm and bounds the discarded old tail by $C\varepsilon a$ independently of the later magnitude of the undilated integer family. W_j commutes with the old self-adjoint normal operator on that row, and w_j is constant, with value $1 + j^2$, on both ends of every retained link of the normal block and every exceptional pair. Thus conjugation by W changes the retained non-old coefficients by $O_{H_0}(aR^{-2})$ up to the accounted tail. Applying the remainder of the proof in L^2 to the field multiplied by W gives the weighted assertion. No uniform equivalence constant independent of the eventual fixed configuration is required.

In particular the propagator U_0 of the bare old variational transverse operator alone (including its old anomalous term) satisfies

$$\|U_0(t, r)\|_{X_R \to X_R} \leq C_{\mathcal{G}} \exp(C_{\mathrm{old}} R^{-2}|t - r|).$$

Indeed remove only the old anomalous links with $I + S_0$, and apply W to the old normal equation with bounded $O(R^{-2})$ remainder. The self-adjoint part preserves the unweighted transformed norm. The two time-dependent changes and their inverses have bounded norms at each endpoint. Thus this bound applies in both time directions even without restricting to a moving complement. On any fixed slow window of length S its factor is $\exp(C_{\mathrm{old}} R^{-2} S/a)$,

bounded independently of R. This is the propagator required to estimate the arbitrary-error equation on the backward interval of Section 4.4.

Effective resonant pairs. We correct the exceptional Fourier waves for the old scalar normal equation. For each target $n = (k, Rj)$, $|k| \asymp R$, use

$$\phi_n = e^{i(kx+Rjy)-i(|n|^2+2M_A-2M_*)t}(1 + F_1/k + F_2/k^2).$$

Writing $V_0 = 2(|A|^2 - M_A)$, take the periodic zero-mean solutions of $2i(F_1)_x = V_0$ and $2i(F_2)_x = V_0F_1 - i(F_1)_t - (F_1)_{xx}$. The second right-hand side has mean zero because $\langle V_0F_1\rangle = i\langle\partial_x(F_1^2)\rangle = 0$. These waves have residual $O(R^{-2})$ in L^2 uniformly in time for the old normal equation, by direct substitution. Use their conjugates for the hole entries. The corrected column matrix differs from the Fourier columns by $O(1/R)$. Its Gram matrix is $I + O(R^{-2})$ because F_1 is imaginary and distinct centers on the same row differ by order R. If $iQ' = L_{\mathrm{old}}Q + O(R^{-2})$, self-adjointness gives $(Q^*Q)' = O(R^{-2})$. Orthogonalizing by $(Q^*Q)^{-1/2}$ preserves the residual. Thus the old operator couples the moving target to its complement with size $O(R^{-2})$ both ways, with unitary old normal evolution on the complement. This follows either by differentiating the norm of the moving orthogonal complement or by unitary transport. The retained Hermitian parent links have no bare target columns or rows and hence after dressing have cross leakage $O(a/R)$. Their complementary compression stays Hermitian. The exceptional anomalous links change by $O(a/R)$ upon dressing. This applies to the weighted conjugated equation as well. The conjugated evolution acts in a time-dependent norm uniformly equivalent to X_R.

Each exceptional channel on the moving target is now the real-linear Hamiltonian pair

$$i\dot z = 2a\beta_P m(t)e^{i\delta_P t}\overline{w}, \qquad i\dot w = 2a\beta_P m(t)e^{i\delta_P t}\overline{z}, \qquad |\delta_P| \le C(R^{-2} + a^2),$$

up to constant phases and the preceding global bounded remainder. In fact $\delta_P = (RP)^2 + 2(2M_A - 2M_*) - \lambda_P^*$. Every pair has only one parent, so independent rotations by half $\delta_P t$ remove the drifts at the cost of a bounded $O(R^{-2} + a^2)$ diagonal. By (3.4) the whole oscillatory part of m is removed by $I + O(C_{\mathrm{old}}a|\beta_{P,R}|)$ on each pair: in its time-generator equation use the bounded matrix primitive of the coefficient minus its constant part. Conjugation leaves the generator of the constant pair plus $O_{\mathrm{old}}(a^2 + R^{-2})$ in physical units. The constant generator is Hermitian after doubling $(z, \bar w)$, with eigenvalues $\pm 2a|\beta_{P,R}c_R|$. In the half-detuning frames the isolated pair projections therefore converge uniformly in time to those of the constant pair, up to the $O(a)$ bounded change of variables, whenever its remaining block error is $o(a)$. This applies in particular to the designated marked subsystem below. The remaining normal complement is unitary. For the direct sum choose two separating real shifts in slow time, one in each gap between zero and the positive or negative rates, and a small open interval about each. The remaining interblock error is bounded by $C\varepsilon + o_R(1)$ in slow units. Choose ε small relative to the least fixed rate and the direct-sum dichotomy constants, then apply the sieve with the corresponding cutoff and take R large. The elementary roughness estimate follows by a Neumann series for the shifted whole-line Green operator. A small tolerance for the central two-sided bound can be obtained by taking the analogous shifts closer to zero and choosing the error accordingly. This proves the lemma. □

A uniform dichotomy in X_R gives the spectral gap at the endpoint. The nonlinear orbit will use different amplitude units on the finite family and its complement. A second estimate controls the effect of this change of norm on the stable projection.

3.3. **The mixed norm and the stable projection.**

Lemma 3.3 (Splitting in the mixed norm)**.** *Use the normalized chart of Section 4, with amplitudes a on the designated marked modes, ah on other transverse and inactive horizontal modes, a^2 on the old actions, and a on the old angles. Its exact terminal linearization has*

a uniform slow-time exponential trichotomy with a strictly positive strong gap. This holds with a chosen strong-stable parameter on the marked pairs and zero independent strong-stable parameters on all additional resonant pairs.

Proof. At the horizontal terminal the exact zero-row and transverse variations are independent. The corrected old actions differ from the raw ones by a quadratic transverse form. The horizontal particular field has zero transverse derivative there. The transverse particular field can, and generally does, have a nonzero derivative at the nonzero terminal horizontal parents.

We prove the transverse splitting by estimating directly the rows and columns incident to the designated modes. A fixed error of size εa in the physical generator may be small in X_R, yet its contribution from a designated mode into the complement has size ε/h in the mixed slow-time equation. The stronger column estimate below removes this loss.

In the doubled variables let $\mathcal{M}$ contain every designated transverse mode, with both real degrees of freedom. Let $\mathcal{P}$ be its union with both ends of every exact anomalous pair in (3.5), including the additional resonant pairs. These are finite sets of rescaled labels (x, m), independent of R after the undilated integer family has been chosen. An exceptional pair meeting $\mathcal{M}$ belongs entirely to $\mathcal{M}$. Indeed its horizontal parent and one daughter are present, so that the other daughter is present by reverse closure. Every $(x, m) \in \mathcal{P}$ has $x \neq 0$ and $x \neq Q$ for every terminal parent Q. For additional resonant pairs this was proved by the sieve. For a designated mode with $x = Q$, the mode (Q, m) and the parent $(Q, 0)$ would require the additional vertical daughter $(0, m)$, contrary to the closure of the finite network.

Work first with the flattened weight W_j of Section 3, increasing L_0 to exceed the longitudinal size of all protected points as well as all horizontal parents. In the W_j metric the old spatial tail has size $C\varepsilon a$, with C independent of these sizes. Fix a sufficiently small $\kappa > 0$ after the undilated integer family. Modify the homological changes there as follows. In addition to the global $|b| \leq H_0$ elimination, eliminate every nonexceptional anomalous and every normal parent entry with a protected endpoint and $|b| < \kappa R$. Keep the unchanged conjugate and reversed entries in the same paired conventions as in the original elimination.

For example, with protected input $k = Rx$ the anomalous divisor for parent P is

$$2R^2\{m^2 - x(P - x)\} + 2R(P - x)b + b^2 + O(R^{-2} + a^2).$$

If the brace is nonzero its modulus is bounded below by cR^2 on this range. If the brace vanishes, $P - x \neq 0$, and every nonzero b has divisor of modulus at least $cR|b|$. The zero shift in this case is precisely the retained exception. Reversing the entry gives the same alternatives. For a normal shift the leading coefficient at $b = 0$ is, up to sign, $2Px$ or $2P(x - P)$. These are nonzero at protected endpoints, and the $O(R|b| + b^2)$ corrections leave a lower bound cR^2 after shrinking κ. Finiteness of $\mathcal{P}$ and the parent set gives a common κ, c. The added matrices and their slow-coefficient time derivatives cost at most $C_{\mathcal{G}}a/R$ by Schur's test. As before, the rapid exponential of the parent is differentiated together with the free commutator.

Even a bounded Fourier shift can destroy the required lower bound for the divisor. For instance, without its coprimality restriction one could take $P = 5$, $m = 2$, $b = 3$, $k = R + 4$. Then $q = 4R - 1$ and the kinetic divisor is identically 17. When this shift is within the old cutoff, the modulus used in Lemma 2.6 contains $1^2 + 2^2 = 5$, so this parent is excluded. In choosing the cutoff, the Green norm for the unperturbed orthogonal sum at a shift a distance g from all center and pair rates is at most $2/g$. The flattened weight ratio for any parent shift is at most $C\langle b\rangle^2$ with C absolute. Thus the smallness of the old Fourier tail can be imposed using only the old reference, the number and amplitudes of terminal parents, and the desired gap. The integer prototype, the protected-endpoint cutoff κ, and the constants in their $o_R(1)$ errors are chosen subsequently.

Let $P_{\mathcal{P}}$ denote bare orthogonal projection in the unweighted Fourier realization of the W_j metric, and put

$$\eta_R = R^{-2} + a/R + a^2.$$

After these changes the part of the old-plus-parent generator outside the retained old normal operator, Hermitian normal block, and bare exceptional pair matrix is a bounded error $\mathcal{E}$ satisfying

$$\|\mathcal{E}\| \le C_0\varepsilon a + C_{\mathcal{G}}\eta_R, \tag{3.6}$$

$$\|\mathcal{E}P_{\mathcal{P}}\| + \|P_{\mathcal{P}}\mathcal{E}\| \le C_{\mathcal{G}}\eta_R. \tag{3.7}$$

Here and below C_0 may depend on the old reference and the number and amplitudes of the intended horizontal parents, all fixed before the sieve. To verify the second bound directly, the discarded spatial tail incident to a protected index now has $|b| \ge \kappa R$. In the weighted Schur test its row and column sums are bounded by $Ca\sum_{|b|\ge\kappa R}\langle b\rangle^2|\widehat{A}_b| = O_{N,\mathcal{G}}(aR^{-N})$ for each fixed N. The residual from the old anomalous removal is $O(R^{-2})$ on the entire space. The slow derivative and bounded-coefficient products of the parent correctors are $O_{\mathcal{G}}(a/R)$. Multiplying the error by the inverse of a change $1 + O_{\mathcal{G}}(R^{-2} + a/R)$ preserves both estimates. Conjugating by W first with the free diagonal weight, and then using $Ww^{-1} = 1 + O(R^{-2})$, has the same property. Its time derivative is $O(R^{-2})$. The Hermitian normal block can be taken to have no entries at protected indices, since those entries were removed. The constants in (3.7) may depend on the fixed integer configuration.

