hexagon:2611.00577v2

INFINITE CASCADES FOR THE DEFOCUSING CUBIC NONLINEAR SCHRÖDINGER EQUATION ON THE SQUARE TORUS

XI CHEN

ABSTRACT. For each $1 < s < 2$ we construct a solution of the defocusing cubic nonlinear Schrödinger equation on the square two-torus with initial data in H^s whose H^s norm tends to infinity as $t \to +\infty$. The growth can be arbitrarily slow. More precisely, the squared norm agrees, up to a bounded error at all positive times, with the number of selected axial frequencies K_j for which $K_j^s \le t$. The central construction couples a finite tree of orthogonal frequency pairs to a periodic one-dimensional solution of the same equation. Each transfer leaves a nonvanishing Fourier coefficient on the horizontal axis. Control on an entire half-line allows successive transfers to retain the coefficients produced earlier.

1. INTRODUCTION

On the square torus $\mathbb{T}^2 := (\mathbb{R}/2\pi\mathbb{Z})^2$ consider

$$i\partial_t u = -\Delta u + |u|^2 u, \qquad u(0) = u_0. \tag{1.1}$$

Here $u(t, \mathbf{x})$ is complex valued, $\mathbf{x} := (x, y)$, and $\Delta := \partial_x^2 + \partial_y^2$. We write $d\mu := dx\, dy/(2\pi)^2$ for normalized Haar measure. For $k = (k_x, k_y) \in \mathbb{Z}^2$ put $\langle k \rangle := (1 + |k|^2)^{1/2}$, where $|k|$ is the Euclidean length, and use the same bracket for integer indices in one dimension. Our Fourier and Sobolev conventions are

$$\widehat{u}(k, t) := \int_{\mathbb{T}^2} u(t, \mathbf{x}) e^{-ik\cdot\mathbf{x}}\, d\mu, \qquad \|u(t)\|_{H^s}^2 := \sum_{k\in\mathbb{Z}^2} \langle k \rangle^{2s} |\widehat{u}(k, t)|^2.$$

We call $(n, 0)$ an axial frequency and a field independent of y an axial field. The conserved mass $M(u)$ and energy $E(u)$ are the integrals against $d\mu$ of $|u|^2$ and $\frac{1}{2}|\nabla u|^2 + \frac{1}{4}|u|^4$, respectively. They bound the H^1 norm along a trajectory. Above H^1, a sequence of smaller and smaller amounts of mass can carry a growing Sobolev norm. The question is whether the autonomous equation can produce this behavior from a single initial datum.

Theorem 1.1. *For every $1 < s < 2$ there is $u_0 \in H^s(\mathbb{T}^2)$ such that the global solution of (1.1) satisfies*

$$\lim_{t\to+\infty} \|u(t)\|_{H^s} = +\infty.$$

The datum can be chosen with arbitrarily small H^s norm.

The growth in Theorem 1.1 comes from distinct axial Fourier coefficients that persist once they have been formed. The next result describes how these contributions build up in physical time. The quantity $\langle k \rangle^s |\widehat{u}(k, t)|$ is the normalized Fourier modulus at k. Its square is precisely the contribution of this frequency to $\|u(t)\|_{H^s}^2$. In the formula below, # counts the indices $j \in \mathbb{N} := \{1, 2, \ldots\}$ in the indicated set.

Theorem 1.2. *Fix $1 < s < 2$, and let $G : [0, \infty) \to (0, \infty)$ be nondecreasing with $G(t) \to \infty$. One can choose a datum as in Theorem 1.1 and increasing positive integers $K_j \to \infty$ so that*

$$\sup_{t\ge0} \Big| \|u(t)\|_{H^s}^2 - \#\{j : K_j^s \le t\} \Big| < \infty, \qquad \|u(t)\|_{H^s} = o(G(t)). \tag{1.2}$$

For each j, there is a time t_j such that $\langle K_j\rangle^s|\widehat{u}((K_j,0),t)|\geq 1/2$ for every $t\geq t_j$. Given any nondecreasing function $\ell:[1,\infty)\to(0,\infty)$ tending to infinity, the choices can be made with $t_j\leq K_j^s\ell(K_j)$ for all j.

In Section 5 of his survey of Hamiltonian PDE, Bourgain asked whether a solution of the periodic equation with data in H^s can exhibit unbounded Sobolev growth and how fast the growth can be [6, p. 43]. Theorem 1.1 answers the existence question for every exponent in $(1,2)$. The limit in the theorem also specifies what happens between successive visits to high frequencies.

On the square torus, Colliander, Keel, Staffilani, Takaoka and Tao [11] showed that a solution can start arbitrarily small in H^s, $s>1$, and reach any prescribed size. The geometry of their construction comes from the rectangles formed by four distinct Fourier modes satisfying the cubic resonance relation. By arranging these rectangles in generations, they made the output of one exchange available for the next. Guardia and Kaloshin [30] obtained a transfer time polynomial in the desired amplification factor.[1] In both results the initial datum may depend on the target. Bourgain's question asks what happens when the datum is chosen once. The limit in Theorem 1.1 requires control even between successive transfers, when energy can flow back.

The beating solutions of Grébert and Villegas-Blas [26] and Grébert and Thomann [25] exhibit this return explicitly. A mode cedes energy to its partner and recovers it later in the same orbit. One can also follow the creation of high modes from finitely supported data by multiphase geometric optics [7, 8]. A different picture emerges when the size of the torus tends to infinity. In the weakly nonlinear large-box limit the sum over discrete resonances gives rise to a continuous resonant equation [16], with its own stationary states and Hamiltonian dynamics [20]. An earlier statistical approach to high frequencies is due to Kuksin [40]. For a randomly forced equation on a cube he obtained lower bounds on spatial derivatives averaged over realizations and time as the dispersion and diffusion tend to zero.

For one fixed orbit, a particularly close precedent is Hani's theorem for a family of cubic equations on $\mathbb{T}^2$ that includes the resonant system [31]. On the waveguide $\mathbb{R}\times\mathbb{T}^2$, Hani, Pausader, Tzvetkov and Visciglia carried unbounded growth from the resonant system to the cubic NLS by modified scattering [32]. Partial harmonic trapping leads to another setting in which the long-time dynamics of small solutions are governed by a resonant equation [33]. On the circle, the cubic Szegő equation has unbounded Sobolev orbits accessible through inverse spectral theory [17, 19]. Its flow approximates small solutions of the cubic half-wave equation over a long finite interval [18]. For the half-wave equation itself, Chen obtained exponential growth for $1/2<s<3/2$ [10].

Fourier restriction estimates gave Bourgain a local theory for the periodic cubic equation [3, 5]. The same method also yields bilinear estimates for quadratic Schrödinger equations in two dimensions [48]. On the square torus, estimates based on almost conservation extended the global cubic theory below the energy space [12], and Herr and Kwak proved that the defocusing cubic equation is globally well-posed in $H^s(\mathbb{T}^2)$ for every $s>0$ [35, 36]. For more regular data one can ask how large a higher Sobolev norm can become by time t. Bourgain and Staffilani obtained polynomial upper bounds [4, 49], and subsequent work on flat tori [9] and compact manifolds [46] showed how the bounds depend on the geometry. For example, Deng and Germain improved the polynomial H^2 bound on generic irrational three-tori for powers $3<p<5$ [13]. If the norm does tend to infinity, how slowly can it do so? For every $1<s<2$ and every

[1] For large finite growth with quintic or analytic nonlinearities on $\mathbb{T}^2$, see [34, 29], and for a convolution potential, see [27]. On irrational two-tori, rational approximations supply nearly resonant rectangles for the cubic equation [22] and for its perturbation by a convolution potential [21].

nondecreasing $G(t) \to \infty$, Theorem 1.2 gives a solution on the square torus whose H^s norm tends to infinity and is $o(G(t))$.

The connection with the equation on the circle is particularly useful on $\mathbb{T}^2$, since any one-dimensional solution remains an exact solution when viewed as a function of (x, y). In particular, the finite-gap tori of the one-dimensional equation sit inside the two-dimensional phase space. Small invariant tori for NLS have also been constructed by KAM methods [41, 14, 47]. Long-time control of motion near small tori goes back to Bambusi [1], and the normal forms of Bambusi and Grébert apply to a broad class of Hamiltonian PDE [2]. For the cubic equation, long-time stability is known near plane waves [15] and near suitable small finite-gap tori [43]. Yet one-dimensional tori can also support large transverse excursions. For $s > 1$, Guardia, Hani, Haus, Maspero and Procesi [28] obtained any prescribed amplification in H^s from data arbitrarily close in L^2 to such a torus.

Here the one-dimensional solution supplies a Fourier harmonic whose phase must agree with that of the transverse interaction. In global Birkhoff coordinates the actions of the one-dimensional solution are constant, and each angle advances at a frequency determined by the actions [24]. Grébert and Kappeler described the symmetries of this flow in action-angle and spectral variables [23]. To use it in a two-dimensional rectangle we need both its frequencies and its ordinary Fourier coefficients. For the latter, the nonlinear part of the Birkhoff map gains one derivative [38]. Quantitative estimates for the map and for the Fourier coefficients of the flow are available in [44, 42, 39]. For the frequencies we use the formulas for the NLS hierarchy [37] and Molnar's local inversion of the renormalized frequency map [45].

1.1. **Strategy of the proof.** Suppose that, for each j, the normalized Fourier modulus at $(K_j, 0)$ stays above $1/2$ from some time t_j onwards. Once the first m such times have passed, these frequencies alone contribute at least $m/4$ to the squared H^s norm. We build towards this conclusion one frequency at a time. At each stage we perturb the initial datum by an arbitrarily small amount in H^s. The new mode eventually contributes about one to the squared norm, while the change at each of the finitely many earlier modes stays small for all future times. Both the permanence of the contribution and the small cost at the initial time come from the form of a single transfer.

The starting point is a one-dimensional solution, viewed on $\mathbb{T}^2$ as a field independent of y. Multiplying it by e^{2iMt}, where M is its mass, puts it in the mass gauge. Proposition 2.1 chooses the solution so that in this gauge it is periodic and the coefficient of e^{ix-it} in its space-time Fourier series is nonzero. Why this particular harmonic? Put $e \coloneqq (1, 0)$ and $k_v \coloneqq e + Rv$ for integer vectors v and a large integer R. If p, q are nonzero and perpendicular, then

$$k_p + k_q = k_{p+q} + e, \qquad |k_p|^2 + |k_q|^2 = |k_{p+q}|^2 + 1.$$

The cubic product of the waves at k_p, k_q with the conjugate of the chosen harmonic has exactly the spatial index and the free time frequency of the wave at k_{p+q}. The harmonic's temporal frequency is as important here as its spatial index. We obtain it by tuning the actions of the one-dimensional solution.

We arrange these rectangles in a binary tree with root velocity $(P, 0)$, so the root has axial Fourier index $(K, 0)$ with $K \coloneqq 1 + RP$. If the two children carry fractions $\zeta, 1 - \zeta$ of the squared length of their parent velocity, their combined homogeneous Sobolev weight is a fraction $\zeta^s + (1 - \zeta)^s$ of the parent's weight. We keep ζ away from zero and one, so for $s > 1$ this fraction is bounded by a constant $c_s < 1$ at every split. At depth d the sum of the primary leaf weights is therefore at most c_s^d times the root weight. The geometry also produces auxiliary resonances. Section 2 includes them in the finite system and chooses a trajectory on which the transverse amplitudes decay and the root tends to one in model units. The auxiliary amplitudes can be

chosen small after the tree is fixed. This is how a contribution of order one at late times can have arbitrarily small initial cost.

To realize this transfer in the full equation, fix the tree and take the physical amplitude $r \coloneqq (K-1)^{-s}$, so that $r\langle K\rangle^s \to 1$ as $R \to \infty$. Two scale comparisons enter Proposition 3.1. The new axial coordinates give an H^1 perturbation of order Rr, which tends to zero. A transverse frequency correction of order R^{-2} accumulates a phase of order R^{-2}/r over the exchange time r^{-1}. This too tends to zero, since $s < 2$. A surviving axial mode must remain in phase on an entire half-line, so we prescribe its exact frequency on an enlarged periodic axial solution. In solving for the difference, we propagate the rapidly decaying transverse directions forwards and prescribe terminal conditions on the complementary directions. Bounds independent of the terminal time give an exact solution approaching the enlarged axial solution. We then follow it back to time zero.

After the first transfer the previous solution is already two-dimensional. If U is that solution and V its periodic axial limit, inserting a new tree at V accounts only for the limiting background. Proposition 4.2 restores the decaying difference $U - V$ by a correction to the newly constructed solution. Section 4 can then choose the successive changes of initial datum summably small in H^s. Uniform control in a lower Sobolev norm protects each previously chosen Fourier coefficient at all later stages. Continuous dependence at each fixed time transfers those bounds to the solution from the limiting datum, giving the divergence in Theorem 1.1.

For Theorem 1.2 we also keep track of the norm during each exchange. The new frequencies lie where the preceding finite-stage orbit has a uniformly small Fourier tail, so cross terms in the squared norm can be made summable. We space the K_j far enough apart that the time windows for successive transfers are disjoint. At any time, all but at most one transfer have either negligible contribution or a contribution within a summable error of one. Counting the thresholds K_j^s therefore gives (1.2) with bounded error even during an exchange. The counting threshold K_j^s lies inside that window. The time by which the root stays above the required height has a separate bound from the finite system. At fixed depth this bound is a constant times $(K_j - 1)^s$, so we can choose K_j large enough to absorb the constant into the prescribed $\ell(K_j)$. Requiring also $G(K_j^s) \geq j$ gives the arbitrarily slow divergence in (1.2). The final two corollaries use the same construction to identify the limiting axial motion in lower Sobolev norms and to allow the contribution of an individual root to increase with depth.

Conventions. We use $\|\cdot\|_b$ for the H^b norm, and $dx/(2\pi)$ for measure on $\mathbb{T} \coloneqq \mathbb{R}/(2\pi\mathbb{Z})$. On the underlying real spaces we use the real part of the usual inner product. Bars denote complex conjugates, A^* denotes the adjoint for the unweighted L^2 or ℓ^2 inner product, and $[A, B] \coloneqq AB - BA$. In estimates $A \lesssim B$ and $A = O(B)$ mean $|A| \leq CB$ for $B \geq 0$. We write $A \asymp B$ when both inequalities hold for nonnegative A, B, and $A = o(B)$ when $A/B \to 0$ in the indicated limit. A subscript indicates allowed dependence of the constant. At a single insertion the tree and the old solution are fixed before $R \to \infty$. Section 4 specifies the order of choices when the tree also varies.

For a constant $c \in \mathbb{R}$, the field $e^{-ict}u(t)$ solves (1.1) with the additional potential c. The choice $c = -2M(u)$ is the *mass gauge*. In that gauge the same Fourier moduli describe the solution, while a common temporal frequency is removed from the equations for the axial modes.

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. An axial solution and a tree of right triangles

Write x, y for the two spatial variables. The subspace of functions of x alone is invariant under the cubic equation. Our construction starts with an orbit in this subspace that is periodic in the mass gauge. We shall use one Fourier coefficient of the orbit to couple two short sides of a right triangle of frequencies to its hypotenuse. Both the spatial and the temporal frequency of this coefficient matter.