Dress every member of $\mathcal{P}$ with the two same-sign correctors $1+F_1/k+F_2/k^2$ and orthonormalize the columns. Use the physical phase of the nonlinear chart on $\mathcal{M}$, half the parent detuning on additional resonant pairs, and zero detuning on other centers. Their difference from the old scalar mean contributes a bounded $O_{\mathcal{G}}(\eta_R)$ zeroth-order term. In these frames the retained pair's phase cancellation is exact. The orthonormal column map Q differs by $O_{\mathcal{G}}(R^{-1})$ from the phase-modulated bare columns. For the skew generator G consisting of the old normal operator and the retained normal block it satisfies $K = Q' - GQ = O_{\mathcal{G}}(\eta_R)$ in physical time: the retained normal block annihilates the bare protected waves, and acts by $O(a/R)$ on their dressing. The identity $Q^*Q = 1$ implies Q^*K is skew. The bounded skew operator

$$KQ^* - QK^* - Q(Q^*K)Q^*$$

maps Q to K. Adding it to G gives an exact unitary transport of the moving orthogonal complement and the prescribed column frame. Subtract it again as a bounded error of size $C_{\mathcal{G}}\eta_R$. Replacing bare anomalous entries by their dressed copies costs $O_{\mathcal{G}}(a/R)$. Replacing $P_{\mathcal{P}}$ by the dressed projection in (3.7) costs at most $O_{\mathcal{G}}(R^{-1})\|\mathcal{E}\|$.

Consequently, in common amplitude units, the protected-to-bulk and bulk-to-protected errors are $O_{\mathcal{G}}(\eta_R)$. The same is true of the error inside the protected block. The leading protected block consists of independent anomalous pairs and center directions. The change removing the oscillatory part of the entire old mean is $1 + O_{\mathrm{old}}(a)$ on each pair, and is internal to that pair. In slow time the pair rates are $\pm 2|\beta_{P,R}c_R|$. Give $\mathcal{M}$ amplitude unit one, and its entire remainder, including the additional resonant pairs, unit h. In slow time the marked-to-remainder and remainder-to-marked errors are at most, respectively,

$$C_{\mathcal{G}}\frac{\eta_R}{ah} \longrightarrow 0, \qquad C_{\mathcal{G}}\frac{h\eta_R}{a} \longrightarrow 0.$$

The diagonal remainder sector alone is a direct sum of additional resonant pairs and the transported bulk, plus a perturbation of size $C_0\varepsilon + o_R(1)$. Before the arithmetic choices the unperturbed shifted Green bounds here can be taken in terms of the gap alone: the bulk transport is unitary, the pair blocks are Hermitian, and the norm of their orthogonal direct sum

is the supremum of their norms. Choose ε smaller than the corresponding roughness threshold. The exact remainder sector therefore has its own uniformly bounded strong stable projection. The $o_R(1)$ off-diagonal errors then give a strong stable projection differing by $o_R(1)$ from the direct sum of that remainder projection and the constant model projection on $\mathcal{M}$, by the shifted Green operator estimate detailed below.

The same estimates for the opposite separating shift give the unstable projection because the bulk transport is unitary in both directions, the pair spectra are symmetric about zero and both scaled off-diagonal bounds tend to zero. The intermediate invariant space has the two-sided central bound obtained from the two separating shifts. We now compare these coordinates with the nonlinear chart differentiated at the terminal. Its transverse marked columns have the same prescribed bare phases, a WKB correction of size $O_{\mathcal{G}}(1/R)$ and a derivative of the quadratic particular of size $O_{\mathcal{G}}(a/R)$ in unit amplitude. Its complement is the kernel of their terminal reference dual as specified in Section 4. Both this complement and the dressed orthogonal remainder are graphs of maps of norm $O_{\mathcal{G}}(1/R)$ over the bare marked complement (in their respective equivalent Hilbert norms). Identify the generic spaces by the free weight w followed by projection along the marked columns. For each fixed configuration this gives a uniformly bounded isomorphism, since after projection onto the bare complement its composition with the reverse identification is $1 + O_{\mathcal{G}}(1/R)$. On a designated pair w is the same scalar $1 + m^2$ on both members and commutes with its model projector. Relative to this block-diagonal identification the two chart maps differ by at most $C_{\mathcal{G}}/(Rh) = o(1)$ in the mixed norm, since the physical normal forms and Ww^{-1} are smaller than $O_{\mathcal{G}}(1/R)$. This is a comparison of evolution operators by time-dependent isomorphisms, so no separate bound on the derivative of the rapid bare phases is needed. More explicitly, write $T(t)$ for the full map from the terminal nonlinear coordinates to these effective coordinates, and $J(t)$ for the block-diagonal identification just defined. Both maps and their inverses are uniformly bounded for a fixed configuration, and $\|T - J\| + \|T^{-1} - J^{-1}\| = O_{\mathcal{G}}((Rh)^{-1})$. With P_G the exact stable projection of the effective remainder, use the boundary projection

$$\Pi_R(t) = J(t)^{-1}(P_{\text{model}} \oplus P_G(t))J(t).$$

The true transverse projection in the nonlinear chart is $T^{-1}P_{\text{eff,full}}T$, hence differs from Π_R by $o_R(1)$. The marked restriction of J commutes with P_{model}. This identifies the zero generic boundary condition directly, even if the moving change of variables has a large derivative.

Green operators and exponential bounds. For the unperturbed direct sum the bounded inverse to the shifted inhomogeneous equation is given by integration forward on the stable sector and backward on its complement. If this inverse is G and the bounded perturbation is D, then $\|GD\| < 1$ gives the new inverse $(1 - GD)^{-1}G$. To recover exponential bounds and compare the associated projections, sample the shifted evolution at integer slow times and set $T_n = U(n, n-1)$. Uniform bounds on both directions on intervals of length one are already available. A bounded source sequence f_n can be realized on $(n-1, n)$ by the forcing $\chi(t - n + 1)U(t, n)f_n$, where χ is supported in $(0, 1)$ and has integral one. Thus $x_n - T_n x_{n-1} = f_n$ has a unique bounded solution with a uniform bound. For an impulse at m, conjugate the difference operator by $w_n = \exp(\delta \min\{|n - m|, N\})$. Its change in operator norm is at most $(e^{\delta} - 1) \sup_n \|T_n\|$. A Neumann series bounds the inverse uniformly in m, N for small $\delta > 0$. Taking $N \to \infty$ gives exponential localization of the impulse response. Its value at m projects onto the forward bounded data along the backward bounded data: the two pieces of the impulse solution exhibit the decomposition and uniqueness of a bounded whole-line homogeneous solution gives trivial intersection. The short-time bounds extend these dichotomies to real time. For a bounded slow-time perturbation of norm δ, Duhamel's formula and uniform unit-interval bounds give $\sup_n \|T_n - T_n^{(0)}\| = O(\delta)$ for the perturbed and unperturbed sampled maps. The resolvent

identity for the two difference operators gives an $O(\delta)$ change of their impulse responses in supremum norm. Their values at the impulse time are the corresponding projections. Translating the sampling grid gives this estimate uniformly at every real time. This argument applies in time-dependent closed subspaces as the scalar weights act in each individual fibre.

On the horizontal variational block the Birkhoff equation is exactly diagonal in inactive modes. The old actions have zero derivative and the normalized old angles have at most a constant shear from these actions and from the real amplitudes of the horizontal parents. In rotating Cartesian coordinates the small high-parent frequency shear has uniformly bounded coefficients (indeed $O(a)$ in their slow amplitude equation): this follows by the action-frequency derivative bounds and $O(a^2)$ physical action deviations. The horizontal variational evolution therefore has at most a uniform polynomial bound in slow time. More explicitly every action variation is constant in the linearized rotating coordinates and its contribution to each angle is linear in time. It joins the center. The common mass-dependent gauge in (1.2) changes its finite shear by the conserved linear mass variation, acting only in the horizontal block. The strong transverse spaces and their compatible estimates are unchanged. This completes the proof. □

For use at a fixed R in smaller Sobolev spaces, the L^2 and X_R splittings are compatible. The bounded shifted solution for a smooth compactly time-supported forcing in X_R is the same in L^2 by uniqueness. Approximating a time-point source by smooth pulses at fixed R identifies the two projection kernels on their common dense domain. Interpolating the resulting propagator bounds gives the physical H^q, H^s bounds for $0 \le q, s \le 2$ required at that fixed stage. No uniform equivalence between an unscaled physical H^s norm and X_R as $R \to \infty$ is asserted.

4. A transfer orbit for the nonlinear equation

Fix the reference and the frequencies selected in Sections 2.1 and 2.2. Write $M_T = M_*$ for the mass of the terminal horizontal solution. We first solve the equations (3.1) and (3.2) with the fixed scalar potential $-2M_T$. Strong convergence at the endpoint will identify the mass of the solution obtained this way and hence give a solution of (1.2).

We use $X = X_R$ and $Y = Y_R$ from Section 3. On horizontal sequences, h^2_R denotes the norm with Fourier weight $\langle k/R\rangle^2$. Weighted convolution on $\mathbb{Z}\times R\mathbb{Z}$ and $\sum_{k,j}(1+k^2/R^2+j^2)^{-2} \le CR$ give

$$\|vw\|_X \le C\|v\|_Y\|w\|_X, \qquad \|vw\|_X \le C\sqrt{R}\,\|v\|_X\|w\|_X. \tag{4.1}$$

For the second bound, use $\langle (n+m)/R\rangle^2 \le 2(\langle n/R\rangle^2 + \langle m/R\rangle^2)$ and $\|\widehat{v}\|_{\ell^1} \le C\sqrt{R}\|v\|_X$. The largest losses will occur in the complementary field and the old actions, which have different amplitude units. The following quantity bounds the losses in both equations.

$$\rho_R = \frac{R^{-2}}{ah} + \frac{1}{Rh} + \frac{a}{h} + \sqrt{R}\,h + \frac{h^2}{a} + \frac{R^{-2}}{a} = o(1). \tag{4.2}$$

For instance, $R^{-2}/(ah) = d^{-1}R^{-1/20}$ and $h^2/a = d^{-1}R^{-3/10}$. In this section C_M denotes a constant on the bounded set in question. We write Poly_M for a polynomial majorant in the displayed nonnegative arguments whose coefficients may depend on a bound M for the designated amplitudes. All small factors refer to the last dilation at this fixed stage.

4.1. **Wave profiles and quadratic corrections.** Write u temporarily for the field in the fixed-mass gauge. Let $U = P_0u$ be its horizontal row and $V = (1-P_0)u$ its transverse part. Here P_0 is spatial averaging in the transverse variable only. Split $\Omega(U)$ into the old positive actions and angles $y = (I,\theta)$, the finite horizontal marks, and the remaining horizontal coordinates. Here and below A in a subscript denotes the finite set of old labels. Write $U_A(y) = \Omega^{-1}(z_A, 0)$ and $M_A = \|U_A\|^2_{L^2}$. The old parameters of the exact terminal are denoted by $y_*(t)$. The kinetic frequency of $p = (K,L)$ is $\lambda_p = K^2 + L^2$. Each physical marked vector is R times a fixed vector.

For a horizontal mark $p=(RP,0)$ set $d_p=\lambda_P^*-(RP)^2$. On the two members of a resonant transverse pair use half its parent detuning on each member. Set $d_p=0$ on all other transverse marks. The resulting frequencies are fixed by the terminal trajectory and $|d_p|\le C(R^{-2}+a^2)$. Write x_2 for the second spatial variable to distinguish it from the old action-angle coordinates y. Set $E_p=\exp(i(Kx+Lx_2)-i(\lambda_p+d_p)t)$.

The real-linear map representing one transverse wave is

$$Q_pz=E_pA_p^{\mathrm{w}}z+\overline{E_p}C_p^{\mathrm{w}}\bar z,\qquad A_p^{\mathrm{w}}=1+F_1/K+F_2/K^2,$$
$$C_p^{\mathrm{w}}=-\frac{U_A^2}{2\lambda_p}+\frac{iK(U_A^2)_x}{2\lambda_p^2}-\frac{U_A^2\bar F_1}{2\lambda_pK}.$$

With $W=2(|U_A|^2-M_A)$ and $\mathcal{D}$ the unforced rotating old angle derivative at y, use the zero-mean solutions of

$$2i(F_1)_x=W,\qquad 2i(F_2)_x=WF_1-(i\mathcal{D}+\partial_x^2)F_1.$$

The mean in the second equation vanishes since $WF_1=i\partial_x(F_1^2)$. Substitution in $i\partial_t+\Delta-2|U_A|^2+2M_T-U_A^2\mathcal{C}$, with $\mathcal{C}$ complex conjugation, leaves a defect in Y bounded by

$$C\big(R^{-2}+|M_A-M_T|+|d_p|+R^{-1}|\dot y-Y_{\mathrm{fr}}(y)|\big)|z| \tag{4.3}$$

at fixed z. Here $Y_{\mathrm{fr}}(I,\theta)=(0,(\nu_i(I,0)-2M_T)_{i\in A})$ is the old action-angle velocity with all other coordinates set to zero, and $\mathcal{D}=Y_{\mathrm{fr}}\cdot\partial_y$. Indeed $2iK\partial_x$ eliminates the two successive same-sign errors. In the other column the principal operator is $-2\lambda_p-2iK\partial_x$. The three displayed terms in C_p^{w} cancel the order-one source, its transport error and the contribution of $\bar F_1/K$. All other coefficients are bounded smooth old functions. In particular $\|\partial_yQ_p\|_Y=O(1/R)$ after removing the bare plane.

We specify all quadratic particular fields, including their entries in the finite marks. At each fixed R write $b_*=b_{*,R}$ for the actual terminal marked array, zero on transverse marks, and $B=b_*+b$ for the total amplitude. The small action choices of Lemma 2.2 give $b_{*,R}\to b_{*,\infty}$, the prescribed model array. Every purely horizontal array is an equilibrium of the quadratic finite equations. Signs and factors in the following source rules are always those of the expanded cubic polynomial. Thus a transverse equation uses $\bar U_Av^2+2U_Av\bar v$ on the bare marked waves, retaining terms with a transverse output. The corresponding horizontal physical forcing uses only transverse inputs. Pure horizontal dynamics is already included in the Birkhoff frequencies. In a source from $p+q$ or $p-q$, put

$$r=p\pm q=(K,L),\qquad D=\lambda_p\pm\lambda_q-\lambda_r.$$

After stripping $a^2e^{i(Kx+Lx_2)-i(\lambda_r+D)t}$ the source coefficient $H(t,x;y,B)$ is a smooth old function times a quadratic amplitude and the prescribed small detuning phases. All its frozen physical time derivatives used below are bounded. In particular this assertion does not differentiate an unknown b'. Such derivatives are accounted for in the exact extraction identity below. We use the convention $i\dot w-\mathcal{L}w=$ source when defining the following particular fields:

(1) For $L\neq0$ and $D\neq0$, insert the stripped coefficient H/D with the same output exponential. Since $|D|\gtrsim R^2$, the transport $2iKH_x/D$ and all other residuals are $O_M(1/R)$, before multiplication by a^2. This includes $K=0$.

(2) For $L\neq0$ and $D=0$, one has $K\neq0$ by the closure assumption and the exclusion of vertical marks. Retain the spatial mean of H in the marked equation. For the rest insert the stripped particular coefficient $(2iK)^{-1}\partial_x^{-1}(H-\langle H\rangle_x)$. Its defect outside the desired source is $O_M(1/R)$. The old conjugating potential on either of these particular fields costs at most their size.

(3) For $L = 0$ and $K \neq 0$, define $\mathcal{T}_y f = iD\Omega(U_A)[-if]$. On $e^{iKx}H$ the difference between $\mathcal{T}_y$ and ordinary Fourier transformation is $O_M(R^{-1})$ in the scaled sequence norm, also with an arbitrary scalar phase of modulus one. To see this, symplecticity identifies the real adjoint of the differential of a Birkhoff map with the inverse differential, up to the fixed complex structures. The one-smoothing estimates and duality give the correction bounds $H^{-2} \to h^{-1}$ and $H^1 \to h^2$, hence $H^{-1} \to \ell^2$ by interpolation. Also $R^{-2}\|e^{iKx}H\|_{H^1} = O_M(R^{-1})$, and $\|e^{iKx}H\|_{H^{-1}} = O_M(R^{-1})$ by the rapid old spatial tail. These prove the assertion.

Writing H_j for the spatial Fourier coefficients, insert in coordinate $n = K + j$ the following full physical particular, restricted to the horizontal indices outside the old active set:

$$a^2 e^{-i(K^2+D)t} \begin{cases} H_j/D, & D \neq 0, \\ -H_j/(2Kj), & D = 0,\ j \neq 0, \\ 0, & D = 0,\ j = 0. \end{cases}$$

In the last case the retained H_0 belongs to the marked equation. In the fixed-mass gauge the terminal diagonal is $\nu_n^* = \nu_n(I_*) - 2M_T = n^2 + O(1)$ uniformly in n, where I_* is the full terminal action sequence. For $D \neq 0$ the error from replacing the divisor by D is bounded by $C(R|j| + j^2 + 1)/R^2$ times $|H_j|$. For $D = 0$, $j \neq 0$ the relative error of the divisor is bounded by $C(|j| + 1)/R$. This is valid even near $j = -2K$. No division by $K^2 - (K + j)^2$ is made. Smooth summability and one frozen derivative give a total defect $O_M(a^2/R)$.

(4) For $L = K = 0$, $D \neq 0$, apply the real-linear $\mathcal{T}_y$ to the actual smooth old forcing. There are two sequences S_n^+, S_n^- such that

$$\mathcal{T}_y(e^{-iDt}H) = e^{-iDt}S^+ + e^{iDt}S^-.$$

They and their frozen derivatives are uniformly smooth in unscaled sequence norms. Insert, for every horizontal index n outside the old active set, the entry

$$a^2 \sum_{\epsilon=\pm 1} \frac{e^{-i\epsilon Dt} S_n^\epsilon}{\epsilon D}.$$

The residual is $O_M(a^2/R^2)$ since $n^2 S_n^\epsilon$ is bounded in the required norm. The omitted old active entries are treated below at the level of their *actual* action and angle equations.

(5) For $L = K = D = 0$ the only normal source is the diagonal density. Writing $Q = \int |V|^2$, its scalar contribution $2QU$ is exactly $2Qz$ in the horizontal Birkhoff equation, by gauge covariance. The difference between this density and its bare marked counterpart in the remaining equation has the product estimates below. Two nonzero opposite waves in the conjugating product have $D = \lambda_p + \lambda_{-p} > 0$ and belong to the preceding case.

Let $\Psi(t, y, B)$ be the sum of these fields. Their bounded coefficient and first amplitude derivative have Y size $O_M(a^2/R)$ in the transverse rows and the analogous scaled sequence bound in the horizontal rows. One can use a scaled Wiener bound for the latter too, since their entries have the smooth summable envelopes displayed above. The fields vanish on a purely horizontal array. To specify the finite vector field $F_0(t, y, b)$, put $m = \langle U_A(y)\rangle_x$ and $B = b_* + b$. Sum over the unordered orthogonal pairs p, q in the family, with $r = p + q$. If $\delta = d_r - d_p - d_q$, the contributions to iF_0 in slots r, p, q are respectively

$$2\bar{m}B_pB_qe^{i\delta t}, \qquad 2mB_r\bar{B}_qe^{-i\delta t}, \qquad 2mB_r\bar{B}_pe^{-i\delta t}.$$

Both orders of p, q occur in the physical square. This definition retains the full temporal dependence of the spatial mean. Thus no divisor involving an old temporal harmonic occurs in Ψ.

4.2. **Equations for the amplitudes and the remainder.** The corrected waves depend on the moving horizontal background. We extract their amplitudes before estimating any error, so that the derivatives of the moving coordinates appear in the evolution equation.

In addition to y, use as raw state x the remaining horizontal Birkhoff slots and the whole transverse field, measured relative to the exact terminal state in these slots. Write $\dot{x} = \mathcal{F}(t, y, x)$ and $\dot{y} = \mathcal{Y}(t, y, x)$ in physical time. In particular $\mathcal{F}(t, y_*, 0) = 0$. Write b_H and $b_\perp$ for the horizontal and transverse parts of the marked displacement b. Let m_0 be the number of designated complex coordinates, horizontal and transverse. Let $\Gamma(t, y)$ consist of the rotating reference coordinate vectors for the horizontal marks and the columns Q_p for the transverse ones. Let $L_*(t)$ be the terminal left inverse, given by the same rotating coordinate tests horizontally and the real symplectic dual to the terminal columns transversely. More explicitly, if J represents the symplectic form and Q is the real matrix of those columns, this dual is $(Q^{\mathsf{T}}JQ)^{-1}Q^{\mathsf{T}}J$. Its Gram matrix is an invertible perturbation of that of the distinct bare modes. Write $L_* = L_*(t)$. At each time we take the complement $\ker L_*$ and extract the normalized variables from

$$x = G(t, y, b) + ahe, \qquad L_* e = 0, \qquad G = a\Gamma(t, y)b + \Psi(t, y, b_* + b).$$

Here $G(t, y, 0) = 0$. The independent additional resonant stable parameters are represented in this complement after projection. Bare additional resonant Fourier vectors need not lie exactly in the kernel of the dressed dual. Put $D_b = G_b/a$, $H_b = L_* D_b$, $\Pi_b = 1 - D_b H_b^{-1} L_*$. On fixed bounded normalized sets

$$H_b = 1 + O_M((|y - y_*| + a)/R).$$

Consequently the displayed equations determine b, e uniquely locally with uniformly bounded normalized inverse.

Let $\mathcal{A}_0$ be the terminal-old transverse Hamiltonian generator and the exact terminal horizontal Birkhoff diagonal. Then

$$T_0 = (\dot{L}_* + L_* \mathcal{A}_0)|_{\ker L_*}, \qquad \|T_0\|_{X \to \mathbb{R}^{2m_0}} \le C(R^{-2} + a^2). \tag{4.4}$$

The horizontal part is zero. In the transverse part the derivative of the Gram inverse vanishes on $\ker L_*$, and the remaining expression tests the error against $\dot{Q} - \mathcal{A}_0 Q$ by the Hamiltonian identity $\mathcal{A}_0^{\mathsf{T}} J + J\mathcal{A}_0 = 0$. Column-space terms give zero and (4.3) controls the rest. We take the fixed terminal old angular velocities here. Their mismatch from the free truncated old velocities is $O(a^2)$.