2.1. **A periodic axial solution.** Write $U(t,x)$ for an axial solution before making the mass gauge transformation. In Birkhoff coordinates $\xi := \Phi(U) = (\xi_n)_{n\in\mathbb{Z}}$, the action $I_n := |\xi_n|^2$ is constant and the frequency $\omega_n(I)$ depends only on the action vector $I := (I_n)_{n\in\mathbb{Z}}$. Thus $\xi_n(t) = \xi_n(0)e^{-i\omega_n(I)t}$. The inverse map is denoted by $\Psi := \Phi^{-1}$, the differential of Φ at zero is the Fourier transform, and the mass in normalized Haar measure is $M = \sum_n I_n$. We subtract the common frequency $2M$ by multiplying the solution by e^{2iMt}. In this gauge the frequency at n is $n^2 + \chi_n(I)$, where

$$\chi_n(I) := \omega_n(I) - n^2 - 2M. \tag{2.1}$$

The same canonical chart diagonalizes the Hamiltonians of the NLS hierarchy [24, 37]. Translation by x_0 and multiplication by $e^{i\vartheta}$ act on ξ_n by multiplication by e^{inx_0} and $e^{i\vartheta}$, respectively, as follows from the action formulas for momentum and mass. The chart and its inverse are real analytic on L^2 and ℓ^2. For $1 \le q < \infty$, ℓ^q denotes the usual sequence space on the indicated indices, with counting measure, and ℓ^∞ has the supremum norm. Use h^b for the space of sequences with norm $(\sum_{n\in\mathbb{Z}}(1+n^2)^b|\xi_n|^2)^{1/2}$. The nonlinear part of each chart means its difference from its differential at zero. These differences gain one derivative as analytic maps from H^1 to h^2 and from h^1 to H^2, respectively [38, Theorem 1.1 and Corollary 1.1]. This also gives the chart regularity at every real order $1 \le b \le 2$ used below. For example, the nonlinear part of Ψ is analytic from h^b to H^b by the continuous inclusions $h^b \subset h^1$ and $H^2 \subset H^b$. Its linear part is the inverse Fourier transform. The same argument applies to Φ. We call $F := \{n : I_n > 0\}$ the Birkhoff support. For fixed actions with finite support, the axial torus is the image under Ψ of all choices of $\xi_j = \sqrt{I_j}e^{i\varphi_j}$, $j \in F$, with all other coordinates zero. A subscript F on a sequence denotes its restriction to these indices.

Another property we shall need is analyticity of $I \mapsto \chi(I)$ on an open ℓ^2 neighborhood of zero, with $d\chi(0) = -\mathrm{Id}$ [45, Theorem 18.1 and Corollary 18.2, with $p = 4$]. To check the normalization, let v solve the equation of period one with cubic coefficient two, as in these references. Then $U(x,t) = \sqrt{2}\,(2\pi)^{-1}v(x/(2\pi), t/(2\pi)^2)$ solves our equation. The actions of v are $(2\pi)^2 I_n/2$, after indexing by the physical Fourier mode, and its frequencies are multiplied by $(2\pi)^{-2}$ when passing to U. Thus the derivative $-2\mathrm{Id}$ of the renormalized frequency map in the cited normalization becomes $-\mathrm{Id}$.

Proposition 2.1. *There are smooth axial solutions U with arbitrarily small H^1 norm and finite Birkhoff support whose mass-gauged fields $V(t,x) := e^{2iMt}U(t,x)$ have period $2\pi Q$ for some positive integer Q and satisfy*

$$g := \frac{1}{2\pi Q}\int_0^{2\pi Q}\int_{\mathbb{T}} e^{it-ix}V(t,x)\,\frac{dx}{2\pi}\,dt \neq 0. \tag{2.2}$$

The initial solution can be chosen with Birkhoff support $F = \{0,4,5,2,6\}$. The Birkhoff coordinate at the spatial frequency 1 vanishes identically on this torus.

Proof. Give F the order $S := (0,4,5,2,6)$ and put $\ell := (1,1,1,-1,-1)$. Write $\mathbf{1} := (1,1,1,1,1)$ and take powers on S componentwise. Then

$$\ell\cdot\mathbf{1} = \ell\cdot S = \ell\cdot S^2 = 1, \qquad \ell\cdot S^3 = -35.$$

We compute the coefficient of $\xi_0\xi_4\xi_5\overline{\xi_2\xi_6}$ in the spatial Fourier coefficient at 1 of $\Psi(\xi)$. Its powers of the five phases are encoded by the character $e^{i\ell\cdot\varphi}$, where $\varphi \coloneqq (\varphi_j)_{j\in S}$ in the indicated order. This is the unique monomial of lowest degree with that character. For a cubic product with unconjugated inputs at a, c and conjugated input at b, put

$$k \coloneqq a - b + c, \qquad D(a,b,c) \coloneqq a^2 - b^2 + c^2 - k^2.$$

When $D \neq 0$, comparison of the cubic terms in the physical equation and its diagonal Birkhoff equation gives $1/D(a,b,c)$ per ordered triple in the inverse map. Indeed, temporal differentiation acts on the monomial at linear order by its signed input energy, whereas the output Laplacian has eigenvalue k^2. The resonant cubic monomials contain an action pair, since $D = -2(a-b)(c-b)$. Such a pair cannot occur in our quintic monomial with five distinct indices.

At fifth order the divisor for the NLS flow vanishes. We compute the same coefficient with the commuting flow $U_\tau = U_{xxx} - 3|U|^2U_x$, obtained from [37, (6)] by the same spatial and amplitude scaling, with the third time of the period-one equation set to $-\tau/(2\pi)^3$. The linear term in the equation for $i\partial_\tau U$ acts on e^{inx} by multiplication by n^3. Symmetrizing the cubic term in its two unconjugated inputs A, C gives the multiplier $3(A+C)/2$ per ordered triple. In the outer cubic $U_A\bar{U}_BU_C$, where U_A denotes a spatial Fourier coefficient, insert the cubic part of Ψ in one slot. The other two slots use its linear part. The inner triple (a,b,c) supplies the index $a-b+c$ for the chosen slot. Inserting it into the conjugated slot reverses all three signs. Summing over the assignments of the three positive indices $0, 4, 5$ and the two negative indices $2, 6$ gives

slot of the inner triple	1	2	3
$\sum 1/D(a,b,c)$	$-1/2$	1	$-1/2$
$\sum 3(A+C)/(2D(a,b,c))$	$-3/4$	$11/2$	$-3/4$

For example, in the middle slot the two negative indices $2, 6$ play the roles of a, c, and b ranges over $0, 4, 5$. The sum of the outer positive indices is $A + C = 9 - b$. The reciprocals $1/D$ for the three choices of b are $-1/24, 1/8, 1/6$. Each occurs four times, since a, c and A, C can be permuted independently. This gives 1 in the first row and $4(-9/16 + 15/16 + 1) = 11/2$ in the second. In the first slot the inner conjugated index is $b = 2$ or 6. For these choices the sums of $1/D$ over the positive assignments are $1/4$ and $-3/4$, respectively, with $A+C = 9-b$. The third slot gives the same sums by symmetry. The forcing for the NLS flow therefore vanishes, whereas the third-flow forcing equals four. Corrections to the linear angle velocities carry at least one action pair and have zero coefficient at this monomial. Its third-flow divisor is $\ell \cdot S^3 - 1^3 = -36$, so its coefficient is $-1/9$. Writing $\xi_j = \sqrt{I_j}e^{i\varphi_j}$, analyticity shows that the full angular coefficient equals

$$\left(-\tfrac{1}{9} + O\Big(\sum_{j\in F} I_j\Big)\right) \prod_{j\in F} \sqrt{I_j}.$$

Every higher monomial with the same angle character is obtained from the lowest monomial by adding action pairs.

We now impose $\ell \cdot \chi_F = 0$ so that this harmonic has time frequency 1. The derivative of the left side at the origin is $-\ell$. Set $I_0 = I_4 = I_5 = I_2 = \epsilon$. The derivative with respect to I_6 is 1, so the implicit function theorem gives $I_6 = 2\epsilon + O(\epsilon^2)$. All five actions are positive for small $\epsilon > 0$. On a relatively open patch about this point the angular coefficient computed above stays nonzero. The local coordinate map $I_F \mapsto \chi_F$ sends the patch to an open subset of $\ell \cdot \chi_F = 0$. This hyperplane has rational coefficients, so we may fix a point in the patch with rational χ_F.

It remains to choose an orbit on this torus. Let a_m, $m \in \mathbb{Z}^5$, be the angular Fourier coefficients of the spatial Fourier mode at 1 on the torus with the chosen actions. We have proved that

$a_\ell \neq 0$. From initial angles φ, the mass-gauged angles at time t are $\varphi - (S^2 + \chi_F)t$. Inserting their Fourier series into (2.2) and averaging over the common period $2\pi Q$ gives

$$g(\varphi) = \sum_{m\cdot(S^2+\chi_F)=1} a_m e^{im\cdot\varphi}.$$

Here Q is a common denominator of the frequencies, and the series is absolutely convergent by analyticity in the angles. Several characters can survive this time average. Since ℓ is among them, orthogonality in the initial angles gives $\int_{\mathbb{T}^5} |g(\varphi)|^2 \, d\varphi/(2\pi)^5 \geq |a_\ell|^2 > 0$. This supplies a single periodic orbit with $g \neq 0$. □

At each subsequent insertion we shall add finitely many axial coordinates and keep the angle velocities already used. This is why we tuned the frequencies through the actions in the first construction.

Lemma 2.2. *Fix a torus of support F in a sufficiently small ℓ^2 action ball, and let G be a disjoint finite set. Sufficiently small positive actions I_G can be added while preserving χ_F by a change of size at most $\|I_G\|_{\ell^2}/3$ in the old actions. By allowing the new positive radii to vary in arbitrarily small open neighborhoods of their prescribed values, one may in addition take all new χ_j, $j \in G$, rational.*

Proof. Take the action ball so small that $\|d\chi + \mathrm{Id}\| \leq 1/4$. The compression of $d\chi$ to F has an inverse of norm at most $4/3$, and the derivative from G to F has norm at most $1/4$. We need the allowed size of I_G to be independent of the locations of the new indices. Identify both coordinate subspaces with their zero extensions and put $x_0 := I_F^{\mathrm{old}}$, $y := I_G$. For fixed y, the equation for the new old actions is the fixed point equation

$$x = x + \chi_F(x + y) - \chi_F(x_0).$$

The right-hand side has Lipschitz constant at most $1/4$ as a function of x. At $x = x_0$ it differs from x_0 by at most $\|y\|_{\ell^2}/4$, since y is supported off F. It thus maps the closed ball of radius $\|y\|_{\ell^2}/3$ about x_0 into itself. For small enough $\|y\|_{\ell^2}$ this ball keeps all old actions positive and $x + y$ inside the action chart. Both restrictions depend only on the old point and the chart neighborhood. The contraction theorem gives the required old actions. The full derivative on $F \cup G$ is likewise within $1/4$ of minus the identity and hence invertible. In its local frequency coordinates, fixing χ_F therefore leaves an open set of χ_G values. This remains true in any open neighborhood where the prescribed new radii are positive. The new χ_G may therefore be taken rational, with the new positive radii in any prescribed sufficiently small open neighborhood of their nominal values. □

Keep the initial old angles fixed. Their subsequent trajectories in the two mass gauges agree in every old angular coordinate. By continuity of Ψ on the compact old angle torus, the change of the axial field is uniformly small in H^1 whenever the new coordinates are sufficiently small in h^1 and the old radii change sufficiently little. When the old occupied frequencies are rational, as in every application below, the two mass-gauged fields have a common period after the new frequencies are chosen rational. Averaging over this common period shows that their coefficients in (2.2) differ by at most their uniform L^2 distance.

We also need the unoccupied frequencies of a fixed smooth axial torus. Here a bound at an individual mode will be more useful than an asymptotic in a sequence norm.

Lemma 2.3. *For a fixed torus of finite Birkhoff support, $\chi_N(I) = O(|N|^{-2})$ as $|N| \to \infty$ through indices outside its support. The constant is locally uniform on compact families with fixed finite support.*

Proof. Use the ungauged field U on the torus, let $\mathcal{L}$ be differentiation along its ungauged angle flow, and put $D \coloneqq -i\partial_x$, $B \coloneqq 2(|U|^2 - M)$, $C \coloneqq U^2$. In the angles just introduced, $\mathcal{L} = -\sum_{j\in F} \omega_j(I)\partial_{\varphi_j}$. The linearized equation for a variation v has right-hand side $(D^2+2|U|^2)v+C\bar{v}$ in its $i\partial_t v$ equation. In its complexification, v and $\bar{v}$ can be varied independently. After extracting the same carrier $e^{iNx-i(N^2+2M)t}$ from both components, the equations for their envelopes f, h are

$$(2ND + D^2 + B - i\mathcal{L})f + Ch = 0, \qquad (2N^2 + 2ND + D^2 + 4M + B + i\mathcal{L})h + \bar{C}f = 0.$$

A column with residual $O(|N|^{-2})$ in L^2, uniformly on the angle torus, is given by

$$f \coloneqq 1 + f_1/N + f_2/N^2, \quad f_1 \coloneqq -\tfrac{1}{2}D^{-1}B, \quad f_2 \coloneqq -\tfrac{1}{2}D^{-1}(D^2 + B - i\mathcal{L})f_1,$$
$$h \coloneqq -\bar{C}/(2N^2) + (D\bar{C} - \bar{C}f_1)/(2N^3).$$

Here D^{-1} divides the k-th Fourier coefficient by k for $k \neq 0$ and returns a function of zero mean. Its first argument B has zero mean by mass conservation. In the second application the only term whose mean might survive is Bf_1. Since B is real, the mean of $BD^{-1}B$ is $\sum_{k\neq 0} |\widehat{B}(k)|^2/k = 0$. The two equations defining f_1, f_2 cancel orders $1, N^{-1}$ in the first row. In the second row the leading balance is $2N^2h + \bar{C} = 0$, and the displayed N^{-3} term cancels the remaining N^{-1} coefficient. All residual coefficients are smooth in x, φ on the fixed compact torus.

Apply the complexified derivative $d\Phi(U)$ to the column. The linearized conjugacy carries its residual into an $O(|N|^{-2})$ sequence in ℓ^2. After extracting the time phase, its N-th positive coordinate a_N equals $1 + O(|N|^{-1})$. Indeed its Fourier contribution is the mean of f, and the nonlinear differential maps the column, of H^1 size $O(|N|)$, into h^2 with the same bound. Since $\xi_N = 0$, variation of its frequency contributes no term to its linearized Birkhoff equation. Consequently $i\mathcal{L}a_N - \chi_N a_N = O(|N|^{-2})$. This holds for every initial angle. Integration over the old angle torus annihilates $\mathcal{L}a_N$. Its other mean factor is $1+O(|N|^{-1})$, proving the assertion. □

2.2. **The resonance geometry.** For a velocity $v \in \mathbb{Z}^2$ the spatial frequency of a new wave will be $k_v \coloneqq (1, 0) + Rv$, with R a positive integer. If $p + q = h$, then

$$|k_p|^2 + |k_q|^2 - |k_h|^2 = 1 - 2R^2 p \cdot q.$$

Thus the coefficient (2.2) couples the three waves exactly at the level of the free dispersion when $p \cdot q = 0$. We call (p, q, h) a triangle in this case and regard its first two entries as unordered. A pair p, h with $p \cdot (h - p) = 0$ can also produce $q = h - p$. We shall keep track of both operations when choosing the velocities.

Proposition 2.4. *For every integer $d \geq 1$ there is a finite set of integer velocities containing a full binary tree of depth d, with two children at every internal vertex and d edges from the root to each leaf. Its root $(P, 0)$ has $P > 0$, and the remaining primary vertices (the vertices of the binary tree) are off the horizontal axis. Each parent is the sum of its two orthogonal children. There is an absolute $p_0 > 0$ such that the two squared child lengths lie between p_0 and $1 - p_0$ times the squared parent length. All horizontal components are nonzero and have distinct absolute values. The following properties hold.*

(1) *At a split $p + q = m$ with m off the axis, write $v = (v_x, v_y)$ for either child and include the points*

$$J(v) \coloneqq (v_y^2/v_x, -v_y), \quad A(v) \coloneqq v + J(v) \quad (v = p, q), \qquad T_m \coloneqq J(p) + J(q).$$

The four pairs of legs at this split are (p, q), $(p, J(p))$, $(q, J(q))$, and $(J(p), J(q))$, with respective hypotenuses m, $A(p)$, $A(q)$, and T_m. The root split consists of a single triangle. These are all nondegenerate forward and reverse quadratic resonances among pairs of the selected velocities, including those whose output is initially sought outside the set.

(2) *If P_i, P_j are distinct horizontal coordinates of axial vertices, then $P_i(P_j - P_i)$ is never a positive rational square.*

Proof. We construct the tree over the rationals, starting with root $(1, 0)$. For an off-axis parent $m = (X, Y)$, take a rational parameter $t \neq 0$ and put

$$p \coloneqq \frac{X + tY}{1 + t^2}(1, t), \qquad q \coloneqq \frac{Xt - Y}{1 + t^2}(t, -1).$$

These are the two orthogonal projections of m. The vectors $p, J(q)$ lie on one of the perpendicular lines and $q, J(p)$ on the other. Their four possible sums are precisely the prescribed hypotenuses. In particular, $J(p)$ is the unique partner of p that is orthogonal to p and has vertical component $-p_y$. Thus both $A(p)$ and $A(q)$ lie on the axis. At the horizontal root $Y = 0$ the projection formulas give $J(p) = q$ and $J(q) = p$, so only p, q are added.

We give the exclusions used at each step. For a previously chosen u and a moving vertex v, a pair can generate a nondegenerate source only on one of the three loci

$$v \cdot u = 0, \qquad v \cdot u = |u|^2, \qquad |v|^2 = v \cdot u. \tag{2.3}$$

The first two are lines perpendicular to u, and the third is the circle with diameter $[0, u]$. Each of p, q traverses the circle with diameter $[0, m]$, so the only persistent relation of this kind for an old point is the prescribed one with $u = m$. Each of $J(p), J(q)$ traces an unbounded curve contained in no line. For example, writing $U \coloneqq (X + tY)/(1 + t^2)$, one has $J(p) = (t^2U, -tU)$. As $t \to \infty$ the first coordinate is unbounded and the second has a finite limit and is nonconstant. The calculation for $J(q)$ uses $t \to 0$. The remaining three new points traverse horizontal lines, and $u_x \neq 0$ rules out either line in (2.3). Consequently each unwanted old–new condition eliminates finitely many values of t.

For pairs within a single split the equations, after clearing denominators and scaling by X, are polynomials in $Y/X, t$. The following one rational evaluation tests all of them simultaneously. For $m = (1, 2)$, $t = 3$, multiply the ordered list $(m, p, q, J(p), J(q), A(p), A(q), T_m)$ by 30 to obtain

$$(30, 60),\ (21, 63),\ (9, -3),\ (189, -63),\ (1, 3),\ (210, 0),\ (10, 0),\ (190, -60).$$

Numbering from zero, the pairs with vanishing dot product are $(1, 2), (1, 3), (2, 4), (3, 4)$. The ordered pairs (i, j), $i \neq j$, with $v_i \cdot (v_j - v_i) = 0$ are $(1, 0), (2, 0), (1, 5), (3, 5), (2, 6), (4, 6), (3, 7), (4, 7)$. All horizontal components in this test have different absolute values. For each unwanted relation, write its cleared polynomial in powers of t, with coefficients polynomial in the parent slope Y/X. The test just given makes at least one coefficient a nonzero polynomial. Its finitely many roots contain all the slopes for which the relation could hold for every t. Taking the union over the finite list of unwanted relations still excludes only finitely many slopes. At the birth of every primary vertex avoid these exceptional slopes, as well as vanishing coordinates and coincident or opposite horizontal components. The horizontal components of the new points are nonconstant rational functions of t, so this also avoids coincidences with older vertices. The two child slopes are t and $-1/t$, which lets us impose in advance the slope exclusions needed when each child later becomes a parent. At the first split the children move on the root circle, and the same finite exclusions on their slopes and coordinates apply.

To ensure the arithmetical condition on axis vertices, note that the two new horizontal outputs have abscissae

$$U_1 \coloneqq X + tY, \qquad U_2 \coloneqq X - Y/t.$$

For a prime π and a nonzero rational number w, write $\operatorname{ord}_\pi(w)$ for the exponent of π in its numerator minus that in its denominator. Rational squares have even valuation at every prime. For each old axial abscissa $C \neq 0$, choose a fresh odd prime π for which Y, C are units (valuation

zero) and all fixed coordinates are integral. Require $t \equiv (C-X)/Y + \pi \pmod{\pi^2}$, with rational denominators read modulo π^2. Then $\operatorname{ord}_\pi(U_1 - C) = 1$, while U_1 and C are units. Both $C(U_1 - C)$ and $U_1(C - U_1)$ therefore have odd valuation. At another fresh prime use the same argument for $U_2 = X - Y/t$, with $1/t$ as parameter. For U_1, U_2 themselves take a fresh prime congruent to 1 mod 4 avoiding $X^2 + Y^2$. Choose t with $t^2 + 1$ of valuation one there. Both U_i are units, and $U_2 - U_1 = -Y(t^2+1)/t$ has valuation one. These finitely many conditions are compatible with prescribing the direction of p in any real open interval. Indeed each finite-place condition is open in the projective parameter line, so weak approximation over $\mathbb{Q}$ gives rational parameters satisfying them all. Concretely, one may impose congruences on the numerator and denominator separately, and then take their representatives sufficiently large to realize the real interval.

Choose the angle between the parent and the line of p in a fixed small interval about $\pi/4$. This gives a common lower bound $p_0 > 0$ for both relative squared lengths by the cosine and sine formulas. Removal of the finitely many forbidden directions leaves an open subinterval. Continuing to depth d proves the construction over the rationals. Multiplication of every velocity by a common denominator preserves all incidence conditions and multiplies each expression $P_i(P_j - P_i)$ by a square. □

2.3. **An orbit of the finite system.** Fix the old coefficient g of (2.2). Write the new wave at k_v, after removal of its free oscillation, as ra_v, where $r > 0$ is the common amplitude scale and a_v is the complex model amplitude. For the quadratic terms induced by the old wave, the two factors r on the right and the one on the left leave the time scale r^{-1}. With a dot denoting differentiation with respect to rt, the contribution to $i\dot{a}_h$ of $p + q = h$ is $2\bar{g}\, a_p a_q$. The contributions to $i\dot{a}_p$ and $i\dot{a}_q$ are $2g\, a_h \bar{a}_q$ and $2g\, a_h \bar{a}_p$. The factor two counts the two orders of the positive inputs at the output h, and is the coefficient of the mixed density term at the other outputs. Writing $a_v = i(g/|g|)X_v$, the real amplitudes X_v satisfy in the model time $z := 2|g|rt$

$$X_h' = X_p X_q, \qquad X_p' = -X_q X_h, \qquad X_q' = -X_p X_h. \tag{2.4}$$

At a vertex in several triangles, all of the displayed contributions are added. From here to the end of this section a prime denotes differentiation with respect to z.

A primary vertex at height j above the leaves has the same incidences as any other primary vertex at this height. We give these vertices a common amplitude x_j, and denote the root by $H := x_d$. For $0 \le j \le d-2$, let w_j, h_j, t_j denote the common amplitudes of the vertices $J(v), A(v), T_m$, respectively, for the split in which v is a child of height j. Thus t_j is one real amplitude per sibling pair. Equation (2.4) preserves this symmetry. In these variables it reads

$$x_j' = x_{j-1}^2 - x_j x_{j+1} - w_j h_j \quad (0 \le j < d), \qquad H' = x_{d-1}^2, \tag{2.5}$$

$$h_j' = x_j w_j, \qquad t_j' = w_j^2, \qquad w_j' = -x_j h_j - w_j t_j \quad (0 \le j \le d-2). \tag{2.6}$$

Here $x_{-1} := 0$, and the term $w_j h_j$ is omitted at $j = d-1$. An empty sum is zero, so the formulas also cover $d = 1$.

For a single split ($d = 1$) the equations can be solved explicitly. If $x_0(0) = a > 0$ and $H(0) = 0$, then $x_0^2 + H^2 = a^2$, so $H' = a^2 - H^2$. Hence

$$H(z) = a \tanh(az), \qquad x_0(z) = \frac{a}{\cosh(az)}.$$

The two leaves decay and the root increases to a. At larger depth the auxiliary triangles also exchange amplitude with the primary vertices. The next lemma chooses their initial amplitudes so that the same final root value is attained in the limit.

Lemma 2.5. *For every $a > 0$ and $b_0, \dots, b_{d-2} > 0$, the system (2.5)–(2.6) has a solution starting with $x_0 > 0$ and $x_1 = \cdots = x_d = 0$. On it, $0 \le H \le a$ and $-b_j \le t_j \le 0$. The solution converges exponentially to an equilibrium at which $H = a$, $h_j = \pm b_j$, and every other amplitude is zero. Its first two derivatives decay exponentially as well.*

Proof. Direct differentiation gives the conserved quantities

$$w_j^2 + h_j^2 + t_j^2 = b_j^2, \qquad \sum_{j=0}^{d} x_j^2 + \sum_{j=0}^{d-2} h_j^2 = a^2 + \sum_{j=0}^{d-2} b_j^2. \tag{2.7}$$

To see the cancellation across levels, sum only up to height $j < d$. Equations (2.5)–(2.6) give

$$\frac{d}{dz}\left(\sum_{k=0}^{j} x_k^2 + \sum_{k=0}^{\min(j,d-2)} h_k^2 \right) = -2x_j^2 x_{j+1}.$$

At each level $2h_k h_k'$ cancels the term $-2x_k w_k h_k$. The terms linking consecutive primary levels telescope, leaving only the displayed boundary term. At $j = d-1$ this is $-(H^2)'$. The first invariant follows by differentiating the three squares on one auxiliary sphere. These identities are for the reduced amplitudes, one per level. The multiplicities of the individual Fourier modes enter later in the Sobolev norm.

For each initial point on the product of the indicated spheres, set all primary amplitudes above height zero equal to zero, and set $x_0(0) = (a^2 + \sum_j (b_j^2 - h_j(0)^2))^{1/2}$. This gives a continuous choice of initial conditions with $x_0(0) \ge a$. The trajectories are global by (2.7). Let $\mathcal{P} := \prod_{j=0}^{d-2} \{(w,h,t) \in \mathbb{R}^3 : w^2 + h^2 + t^2 = b_j^2\}$. For $Z > 0$ let $F_Z : \mathcal{P} \to \mathcal{P}$ map these initial sphere coordinates to their values at time Z. The maps F_{uZ}, $0 \le u \le 1$, give a homotopy to the identity. Thus F_Z has degree one and is onto. Indeed a map of a closed oriented manifold missing a point has degree zero at that point. Take a preimage of the point with coordinates $(0, b_j, 0)$ for every j. At time Z the second invariant reduces to $H(Z)^2 + \sum_{k<d} x_k(Z)^2 = a^2$, where the sum starts at zero. Since $H(0) = 0$ and $H' \ge 0$, we have $0 \le H \le a$ on $[0, Z]$. Likewise $t_j' \ge 0$ and $t_j(Z) = 0$, while the sphere bounds t_j below by $-b_j$. For a sequence $Z \to \infty$ these preimages have a convergent subsequence in $\mathcal{P}$. Fix any finite observation time. Continuous dependence passes all the inequalities at that time to the trajectory from the limiting initial point. The same initial point therefore works on the whole half-line. For $d = 1$ the product of spheres consists of a single point and the same argument uses the primary invariant alone.

The omega limit set consists of limits of states at times tending to infinity. It is invariant under this finite-dimensional flow. The bounded monotone coordinates t_j, H each have a limit, so these coordinates are constant on the limit set. Their derivatives w_j^2, x_{d-1}^2 therefore vanish there. Invariance and (2.5) then give $x_{d-2} = \cdots = x_0 = 0$. The two identities in (2.7) give $H^2 = a^2 + \sum_j t_j^2$, so $H = a$, all $t_j = 0$, and $h_j = \pm b_j$. The omega limit set is connected, being the intersection of the nested compact connected closures of trajectory tails. It therefore consists of a single point. If the trajectory left some fixed neighborhood of this point at arbitrarily late times, compactness would supply a second limit point. This proves convergence.

For the decay rate, the vector of the lower x_j's and the w_j's solves a linear equation $Y' = (A + E(z))Y$ along this trajectory, with $E(z) \to 0$. The matrix A consists of the scalar $-a$ on x_{d-1} and the symmetric blocks $\left(\begin{smallmatrix} 0 & -h_j(\infty) \\ -h_j(\infty) & 0 \end{smallmatrix}\right)$ on (x_j, w_j). For example, write x_{j-1}^2 as the unknown x_{j-1} times a coefficient tending to zero. Fix $\lambda > 0$ less than the smallest absolute value of the eigenvalues of A. We use the supremum norm weighted by $e^{\lambda(z-Z)}$ on a late half-line $z \ge Z$. Write P_-, P_+ for the orthogonal projections onto the negative and positive eigenspaces

of A, and write $\lambda_0 \coloneqq \min_{\mu\in\operatorname{spec}(A)} |\mu|$. The equation on $z \geq Z$ takes the form

$$Y(z) = e^{(z-Z)A}P_-Y(Z) + \int_Z^z e^{(z-v)A}P_-E(v)Y(v)\,dv - \int_z^\infty e^{(z-v)A}P_+E(v)Y(v)\,dv.$$

In obtaining the last integral, the boundary term at infinity vanishes because Y is bounded. In the weighted norm the two integrals have combined operator norm at most $2(\lambda_0-\lambda)^{-1}\sup_{v\geq Z}\|E(v)\|$. They also contract in the unweighted norm when Z is sufficiently large. Uniqueness in the unweighted space identifies the bounded solution with the weighted one. Integrating the equations for H, h_j, t_j from infinity gives decay of the remaining differences. The polynomial equations and their first derivatives give the assertion about derivatives. □

The Sobolev cost of this orbit at its initial time is particularly simple. At each primary split, write $\zeta, 1-\zeta$ for the two ratios of squared lengths. For $s > 1$ define

$$c_s \coloneqq \max_{p_0\leq\zeta\leq 1-p_0} \left(\zeta^s + (1-\zeta)^s\right) < 1.$$

If the tree has root $(P, 0)$, iteration of the split gives

$$\sum_{\text{primary leaves } v} |v|^{2s}|x_0(0)|^2 \leq |P|^{2s}\Big(a^2 + \sum_j b_j^2\Big)c_s^d. \tag{2.8}$$

All higher primary levels are initially zero. Each auxiliary amplitude at height j is bounded by b_j for all time. Since the geometry is fixed before choosing the positive b_j, the total weighted cost of all the auxiliary vertices can be made arbitrarily small.

The identity $H' = x_{d-1}^2$ gives a quantitative estimate on the time spent below a prescribed fraction of the terminal amplitude. The estimate is uniform over the trajectories selected by the degree argument.

Proposition 2.6. *Take $a = 1$, $0 < \theta < 1$, and put*

$$C_\theta \coloneqq \left(\frac{100}{1-\theta^2}\right)^4, \qquad Z \coloneqq C_\theta d^{3/2}.$$

If $\sum_{j=0}^{d-2} b_j^{3/2} \leq Z^{-1/2}$, then $H(z) \geq \theta$ for every $z \geq Z$. In physical time the root has thus reached this fraction of its terminal amplitude by $C_\theta d^{3/2}/(2|g|r)$.

Proof. Write $E_j \coloneqq Z^{-1}\int_0^Z x_j^2\,dz$, including $x_d = H$. From (2.7) and monotonicity,

$$1 - H(Z)^2 \leq \sum_{j<d} E_j, \qquad E_{d-1} \leq Z^{-1}, \qquad E_d \leq 1. \tag{2.9}$$

Indeed at each time $1 - H^2 = \sum_{j<d} x_j^2 - \sum_j (w_j^2 + t_j^2)$, and $H' = x_{d-1}^2$. For $1 \leq j < d$, integrate (2.5), using $x_j(0) = 0$, $\int_0^Z w_j^2 \leq b_j$, and $|h_j| \leq b_j$, to get

$$E_{j-1} \leq (E_jE_{j+1})^{1/2} + e_j, \qquad e_j \coloneqq \frac{|x_j(Z)|}{Z} + \frac{b_j^{3/2}}{\sqrt{Z}}, \tag{2.10}$$

where $b_{d-1} \coloneqq 0$. For example, the mixed term in that integration is bounded by

$$Z^{-1}\left|\int_0^Z w_jh_j\,dz\right| \leq b_jZ^{-1/2}\left(t_j(Z) - t_j(0)\right)^{1/2} \leq b_j^{3/2}Z^{-1/2}.$$

At height $d-1$ this term is absent. As $Z \geq 1$, our choice of radii implies $\sum_j b_j^2 \leq 1$. Hence

$$e_j \leq (\sqrt{2}+1)/Z, \qquad \sum_{j=1}^{d-1} e_j \leq (\sqrt{2d}+1)/Z.$$

The second estimate uses Cauchy–Schwarz at the single endpoint Z.