Lemma 4.1 (Evolution in the moving coordinates)**.** *On smooth states set*

$$K = \mathcal{F}(t, y, G + ahe) - G_t - G_y \mathcal{Y}(t, y, G + ahe) - ah\mathcal{A}_0 e - a^2 D_b F_0.$$

Primes below refer to $\sigma = at$. The exact equations are

$$b' = F_0 + H_b^{-1}(a^{-2} L_* K + (h/a) T_0 e), \tag{4.5}$$

$$e' = a^{-1}(\mathcal{A}_0 - D_b H_b^{-1} T_0) e + (a^2 h)^{-1} \Pi_b K. \tag{4.6}$$

Proof. Differentiate $x = G + ahe$ and use $L_* e' = -\dot{L}_* e/a$. Taking L_* gives (4.5). Substitution in the original differentiated identity gives (4.6). In particular the identity uses only $L_* D_b = H_b$. The inhomogeneous term in the complementary equation belongs to the required moving fibre: $L_* \Pi_b = 0$ and $L_* D_b H_b^{-1} = 1$. Subtracting the linearization at $b = 0$ therefore leaves a forcing in $\ker L_*$ even in the term containing $T_0 e$. □

Bounds for the residual. We first evaluate K when the complementary field vanishes, and then estimate its change when that field is restored. The latter subtraction cancels the unbounded

generator exactly. Normalize old coordinates temporarily by $I - I_* = a^2 J_{\mathrm{raw}}$, $\theta - \theta_* = a\Theta$. On a bounded normalized ball the raw physical velocity mismatch is

$$|\mathcal{Y} - Y_{\mathrm{fr}}| \le C_M a^2. \tag{4.7}$$

Indeed the unforced old frequency change from all high actions is $O_M(a^2)$ and the transverse forcing of the old finite Birkhoff tests is quadratic. With U_A fixed, write the true horizontal field as its value with arbitrary horizontal error removed, plus an X error of size $O_M(ah\|e\|_X)$. For the former value one has

$$U = U_A + a \sum_{P \text{ horizontal}} B_P E_P + O_Y(a/R) + O_X(a^2/R + ah\|e\|_X). \tag{4.8}$$

The same estimates hold for bounded old parameter derivatives. To justify the Wiener remainder for the finite parent profiles, integrate the $h^{-1} \to L^2$ bound for $D\Omega^{-1} - \mathcal{F}^{-1}_{\mathrm{Fourier}}$ along those finitely many high coordinates. They have h^1 size $O(aR) = o(1)$ and h^{-1} size $O(a/R)$. This estimate holds on a fixed complex strip of the old angles and of the finitely many normalized new phase parameters. Translation covariance fixes the spatial index of each angular Fourier coefficient as its momentum. Exponential angular decay therefore gives the same $O(a/R)$ bound in Y, with all fixed normalized amplitude derivatives, including at zero amplitude by analytic continuation. The arbitrary error bound follows by the uniform scaled chart estimate: differentiating translation covariance twice gives X_R bounds for each prescribed derivative of $\Omega^{\pm 1}$ on a bounded scaled neighborhood of the old torus. For instance each spatial first derivative applied to an input costs at most $\|w\|_{L^2}^{1/2}\|w\|_{H^2}^{1/2}$, and two derivatives are accompanied by R^{-2}. The other factors are controlled by the local L^2 analytic derivatives. More explicitly, if G is a translation-equivariant map and $G_k = D^k G(U)[v_1, \dots, v_k]$, differentiate the identity obtained by translating U and every v_j. The first scaled derivative is the sum obtained by applying $R^{-1}\partial_x$ to one input v_j, or by adjoining $R^{-1}U_x$ to $D^{k+1}G$. The second derivative repeats this rule. There is either one scaled second derivative or two scaled first derivatives. Both are bounded in L^2 by the input scaled H^2 norms, using interpolation for each first derivative. These identities apply first to smooth inputs, then by density. On a sufficiently small L^2 tube around the compact old torus the multilinear L^2 bounds are uniform. Applying the resulting bound to D^2G along the segment between two points yields the Lipschitz estimate on bounded scaled balls. This also gives $\|D\Omega(U) - D\Omega(U_A)\|_{X \to h^2_R} \le C_M a$. The horizontal particular fields contribute $O_M(a^2/R)$ to the arbitrary-error-free estimate, either by the same translation argument on their smooth envelopes or as X errors in products with marked multipliers. The latter suffices throughout.

Set $K_p = K|_{e=0}$ at the *same* raw y, b. At bounded b the explicit pure-source calculation above yields

$$\|K_p\|_X \le C_M\{aR^{-2} + a^2/R + a^3(1 + |J_{\mathrm{raw}}|)\}. \tag{4.9}$$

We verify this before applying either extraction projection. Write $V_H = a\sum_{\mathrm{hor}} E_p B_p$, $V_b = a\sum_{\mathrm{tr}} E_p B_p$, $Q_b = \int |V_b|^2$ and $\mathcal{T}_U = iD\Omega(U)(-i\,\cdot)$. All errors in the next two displays are in X or its horizontal sequence counterpart. At $e = 0$ the transverse polynomial remaining after removing its old real-linear part is

$$\begin{aligned}(1 - P_0)\big[2(|U|^2 - |U_A|^2)V + (U^2 - U_A^2)\bar V + 2U|V|^2 + \bar U V^2 + |V|^2 V\big] \\ = (1 - P_0)\big[2(U_A\bar V_H + \bar U_A V_H)V_b + 2U_A V_H \bar V_b + 2U_A|V_b|^2 + \bar U_A V_b^2\big] + O_M(a^2/R + a^3).\end{aligned}$$

For example, replacing V by V_b in a quadratic term places $V - V_b = O_Y(a/R + a^2/R)$ beside a marked factor of size $O_Y(a)$. Replacing $U - U_A$ by V_H in a linear multiplier of V costs $O_X(a/R + a^2/R)$ beside $V = O_Y(a)$. A product of two high horizontal factors beside V has size $O_M(a^3)$, including the small particular part of (4.8). Indeed an $O_X(a^2/R)$ factor paired with

another such factor costs only $\sqrt{R}$ in the product. The three remaining monomials with three bare high factors also cost $O_M(a^3)$.

In the horizontal coordinates separate the exact scalar gauge first. The forcing outside the old slots is the restriction of

$$\begin{aligned} &2Qz + \mathcal{T}_U\Big[2U(P_0|V|^2 - Q) + \bar{U}P_0(V^2) + P_0(|V|^2V)\Big] \\ &\qquad = 2Qz + \mathcal{T}_y\Big[2U_A(P_0|V_b|^2 - Q_b) + \bar{U}_A P_0(V_b^2)\Big] + O_M(a^2/R + a^3), \end{aligned}$$

where $z = \Omega(U)$ and $\mathcal{T}_y = \mathcal{T}_{U_A}$. Here $|Q - Q_b| = O_M(a^2/R)$, the chart differential changes by $O_M(a)$ on a forcing of size $O_M(a^2)$, and the factor $2Qz$ in the non-old slots costs $O_M(a^3)$. These exact polynomials show which monomials remain after replacing the physical fields by bare ones. In the transverse display they are precisely the square and density interactions of $V_H + V_b$ with nonzero output row. The horizontal display has only the two transverse-input interactions already treated by the longitudinal primitives. The shifted frequency of each exact horizontal terminal slot acts on its prescribed terminal amplitude as well as on its displacement. Their combined amplitude is $O_M(a)$.

The uncancelled column defect is (4.3) times a. Let Γ_{bare} denote the same column map with the transverse profiles replaced by E_p. The retained quadratic source is $a^2\Gamma_{\mathrm{bare}}F_0$, so replacing the bare columns by Γ costs $O_M(a^2/R)$. The spatial and kinetic primitives have just this total residual. The differentiation of G uses its actual Jacobian and contributes the further term $-a\Psi_b F_0$, of size $O_M(a^3/R)$. The difference between frozen and actual old velocities costs $O_M(a^3/R)$ by (4.7) and $\|G_y\| \le C_M a/R$. The change of a horizontal diagonal frequency from its reference value and the old scalar mass mismatch acting on marked waves cost $O_M(a^3(1 + |J_{\mathrm{raw}}|))$. The remaining cubic expressions on finite smooth marked profiles cost $O_M(a^3)$ by (4.1). In particular $K_p(t, y_*, 0) = 0$ exactly. Its outgoing marked linear column in (4.6) is $\Pi_0(\partial_b K_p)_*/(a^2h)$ since $(\partial_b\Pi_b)K_p = 0$ there. The derivative of Ψ itself generally *does not* vanish at the terminal parent. Differentiating the term above gives $-a(\Psi_b)_*(\partial_b F_0)_*$, whereas the term with Ψ_{bb} vanishes because $F_0|_* = 0$. The surviving term has normalized size $O_M(a/(Rh))$.

The arbitrary-error difference has the exact factorization

$$\begin{aligned} K - K_p = {}& \mathcal{F}(t, y, G + ahe) - \mathcal{F}(t, y, G) - ah\mathcal{A}_0 e \\ & - G_y(\mathcal{Y}(t, y, G + ahe) - \mathcal{Y}(t, y, G)). \end{aligned} \tag{4.10}$$

The free unbounded generator has cancelled before estimation. In the first line the change of $2|U_A|^2$ or of U_A^2 from its terminal value costs $O(a|\Theta| + a^2|J_{\mathrm{raw}}|)$ as a multiplier of ahe. An $O(a)$ terminal parent multiplier is part of the exact linearization. With a single arbitrary input, a bare marked factor of size $O_M(a)$ is put in Y. In particular in the product of an arbitrary horizontal error and a transverse marked wave it is the *wave* that is used as multiplier. Two arbitrary inputs cost $C_M a^2\sqrt{R}\, h^2\|e\|_X^2$. For the horizontal equation the forcing before $D\Omega$ is

$$2U(P_0|V|^2 - Q) + \bar{U}P_0(V^2) + P_0(|V|^2V)$$

in addition to the gauge term $2QU$. One arbitrary transverse input costs $C_M a^2 h\|e\|_X$, and two cost $C_M a^2\sqrt{R}\, h^2\|e\|_X^2$. All cubic terms obey the same bounds with an extra marked amplitude or a factor bounded by $a(1 + \sqrt{R}h\|e\|_X)$. The horizontal diagonal difference is $O_M(a^2(1 + |J_{\mathrm{raw}}| + h^2\|e\|_X^2))$ uniformly in its index. There is no loss of derivatives in this term. Testing the same expressions on finite old coordinates bounds the velocity difference in the second line of (4.10) by $C_M a^2 h(1 + \sqrt{R}h\|e\|_X)\|e\|_X$ on bounded sets. It follows that $(K - K_p)/(a^2h)$ is uniformly bounded on the normalized balls and vanishes at $e = 0$. On a finite interval where only the marks have been bounded a priori its upper bound has the form

$$C_M(1 + |\Theta|)\|e\|_X + o_R(1)\,\mathrm{Poly}_M(|J_{\mathrm{raw}}|, |\Theta|, \|e\|_X)$$

on every fixed larger ball, with a fixed-degree local majorant. In (4.5) all these incoming-error coefficients acquire an additional h.

4.3. **The old action equations.** The independent old action and angle equations have different required physical sizes. The angle velocity can change by $O(a^2)$. The required bound for the corrected action velocity is $O(a^3)$.

Lemma 4.2 (Quadratic correction of the old actions)**.** *There are real quadratic forms $B_i(t, y; V)$ depending smoothly on the old parameters, with $|B_i| \le CR^{-1}\|V\|_{L^2}^2$, such that*

$$J_i = (I_i - I_{*,i} + B_i)/a^2$$

gives a uniformly invertible local change of normalized coordinates. At the horizontal terminal J_i' has zero value and zero first derivative. On any bounded normalized ball its second derivative is uniformly bounded, and on a fixed finite arc with bounded marks one has

$$|J'| \le C_M + o_R(1)\,\mathrm{Poly}_M(|J|, |\Theta|, \|e\|_X), \tag{4.11}$$

$$|\Theta'| \le C_M(1 + |J|) + o_R(1)\,\mathrm{Poly}_M(|J|, |\Theta|, \|e\|_X). \tag{4.12}$$

Proof. At fixed physical V and fixed other horizontal Birkhoff coordinates, the contribution to $\dot I_i$ is minus the θ_i derivative of the transverse Hamiltonian. Its quadratic and cubic parts which depend on the horizontal field are

$$\int \left(2|U|^2|V|^2 + \mathrm{Re}(\bar U^2 V^2) + 2\,\mathrm{Re}(\bar U V)|V|^2\right).$$

The horizontal Hamiltonian is action dependent and makes no contribution. This identity is exact even though the auxiliary extraction G depends on the instantaneous old parameters: the Hamilton derivative is first taken at *fixed physical* V. Replace U by U_A in the quadratic term. By (4.8), including its old angular derivative, the physical error in the resulting action equation is bounded by

$$C_M a(1 + \sqrt{R}\,h\|e\|_X)\|V\|_{L^2}^2.$$

The cubic derivative is bounded by $C_M(1 + o(1)\|e\|_X)\|V\|_\infty\|V\|_{L^2}^2$ on bounded balls, and $\|V\|_\infty \le C_M a(|b_\perp| + \sqrt{R}h\|e\|_X)$.