To exploit the geometric mean in (2.10), order the levels downward, putting $m_k := E_{d-1-k}$ for $-1 \leq k < d$. The starting bounds are $m_{-1} \leq 1$ and $m_0 \leq Z^{-1}$. If the two preceding averages are bounded by powers of Z^{-1}, their geometric mean has the average of those exponents. This suggests $\alpha_{-1} := 0,\ \alpha_0 := 1,\ \alpha_k := (\alpha_{k-1}+\alpha_{k-2})/2$, giving $\alpha_k = 2/3 + (-1/2)^k/3$. For $k \geq 1$ they belong to $[1/2, 3/4]$. Multiplying (2.10) by Z^{α_k} shows that $m_k Z^{\alpha_k}$ exceeds the larger of the preceding two scaled averages by at most $(\sqrt{2}+1)Z^{-1/4}$. Let $n := 1 + \lceil \log_2(2 + \log Z) \rceil$, with log the natural logarithm. Since $Z \geq 100^4$, we have $n+1 \leq 4Z^{1/4}$ and $nZ^{-1/4} < 5$ (use $n \leq 3 + \log Z$). Induction gives $m_k \leq 15 Z^{-\alpha_k}$ up to $k = \min(n, d-1)$. These levels contribute at most $60Z^{-1/4}$ to (2.9).

If there are further levels, the last two exponents already satisfy $|\alpha_k - 2/3| \log Z \leq 1/3$, for $k = n-1, n$. Thus both averages are at most $21Z^{-2/3}$. For the remaining levels use $\sqrt{uv} \leq \max(u, v)$ in (2.10). Each is bounded by $21Z^{-2/3} + \sum_j e_j$. Substitution into (2.9) yields

$$1 - H(Z)^2 \leq 60Z^{-1/4} + 21dZ^{-2/3} + \frac{d(\sqrt{2d}+1)}{Z} \leq 60C_\theta^{-1/4} + 21C_\theta^{-2/3} + 3C_\theta^{-1} < 1 - \theta^2.$$

Monotonicity of H proves the proposition. □

There is also a lower bound within the finite system. The moment

$$\sum_{j=0}^{d} j x_j^2 + \sum_{j=0}^{d-2} j h_j^2$$

has derivative $2\sum_{j<d} x_j^2 x_{j+1} \leq 2B^3$, where $B^2 := 1 + \sum_j b_j^2$. Its initial value is at most $d\sum_j b_j^2$. Thus, if $\sum_j b_j^2 \leq \theta^2/2$, the model time of the first crossing $H = \theta$ is at least $d\theta^2/(4B^3)$. Both estimates concern a single choice of tree, radii, and initial angles.

3. Insertion of a tree into the two-dimensional flow

Let V be any smooth axial solution with finite Birkhoff support whose action vector lies in the ball of Lemma 2.2. Assume its occupied frequencies in the mass gauge are rational and its space-time coefficient (2.2) is nonzero. Proposition 2.1 provides an initial solution with these properties. Throughout this section axial solutions are written in their mass gauge. For mass M the equation has the additional constant potential $-2M$. Write g for the $(1,1)$ space-time Fourier coefficient of V, and F for its finite set of occupied Birkhoff indices. Fix the tree of Proposition 2.4 and an orbit given by Lemma 2.5. Let $\mathcal{V} \subset \mathbb{Z}^2$ be its full set of velocities, including the auxiliary vertices. For $v = (v_x, v_y) \in \mathcal{V}$, write $x_v(z)$ for the real amplitude assigned to that vertex in (2.5)–(2.6). We use three time variables. The equation is solved in physical time t, the error estimates use slow time $\tau := rt$, and the tree uses time $z := 2|g|\tau$, in which its interaction coefficient equals one. The positive amplitude scale r is specified in the next proposition. In complex amplitudes these functions all carry the same constant unit factor $\gamma := ig/|g|$. Enumerate the horizontal vertices as $(P_i, 0)$, and let $c_i := \lim_{z\to\infty} x_{(P_i,0)}(z)$. At the main root $P_i = P$, this limit is $a > 0$. At each other horizontal vertex it is one of the signed auxiliary radii. Both the tree and its positive auxiliary radii are fixed throughout the proof below.

Proposition 3.1. *Let $1 < s_0 < s < 2$ and $a_0 > 0$. Set $r := a_0 R^{-s}$ and $k_v := (1,0) + Rv$. For all sufficiently large integers R there are a periodic axial solution V_R and a solution W_R of the full two-dimensional equation in the mass gauge of V_R, with the following properties.*

(1) *On F, V_R and V have the same mass-gauged frequencies and the same Birkhoff angles at $t = 0$. The newly occupied Birkhoff indices of V_R are $K_i := 1 + RP_i$. At K_i the modulus of the Birkhoff coordinate is $r|c_i + e_i|$ for some real e_i with $|e_i| \le R^{-10}$.*

(2) *$W_R(t) - V_R(t)$ tends to zero exponentially in H^s as $t \to +\infty$. In the respective gauges of W_R and V,*
$$\sup_{t\ge 0} \|W_R(t) - V(t)\|_{H^{s_0}} = o(1).$$

(3) *There are real numbers $\kappa_v = |k_v|^2 + O(R^{-2} + r^2)$ such that*
$$\sup_{t\ge 0} \Big\| W_R(t) - V(t) - r\gamma \sum_{v\in\mathcal{V}} x_v(2|g|rt) e^{ik_v\cdot(x,y) - i\kappa_v t} \Big\|_{H^s} \longrightarrow 0. \tag{3.1}$$
In particular
$$\|W_R(0) - V(0)\|_{H^s} \le a_0 \Big(\sum_v |v|^{2s} |x_v(0)|^2 \Big)^{1/2} + o(1).$$
At the main root $K := 1 + RP$ one has, writing $H := x_{(P,0)}$ and using $\widehat{W}_R(K,0,t)$ for the coefficient at $(K,0)$ at time t,
$$\sup_{t\ge 0} \Big| \, |\widehat{W}_R(K,0,t)| - rH(2|g|rt) \, \Big| = o(r).$$

All limits in this statement are taken after the finite tree and the old orbit have been fixed.

The comparison in (3.1) involves three exact solutions. The old periodic field V supplies the coefficient g. The enlarged periodic field V_R prescribes the final axial amplitudes and phases. It is W_R that undergoes the two-dimensional transfer and converges to V_R. In particular the reference field in the finite-plane expansion (3.1) is V.

There are two different roles for the linear frequencies in this proof. An error in the frequency of a persistent axial mode accumulates for all future time. We therefore evolve these modes at their exact Birkhoff frequencies. Off the axis, the limiting field vanishes, but its linearized equation has both growing and decaying directions. We identify these directions explicitly and use them to prescribe the boundary conditions for the error. Unless indicated otherwise, constants in this section may depend on the fixed old orbit, tree, and a_0, s_0, s. They are independent of R and of the terminal time.

3.1. **The norm and the axial equations.** Use Lemma 2.2 to adjust the old actions by $O(r^2)$ and to choose the new radii in the intervals of part (1). Only finitely many new actions are involved, and the old ones stay positive. Let I^∞ denote the resulting action vector, let $c_R := -2\sum_j I_j^\infty$, and take the new initial phases so that the complex terminal amplitudes are $r\gamma(c_i + e_i)$. All renormalized occupied frequencies are rational, so the resulting axial solution V_R is periodic in this gauge. In particular its constant potential differs from that of V by $O(r^2)$.

For one insertion it suffices to work in the exact invariant subspace with vertical Fourier indices in $R\mathbb{Z}$. These are the functions fixed by translation of y by $2\pi/R$. Write $p := (s+2)/2$ and set
$$\alpha := \tfrac{1}{2}\min(2 - s, p + s - 2), \qquad L := R^\alpha, \qquad \rho := R^{-3\alpha/4}. \tag{3.2}$$

Here $s < p < 2$, $L \to \infty$, and $L/R \to 0$. For $n, q \in \mathbb{Z}$ the integer q labels the vertical frequency qR. We weight the Fourier coefficient at (n, qR) by

$$w(n,q) \coloneqq (1 + |n|/R + |q|)^p (1 + L^{-1} \operatorname{dist}(n, R\mathbb{Z})).$$

Thus $\operatorname{dist}(n, R\mathbb{Z}) \coloneqq \min_{j\in\mathbb{Z}} |n - Rj|$. The first factor measures the frequency in units of R. The second measures its horizontal displacement from a multiple of R. For $f_{n,q} \coloneqq \widehat{f}(n, qR)$ put

$$\|f\|_Z^2 \coloneqq \sum_{n,q} w(n,q)^2 |f_{n,q}|^2, \qquad \|f\|_B \coloneqq \sum_{n,q} w(n,q) |f_{n,q}|.$$

For sequences on the horizontal axis only the terms with $q = 0$ occur. Reflection, conjugation and Fourier coordinate projections are bounded in both norms. Since $\langle (n, qR)\rangle^b \le CR^b w(n,q)$ for $0 \le b \le p$, we have $\|f\|_{H^b} \le CR^b \|f\|_Z$.

Lemma 3.2. *Uniformly in R,*

$$\|fh\|_Z \le C\|f\|_B \|h\|_Z, \qquad \|fh\|_Z \le C\sqrt{L}\, \|f\|_Z \|h\|_Z.$$

A plane wave at k_v times an axial function in a bounded smooth family has bounded B norm.

Proof. The first assertion follows by submultiplicativity and Young's inequality. For the second, write $k = (n, q)$ for the two indices of the weight. Cauchy–Schwarz in each convolution output gives

$$\|fh\|_Z^2 \le \left(\sup_k \sum_l \frac{w(k)^2}{w(l)^2 w(k-l)^2} \right) \sum_{k,l} w(l)^2 |f_l|^2 w(k-l)^2 |h_{k-l}|^2.$$

The second factor is $\|f\|_Z^2 \|h\|_Z^2$. To bound the first, write $n = Rj + d$, $-R/2 < d \le R/2$. The first factor in w is comparable to $a(j,q) \coloneqq (1 + |j| + |q|)^p$. If $k = l + (k - l)$, the two input (j, q) indices add to the output up to a carry of $0, 1$, or -1 in j. For each fixed input remainder, the carry is fixed by the output. Consequently the squared ratio of the three a weights is bounded by a constant times the sum of the two reciprocal squared input weights, by $(A + B)^p \le 2^{p-1}(A^p + B^p)$ for $A, B \ge 0$. Summing over the first input (j, q) costs a constant since $\sum_{j,q} a(j,q)^{-2} < \infty$. For the second factor of w, the same argument with circular distance on $\mathbb{Z}/R\mathbb{Z}$ bounds the sum over the input remainder by $C \sum_{|d|\le R/2} (1 + |d|/L)^{-2} \le CL$. Finally the velocity components are integers, so shifting an axial envelope by k_v changes its remainder by one and its large-scale weight by at most a fixed polynomial in the envelope index. □

Let P_0 denote averaging in y, equivalently the Fourier projection onto $q = 0$, and let $P_\perp \coloneqq 1 - P_0$. For a solution u in gauge c_R, write $U \coloneqq P_0 u$, $w_\perp \coloneqq P_\perp u$, and $z_j \coloneqq (\Phi(U))_j$, where Φ is the axial Birkhoff map and $\Psi = \Phi^{-1}$. The index on z_j distinguishes the Birkhoff variables from the model time $z = 2|g|rt$. In the expansion of $|U + w_\perp|^2 (U + w_\perp)$, the two terms linear in $w_\perp$ are $2|U|^2 w_\perp + U^2 \overline{w_\perp}$. Both have zero y-average. Projecting the remaining terms and differentiating the real map Φ on the axial component gives the exact equations

$$\dot{z}_j = -i(\omega_j(I) + c_R) z_j + D\Phi_j(U)[-if], \tag{3.3}$$

$$f \coloneqq 2UP_0|w_\perp|^2 + \overline{U} P_0(w_\perp^2) + P_0(|w_\perp|^2 w_\perp),$$

$$i\dot{w}_\perp = (-\Delta + c_R + 2|U|^2) w_\perp + U^2 \overline{w_\perp} + N(U, w_\perp), \tag{3.4}$$

$$N(U,w) \coloneqq P_\perp(2U|w|^2 + \overline{U} w^2 + |w|^2 w).$$

Dots denote physical time derivatives, $I_j = |z_j|^2$, and ω_j are the ungauged frequencies. To justify (3.3) for H^p solutions, first prescribe a continuous axial H^p forcing. Approximate the forcing and initial data by smooth functions and use the integral equation in Birkhoff variables after removing the fixed diagonal j^2. The remaining frequency corrections are locally Lipschitz in ℓ^∞,

by the explicit mass term and the ℓ^2 action frequency map [45]. The chart and its differential are locally Lipschitz in the indicated spaces [24, 38]. This gives (3.3) and uniqueness for the forced integral equation. The same reasoning allows reconstruction by Ψ.

On F write $z_j = \sqrt{I_j}\exp(-i\theta_j)$, and put

$$J := (I_F - I_F^\infty)/r, \qquad \Theta := \theta_F - \theta_F^\infty(t), \qquad V_* := \Psi(z|_F). \tag{3.5}$$

Here $I_F = (I_j)_{j\in F}$, and similarly for the other restrictions. The terminal angles are the continuous real lifts $\theta_j^\infty(t) := \theta_j^\infty(0) + (\omega_j(I^\infty) + c_R)t$ on F. We use the lift of the actual angles near these nominal angles, so that Θ measures their displacement. In $z|_F$ we set all coordinates outside F to zero. The profile V_* uses the actual old coordinates and is smooth on a fixed compact parameter neighbourhood.

Two chart estimates will be used repeatedly. Write $\mathcal{A}$ for either chart and $\mathcal{F}_\mathcal{A} := D\mathcal{A}(0)$ for its linear Fourier part. We identify spatial functions with their Fourier coefficients in norms on either side of the chart. For a point x_* on the old smooth family and an axial perturbation y_* of Z norm at most Cr, we have

$$\|D\mathcal{A}(x_* + y_*) - D\mathcal{A}(x_*)\|_{Z\to Z} \le Cr(1 + R^{2-p}), \tag{3.6}$$

$$\|(D\mathcal{A}(x_*) - \mathcal{F}_\mathcal{A})h\|_Z \le C(R^{-1} + L^{-1}) \tag{3.7}$$

in the second line for $h = e^{iRv_x x}h_0(x)$ with fixed $v_x \neq 0$ and h_0 in a bounded smooth family. The same bound applies to finite sums with summable smooth-envelope bounds, by linearity over the reals.

For clarity, first drop the residue factor. On the axis the resulting norm Z_0 is equivalent to $\|\cdot\|_{L^2} + R^{-p}\|\cdot\|_{H^p}$ (and to the corresponding sum with ℓ^2 and h^p norms for coordinate sequences). Thus $\|h\|_{H^1} \lesssim R\|h\|_{Z_0}$, and $Rr = o(1)$. The full chart is analytic in L^2, while its difference from Fourier is analytic $H^1 \to H^2$ in both directions [38]. Integrating the second differential along $x_* + ty_*$ gives the first estimate in Z_0. The two terms are $O(r)$ and $O(R^{-p}(Rr)R)$, respectively. To restore the residue factor, note that $\|h\|_Z \asymp \|h\|_{Z_0} + \|\delta h\|_{Z_0}$, where $\delta := (R/L)(T_{2\pi/R} - 1)$ and T_ℓ is spatial translation, or its diagonal action on Birkhoff coordinates. Indeed the modulus of the Fourier multiplier of δ is $2(R/L)|\sin(\pi n/R)|$, comparable to $\operatorname{dist}(n, R\mathbb{Z})/L$. The charts commute with translation. Applying the finite difference to $D^2\mathcal{A}(x_* + ty_*)[y_*, h]$ therefore puts it on y_*, on h, or, by the mean value formula, on the base of the second differential. The last term uses $D^3\mathcal{A}$ with the additional argument $\delta(x_* + ty_*)$, bounded in H^1 by $C(L^{-1} + Rr)$ and in L^2 by $C(L^{-1} + r)$. This proves (3.6). At the smooth base x_*, $D\mathcal{A}(x_*)$ is bounded on L^2 and on H^p, hence also on Z_0. Applying the same translation identity to the first differential adds a term involving $D^2\mathcal{A}$ and $\delta x_* = O(L^{-1})$ in H^1. Its scaled H^p norm is $O(R^{1-p}/L)$ for an input of unit Z norm. This also gives $\|D\mathcal{A}(x_* + y_*)\|_{Z\to Z} \le C$, since $rR^{2-p} = o(1)$.

To prove (3.7), at a smooth base the differential minus its Fourier part maps $H^1 \to H^2$. Its real Hilbert adjoint gains the same derivative. Indeed if S is the full chart differential and $\mathcal{J}_{\text{in}}, \mathcal{J}_{\text{out}}$ are the constant symplectic matrices in the two sets of coordinates, then $S^*\mathcal{J}_{\text{out}}S = \mathcal{J}_{\text{in}}$. Applying this formula also to the linear Fourier part expresses the adjoint of the difference through the difference of the inverse differentials at corresponding points. Each $\mathcal{J}$ and its inverse preserve Sobolev norms. Duality first gives $H^{-2} \to H^{-1}$ for the chart differential minus Fourier. Interpolation with $H^1 \to H^2$ gives both $H^{-1} \to L^2$ and $H^{p-1} \to H^p$. A shifted smooth envelope h as in (3.7) has $\|h\|_{H^{-1}} = O(R^{-1})$ and $\|h\|_{H^{p-1}} = O(R^{p-1})$, by rapid Fourier decay of the envelope. In the two terms of Z_0 these bounds both give $O(R^{-1})$. If $A(x_*) := D\mathcal{A}(x_*) - \mathcal{F}_\mathcal{A}$, equivariance gives

$$\delta(A(x_*)h) = A(T_{2\pi/R}x_*)\delta h + (R/L)(D\mathcal{A}(T_{2\pi/R}x_*) - D\mathcal{A}(x_*))h.$$

The first term costs $O(L^{-1})$ since $\|\delta h\|_{Z_0} = O(L^{-1})$. In the second term $\|\delta x_*\|_{H^1} = O(L^{-1})$. The L^2 and scaled H^p estimates cost $O(L^{-1})$ and $O(R^{1-p}L^{-1})$. This proves the claimed bound.

3.2. **The transverse propagator.** We first retain from the linear part of (3.4) the old coefficients at V_*, together with all terms linear in the terminal axial planes rY_∞. Here

$$Y_\infty(t,x) \coloneqq \gamma \sum_i (c_i + e_i) e^{iK_i x - i(K_i^2 + \chi_{K_i}(I^\infty))t}, \qquad \chi_j(I) = \omega_j(I) - j^2 - 2\sum_k I_k.$$

By Lemma 2.3 and the $O(r^2)$ change of the actions, $\max_i |\chi_{K_i}(I^\infty)| = O(R^{-2} + r^2)$. The scalar $2\sum_{j\in F} I_j + c_R$ goes into the remainder because it equals $2r\sum_{j\in F} J_j - 2\sum_{j\notin F} I_j^\infty$ and is $O(r|J| + r^2)$. For each $m \coloneqq qR > 0$, use the two horizontal sequences

$$\big(\widehat{w_\perp}(n,m),\ \overline{\widehat{w_\perp}(n,-m)}\big)_{n\in\mathbb{Z}}.$$

Thus the label in the second copy is the index *before* conjugation. These two copies, taken over $q = 1, 2, \ldots$, represent all transverse Fourier coefficients. Multiplication by an axial potential acts within a copy. The term $U^2\overline{w_\perp}$ acts between copies and its upper-right entry at (n, n') is $\widehat{U^2}(n + n')$. In particular the blocks for different $m > 0$ evolve independently under the axial linearization.

We make the following three changes to the transverse linear equation. First remove the off-diagonal term formed solely from V_*. Next diagonalize the scalar old potential up to a prescribed order at large horizontal indices. Finally use the space-time harmonics of the old orbit to separate the terms linear in rY_∞. All three steps are performed in a single spatial norm.

We use the following version of the Schur test. Given a positive Hilbert weight v_ι, let the size of a matrix be the larger of the two suprema of its absolute row and column sums, after multiplying entry $(\iota, \jmath)$ by $\max(v_\iota/v_\jmath, v_\jmath/v_\iota)$. These sizes are submultiplicative, are unchanged by adjoints, and bound the operator norms on weighted ℓ^1 and ℓ^2. It is convenient to replace the weight in Lemma 3.2 by the equivalent weight

$$\widetilde{w}(n,q) \coloneqq (1 + |q| + (|n|/R - C_*)_+)^p (1 + (\operatorname{dist}(n, R\mathbb{Z})/L - 2)_+), \tag{3.8}$$

where $a_+ \coloneqq \max(a, 0)$. We take $C_* \coloneqq 2 + \max_{v\in\mathcal{V}} |v_x|$, so all axial indices $1 + RP_i$ lie in the flat part of the first factor for large R. Under a shift or reflection $n \mapsto \pm n + RP_i + b$, its ratios are bounded by $C(1 + |b|)^{p+1}$, and by a constant independent of R when $|b| \le C' \log R$. The same holds with $P_i = 0$. This bound covers both factors of the product weight in all Schur estimates below.

We begin with the order-one terms from V_*. In the doubled generator the two free blocks are $A \coloneqq D_x^2 + m^2$ and $-A$, $D_x \coloneqq -i\partial_x$. Write $C_{nn'} \coloneqq \widehat{V_*^2}(n + n')$, $D_{\text{free}} \coloneqq \operatorname{diag}(A, -A)$, and put

$$X_{nn'} \coloneqq -\frac{C_{nn'}}{n^2 + (n')^2 + 2m^2}, \qquad J_{\text{off}} \coloneqq \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix}.$$

A star denotes the ordinary Hilbert space adjoint on the Fourier sequences. Since $AX + XA = -C$, the off-diagonal entries of $[D_{\text{free}}, J_{\text{off}}]$ are $-C, C^*$, where $[D, J] = DJ - JD$. The denominator is at least $2R^2$, and the smooth old coefficients decay rapidly in $n + n'$. It follows from the Schur test above that J_{off} and its derivatives in the finitely many old angles and actions have size $O(R^{-2})$, uniformly in q. Use $e^{J_{\text{off}}}$ as the map from new to physical variables. In $e^{-J_{\text{off}}} D_{\text{free}} e^{J_{\text{off}}}$ the first commutator cancels the old off-diagonal pair $C, -C^*$. Each further commutator has an additional factor J_{off}. The time derivative costs $O(R^{-2})$ at bounded parameter speeds. The resulting scalar potential, after removal of the scalar specified above, is $B_0 \coloneqq 2(|V_*|^2 - \sum_{j\in F} I_j)$ in the upper copy and its signed conjugate in the lower copy. It is real and has zero spatial mean, by the mass identity for the axial chart.

For this scalar potential we construct an orthonormal basis depending on the old angles and actions. Holding the actions fixed gives the nominal angular derivative $\partial_* \coloneqq \sum_{j\in F}(\omega_j(I^\infty)+c_R)\partial_{\theta_j}$. Fix an integer k_* with $k_*\alpha \geq s+2\alpha$. For $|n|\geq L$ seek a column $e^{inx}f_n$ satisfying, to order $O(|n|^{-k_*})$,

$$(2nD_x + D_x^2 + B_0 - i\partial_*)f_n = d_n f_n.$$

To construct the column, write $\eta \coloneqq (2n)^{-1}$, $\mathcal{H} \coloneqq D_x^2 + B_0 - i\partial_*$, and use the finite expansions $f_n \coloneqq \sum_{j=0}^{k_*}\eta^j F_j$, $d_n^{\mathrm{app}} \coloneqq \sum_{j=0}^{k_*-1}\eta^j\lambda_j$, with $F_0 = 1$. Let $\langle h\rangle_x$ denote the spatial mean. If the coefficients through F_j have been found, set

$$S_j \coloneqq \mathcal{H}F_j - \sum_{a=0}^{j-1}\lambda_a F_{j-a}, \qquad \lambda_j \coloneqq \langle S_j\rangle_x, \qquad F_{j+1} \coloneqq -D_x^{-1}(S_j - \lambda_j).$$

The sum is empty for $j=0$. We take the inverse on mean-zero functions, with zero-mean output. Thus the coefficient of η^j in $(\eta^{-1}D_x + \mathcal{H})f_n - d_n^{\mathrm{app}}f_n$ vanishes for $0\leq j<k_*$. In particular $\lambda_0 = 0$, $F_1 = -D_x^{-1}B_0$, and $\lambda_1 = -\langle B_0 D_x^{-1}B_0\rangle_x = 0$. For the last identity pair the Fourier labels $l, -l$, whose denominators have opposite sign. Only finitely many smooth coefficients are needed, so the same bounds hold after any fixed number of old-parameter derivatives. Normalize each column in L^2. Its component of the defect perpendicular to the column is still $O(|n|^{-k_*})$. The parallel component of the normalized $(D_x^2 + B_0 - i\partial_*)$-column is real since the derivative of its squared norm vanishes. Taking this real expectation minus n^2 as d_n gives the asserted residual and $d_n = O(|n|^{-2})$. Indeed the normalizing factor is $1+O(n^{-2})$ since the first correction to f_n had zero mean. The residual Fourier kernels are bounded by $C_j|n|^{-k_*}(1+|n'-n|)^{-j}$ for every fixed j. The same argument applies in the conjugate copy.

The approximate columns need not be mutually orthogonal. We correct their Gram matrix, including at $|n| = L$. In the next paragraph P is the projection onto the horizontal indices $|n|\geq L$, $Q \coloneqq 1-P$, and $T : P\ell^2 \to \ell^2$ maps the n-th unit sequence to the normalized column just constructed. Set $D_h \coloneqq (D_x^2 + \operatorname{diag} d_n)|_{P\ell^2}$ and $H_* \coloneqq D_x^2 + B_0$. We have $H_*T - i\partial_* T - TD_h = E_*$ with Schur size $O(L^{-k_*})$. Multiplication by T^* and subtraction of the adjoint identity give

$$[D_h, T^*T] - i\partial_*(T^*T) = T^*E_* - E_*^*T.$$

Put $G \coloneqq T^*T$ on $P\ell^2$, and let $\mathfrak{D}B \coloneqq [D_h, B] - i\partial_* B$. This is a derivation, so $\mathfrak{D}(G-1)^j$ is the sum of j terms with one factor $\mathfrak{D}G$. Since $G-1 = O(L^{-1})$ in Schur size, the binomial series gives $\mathfrak{D}(G^{-1/2}) = O(L^{-k_*})$. The high-column map $U_h \coloneqq TG^{-1/2}$ is an isometry in unweighted ℓ^2, and its residual equals $E_*G^{-1/2} + T\mathfrak{D}(G^{-1/2})$. It is still $O(L^{-k_*})$. Extend it by the isometric low-column map

$$(1-U_hU_h^*)Q\left[Q(1-U_hU_h^*)Q\right]^{-1/2},$$

where the inverse is on Q. The resulting map is unitary since its two ranges are orthogonal and it is within $O(L^{-1})$ of the identity. Commuting D_x^2 through a high column of $T-P$ costs $O(1)$ in Schur size, since the factor $2n(n'-n)+(n'-n)^2$ multiplies an $O(1/n)$ smooth envelope. Product rules and the same binomial series bound the commutator with the full extension by $O(1)$. In the new basis the scalar operator, less D_x^2 and including the nominal connection, is Hermitian. Its high columns, and hence high rows, are the real diagonal d_n plus $O(L^{-k_*})$. The remaining bounded Hermitian matrix acts solely on $|n|<L$.

Denote the full scalar column map in the two copies by $\mathcal{U}$, with the conjugate construction in the second copy. Then $\mathcal{S} \coloneqq e^{J_{\mathrm{off}}}\mathcal{U}$ maps the resulting coordinates back to the original field. In an equation $i\dot{w} = \mathcal{L}w$, the substitution $w = \mathcal{S}a$ gives the generator $\mathcal{S}^{-1}\mathcal{L}\mathcal{S} - i\mathcal{S}^{-1}\dot{\mathcal{S}}$. Both $\mathcal{S}-1$ and its old-parameter derivatives have Schur size $O(L^{-1})$. Along nonnominal old parameters

the difference of the actual and nominal connections has size

$$CL^{-1}(|\dot I_F| + |\dot\theta_F - \dot\theta_F^\infty|). \tag{3.9}$$

In particular variations of the order-one low Hermitian matrix may be retained without any loss on the slow time scale.

It remains to decide which terms of order r are resonant. Use the Fourier series $\sum_{b,\nu} v_{b,\nu} e^{ibx - i\nu t}$ of the old axial profile with its retuned old actions and nominal angles. Here $b \in \mathbb{Z}$ is a spatial frequency and $\nu \in Q_F^{-1}\mathbb{Z}$ a time frequency, with Q_F a common denominator of the old occupied frequencies in the mass gauge. Thus $2\pi Q_F$ is an old period. The coefficient $v_{b,\nu}$ is the sum of all angular harmonics with that space-time frequency. A displacement $O(r^2)$ of the actions changes the summed coefficients by $O(r^2)$ in every fixed Fourier moment used here. Indeed the finite-angle parametrization is analytic into each fixed Sobolev space on a complex angular strip of positive width. For a multi-index $l = (l_j)_{j\in F} \in \mathbb{Z}^F$, the angular harmonic $e^{-il\cdot\theta}$ has the single spatial frequency $b = \sum_{j\in F} j l_j$, by translation covariance, and time frequency $\nu = \sum_{j\in F} l_j(\omega_j(I^\infty) + c_R)$. Its coefficient decays exponentially in $|l|_1 := \sum_j |l_j|$. Truncation at $|l|_1 \le C_A \log R$ leaves an $O(R^{-A})$ tail in Schur and multiplier norms, with $|b| + |\nu| \le C'_A \log R$ in the truncation. Choose A larger than every fixed power needed below. Changing the old phases and actions to their actual values in a term already of order r costs $O(r(|\Theta| + r|J|))$.

Pairing the terminal plane at K_i with $v_{b,\nu}$ gives an opposite-copy entry between n and $n' := K_i + b - n$. Here and below a bare divisor uses the free root frequency K_i^2, with its shift $\chi_{K_i}(I^\infty)$ treated separately. For this entry the divisor is

$$D := n^2 + (n')^2 + 2(qR)^2 - K_i^2 - \nu.$$

If $D \neq 0$, then $|D| \ge Q_F^{-1}$. If an endpoint has $|n| < L$ or $|n'| < L$, then $|D| \gtrsim R^2$, since $|q| \ge 1$, $L = o(R)$, and the other labels differ from the root by $O(L + \log R)$. On a zero-divisor edge q is bounded and both $|n|$ and $|n'|$ are between fixed positive multiples of R. Indeed, substituting $n' = K_i + b - n$ gives

$$2\big(n - (K_i + b)/2\big)^2 + 2(qR)^2 = K_i^2 + \nu - (K_i + b)^2/2.$$

The right-hand side is $O(R^2)$, so n/R and q are bounded. Along a subsequence q is a fixed nonzero integer and n/R has a limit X. Dividing by R^2 gives $q^2 = X(P_i - X)$, so X and $P_i - X$ are both nonzero. In particular $P_i = X + q^2/X$. Since the P_i are distinct, a row cannot serve two roots for arbitrarily large R. For the same root changing b changes $(n')^2$ by at least cR, whereas two retained temporal labels differ by $O(\log R)$. Thus the exact resonances form disjoint pairs.

The same-copy terms and their adjoints take n to $n + K_i - b$. Their bare divisor is the difference

$$(n + K_i - b)^2 - n^2 - (K_i^2 - \nu) = 2(K_i - b)(n - b) + \nu - b^2.$$

It has modulus $\gtrsim R$ unless $n = b$. Keep every entry $b \leftrightarrow K_i$ and its adjoint as a Hermitian coupling, with their actual time oscillations. The vertex b cannot meet an opposite-copy resonance. Intersection at K_i would require $P_i(P_j - P_i) = q^2$ in the limit, excluded by Proposition 2.4. This separates all opposite-copy pairs from the low Hermitian matrix and the retained same-copy terms.