In the normal quadratic term write $f = -2\partial_{\theta_i}|U_A|^2$. Its spatial mean is zero since $\int |U_A|^2 = M_A$ is angle independent. In the row of transverse frequency ℓ define a correction by the Hermitian matrix

$$T_{kq} = -\frac{f_{k-q}\chi((k+q)/(2R))}{i(k^2 - q^2)}\mathbf{1}_{0<|k-q|<cR}.$$

Set the entry to zero whenever the cutoff vanishes, before taking the quotient. Here χ is a real smooth cutoff supported where the longitudinal velocity is separated from zero, equal to one on enlarged neighborhoods of all marked longitudinal velocities, and $c > 0$ is chosen much smaller than those neighborhood sizes. The correction added to I_i is $\sum_\ell \langle V_\ell, TV_\ell\rangle$ with the convention $\langle v, Tv\rangle = \sum_{k,q}\bar v_k T_{kq} v_q$. Its free derivative produces $i(k^2 - q^2)T_{kq}$. For $j = k - q \neq 0$ the midpoint symbol is a function of k with derivative bounded by $CR^{-2}|f_j|/|j|$, and with size $CR^{-1}|f_j|/|j|$. Indeed the denominator is $ij(2k - j)$ and the cutoff is supported at $|2k - j| \gtrsim R$. The commutator of T with any fixed smooth normal old potential is $O(R^{-2})$ on ℓ^2: expand the potential in shifts $v \mapsto v_{k-m}$, commute it past the j-shift, and sum the symbol differences bounded by $CR^{-2}|m|\,|f_j|/|j|$. The frozen time derivative of T is $O(R^{-1})$ and has no diagonal. Remove it with a second division by $i(k^2 - q^2)$ on the very same support. Its frozen derivative and all commutators now cost $O(R^{-2})$. The difference between actual and frozen old velocities costs $O_M(a^2/R)$ by (4.7).

It remains to eliminate the anomalous quadratic form, including the anomalous form produced by differentiating the first correction in the old conjugating potential. For a current term $\mathrm{Re}\sum_{k,q,\ell} g_{kq\ell}V_{k,\ell}V_{q,-\ell}$ add the form with coefficients

$$\frac{g_{kq\ell}}{i(k^2+q^2+2\ell^2)}.$$

Its free derivative cancels the displayed term. The denominator is at least $2R^2$ on these rows, and the old coefficients, their first frozen time derivatives and their products with the preceding smooth Fourier matrices have uniform row and column Schur bounds. This correction, its frozen derivative and its commutators with bounded old potentials are consequently $O(R^{-2})$. This argument handles the old entries excluded in the fourth particular-field case above: the density output $p-q=0$ implies $p=q$, $D=0$. The remaining fast zero output is precisely the opposite-wave density V_pV_{-p} in the anomalous term. Its raw old-action projection can have physical size a^2. Canonicality identifies that projection with the actual θ_i derivative of $\mathrm{Re}\int \bar{U}_A^2V^2$, and the displayed sum-frequency primitive eliminates it. Its raw angular projection of size a^2 already has the allowed scale.

For the remaining normal quadratic expectation the symbol $1-\chi((k+q)/(2R))$ can meet no marked longitudinal core for $|k-q|<cR$. Increase the marked neighborhoods before fixing c. The old Fourier tail for $|k-q|\geq cR$ decays to arbitrary order. The enlarged core also contains all but a rapidly small tail of the same-sign smooth WKB envelopes. Opposite-sign column errors are $O_M(a/R^2)$, and all particular fields are $O_M(a^2/R)$. Thus the uncancelled normal expectation is bounded by

$$CR^{-2}\|V\|_{L^2}^2 + C_M(ah\|e\|_X + a/R^2 + a^2/R)^2 + O_{M,N}(a^2R^{-N}).$$

For its local quadratic behavior, the last three pure marked terms carry at least a factor $|b_\perp|$ before squaring (and the rapid-tail term carries $|b_\perp|^2$). The perturbation of the old linear transverse operator and the nonlinear transverse equation acting on the correction are bounded on this chart by $C_Ma(1+\sqrt{R}h\|e\|_X)^2\|V\|_{L^2}$ in L^2. They incur an extra factor $O(R^{-1})\|V\|_{L^2}$. Dividing all these physical estimates by a^3 gives (4.11), using $h^2/a=o(1)$ and $R^{-2}/a=o(1)$. Every surviving term in the old action derivative has at least two transverse factors, hence has zero value and linearization at the horizontal terminal. The normalized coordinate change is $O_M(R^{-1})$ with bounded first derivatives relative to identity, and is invertible uniformly on smaller balls.

For Θ' the physical angular perturbation from the old Hamiltonian frequencies is bounded by the action difference

$$Ca^2\big(|J|+|b_H|+|b_H|^2+h^2\|e\|_X^2 + O_M(R^{-1})(|b_\perp|+h\|e\|_X)^2\big).$$

This uses the uniform action Lipschitz estimate, Cartesian coordinates at high horizontal marks, and the just obtained change in old actions. The transverse force projected onto an old positive angular coordinate costs $O_M(a^2)$ on bounded balls, and the parts involving arbitrary e acquire a factor h or $\sqrt{R}h^2$. This proves (4.12). The local differentiability assertions also follow from the uniform analytic bounds detailed next. □

We can now combine the field and action estimates. Give $Z=(b,e,J,\Theta)$ the sum of the Euclidean norms on its finite components and the X norm on e.

Proposition 4.3 (Uniform nonlinear estimate)**.** *In the above normalized coordinates the terminal is the zero displacement. After subtracting its exact linearization, the full bounded nonlinear remainder obeys*

$$\|\mathcal{N}(t,Z)-\mathcal{N}(t,\widetilde{Z})\| \leq C(\|Z\|+\|\widetilde{Z}\|)\|Z-\widetilde{Z}\|$$

on a fixed small ball, uniformly for all real t and sufficiently large admissible R. The projected marked equation on bounded balls equals F_0 with an $o(1)$ error. On a fixed bounded slow-time interval, while the marks are bounded by M, the remaining equations obey the triangular bounds (4.11)–(4.12) *and the corresponding variation-of-constants bound for e with coefficient $C_M(1+|\Theta|)$ and an additive $C_M + o_R(1)\,\mathrm{Poly}_M$.*

Proof. Use the real complexification (the two conjugate slots are independent) on a fixed normalized ball. All prescribed time exponentials retain their real reference frequencies. The inverse Birkhoff map, finite old active chart, and homological coefficients are analytic on fixed complex neighborhoods. The form and transpose identities just proved are algebraic in these slots and their bounds persist there. Formulas (4.9) and (4.10) apply on those balls: $K_p/(a^2h) = O_M(\rho_R)$, while $(K - K_p)/(a^2h)$ is bounded. The matrices H_b^{-1} and Π_b are uniformly bounded analytic. By (4.4), the additional bounded factor T_0/a is $o(1)$. In (4.6) the only unbounded operator is the fixed reference generator $\mathcal{A}_0/a$. Composing the remaining bounded expressions with the corrected-action inverse makes no derivative act on this generator. Banach-space Cauchy estimates on a smaller ball now bound all second derivatives of the bounded vector fields uniformly. In particular, for $k = 1, 2$ they give the pure-column estimate

$$\frac{\|\partial_b^k K_p\|_X}{a^2h} \le C_M \left(\frac{R^{-2}}{ah} + \frac{1}{Rh} + \frac{a}{h}(1 + |J_{\mathrm{raw}}|) \right)$$

on a smaller compactly contained amplitude ball. All Cauchy estimates use a fixed complex radius in *normalized* variables. The Gram matrix and symplectic transpose are continued algebraically from the underlying real coordinates. The large free generator has already cancelled in the bounded differences before Cauchy's formula is applied. Their value at zero is exactly zero. Taylor's formula gives the asserted estimate after subtracting the *entire* variational part. Equation (4.5) gives the stated marked error. Identities on smooth states extend to the equivalent mild equations by density and these bounded remainder estimates.

For the last assertion take variation of constants in (4.6) about the old high-row linear propagator and the horizontal terminal diagonal. On a fixed slow-time interval the old high-row propagator is bounded in X by the high-row estimate of the endpoint analysis. Equivalently one may use the terminal complementary Green estimates. The remaining terms obey (4.9)–(4.10) and (4.4). The corrected raw action differs from J only by the bounded quadratic change in Lemma 4.2. This gives the claimed triangular bound in either time direction. These estimates are the nonlinear inputs to the stable selection argument in Section 4.4. □

4.4. **Selecting the orbit in the stable fibre.** The endpoint has several stable directions. We must choose stable data whose backward evolution reaches the low-norm section of the finite model. We now impose that condition at a late, fixed slow time S. In this subsection the time variable is $\sigma = at$ unless otherwise specified. Let $b_\star = b_{*,\infty}$ denote the limiting model terminal array, and let $\beta(\sigma)$ be the displacement from $b_\star$ of the orbit in Section 2.2, including all marked coordinates. In the coordinates of the previous subsection, a triad $p+q = r$, $p \cdot q = 0$, contributes

$$i\partial_\sigma B_r = 2\bar{m} B_p B_q e^{i(d_r - d_p - d_q)t},$$
$$i\partial_\sigma B_p = 2m B_r \bar{B}_q e^{-i(d_r - d_p - d_q)t},$$

and the term with p, q interchanged in the daughter equation. The parent factor two counts both orders of its distinct inputs in the physical square. With $c = |c|e^{i\phi}$ choose all newly added terminal horizontal phases to be $e^{i\phi}$ times their real model values. The old phases stay at their prescribed physical launch values. Substituting $B = e^{i\phi}\widetilde{B}$ and using time $2|c|\sigma$ gives exactly the model Hamiltonian of Section 2.2 after averaging. Use β in these physical amplitude and slow-time conventions, so that $\dot{\beta} = \mathcal{B}_c(\beta)$ is the full averaged marked equation. Set $\sigma = 0$

at its low-norm section. The terminal frequencies of the horizontal Birkhoff waves have been removed in the definition of the displacements. On each anomalous transverse pair use half the exact parent detuning on each daughter. Thus the linearized pair has no running relative phase even on the whole line. For other quadratic interactions the remaining phase e^{idt} has $d = O(R^{-2} + a^2) = o(a)$.