For the discarded entries solve the homological equation using the bare divisors above. More explicitly, an entry $B_{uv}e^{-i\zeta t}$ in the generator with free energies E_u, E_v is cancelled to first order by the entry $-B_{uv}e^{-i\zeta t}/(E_u - E_v - \zeta_{\mathrm{bare}})$ in the exponent of the change of variables. The omitted root shift $\zeta - \zeta_{\mathrm{bare}}$ is $O(R^{-2} + r^2)$. Each Fourier harmonic defines a partial shift or reflection with at most one entry per row and column, so the summed generator has Schur size $O(r)$. Its terms touching the low subspace have size $O(r/R)$ for same-copy entries and $O(r/R^2)$ for opposite-copy entries. Its commutator with the bounded low matrix therefore costs

$O(r/R)$, and that with the high real diagonal costs $O(rL^{-2})$. The terminal planes oscillate at frequencies of order R^2. To keep those frequencies out of the error bound, combine the free commutator with the time derivative before estimating either one. Write the change of variables as $\mathcal{M} \coloneqq e^{\mathcal{B}(t)}$, where $\mathcal{B}$ is the exponent just constructed. Differentiating the conjugation with respect to θ gives the exact identity

$$\mathcal{M}^{-1} D_{\mathrm{free}} \mathcal{M} - i\mathcal{M}^{-1}\partial_t \mathcal{M} - D_{\mathrm{free}} = \int_0^1 e^{-\theta\mathcal{B}}([D_{\mathrm{free}}, \mathcal{B}] - i\partial_t \mathcal{B}) e^{\theta\mathcal{B}}\, d\theta.$$

On a Fourier entry the factor in parentheses multiplies the exponent's coefficient by its actual divisor. It equals the negative of the discarded entry up to the root shift estimated above, and has total Schur size $O(r)$. Since $\mathcal{B}$ itself has size $O(r)$, conjugation inside the integral changes this factor by $O(r^2)$. This gives $\mathcal{M} = 1 + O(r)$ and the required higher-order bound. Conjugation of the order-r terms by the old frame costs $O(r/L)$.

These constructions give a bounded inverse for the transverse inhomogeneous problem with mixed initial and terminal data. On the complement of the opposite-copy pairs, the free diagonal, the old low matrix, the real $d_n(t)$ and the retained same-copy terms generate an isometry in (3.8). Indeed each coupling within a given q and copy touches only indices where the weight equals $(1 + |q|)^p$. This holds on $|n| < L$, at $n = b$ with $|b| = O(\log R)$, and at $n = K_i$ with $\operatorname{dist}(K_i, R\mathbb{Z}) \le 1$. Every other entry is real diagonal. This holds for the old matrix evaluated at the actual, time-dependent parameters. On each opposite-copy pair remove the two free phases. Let $\delta_i \coloneqq \chi_{K_i}(I^\infty)$. Since the bare divisor vanishes, the upper-right entry in the equation for $i\partial_t$ is now $r\eta e^{-i\delta_i t}$, where $\eta \coloneqq 2\gamma(c_i + e_i)v_{b,\nu}$ is the coupling coefficient of this pair. The lower-left entry is its negative conjugate. The coordinate change $\operatorname{diag}(e^{-i\delta_i t/2}, e^{i\delta_i t/2})$ removes these phases and adds a diagonal of size $|\delta_i|/2$ in physical time. Thus in slow time $\tau = rt$ the principal matrix for ∂_τ is

$$\begin{pmatrix} 0 & -i\eta \\ i\bar{\eta} & 0 \end{pmatrix}.$$

It is Hermitian and its square equals $|\eta|^2$ times the identity, so its eigenvalues are $\pm|\eta|$. The rotations and spectral projections are uniformly bounded in the flattened weight because the two endpoint weights are comparable. Discarding the half-angle diagonal correction and the d_n's on these pairs costs $O((R^{-2} + r^2)/r)$.

The set of exact pairs can change with R. To choose one time weight for all large R, take the coefficients $v^0_{b,\nu}$ of the original periodic field V, before retuning. For every $a > 0$ the set of numbers

$$2|c_i v^0_{b,\nu}|, \qquad i \text{ an axial vertex}, \quad (b, \nu) \in \mathbb{Z} \times Q_F^{-1}\mathbb{Z},$$

contains only finitely many values above a. This follows from absolute summability and the finiteness of the tree. Choose $\sigma > 0$ below the model decay rate in τ and at positive distance from this set and zero. Each rate on a retained pair differs from its listed value by $O(r^2 + R^{-10} + R^{-A})$, by the retuning and angular tail estimates. Thus the distance of all signed rates from $-\sigma$ has a positive lower bound independent of R. For any $T < S$ measure inhomogeneities in the norm

$$\|f\|_{\sigma,Z} \coloneqq \sup_{T \le \tau \le S} e^{\sigma(\tau - T)} \|f(\tau)\|_Z.$$

Write $\mathcal{U}_0(\tau, u)$ for the propagator of this principal system, including the changes of phase on the pairs, from slow time u to τ. Denote by $\Pi_f(u)$ the projection at time u onto the decaying pair directions whose exponents are less than $-\sigma$, and put $\Pi_c(u) \coloneqq 1 - \Pi_f(u)$. These projections commute with propagation in the sense that $\mathcal{U}_0(\tau, u)\Pi_f(u) = \Pi_f(\tau)\mathcal{U}_0(\tau, u)$. For forcing f in

the equation for ∂_τ, define

$$(\mathcal{G}_\perp f)(\tau) \coloneqq \int_T^\tau \mathcal{U}_0(\tau,u)\Pi_f(u)f(u)\,du - \int_\tau^S \mathcal{U}_0(\tau,u)\Pi_c(u)f(u)\,du.$$

This prescribes zero fast stable component at T and zero complementary component at S. A scalar exponent μ contributes the time-weighted factor $e^{(\mu+\sigma)(\tau-u)}$. The first integral has $\mu+\sigma<0$ and $\tau-u\geq 0$, while the second has $\mu+\sigma>0$ and $\tau-u\leq 0$. The isometric part obeys the second estimate with $\mu=0$. Since all the exponents stay a fixed distance from $-\sigma$, both weighted kernels are bounded in operator norm by $Ce^{-\delta_{\mathrm{gap}}|\tau-u|}$ for some $\delta_{\mathrm{gap}}>0$ independent of R,T,S. Here δ_{gap} can be any positive lower bound for σ and the distances of the signed pair rates from $-\sigma$. Taking the Hilbert direct sum of the blocks before time integration gives $\|\mathcal{G}_\perp f\|_{\sigma,Z}\leq 2C\delta_{\mathrm{gap}}^{-1}\|f\|_{\sigma,Z}$, uniformly in the length of the interval. On a fixed bounded slow interval the same principal propagator is bounded in both time directions. Large unitary rotations used to obtain the pairwise estimate are needed only to define these propagators and projections.

3.3. **Resonant forcing and the nonlinear terminal problem.** We first choose the time phase of each carrier. A transverse wave has a decaying amplitude, but a new axial wave survives at positive infinity. We therefore use its exact terminal frequency, including the shift due to all occupied axial modes. Put $\kappa_v\coloneqq|k_v|^2+\varepsilon_v$, with $\varepsilon_{(P_i,0)}\coloneqq\chi_{K_i}(I^\infty)$ at every axial vertex. For each selected triangle $v+v'=u$ impose $\varepsilon_v+\varepsilon_{v'}=\varepsilon_u$. These conditions have a common solution. Starting from the root, split the assigned shift of each primary parent $m=p+q$ between its children in any fixed way, for instance $\varepsilon_p=0$ and $\varepsilon_q=\varepsilon_m$. When m is off the axis its split contains three additional triangles. Their equations read

$$\varepsilon_{J(p)}=\varepsilon_{A(p)}-\varepsilon_p,\qquad \varepsilon_{J(q)}=\varepsilon_{A(q)}-\varepsilon_q,\qquad \varepsilon_{T_m}=\varepsilon_{J(p)}+\varepsilon_{J(q)}.$$

Here $A(p),A(q)$ are on the axis and their shifts have already been specified. The auxiliary vertices are distinct at different splits by Proposition 2.4. This recursion gives $\varepsilon_v=O(R^{-2}+r^2)$ on the fixed tree.

Write $\mathbf{x}=(x,y)$, $\tau=rt$, and separate the known waves into their transverse and axial parts. In the following formulas we identify the axial polynomial Y with its sequence of Fourier coefficients.

$$W(t,\mathbf{x})\coloneqq\gamma\sum_{v_y\neq 0}x_v(2|g|\tau)e^{ik_v\cdot\mathbf{x}-i\kappa_v t},$$
$$Y(t,x)\coloneqq\gamma\sum_i\big(x_{(P_i,0)}(2|g|\tau)+e_i\big)e^{iK_ix-i\kappa_{(P_i,0)}t}.$$

The first sum is over the transverse vertices of $\mathcal{V}$. The retuning errors e_i put the limiting axial amplitudes at their prescribed values in Y_∞. By Lemma 2.5, for some $\lambda>\sigma$ and every $\epsilon>0$ one can choose T so that the amplitudes of $W, Y-Y_\infty$ and their first two derivatives in τ are bounded by $\epsilon e^{-\lambda(\tau-T)}$ for $\tau\geq T$. Such an amplitude derivative keeps the rapid carrier phase fixed.

In expanding $|V_*+r(Y+W)|^2(V_*+r(Y+W))$, the coefficient of r^2 is $2V_*|Y+W|^2+\overline{V_*}(Y+W)^2$. Terms involving only Y are axial and are already included in the Birkhoff frequencies. Terms involving just one transverse factor also have zero axial projection. Thus the order-r^2 transverse forcing and the axial forcing after subtracting the mean transverse density are given by

$$Q_w\coloneqq P_\perp\{2V_*(|Y+W|^2-|Y|^2)+\overline{V_*}((Y+W)^2-Y^2)\},$$
$$Q_z\coloneqq P_0\{2V_*(|W|^2-S_W)+\overline{V_*}W^2\},\qquad S_W\coloneqq\sum_{v:\,v_y\neq 0}|x_v(2|g|\tau)|^2. \tag{3.10}$$

In fact S_W is the spatial mean of $|W|^2$, by orthogonality of its distinct Fourier modes. To handle this mean exactly, take $2r^2 S_W U$ out of the full axial force f in (3.3). Gauge equivariance, differentiated along the curve $a \mapsto e^{-ia}U$ at zero, gives

$$D\Phi(U)[-2ir^2 S_W U] = -2ir^2 S_W z, \qquad 2\,\mathrm{Re}(\bar{z}_j(-2ir^2 S_W z_j)) = 0.$$

Thus the mean changes each angular speed by $2r^2 S_W$. Its contribution on a new axial mode of size $O(r)$ is $O(r^3 S_W)$. The remaining known force has leading term $r^2 Q_z$. Its outputs have nonzero horizontal velocity. Indeed a zero difference in horizontal velocity would require the same vertex, whose diagonal density has just been removed, and a zero sum would require opposite horizontal velocities. Both statements use the distinct absolute horizontal components in Proposition 2.4. This observation permits use of (3.7) before projecting to the old coordinates. In extracting the prescribed harmonics of Q_w, Q_z we use the nominal old angles. Replacing the actual V_* by that profile costs $C(|\Theta| + r|J| + r^2)$ times a decaying amplitude. The low linear operator continues to use the actual old coordinates.

We determine the time frequency ζ and spatial frequency k of each remaining product. Put $e = (1, 0)$ and write v_x for the horizontal component of v. For two positive carriers and a conjugated old harmonic (b, ν), one has $k = (2 - b)e + R(v + v')$ and $\zeta = \kappa_v + \kappa_{v'} - \nu$. Direct expansion gives

$$\zeta - |k|^2 = -2R^2 v \cdot v' + 2R(b - 1)(v_x + v'_x) + 2 - \nu - (2 - b)^2 + \varepsilon_v + \varepsilon_{v'}.$$

For a carrier times a conjugated carrier and a positive old harmonic, $k = be + R(v - v')$, $\zeta = \kappa_v - \kappa_{v'} + \nu$, and

$$\zeta - |k|^2 = 2R^2(v - v') \cdot v' + 2R(1 - b)(v_x - v'_x) + \nu - b^2 + \varepsilon_v - \varepsilon_{v'}.$$

When the coefficient of R^2 is nonzero it is an integer of modulus at least one. If it vanishes, Proposition 2.4 puts the nonconstant output velocity in the tree, with nonzero horizontal component.

When $b \neq 1$ the linear term then dominates the logarithmically bounded remainder. When $b = 1$, the only desired temporal harmonic is $\nu = 1$. All others have separation at least $1/Q_F$ before detuning. Consider the surviving harmonic $(b, \nu) = (1, 1)$ on one triangle $p + q = h$. The assigned frequency corrections give $\kappa_p + \kappa_q - \kappa_h = 1$, so the product at each output has exactly that carrier's phase. After removing r^2 and the carrier, the two orderings at the parent and the density term at a child have coefficients

$$\begin{aligned} 2\bar{g}\,\gamma^2 x_p x_q &= 2i|g|\gamma\, x_p x_q, \\ 2g\,|\gamma|^2 x_h x_q &= -2i|g|\gamma\, x_h x_q, \end{aligned} \qquad \gamma = ig/|g|.$$

On the left of the evolution equation the amplitude derivative is $2i|g|\gamma\, x'_v$, with the prime taken in $z = 2|g|rt$. The first line gives the parent equation and the second gives the child equation in (2.4). Their contributions add at shared vertices. The change in g from old-action retuning is $O(r^2)$, and the omitted angular tail has the bound above. Every occurrence of a new radius error e_i in these products is paired with a decaying transverse amplitude.

There is also a linear term at each surviving axial index K_i. Its carrier has the exact frequency $\kappa_{(P_i,0)} = \omega_{K_i}(I^\infty) + c_R$. When the actual diagonal operator acts on that carrier, the remaining frequency difference is $\omega_{K_i}(I) - \omega_{K_i}(I^\infty)$. It tends to zero with $I - I^\infty$, at the rate estimated below. Using the free frequency K_i^2 here would leave the constant terminal shift multiplying a nondecaying wave.

For each remaining retained harmonic $a(\tau)e^{ik\cdot(x,y)-i\zeta t}$ in the nominal forcing put $E_{k,\zeta} \coloneqq e^{ik\cdot\mathbf{x}-i\zeta t}$ and use

$$(i\partial_t+\Delta)\frac{a(\tau)E_{k,\zeta}}{\zeta-|k|^2}=a(\tau)E_{k,\zeta}+\frac{ir\,(\partial_\tau a)(\tau)E_{k,\zeta}}{\zeta-|k|^2}.$$

The quotient on the left is the particular solution for a frozen amplitude. The second term displays the gain of r when the amplitude varies. Write P_w, P_z for the resulting sums, where on the axis we only take indices outside F. For large R all retained axial outputs here have $|n|\gtrsim R$. The absolute divisors are at least $1/(2Q_F)$. A transverse output with zero horizontal velocity must have a nonzero R^2 coefficient, so its particular carries $O(R^{-2})$. This pays for applying the entire bounded old low matrix. Every other particular lies at $|n|\gtrsim R$, where the real diagonal shifts are $O(R^{-2})$. On the axis the same bound holds for the empty-site renormalized frequency of the smooth old state. Its further change at the actual actions is $O(r|J|+r^2)$ in the gauged frequency. The scalar mass discrepancy is included in this last estimate. Thus the bounded divisors are used in the dressed transverse coordinates and in the true axial coordinates. The B norms of the particular sums and of the sums with one slow amplitude derivative are bounded by $C\epsilon e^{-\lambda(\tau-T)}$, by absolute old-harmonic summation. This construction leaves errors of size $O(r)$ after division by r^2 when a slow derivative falls on the particular amplitudes.

We now seek the actual field as the sum of these prescribed terms and an error. Write $F^c \coloneqq \mathbb{Z}\setminus F$. Let h be the transverse error after both coordinate changes, and $b=(b_j)_{j\in F^c}$ the axial error in Birkhoff coordinates. Each error is divided by r. With B_w, B_z denoting the prescribed profiles, set

$$w_\perp=\mathcal{S}(B_w+r\mathcal{M}h),\qquad z|_{F^c}=B_z+rb,\qquad B_w\coloneqq rW+r^2P_w,\quad B_z\coloneqq rY+r^2P_z. \tag{3.11}$$

In the first identity the left side denotes the doubled transverse field described above, and W, P_w represent their doubled versions. Every transverse index and every axial index in F^c is retained. In particular $\mathcal{M}$ only acts on the unknown h. The terms of the mixed operator on the known fields were already included in Q_w.

For reference we specify the exact equations that the terminal integral map solves. Reconstruct the old coordinates from (3.5) as $z_j=\sqrt{I_j^\infty+rJ_j}\,\exp[-i(\theta_j^\infty+\Theta_j)]$, and put $U=\Psi(z)$. Let $\Lambda\coloneqq\operatorname{diag}(\omega_j(I)+c_R)$, $D_f\coloneqq D\Phi(U)[-if]$, and let $\mathcal{L}(U)$ be the doubled linear generator in (3.4). Denote the doubled nonlinear term by $\mathcal{N}\coloneqq(N,-\overline{N})$, with the appropriate vertical components taken in each copy. Set $\mathcal{K}\coloneqq\mathcal{S}^{-1}(\mathcal{L}(U)\mathcal{S}-i\dot{\mathcal{S}})$. Write Λ_{F^c} for the restriction of Λ to F^c. All four equations below have the same slow time on the left. The factors of r can be seen, for example, from $\partial_t(B_z+rb)=\dot B_z+r^2\partial_\tau b$. For $j\in F$, $\dot I_j=2\operatorname{Re}(\bar z_j\dot z_j)$ and $\dot\theta_j=-\operatorname{Im}(\dot z_j/z_j)$. Substituting the reconstruction into the equations therefore gives

$$\begin{aligned}
i\partial_\tau h&=r^{-1}\mathcal{M}^{-1}(\mathcal{K}\mathcal{M}-i\dot{\mathcal{M}})h+r^{-2}\mathcal{M}^{-1}(\mathcal{K}B_w-i\dot B_w+\mathcal{S}^{-1}\mathcal{N}),\\
\partial_\tau b&=-ir^{-1}\Lambda_{F^c}b+r^{-2}(-i\Lambda_{F^c}B_z-\dot B_z+(D_f)|_{F^c}),\\
\partial_\tau J_j&=2r^{-2}\operatorname{Re}(\bar z_j(D_f)_j),\qquad j\in F,\\
\partial_\tau\Theta_j&=r^{-1}(\omega_j(I)-\omega_j(I^\infty)-\operatorname{Im}((D_f)_j/z_j)),\qquad j\in F.
\end{aligned} \tag{3.12}$$

On a trial path evaluate $\dot{\mathcal{S}}$ by the chain rule using the instantaneous physical velocities $2\operatorname{Re}(\bar z_j(D_f)_j)$ and $\omega_j(I)+c_R-\operatorname{Im}((D_f)_j/z_j)$. This prescription only uses values of the trial path. At a fixed point these are its actual velocities, so that (3.12) is equivalent to (3.3)–(3.4). The dots on $\mathcal{M}, B_w, B_z$ are ordinary physical-time derivatives of prescribed functions.

For a normed space D put $\|a\|_{\sigma,D}\coloneqq\sup_{T\le\tau\le S}e^{\sigma(\tau-T)}\|a(\tau)\|_D$. Set

$$H_1\coloneqq\|h\|_{\sigma,Z},\quad A_1\coloneqq\|b\|_{\sigma,Z},\quad J_1\coloneqq\|J\|_{\sigma,\mathbb{R}^F},\quad L_1\coloneqq\|\Theta\|_{\sigma,\mathbb{R}^F},\quad X\coloneqq H_1+A_1+J_1+L_1.$$

Here $\mathbb{R}^F$ carries its Euclidean norm. For a vanishing trial error the residual is bounded in these norms by a constant times

$$e_R \coloneqq L^{-1} + R^{-1} + R^{-2}/r + r(1 + R^{2-p}) + L^{-k_*}/r + r = O(R^{-\alpha}). \tag{3.13}$$

Indeed $R^{-2}/r = a_0^{-1} R^{-(2-s)}$, $rR^{2-p} = a_0 R^{-(p+s-2)}$, and $L^{-k_*}/r \le a_0^{-1} R^{-2\alpha}$. The scale $L^{-1} = R^{-\alpha}$ bounds their sum along with the remaining terms.

On $[T, S]$ we define a map from trial errors to new errors. Evaluate the coefficients and remainders in (3.12) at the trial path. Solve for the new h using the transverse Green operator, keeping the Hermitian low matrix at its trial values. For the new b, retain its exact real diagonal and prescribe $b(S) = 0$. Thus a source $f_j(\tau)$ in its equation contributes, at $j \in F^c$, the integral

$$-\int_\tau^S \exp\!\left(\frac{i}{r}\int_\tau^u [\omega_j(I(v)) + c_R]\, dv\right) f_j(u)\, du.$$

The diagonal exponential has modulus one. Multiplication by the time weight therefore bounds this integral in Z by $\sigma^{-1}\|f\|_{\sigma,Z}$. Integrate the J, Θ equations backwards with zero terminal values, with the same bound. Write H_1', A_1', J_1', L_1' for the norms of the four outputs of this map. For $X \le 1$, $\sqrt{L}X \le 1$, they satisfy

$$\begin{aligned}
H_1' &\le C(e_R + \epsilon X + \sqrt{L}X^2 + o(1)H_1),\\
A_1' &\le C(e_R + J_1 + \epsilon X + \sqrt{L}X^2 + o(1)A_1),\\
J_1' &\le C(e_R + \epsilon H_1 + \epsilon^2 A_1 + \sqrt{L}X^2),\\
L_1' &\le C(e_R + J_1 + r(A_1 + H_1) + \sqrt{L}X^2).
\end{aligned} \tag{3.14}$$

The constants here depend on the fixed background, tree and weight σ. They are uniform in R sufficiently large, in $S > T$, and in the starting time for the small tail. The $o(1)$ coefficients tend to zero as $R \to \infty$ for these fixed choices.

For the proof put $E \coloneqq e^{-\sigma(\tau - T)}$. First isolate the error of the axial chart on the prescribed waves. Define

$$e_0 \coloneqq r^{-1}\big[\Psi(z|_F + rY) - V_*\big] - Y, \qquad e_1 \coloneqq r^{-1}\big[U - \Psi(z|_F + rY)\big],$$

using Fourier series on the field side. In the transverse variables set $W_1 \coloneqq (\mathcal{S} - 1)W + r\mathcal{S}P_w$ and $h_1 \coloneqq \mathcal{S}\mathcal{M}h$. Then

$$U = V_* + rY + re_0 + re_1, \qquad w_\perp/r = W + W_1 + h_1,$$

where

$$\|e_0\|_Z \le Ce_R, \quad \|e_1\|_Z \le C(A_1 + r\epsilon)E, \qquad \|W_1\|_B \le C(L^{-1} + r)\epsilon E, \quad \|h_1\|_Z \le CH_1E.$$

Indeed e_0 is the Taylor error in replacing $\Psi(z|_F + rY)$ by $V_* + rY$, divided by r. Equations (3.6)–(3.7) at the actual old state give its bound. For e_1 use the bounded full chart differential on the $O(r)$ tube. For W_1, h_1 use the two Schur bounds in (3.11). All prescribed Fourier polynomials have uniform multiplier norms at their indicated amplitude.

Transverse field. Put $v \coloneqq U - V_*$. The additional linear terms on a physical field w are exactly

$$2(\overline{V_*}v + V_*\bar{v})w + 2V_*v\bar{w} + 2|v|^2 w + v^2\bar{w}.$$

Their order-r^2 contributions on W have already been included in Q_w. Their order-r coefficients on h from Y_∞ are precisely the mixed terms of the preceding subsection. Replacing Y_∞ by Y adds $O(\epsilon)H_1$ in slow units. The factor e_0 against the known multiplier W costs $O(e_R)$. Against a general error it costs $O(\sqrt{L}\, e_R)H_1 = o(1)H_1$. The factor e_1 against W costs $O(\epsilon(A_1 + r\epsilon))$,

while against h_1 it costs $O(\sqrt{L}(A_1 + r\epsilon)H_1)$. For example, two transverse unknowns and a smooth old factor satisfy

$$r^{-2}\|P_\perp(V_*|rh_1|^2)\|_Z \le C\sqrt{L}\,H_1^2E^2, \qquad \|Wh_1\|_Z \le C\epsilon H_1E^2.$$

In the second inequality the known W is a B multiplier. The different estimates for these two products explain the single loss $\sqrt{L}$ at quadratic order. Terms with two factors of v have an additional r. In $N(U,w)$ the two known transverse factors paired with V_* are included in Q_w. A single unknown pairs with a decaying known factor, and two unknowns cost $O(\sqrt{L}X^2)$. The cubic all-error term is $O(rLX^3)$. These bounds follow by applying Lemma 3.2 once for each additional arbitrary factor, using B for the known ones.

The old-frame residual on a known transverse mode costs $O((R^{-2} + L^{-k_*})/r)$. Its horizontal index is comparable to R, so its own scalar diagonal is only $O(R^{-2})$. The scalar $2\sum_F I_j + c_R$ on the known modes costs $C\epsilon(J_1 + r)$, and on h costs $C(J_1 + r)H_1$. The extra connection (3.9), divided by r, is small. The true velocities at the trial point give $|\dot I_F| + |\dot\theta_F - \dot\theta_F^\infty| \le C(rJ_1E + r^2)$ on our tube. To see the r^2 in this estimate, use $\|w_\perp\|_Z \le Cr(\epsilon + H_1)E$, expand products with known multipliers first, and use $\sqrt{L}X \le 1$ on the remaining pair of arbitrary factors. Terms due to freezing the mixed coefficients at the nominal phases carry $|\Theta| + r|J|$. The discarded nonresonant homogeneous terms have slow operator norm $O((R^{-2} + L^{-k_*})/r + L^{-1} + r)$ in addition to these modulation errors. These observations give the first line of (3.14).

Axial modes outside F. The ℓ^1 difference of the actual and terminal actions is bounded by

$$C\big(rJ_1 + r^2(\epsilon + A_1 + r\epsilon)\big)E.$$

Indeed on F we have $I_j - I_j^\infty = rJ_j$. On F^c put $z^\infty := rY_\infty$, with its true time phases, and use

$$\sum_{j\in F^c} \big||z_j|^2 - |z_j^\infty|^2\big| \le \|z|_{F^c} - z^\infty\|_{\ell^2}\big(\|z|_{F^c}\|_{\ell^2} + \|z^\infty\|_{\ell^2}\big).$$

The last factor is $O(r)$, while the first contains $r(Y - Y_\infty)$, r^2P_z, and rb. This bounds the shift of every gauged axial frequency by the mass formula and the local ℓ^2-Lipschitz bound for χ. On a persistent new root the frequency shift multiplies $O(r)$, and division by r^2 accounts for the unsuppressed J_1 in the second line. The diagonal on the error b itself is retained. After removing the exact density rotation, f contains two transverse factors or three. One unknown transverse factor times a known one costs $C\epsilon H_1$, two arbitrary ones $C\sqrt{L}X^2$. On the leading known terms the price for replacing $D\Phi(U)$ by $D\Phi(V_*)$ is $O(r(1 + R^{2-p}))$, and that for replacing the latter by Fourier is $O(R^{-1} + L^{-1})$ on the high horizontal outputs. All these bounds apply to the strong norm Z.

Old actions and angles. For each $j \in F$ the density rotation drops out of the action equation by $\mathrm{Re}(-i|z_j|^2) = 0$. The other known order-$r^2$ terms have high horizontal index after the old angular truncation, so their Fourier projection to F vanishes. The same chart estimate gives $O(e_R)$. This works at the actual old angles, hence produces no order-one coefficient of L_1. An axial error in the old-coordinate forcing has to act on a transverse pair. The latter either consists of two known decaying factors, or has an additional unknown. Differentiating the chart with respect to an axial change of size rA_1E costs at most $Cr(1 + R^{2-p})A_1E$ by the proof of (3.6). These estimates give the third line. For the angular equation division is by r, and the old z_j are bounded away from zero for this fixed insertion. The frequency shift just obtained gives $C[J_1 + r(\epsilon + A_1)]E$. The remaining force gains an additional factor r compared with the action equation. Since $r \lesssim e_R$, this proves the last line. Finally, the bounds for the particulars, their differentiated amplitudes, and the omitted angular tails give contributions $O(e_R)E$ in each line. Every pure error evaluated in (3.12) has a decaying factor because the axial state at infinity is exactly V_R.

3.4. **Passage to the half-line.** The only linear terms without a small coefficient on the right of (3.14) lead from J_1 to A_1' and L_1'. Accordingly give the latter two components more room. Choose constants $M \geq 1$ and $M' \geq M$ with $M' \geq 8CM$, increasing C in (3.14) to at least one. On the set

$$H_1, J_1 \leq M\rho, \qquad A_1, L_1 \leq M'\rho$$

we have $X \leq 4M'\rho$. The terms CJ_1 use at most $M'\rho/8$ of the corresponding bounds. Choose $4C\epsilon M' \leq M/8$, with $\epsilon \leq 1$, and then choose T to ensure the required tail bound for the model. Keep these choices fixed. Since

$$\frac{e_R}{\rho} = O(R^{-\alpha/4}), \qquad \sqrt{L}\,\rho = R^{-\alpha/4},$$

taking R sufficiently large makes $C(e_R + \sqrt{L}X^2) = o(\rho)$, $rX = o(\rho)$, and $\sqrt{L}X \leq 1$. The $o(1)$ coefficients in (3.14) are absorbed at this last step. Thus the terminal map takes the displayed set into itself for every $S > T$.

To apply the compact fixed point theorem, fix R and S. For this fixed R, the space Z is equivalent to H^p on the chosen lattice, with the corresponding sequence interpretation on the axis. Use continuous paths in H^s, and ordinary continuous paths for the two finite-dimensional variables, as the ambient space. Intersect the four bounds above with a common Lipschitz bound in H^{p-2} for the two field variables and in Euclidean norm for the old variables. This intersection contains zero and is convex. The image of the terminal map satisfies such a bound with a constant depending on R and the tree. Indeed an output derivative in slow time is its principal generator times the output plus its source. The free generators divided by r map H^p to H^{p-2} at fixed R, and the remaining bounded operators and sources have uniform bounds on each compact interval. In particular the local bounds can be taken independently of the right endpoint S.

For completeness, spatial tails beyond Fourier radius N of an H^p-bounded path have H^s norm at most $\langle N\rangle^{s-p}$ times the H^p bound. Time increments have H^s norm $O(|\tau - \tau'|^{(p-s)/2})$ by interpolating the uniform H^p bound and Lipschitz H^{p-2} bound. Thus Arzelà–Ascoli gives relative compactness in the path topology. Lower semicontinuity, applied first to finite Fourier sums, shows that the four bounds and the Lipschitz constraint pass to the limit. We have obtained a compact convex domain for the map.

We check continuity on this domain at fixed R, S. Convergence in $C_\tau H^s$ suffices for all the pointwise products since $s > 1$. The chart and its differential are continuous at this regularity. The principal low matrix and $\mathcal{S}$ depend smoothly on the finitely many old parameters. The differentiated frame uses the instantaneous velocities defined after (3.12). There is one unbounded operator to check in the field equation. After separating D_{free}, it enters the remainder through commutators with the changes of coordinates. The preceding Schur bounds for these commutators hold also with unit weight. With the weight w they give bounded operators on H^p at fixed R. Interpolation gives boundedness on H^s and, by the same argument applied to operator differences, continuity there. For $\mathcal{M}$ the computed commutator includes its time derivative, which is bounded for fixed R by the finite angular truncation. Thus convergence of trial paths in $C([T, S], H^s)$ suffices for the bounded remainder. The free generator is kept in the propagator. The differences of axial diagonal generators are bounded operators tending to zero since the actions converge in ℓ^1 and the map χ is locally Lipschitz on ℓ^2. For two principal evolutions with the same free generator, Duhamel's formula bounds the difference of propagators on the finite interval by the integral of the operator-norm difference of their bounded parts times the two evolution bounds. The boundary projections on the opposite-copy pairs are fixed, while the trial-dependent low matrix acts only on their complement. This proves continuity of the terminal integral map.

Schauder's theorem gives a fixed point. Letting $S \to \infty$, take a diagonal subsequence on compact intervals in the same strong path topology and pass to the equations in integral or distributional form. The exponential bounds persist. At the limit the reconstructed field is a solution by (3.12) and the forced axial conjugacy above. For fixed R, the decaying variables then imply exponential convergence in H^s to V_R. In particular the conserved mass of this solution is the mass used to define c_R.