Lemma 4.4 (Stable selection)**.** *For every sufficiently large admissible dilation R there is an exact solution whose total marked amplitudes converge to $b_\star + \beta$ uniformly for $0 \le \sigma < \infty$ as $R \to \infty$. Its remaining normalized coordinates are uniformly bounded there. For some S and $\gamma > 0$, fixed before the final dilation, its full displacement from the terminal trajectory is $O(e^{-\gamma(\sigma - S)})$ for $\sigma \ge S$. Writing b_R for the marked displacement, one has for every $0 < \gamma' < \gamma$*

$$\sup_{\sigma \ge S} e^{\gamma'(\sigma - S)} |b_R(\sigma) - \beta(\sigma)| \longrightarrow 0.$$

Proof. In the normalized local coordinates $z = (b, e, J, \Theta)$ the origin is the exact terminal trajectory. Write $z' = \mathcal{A}_R(\sigma) z + N_R(\sigma, z)$ in the mild sense on the moving linear spaces. The local estimates give $N_R(0) = D_z N_R(0) = 0$ and uniform C^2 norm on a fixed ball. Let U_R and P_R be the exact evolution and strong stable projection. In slow time their bounds have stable rate $\gamma + g$ forward and inverse complementary growth rate $\gamma - g$ backward for some $0 < g < \gamma$. Choose γ strictly below both the known rate of decay of the particular model orbit and the least hyperbolic model rate. The former can be taken as $|c|\delta$ by the finite construction and its time change. Choose g and the central spectral tolerance smaller than the available margins before the cutoff in the sieve. In the norm $\sup_{\sigma \ge S} e^{\gamma(\sigma - S)} \|z(\sigma)\|$, the forward P_R convolution and backward $1 - P_R$ convolution have total norm at most $2K/g$. For $\zeta \in \operatorname{ran} P_R(S)$ the equation

$$\begin{aligned} z(\sigma) = U_R(\sigma, S)\zeta &+ \int_S^\sigma U_R(\sigma, r) P_R(r) N_R(r, z(r))\, \mathrm{d}r \\ &- \int_\sigma^\infty U_R(\sigma, r)(1 - P_R(r)) N_R(r, z(r))\, \mathrm{d}r \end{aligned}$$

is therefore contractive on a fixed small ball, uniformly over the choice of S, R. Its value at S is $\zeta + G_R(\zeta)$, where G_R lies in $\ker P_R(S)$ and has Lipschitz constant $O(r)$ on a radius r ball. The same argument applies to the finite averaged model, including any stable directions lying outside the symmetric ansatz.

Here is the boundary condition in coordinates shared by the PDE and the model. Let Π_R be the direct sum of the constant averaged stable projection on the marked sector and the reference projection on the arbitrary remainder, with zero old neutral stable parameters. In the direct proof of Section 3.3 this denotes precisely the effective remainder projection transported by the block-diagonal identification J. We impose the boundary condition in the original marked coordinates. The block splitting of Section 3 gives, in equivalent uniformly adapted norms, close projections P_R, Π_R. In particular $\Pi_R : \operatorname{ran} P_R \to \operatorname{ran} \Pi_R$ has a uniformly bounded inverse and the converse projections are bounded. One can see this by using the conjugating operator $P_R \Pi_R + (1 - P_R)(1 - \Pi_R)$ when the difference is sufficiently small. Given ξ in the range of Π_R, solve $\Pi_R(\zeta + G_R(\zeta)) = \xi$ by the same small-radius contraction after applying that inverse. Choose S far enough down the *already fixed* model orbit, and prescribe

$$\xi = (P_{\mathrm{model}} \beta(S), 0, 0, 0).$$

This fixes every stable coordinate in the finite system, including zero entries where appropriate. It also selects one exact exponentially convergent full trajectory on $[S, \infty)$.

Comparison on the stable tail. We compare the selected solution with the specified finite orbit on the entire forward half-line from S. Write $Z_R = (b_R, e, J, \Theta)$ for the full displacement in the normalized variables. Each retained oriented triad has a polynomial $Q_{\ell,R}(b)$ obtained by

inserting $b_{*,R}+b$ in one of the quadratic monomials above, including its numerical coefficient and the factor $-i$. Here ℓ runs over the finitely many oriented terms, including their conjugates in the real system. Since the terminal array is horizontal, each such monomial contains a transverse amplitude and hence vanishes at $b=0$. Their coefficients converge to those of $Q_{\ell,\infty}$ and their derivatives are uniformly bounded on every fixed ball. Let $d_{\ell,R}$ be the signed detuning, $D_R=\max_\ell|d_{\ell,R}|$, and $q_{\ell,R}(\sigma)=e^{id_{\ell,R}\sigma/a}$. Write $\mu_{\ell,R}(t)$ for the old reference mean $m_R(t)$ or its conjugate, and $c_{\ell,R}$ for its time average.

In these variables (4.5) takes the form

$$b_R'=\sum_\ell \mu_{\ell,R}(\sigma/a)q_{\ell,R}(\sigma)Q_{\ell,R}(b_R)+r_R(\sigma). \tag{4.13}$$

Here $|r_R|\le\varepsilon_R\|Z_R\|$ in a fixed smaller terminal ball, $\varepsilon_R\to0$. To see this, on a fixed complex ball use (4.9)–(4.10) in (4.5). The pure marked error there is $O(R^{-2}/a+R^{-1}+a)$, the arbitrary-input term is $O(h)$, and $(h/a)T_0$ is $o(1)$. The old mean evaluated at the variable actions and angles differs from the one at the reference by $O(a|\Theta|+a^2|J_{\mathrm{raw}}|)$. The total error vanishes at the exact terminal for every t. The Cauchy estimate on a smaller ball bounds its first derivative by $o(1)$ uniformly in t. On larger bounded boxes only the absolute bound $o(1)$ is needed.

Choose bounded primitives $F_{\ell,R}'=\mu_{\ell,R}-c_{\ell,R}$ as in (3.4), with derivative in physical time, and set

$$\widetilde{b}_R=b_R-a\sum_\ell F_{\ell,R}(\sigma/a)q_{\ell,R}(\sigma)Q_{\ell,R}(b_R).$$

On the tail $|\widetilde{b}_R-b_R|\le Ca|b_R|$. Differentiating cancels the oscillatory part of each mean in (4.13). A derivative on $q_{\ell,R}$ costs at most $CD_R|b_R|$ after multiplication by a, and a derivative on $Q_{\ell,R}$ costs $Ca\|Z_R\|$ since $|b_R'|\le C\|Z_R\|$. After averaging, the limiting vector field is $\mathcal{B}_c(b)=\sum_\ell c_{\ell,\infty}Q_{\ell,\infty}(b)$. Coefficient convergence and $Q_{\ell,R}(0)=0$ give

$$|\widetilde{b}_R'-\mathcal{B}_c(\widetilde{b}_R)|\le o_R(1)\|Z_R\|+C\min(2,|\sigma|D_R/a)|b_R|.$$

First take $\gamma'=\gamma/2$ and write $\|w\|_{\gamma',S}=\sup_{\sigma\ge S}e^{\gamma'(\sigma-S)}|w(\sigma)|$. The constructed PDE trajectory is bounded in the stronger weight γ. The residual in the last display therefore tends to zero in $\|\cdot\|_{\gamma',S}$, since

$$\sup_{\sigma\ge S}e^{-(\gamma-\gamma')(\sigma-S)}|\sigma|D_R/a\le(S+1/(\gamma-\gamma'))D_R/a\longrightarrow0.$$

The boundary condition in these coordinates fixes $P_{\mathrm{model}}b_R(S)=P_{\mathrm{model}}\beta(S)$ exactly. Thus the stable projection of $\widetilde{b}_R(S)-\beta(S)$ is $O(a)$. For the finite matrix $L_c=D\mathcal{B}_c(0)$, convolution of the forward stable and backward complementary kernels is bounded in weight γ', since the stable rates exceed γ and the remaining rates have nonnegative real part. Apply these kernels to $\widetilde{b}_R-\beta$. The nonlinear part $\mathcal{B}_c-L_c$ has Lipschitz constant Cr on a ball of radius r. Take r small relative to the fixed kernel bound and then take S sufficiently large in the boundary construction. Absorption gives

$$\|b_R-\beta\|_{\gamma',S}\longrightarrow0. \tag{4.14}$$

For the same orbit and same boundary time, the bound also holds with any $0<\gamma'<\gamma$. Smaller exponents follow by monotonicity. For larger exponents interpolate the vanishing norm at $\gamma/2$ with the uniform bound at γ for $b_R-\beta$.

Reaching the low-norm section. Continue backward on $[0,S]$. While $|b|\le B$ on a fixed bounded neighborhood of the model arc, the chart estimates have the following triangular form, where $\rho_R\to0$ and P is a fixed polynomial in the absolute values of the normalized coordinates:

$$|J'|\le C_B+\rho_RP,\quad |\Theta'|\le C_B(1+|J|)+\rho_RP.$$

For e the mild propagator of its fixed reference linear part is uniformly bounded on $[0,S]$. The inhomogeneous term costs $C_B[1+(1+|\Theta|)\|e\|]+\rho_R P$. This version is a variation-of-constants bound and avoids differentiating $\|e\|$ under the unbounded free evolution. First discard the $\rho_R P$ terms and add one to each right hand side. Integrate the two scalar bounds and then use Gronwall in the e bound. Fix the resulting enlarged bounds before increasing R so that $\rho_R P<1$ throughout the enlarged box. For example, if K_0 bounds the fixed linear propagator on that interval, successive bounds can be taken to be

$$\begin{aligned} J_1 &= |J(S)| + (C_B+1)S + 1,\\ T_1 &= |\Theta(S)| + [C_B(1+J_1)+1]S + 1,\\ e_1 &= K_0(\|e(S)\| + (C_B+1)S)e^{K_0 C_B(1+T_1)S} + 1. \end{aligned}$$

Use the uniform small-ball bounds at S when choosing the levels. On the resulting stopped interval, the same primitive correction as on the tail has size $O(a)$ and its equation has residual $o(1)+O(SD_R/a)=o(1)$ in supremum norm. Integrate the finite-dimensional difference backward from S. Gronwall bounds it by $e^{C_B S}(|\widetilde{b}_R(S)-\beta(S)|+So(1))$. Choose the tube around the model arc to have fixed positive width, and the three auxiliary bounds strictly larger than their estimates. A first exit in $[0,S]$ contradicts either this difference bound or one of the three strict bounds. Hence the solution reaches the low section at zero. Together with (4.14) this proves uniform convergence of the marked arrays on $[0,\infty)$.

For fixed R the order-two state space equals H^2 with an equivalent norm. The mild solution constructed above is continuous there. Its free diagonal applied to it is continuous into L^2. The other terms on compact intervals are continuous into L^2. The variation-of-constants identity therefore gives an L^2 time derivative. The smooth finite columns and the physical quadratic corrections can be differentiated there. The Birkhoff change of the horizontal row is analytic on L^2, and the old-action correction uses bounded forms there. Undoing the chart by the chain rule in L^2 recovers the original PDE, since the extraction identities invert its vector field on smooth states and extend in these norms by density. This justifies the use of a mild solution of the normalized equations to construct a true PDE trajectory. For fixed R, both the chart and its inverse have bounded derivatives near the compact terminal trajectory (the constants may depend on R). Consequently decay in the chart gives strong convergence to that trajectory in physical H^2 at rate $a\gamma$. Conservation and continuity of the mass identify the frozen scalar M_T with the mass of this exact solution. Rotating both trajectories back by the same scalar phase preserves their H^2 distance and gives a landing for the original NLS. Comparisons with the preceding reference C, whose mass may differ, are made in the common mass-dependent flow (1.2).

The global H^2 flow from Section 1 also extends this solution to every negative physical time. □

It remains to translate the normalized trajectory back into physical Sobolev norms. This is the point where the ratio in Proposition 2.3 determines the initial size and the terminal height simultaneously.

Conclusion of the proof of Proposition 3.1. Fix the admissible reference and its launch angles. Use Lemma 2.2 with the old quotient Diophantine condition at each extension. For the mixed linear input apply Section 3.3 and the horizontal block of Lemma 3.3. For the nonlinear input apply Proposition 4.3. They apply to the same horizontal solution C_R. We spell out the order in which the spatial and amplitude scales are fixed. Choose the finite model's norm ratio greater than $8K/\varepsilon$ (increase K by the fixed old norm if necessary). After all arithmetic choices, put $W^2=\sum_p |p|^{2s}|\beta_p(0)+b_{\star,p}|^2$ at the low section, and $d=\varepsilon/(2W)$ in $a=dR^{-s}$. To compare all

three physical trajectories, let $B_R = b_{*,R} + b_R$ be the total marked array of this solution. For $r = q, s$, uniformly in physical time $t \geq 0$,

$$V_R(t) - C(t) = a \sum_{p \in \mathcal{G}} E_{Rp}(t) B_{R,p}(at) + O_{H^r}(a + aR^{r-1} + a^2 R^r + ahR^r).$$

Indeed the physical old actions differ from their original values by $O(a^2)$, and the old angles by $O(a)$, uniformly in time because the terminal old frequencies match exactly. The old finite embedding is Lipschitz into H^r. Here the sum is indexed by undilated labels p, so the physical wave is E_{Rp}. The remaining terms follow by applying $\|v\|_{H^r} \leq R^r \|v\|_X$ to the particular fields, the generic error and the dressed-column error (or using Y for finite columns). All terms on the right are uniformly $O(1)$ for $r = s$ and $o(1)$ for $r = q$. The display at $r = s$ and compact-interval convergence of B_R yield the initial-window bound since the complex phases preserve the weighted norm of the distinct Fourier coordinates. At the terminal project onto those same high Fourier coordinates. This projection is onto all of $R\mathcal{G}$. The old embedding has rapidly decreasing Fourier tails uniformly in its angles. The projected terminal norm therefore tends uniformly, up to its constant phase factors, to $d|b_\star|_{s,\mathcal{G}} > 4K$, whereas its early counterpart for $V_R - C$ tends to $dW = \varepsilon/2$. Compact slow-time control and the stable tail in Lemma 4.4 give one upper bound L independent of the last dilation. The strong decay and splitting have physical weight $a\gamma$ and positive margin ag. At fixed R their compatibility in H^q, H^s is the interpolation statement of Section 3. Increasing R after the undilated configuration is fixed also makes the added ℓ^2 mass-amplitude budget arbitrarily small. This proves the finite landing assertion for this class of references. □

5. A single orbit and its horizontal limit

Suppose that $B + w$ tends exponentially to a trajectory B already incorporated in the construction. When another transfer changes B, we must reconstruct w above the new trajectory and control its value at time zero. We do this in two norms. In H^q we compare the bases on the entire half-line. In H^s we allow their difference to be large at late times, when the old correction is exponentially small.

For $T \geq 0$ and a physical-time weight η write $\|w\|_{r,\eta;T} = \sup_{t \geq T} e^{\eta(t-T)} \|w(t)\|_r$. The notation $\|w\|_{r,\eta}$ below suppresses T. We denote by $U(t, r)$ the propagator from r to t of the linearized equation and by $P(t)$ its invariant strong stable projection.

5.1. **Transport of a stable correction.**

Lemma 5.1 (Persistence of a strong stable correction)**.** *Let C be a fixed horizontal reference with uniformly bounded H^s norm. Assume its variational equation on the sublattice under consideration has compatible evolutions on H^q, H^s, $1 < q < s$, with*

$$\|U(t, r)P(r)\| \leq Ke^{-(\gamma+g)(t-r)} \qquad (r \leq t),$$
$$\|U(t, r)(1 - P(r))\| \leq Ke^{(\gamma-g)(r-t)} \qquad (r \geq t),$$

where $0 < g < \gamma$. Let $B, B + w$ be fixed exact solutions, with $\sup_{t \geq 0} \|B(t) - C(t)\|_q$ sufficiently small for this splitting, $\sup_{t \geq 0} \|B(t)\|_s < \infty$, and w in the time weight γ in both regularities. Let $B' = B + d$ be another exact solution with finite $\sup_{t \geq 0} \|B'(t)\|_s$. On $[T, \infty)$ assume both

$$\delta_q = \sup \|d(t)\|_q, \qquad \delta_s = \sup e^{-(g/2)(t-T)} \|d(t)\|_s$$

are small. For T sufficiently large for the old correction, B' carries an exact correction w' in the original time weight γ in H^q and H^s. The difference $w' - w$ tends to zero in weight γ in H^q and weight $\gamma - g/2$ in H^s as $\delta_q + \delta_s \to 0$. Its smallness modulus may depend on B, w, C, T. It is independent of the unweighted H^s bound of B'.

Proof. Let F be the vector field of (1.2), and put $N_B(w) = F(B+w) - F(B) - DF(C)w$. The free Laplacian cancels. All remaining maps are cubic polynomials and include the polarized rank-one term obtained from $M(v)v$. Fix the same boundary value $P(T)w'(T) = P(T)w(T)$ in the strong stable integral equation for w' with nonlinearity $N_{B'}(w')$. The forward and backward kernels have bounded convolution norm both at γ and at $\gamma - g/2$. In H^q the Lipschitz factor in addition to this kernel bound is at most $C_C(\|B-C\|_q + \|d\|_q + r)$ for a ball of radius $r \le 1$. Take T large so that $w(T)$ and its shifted weighted radius on this half-line are small. This gives the unique branch and continuous comparison in the strong H^q weight.

For the higher norm use

$$\|fg\|_s \le C_{s,q}(\|f\|_s\|g\|_q + \|g\|_s\|f\|_q).$$

It follows termwise by Fourier convolution and $\|\widehat{f}\|_{\ell^1} \le C_q\|f\|_q$ on the square torus. The mass polynomial obeys the same bound. In Picard iteration any higher-norm unknown multiplied by lower norms of $B'-C$ or a correction has small coefficient. Any high norm on the known base multiplies a lower-norm correction. One first obtains the strong H^s bound, for fixed B', by iteration with the previously established strong H^q bound. Explicitly the successive difference bound has the form $X_{k+1} \le \lambda X_k + L'Y_k$ for strong high and low differences, $\lambda < 1$, Y_k summably geometric and L' finite. Terms in which one previous high iterate multiplies a low difference are controlled by its uniform high bound from the same triangular inequality on the iterates. This proves convergence in the original strong weight with a bound allowed to depend on L'.

To compare, write $x = w' - w$ and expand

$$N_{B'}(w+x) - N_B(w) = N_B(w+x) - N_B(w) + \{F(B+d+w+x) - F(B+d) - F(B+w+x) + F(B)\}.$$

For an explicit bound, polarize the cubic part of F as a symmetric real trilinear map T_3, including the mass term. Then

$$N_B(w) = 3[T_3(B,B,w) - T_3(C,C,w)] + 3T_3(B,w,w) + T_3(w,w,w).$$

In the first difference every monomial contains x. If a derivative falls on x, its coefficient is small by the low bounds on $B-C, w, x$. Otherwise the differentiated factor is old and is multiplied by a low x. In particular no bound for an unknown high derivative is used in the latter estimate. With $v = w + x$, the expression in braces equals

$$6T_3(B,d,v) + 3T_3(d,d,v) + 3T_3(d,v,v).$$

To estimate the expression in braces, place one factor in H^s and all the others in H^q. When d carries the derivative, use δ_s times the strong weighted H^q bound of $w+x$. When a correction carries the derivative, use δ_q times the weaker weighted H^s norm of $w+x$. When an old coefficient carries it, use the fixed old bound. A second factor d is bounded in H^q whenever the first carries the derivative. Thus after absorption the weaker weighted H^s norm of x is $O_{C,B,w,T}(\delta_s + \delta_q + \|x\|_{q,\gamma})$, proving the stated independence. On a finite interval preceding T, closeness of $B' + w'$ to $B + w$ now follows from closeness at T and continuous dependence of the full flow. □

5.2. **Passage to a single initial datum.** We can now apply Proposition 3.1 repeatedly. The reference at the first step is a positive three-action torus on the pinned surface of Lemma 2.1. Choose its rotating frequency Diophantine by Lemma 2.2. At any finite stage the same lemma supplies the next admissible reference. The sum of the allowed changes of the primary actions and of the added ℓ^2 norms can be taken so small that the stationary coefficient stays separated from zero throughout, as shown at the end of Section 2.1.

Proof of Theorem 1.1. Use divisors $d_1 \mid d_2 \mid \cdots$. After $n-1$ stages keep the exact trajectories V_{n-1}^j for $1 \le j \le n$, with $V_{n-1}^n = C_{n-1}$ horizontal. The difference $V_{n-1}^j - V_{n-1}^{j+1}$ belongs to the

original jth strong fibre on the d_j sublattice. Keep its original reference C_j, projection and positive physical rates γ_j, g_j. At each earlier level the base V_{n-1}^{j+1} is uniformly H^q-close to C_j, since all its subsequent increments were chosen within a summable smallness budget for this fixed older splitting.

At the next stage use the proposition about C_{n-1} to construct V_n^n on $d_n\mathbb{Z}$, landing on $V_n^{n+1} = C_n$. For each earlier application of Lemma 5.1 fix its time T_j and required small H^q and damped H^s accuracy using only the existing tower. The early accuracy of the new block can be fixed before its upper norm L is known. Once L is known, choose a physical deadline $D > \max T_j$ so large that $Le^{-(g_j/2)(D-T_j)}$ is smaller than the required tolerances for every j. Now increase the final dilation. The early estimates up to D and the bound L thereafter give all damped hypotheses at the same time. The early comparison uses $V_n^n - C_{n-1}$, while the new endpoint C_n already has large norm.

For $j = n-1, n-2, \dots, 1$ apply Lemma 5.1 to the old correction on $B = V_{n-1}^{j+1}$, attaching it over $B' = V_n^{j+1}$ to form V_n^j. Nested divisibility places B' in the required d_j sublattice. The resulting difference of corrections is small in a positive exponential weight in H^s on its half-line, and is small at fixed earlier times by continuous dependence of the exact solutions. The base increment is already small in every damped norm required by the remaining layers. Their sum is therefore small in each of these norms as well as uniformly small in H^q. There are only finitely many moduli of continuity at this stage, depending on the old configuration. More explicitly, fix $0 < \eta \le \min_j g_j/2$ over the older layers and endow trajectory differences with the norm $\sup_t \|d(t)\|_q + \sup_t e^{-\eta t}\|d(t)\|_s$. When choosing D we also require $Le^{-\eta D}$ to be smaller than the input tolerance for this common norm. The common norm controls each earlier damped high-norm estimate on $[T_j, \infty)$ with factor $e^{\eta T_j}$. The difference of corrections from a single append map is small in this same norm by the positive high comparison weight after T_j and finite-time continuity before T_j. Thus the finite composition is continuous at the existing tower. Pick its input accuracy first, then the early budget, the stage bound L, the deadline D and finally the dilation, in that order. For each earlier j spend no more than 2^{-n} times its original H^q tolerance. This maintains the small-base hypotheses for future stages. The individual corrected trajectories belong to the original strong H^s weight, even though their comparison consumed part of the temporal margin.

The exact trajectory V_n^1 converges to C_n at positive exponential rates, with constants allowed to depend on the finite tower. Choose $K > 2n$ in its construction. There is a time $t_n > \max(n, t_{n-1}+1)$ with $\|V_n^1(t_n)\|_s > 2n$. Let $f_n = V_n^1(0)$. The stage allowed an arbitrarily small $\|f_n - f_{n-1}\|_s$. Use continuity of the global H^s flow at these finitely many fixed times to choose nested closed balls about the f_n, their radii tending to zero, inside the open sets where the finitely many selected hits exceed j, $1 \le j \le n$. This is possible by choosing each new initial increment strictly inside the previous ball and retaining a positive margin at the new hit. Completeness gives f in their intersection. Equation (1.2) and the original equation have identical Sobolev norms on each trajectory. Taking $u_0 = f$ proves the asserted unbounded orbit. □

There is a more precise description of the trajectory at large time. The terminal solutions recur simultaneously on their finite horizontal tori. This allows their uniform limit to be identified in the weaker norm throughout its entire orbit.

Corollary 5.2 (Asymptotic horizontal motion)**.** *Fix $1 < q < 6/5$. There exist an initial datum as in Theorem 1.1 and a horizontal solution $u_\infty \in C(\mathbb{R}; H^q)$ of (1.1) such that*

$$\lim_{t\to+\infty} \|u(t)\|_{6/5} = \infty, \qquad \lim_{t\to+\infty} \|u(t) - u_\infty(t)\|_q = 0.$$

The trajectory u_∞ is almost periodic in H^q, and every point of its H^q orbit closure lies outside $H^{6/5}$.

Proof. The preceding induction permits a summable bound on $\sup_{t\geq 0}\|V_n^1(t)-V_{n-1}^1(t)\|_q$ as well as on the strong initial increments. The uniform H^q limit on the half-line is the global H^s solution V with the strong limit initial value, by uniqueness on compact intervals. For the newest pulse one may also prescribe an arbitrarily small bound $\sup_{t\geq 0}\|V_n^n(t)-C_{n-1}(t)\|_q\leq\eta_n$. Its convergence to C_n gives $\limsup_{t\to\infty}\|C_n(t)-C_{n-1}(t)\|_q\leq\eta_n$. The pair of finite horizontal motions is recurrent simultaneously: approximate the identity translation in its product angle torus by positive times tending to infinity. It follows that the same bound holds on all of $\mathbb{R}$. For summable η_n, C_n converges uniformly on the line to an almost periodic horizontal C_∞, which is an exact H^q trajectory by continuous dependence. Taking first $t\to\infty$ and then $n\to\infty$ in the triangle inequality with V_n^1, C_n shows $\|V(t)-C_\infty(t)\|_q\to 0$. This also gives equality of their masses, momenta and energies. In particular the same scalar phase restores both to the standard NLS.

Choose $K_n>2n$. The C_n phase orbit has compact closure in H^s, so choose an integer N_n so that the Fourier projection $P_{\leq N_n}$ onto $|k|\leq N_n$ satisfies $\inf_t\|P_{\leq N_n}C_n(t)\|_s>3n$. For $k>n$ arrange $\eta_k\leq 2^{-k}n/\langle N_n\rangle^{s-q}$, using the minimum over the finitely many old constraints at stage k. Then $\|P_{\leq N_n}C_\infty(t)\|_s\geq 2n$ for every t, and this inequality extends by H^q continuity to the whole orbit closure. For each n,

$$\liminf_{t\to\infty}\|V(t)\|_s\geq 2n-\langle N_n\rangle^{s-q}\limsup_{t\to\infty}\|V(t)-C_\infty(t)\|_q=2n.$$

This proves the divergence of the full norm. Finally, the common scalar phase in (1.2) is periodic (or constant). Its product with the horizontal almost periodic trajectory is again almost periodic. The same asymptotic comparison therefore holds in the original equation. □

References

[1] D. Bambusi, Long time stability of some small amplitude solutions in nonlinear Schrödinger equations, *Comm. Math. Phys.* **189** (1997), 205–226.

[2] D. Bambusi and B. Grébert, Birkhoff normal form for partial differential equations with tame modulus, *Duke Math. J.* **135** (2006), 507–567.

[3] J. Bourgain, Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations, *Geom. Funct. Anal.* **3** (1993), 107–156.

[4] J. Bourgain, On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE, *Int. Math. Res. Not.* (1996), 277–304.

[5] J. Bourgain, Global solutions of nonlinear Schrödinger equations, *Amer. Math. Soc. Colloq. Publ.* **46** (1999), Amer. Math. Soc., Providence, RI.

[6] J. Bourgain, Problems in Hamiltonian PDE's, *Geom. Funct. Anal., Special Volume (GAFA 2000), Part I* (2000), 32–56.

[7] R. Carles, E. Dumas and C. Sparber, Multiphase weakly nonlinear geometric optics for Schrödinger equations, *SIAM J. Math. Anal.* **42** (2010), 489–518.

[8] R. Carles and E. Faou, Energy cascades for NLS on the torus, *Discrete Contin. Dyn. Syst.* **32** (2012), 2063–2077.

[9] F. Catoire and W.-M. Wang, Bounds on Sobolev norms for the defocusing nonlinear Schrödinger equation on general flat tori, *Commun. Pure Appl. Anal.* **9** (2010), 483–491.

[10] X. Chen, Infinite cascades for the defocusing cubic half-wave equation on the torus, arXiv:2609.24930 (2026).

[11] J. Colliander, M. Keel, G. Staffilani, H. Takaoka and T. Tao, Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation, *Invent. Math.* **181** (2010), 39–113.

[12] D. De Silva, N. Pavlović, G. Staffilani and N. Tzirakis, Global well-posedness for a periodic nonlinear Schrödinger equation in 1D and 2D, *Discrete Contin. Dyn. Syst.* **19** (2007), 37–65.

[13] Y. Deng and P. Germain, Growth of solutions to NLS on irrational tori, *Int. Math. Res. Not.* (2019), 2919–2950.

[14] L. H. Eliasson and S. B. Kuksin, KAM for the nonlinear Schrödinger equation, *Ann. of Math. (2)* **172** (2010), 371–435.

[15] E. Faou, L. Gauckler and C. Lubich, Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus, *Comm. Partial Differential Equations* **38** (2013), 1123–1140.

[16] E. Faou, P. Germain and Z. Hani, The weakly nonlinear large-box limit of the 2D cubic nonlinear Schrödinger equation, *J. Amer. Math. Soc.* **29** (2016), 915–982.

[17] P. Gérard and S. Grellier, The cubic Szegő equation, *Ann. Sci. Éc. Norm. Supér. (4)* **43** (2010), 761–810.

[18] P. Gérard and S. Grellier, Effective integrable dynamics for a certain nonlinear wave equation, *Anal. PDE* **5** (2012), 1139–1155.

[19] P. Gérard and S. Grellier, The cubic Szegő equation and Hankel operators, *Astérisque* **389** (2017).

[20] P. Germain, Z. Hani and L. Thomann, On the continuous resonant equation for NLS. I. Deterministic analysis, *J. Math. Pures Appl. (9)* **105** (2016), 131–163.

[21] F. Giuliani, Sobolev instability in the cubic NLS equation with convolution potentials on irrational tori, *J. Differential Equations* **416** (2025), 1–27.

[22] F. Giuliani and M. Guardia, Sobolev norms explosion for the cubic NLS on irrational tori, *Nonlinear Anal.* **220** (2022), Paper No. 112865.

[23] B. Grébert and T. Kappeler, Symmetries of the nonlinear Schrödinger equation, *Bull. Soc. Math. France* **130** (2002), 603–618.

[24] B. Grébert and T. Kappeler, The defocusing NLS equation and its normal form, *EMS Ser. Lect. Math.* (2014), European Mathematical Society, Zürich.

[25] B. Grébert and L. Thomann, Resonant dynamics for the quintic nonlinear Schrödinger equation, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **29** (2012), 455–477.

[26] B. Grébert and C. Villegas-Blas, On the energy exchange between resonant modes in nonlinear Schrödinger equations, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **28** (2011), 127–134.

[27] M. Guardia, Growth of Sobolev norms in the cubic nonlinear Schrödinger equation with a convolution potential, *Comm. Math. Phys.* **329** (2014), 405–434.

[28] M. Guardia, Z. Hani, E. Haus, A. Maspero and M. Procesi, Strong nonlinear instability and growth of Sobolev norms near quasiperiodic finite gap tori for the 2D cubic NLS equation, *J. Eur. Math. Soc.* **25** (2023), 1497–1551.

[29] M. Guardia, E. Haus and M. Procesi, Growth of Sobolev norms for the analytic NLS on $\mathbb{T}^2$, *Adv. Math.* **301** (2016), 615–692.

[30] M. Guardia and V. Kaloshin, Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation, *J. Eur. Math. Soc.* **17** (2015), 71–149.

[31] Z. Hani, Long-time instability and unbounded Sobolev orbits for some periodic nonlinear Schrödinger equations, *Arch. Ration. Mech. Anal.* **211** (2014), 929–964.

[32] Z. Hani, B. Pausader, N. Tzvetkov and N. Visciglia, Modified scattering for the cubic Schrödinger equation on product spaces and applications, *Forum Math. Pi* **3** (2015), Paper No. e4.

[33] Z. Hani and L. Thomann, Asymptotic behavior of the nonlinear Schrödinger equation with harmonic trapping, *Comm. Pure Appl. Math.* **69** (2016), 1727–1776.

[34] E. Haus and M. Procesi, Growth of Sobolev norms for the quintic NLS on $\mathbb{T}^2$, *Anal. PDE* **8** (2015), 883–922.

[35] S. Herr and B. Kwak, Strichartz estimates and global well-posedness of the cubic NLS on $\mathbb{T}^2$, *Forum Math. Pi* **12** (2024), Paper No. e14.

[36] S. Herr and B. Kwak, Global well-posedness of the cubic nonlinear Schrödinger equation on $\mathbb{T}^2$, *Invent. Math.* **245** (2026), 513–629.

[37] T. Kappeler, B. Schaad and P. Topalov, Semi-linearity of the non-linear Fourier transform of the defocusing NLS equation, *Int. Math. Res. Not.* (2016), 7212–7229.

[38] S. B. Kuksin, On turbulence in nonlinear Schrödinger equations, *Geom. Funct. Anal.* **7** (1997), 783–822.

[39] S. Kuksin and J. Pöschel, Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, *Ann. of Math. (2)* **143** (1996), 149–179.

[40] A. Maspero, Tame majorant analyticity for the Birkhoff map of the defocusing nonlinear Schrödinger equation on the circle, *Nonlinearity* **31** (2018), 1981–2030.

[41] A. Maspero and M. Procesi, Long time stability of small finite gap solutions of the cubic nonlinear Schrödinger equation on $\mathbb{T}^2$, *J. Differential Equations* **265** (2018), 3212–3309.

[42] J.-C. Molnar, New estimates of the nonlinear Fourier transform for the defocusing NLS equation, *Int. Math. Res. Not.* (2015), 8309–8352.

[43] J. Molnar, Features of the nonlinear Fourier transform for the dNLS equation, *Ph.D. thesis, University of Zürich* (2016).

[44] F. Planchon, N. Tzvetkov and N. Visciglia, On the growth of Sobolev norms for NLS on 2- and 3-dimensional manifolds, *Anal. PDE* **10** (2017), 1123–1147.

[45] C. Procesi and M. Procesi, A KAM algorithm for the resonant non-linear Schrödinger equation, *Adv. Math.* **272** (2015), 399–470.

[46] G. Staffilani, Quadratic forms for a 2-D semilinear Schrödinger equation, *Duke Math. J.* **86** (1997), 79–107.

[47] G. Staffilani, On the growth of high Sobolev norms of solutions for KdV and Schrödinger equations, *Duke Math. J.* **86** (1997), 109–142.

Xi Chen, Department of Mathematics and Computer Science, University of Basel, Spiegelgasse 1, CH–4051 Basel, Switzerland

Email address: xi01.chen@unibas.ch