At $t = T/r$, lower semicontinuity gives H^p data with normalized errors of size $O(\rho)$. Solving the Cauchy problem from these data gives a path continuous in H^p, which by uniqueness in H^s agrees with the compactness limit on their common interval. Thus at fixed R its Z norm is continuous, as needed in the following continuation argument. Extend the solution backwards by the defocusing Cauchy flow [36]. On the fixed slow interval $[0, T]$ the model coefficients and their derivatives are bounded by a finite constant. The same calculation along this actual trajectory now has normalized source bounded by $C_T(e_R + X + \sqrt{L}X^2)$, where X here denotes the sum of the instantaneous error sizes. The principal transverse propagator is bounded by C_T in both directions, and the axial diagonal is isometric. On any continuation interval on which $\sqrt{L}X \le 1$, the nonlinear term obeys $\sqrt{L}X^2 \le X$. Duhamel's formula, integrated backwards, gives

$$X(\tau) \le C_T(\rho + Te_R) + C_T \int_\tau^T X(u)\, du.$$

Gronwall yields $X(\tau) \le C_T'\rho$. For large R this gives $\sqrt{L}X < 1/2$ on the whole continuation interval, so by continuity the interval reaches zero. Here persistence and uniqueness in H^p allow continuation of the same field at the auxiliary order.

To read off the physical field, $\mathcal{S} - 1 = O(L^{-1})$ on each known transverse plane, the particulars cost $O(r^2)$, and each free high error is $O_T(r\rho)$ in Z, uniformly in positive time. The axial chart estimates replace the added finite axial planes by their physical Fourier series with error $O(re_R + r\rho)$ in Z. At old indices the action differs from the original one by $O(r^2 + r|J|)$ and the angle by Θ, uniformly in time because the old terminal angular speeds are locked. On this finite-dimensional torus the radii stay positive and the chart is smooth into every fixed Sobolev space. Consequently $\sup_{t\ge 0} \|V_*(t) - V(t)\|_{H^s} = O_T(\rho + r^2)$. The e_i cost $O(rR^{-10})$. Converting high errors to H^s costs at most CR^s, and $rR^s = a_0$ is fixed. This proves (3.1). Conversion at s_0 costs $O(R^{s_0}r) = o(1)$ even for the entire added planes. At the main root the old smooth field has Fourier coefficient $o(r)$, uniformly over its old angle torus. For the remainder f_R in (3.1) the single-coefficient estimate gives

$$|\widehat{f_R}((K, 0), t)| \le \langle K\rangle^{-s} \|f_R(t)\|_{H^s} = o(r)$$

uniformly in t, since $K = 1 + RP$ and $r = a_0R^{-s}$ with P, a_0 fixed. Orthogonality of the finitely many distinct planes at time zero gives the asserted input cost. This completes the proof of Proposition 3.1.

4. Persistence along one trajectory

We now return to all vertical frequencies. To use Proposition 3.1 a second time, we must preserve the part of the previously constructed orbit that led to its periodic asymptote. We first describe the linearized equation at this asymptote.

4.1. **The previous orbit.** Let V be smooth, axial and periodic of period $T_0 > 0$ in the equation with real constant potential c. Linearizing the cubic product at V gives the equation for a perturbation h,

$$i\partial_t h = (-\Delta + c + 2|V|^2)h + V^2\bar{h}. \tag{4.1}$$

Let $L(t,v)$ take the value of h at time v to its value at time t. It is real linear because the equation involves $\bar{h}$. The operator norms below are taken on the indicated Sobolev spaces. The one-period map $P \coloneqq L(T_0,0)$ is its monodromy. If Π is a spectral projection of P, its transport to time t is $\Pi(t) \coloneqq L(t,0)\Pi L(0,t)$. Since Π commutes with P, this transported projection is T_0-periodic.

Lemma 4.1. *On each $H^b(\mathbb{T}^2)$, $b \geq 0$, the spectrum of the complexification of P off the unit circle consists of isolated eigenvalues of finite algebraic multiplicity. These eigenvalues and their generalized eigenspaces are independent of b, and the eigenspaces consist of smooth functions. Finite sums of the associated spectral projections, transported by $L(t,0)$, are bounded from H^{b_0} to H^b uniformly in t, for $0 \leq b_0 \leq b$.*

Proof. Double the equation using $h, \bar{h}$ as independent fields, both expanded in the usual spatial Fourier basis. The generator for $i\partial_t$ is the sum of the free diagonal $D \coloneqq \operatorname{diag}(1-\Delta, -(1-\Delta))$, a diagonal matrix of real smooth potentials, and the off-diagonal entries $V^2, -\bar{V}^2$. In these doubled coordinates the spectra and projections below are taken over the complex numbers. Set $J \coloneqq \left(\begin{smallmatrix} 0 & X \\ X^* & 0 \end{smallmatrix}\right)$. For $k, l \in \mathbb{Z}^2$ the matrix of X in the Fourier basis is

$$X_{kl} \coloneqq -\frac{\widehat{V^2}(k-l)}{\langle k\rangle^2 + \langle l\rangle^2}.$$

Multiplying X_{kl} by $\langle k\rangle^2 + \langle l\rangle^2$ gives the negative of the upper off-diagonal entry. The lower entry cancels by taking adjoints with the opposite sign. For the smoothing bound, $\langle k\rangle^2/(\langle k\rangle^2+\langle l\rangle^2) \leq 1$ and $\langle k\rangle^b/\langle l\rangle^b \lesssim_b \langle k-l\rangle^b$. After inserting the Sobolev weights, every entry is bounded in modulus by a fixed summable function of $k-l$, since V is smooth and periodic. The same calculation with $\partial_t(V^2)$ proves $J, \partial_t J : H^b \to H^{b+2}$ uniformly over a period. In the coordinates given by e^J, the conjugated generator is therefore the scalar diagonal Schrödinger operator plus a remainder gaining two derivatives. Indeed the first commutator with the free diagonal is bounded at every Sobolev order and every further commutator has a factor J. The differentiated coordinate change has the same gain. Duhamel's formula on a period and $e^{\pm J} - 1 : H^b \to H^{b+2}$ now give $P = P_D + K$, where P_D is the scalar diagonal monodromy and $K : H^b \to H^{b+2}$.

Here and below a scalar evolution refers to either of the two diagonal equations. Both scalar evolutions are unitary on L^2. Commuting a spatial derivative of integer order m through the potential produces only derivatives of order at most $m-1$ on the solution. Integration and induction give polynomial bounds in $|t-v|$ on H^m, in either time direction. Interpolation gives the same conclusion for real $b \geq 0$. At integer multiples of T_0 these are polynomial bounds on both P_D^n and P_D^{-n}, $n \geq 1$. The resolvent of P_D thus exists off the unit circle, by the geometric series for P_D or P_D^{-1}. Compactness of K and the analytic Fredholm theorem give the spectral assertion, applied separately inside and outside the unit circle. Indeed, for $|\lambda| \neq 1$,

$$\lambda - P = \left[1 - K(\lambda - P_D)^{-1}\right](\lambda - P_D).$$

The bracket is the identity plus a compact operator. The product is invertible at $\lambda = 0$, since the linear evolution is reversible, and for $|\lambda| > \|P\|$ by a geometric series. These give one invertible point in each of the two connected regions. For an eigenvector the identity $h = (\lambda - P_D)^{-1}Kh$ gains two derivatives. Induction gives the same conclusion along a Jordan chain. The resolvents on different Sobolev spaces agree on their common domains wherever both exist, so the corresponding finite-rank projections agree. A bounded finite-rank projection at the lower order with smooth range is bounded into H^b. Transport and periodicity complete the proof. Projections defined by conjugation-invariant spectral sets preserve the physical real subspace. □

Here is the nonlinear consequence we need. Its weak-norm assumption controls the finitely many rapidly expanding old directions. The strong-norm assumption is made only at the initial time.

Proposition 4.2. *Let $1 < s_0 < s$. Suppose U, V are solutions in the same constant gauge c, V is as above, and $\|U(t) - V(t)\|_s \le Ce^{-\gamma t}$ for $t \ge 0$, with $\gamma > 0$. Let W_n be solutions in gauges c_n, and suppose*

$$\varepsilon_n := |c_n - c| + \sup_{t\ge 0} \|W_n(t) - V(t)\|_{s_0} \to 0, \qquad d_n := \|W_n(0) - V(0)\|_s \to 0.$$

There are solutions $U_n = W_n + b_n$ in the gauges c_n such that

$$\sup_{t\ge 0} \|b_n(t) - (U(t) - V(t))\|_s \to 0, \qquad \|b_n(t)\|_s \le C' e^{-\nu t}$$

for some $\nu > 0$ and all large n. In particular $U_n(0) \to U(0)$ in H^s, and $U_n - U \to 0$ uniformly in H^{s_0} on the positive half-line.

Proof. For an eigenvalue λ call $T_0^{-1} \log|\lambda|$ its growth rate. Choose $0 < \nu < \beta < \gamma$ with $[-\beta, -\nu]$ in a single gap of the rates, using Lemma 4.1. Let $P_f(t)$ be the transported projection onto the generalized eigenspaces of rates below $-\beta$ (the fast stable directions), and $P_c(t) := 1 - P_f(t)$. For some $k > 0$, in each of the two Sobolev spaces we have

$$\|L(t,v)P_f(v)\| \le Ce^{-(\beta+k)(t-v)} \quad (v \le t),$$
$$\|L(t,v)P_c(v)\| \le Ce^{(\nu-k)(v-t)} \quad (t \le v).$$

For integer periods this is the spectral radius formula with a strict gap, applied to the monodromy or its inverse on the indicated subspace. The propagators over the remaining fraction of a period are uniformly bounded.

Fix $0 < \eta < \beta - \nu$. We claim that

$$\sup_{t\ge 0} e^{-\eta t}\|W_n(t) - V(t)\|_s \le C_\eta(\varepsilon_n + d_n). \tag{4.2}$$

Writing $e := W_n - V$, the remainders about the old linearization are $2V|e|^2 + \bar{V}e^2 + |e|^2 e + (c_n - c)(V + e)$, up to a unit factor. Their H^s norm is at most $C\varepsilon_n(1 + \|e\|_s)$. Here and below we use $\|ab\|_s \le C(\|a\|_{s_0}\|b\|_s + \|a\|_s\|b\|_{s_0})$, which follows from weighted Fourier convolution and $\ell^2(\langle k\rangle^{2s_0}) \subset \ell^1$. Project away the finitely many old rates above a threshold in $(0, \eta/3)$ outside the rate set. The removed projection of e is bounded by $C_\eta\varepsilon_n$ in H^s at every time by Lemma 4.1. The complementary propagator grows at most as $C_\eta e^{\eta(t-v)/2}$ forwards. To see explicitly where smallness enters, let $m(t) := \sup_{0\le v\le t} e^{-\eta v}\|e(v)\|_s$. Apply Duhamel's formula to the complementary component, multiply by $e^{-\eta t}$, and use $\int_0^t e^{-\eta(t-v)/2}\,dv \le 2/\eta$. Adding back the removed component yields $m(t) \le C_\eta(d_n + \varepsilon_n) + C_\eta\varepsilon_n m(t)$. Absorb the last term for large n to obtain (4.2).

Write $d := U - V$. Starting at a late time A, prescribe $P_f(A)b(A) = P_f(A)d(A)$ and set

$$(\mathcal{G}f)(t) := \int_A^t L(t,v)P_f(v)f(v)\,dv - \int_t^\infty L(t,v)P_c(v)f(v)\,dv.$$

For $\xi = \beta, \nu$ denote by $\|h\|_{b,\xi} := \sup_{t\ge A} e^{\xi(t-A)}\|h(t)\|_b$. For $\nu \le \xi \le \beta$, multiplying the forward kernel by $e^{\xi(t-v)}$ leaves decay at rate $\beta + k - \xi$. On the backward kernel the corresponding rate is $\xi - \nu + k$. Both are at least k, so integration bounds $\mathcal{G}$ in the two norms $\|\cdot\|_{s_0,\beta}$ and $\|\cdot\|_{s,\nu}$. We seek a fixed point of

$$b \longmapsto L(t,A)P_f(A)d(A) + \mathcal{G}F_n(b),$$

where

$$iF_n(b) = (c_n - c)b + 2(|W_n|^2 - |V|^2)b + (W_n^2 - V^2)\bar{b} + 2W_n|b|^2 + \bar{W}_n b^2 + |b|^2 b.$$

By (4.2), for fixed A the bound $D_* := \sup_{t\geq A} e^{-\eta(t-A)}(1 + \|V(t)\|_s + \|W_n(t)\|_s) \leq 2 + 2\sup_t \|V(t)\|_s$ holds for all sufficiently large n. On a ball of radius $a \leq 1$ in both norms, write $h := b - b'$. The product inequality gives

$$\|F_n(b) - F_n(b')\|_{s_0,\beta} \leq C(\varepsilon_n + a)\|h\|_{s_0,\beta},$$
$$\|F_n(b) - F_n(b')\|_{s,\nu} \leq C(\varepsilon_n + a)\|h\|_{s,\nu} + C(D_* + 1)(1 + a)^2\|h\|_{s_0,\beta}.$$

For instance, put $a_n := |W_n|^2 - |V|^2$. Then $\|a_n(t)\|_{s_0} \lesssim \varepsilon_n$ and $\|a_n(t)\|_s \lesssim (D_* + 1)e^{\eta(t-A)}$. The product estimate gives

$$e^{\nu(t-A)}\|a_n h\|_s \lesssim \varepsilon_n\|h\|_{s,\nu} + (D_* + 1)e^{-(\beta-\nu-\eta)(t-A)}\|h\|_{s_0,\beta}.$$

The strict inequality $\eta < \beta - \nu$ bounds the second term at every time. In the quadratic terms of F_n, expanding the difference between b and b' leaves one factor of b or b', of size at most a. The cubic term leaves two such factors. In the combined norm $\|h\|_{s,\nu} + \Lambda\|h\|_{s_0,\beta}$, first choose Λ large compared with the second constant and the Green bound, then take a small ball, A large so the free datum lies in half the ball, and n sufficiently large. This gives a contraction. For the limiting coefficients V, c, the fixed point is $d = U - V$, since the backward boundary term vanishes by the decay $e^{-\gamma t}$. To compare the two maps at this fixed d, keep A fixed. By (4.2),

$$\sup_{t\geq A} e^{-\eta(t-A)}\|W_n(t) - V(t)\|_s \leq C_\eta e^{\eta A}(\varepsilon_n + d_n) \longrightarrow 0.$$

Factor $W_n^2 - V^2 = (W_n - V)(W_n + V)$ and use the same expansion with a conjugate for $|W_n|^2 - |V|^2$. By the product inequality, their H^{s_0} norms tend uniformly to zero and their H^s norms are $o(1)e^{\eta(t-A)}$, uniformly for $t \geq A$. Every coefficient difference in $F_n(d) - F_\infty(d)$, where F_∞ uses V, c, multiplies at least one d or $\bar{d}$. Its decay absorbs the factor $e^{\eta(t-A)}$ since $\nu + \eta < \gamma$, and also allows the weak weight $\beta < \gamma$. The Green bound now shows that the images of d converge in both weighted norms. The common contraction constant gives $b_n \to d$ in the two weighted norms. On $[0, A]$ the strong convergence $W_n \to V$ follows from continuous dependence. Continuing the exact solution $W_n + b_n$ backwards from A gives strong convergence there to U as well. This proves the proposition. If W_n tends to a periodic solution V_n, then U_n tends to the very same V_n. Equality of their masses follows on taking the limit of the difference, so their mass gauges also agree. □

4.2. **Choosing the data and the times.** Two indices enter the construction. The stage number j counts transfers on the final orbit. Within a single stage we vary the tree depth d to reduce the initial cost and choose a frequency spacing $R = R(d)$ to approximate that particular tree. Only one pair $(d, R(d))$ is retained at stage j.

Fix $1 < s < 2$, put $s_0 := (1 + s)/2$, and fix G, ℓ as in Theorem 1.2. Begin with $U_0 = V_0$ as in Proposition 2.1. The initial field $V_0(0)$ may be taken arbitrarily small in H^s because its Birkhoff indices are fixed, the five actions tend to zero, and the inverse chart is continuous in H^s. Write $g_0 \neq 0$ for its coefficient in (2.2). At any finite stage we have a true solution U exponentially asymptotic to a smooth periodic axial solution V, both in the mass gauge of V, whose coefficient g in (2.2) satisfies $|g| \geq |g_0|/2$.

We explain first how to make one choice with arbitrarily small error. For an increasing sequence of depths d, apply Proposition 2.4 and Lemma 2.5 with main terminal amplitude $a = 1$. Once the geometry is fixed, take the radii $b_j > 0$ so small that $\sum_j b_j^2 \leq 1$, that the hypothesis of Proposition 2.6 holds for $\theta = 3/4$, and that the sum of $|v/P|^{2s} b_j^2$ over all auxiliary vertices of

the respective levels is at most $1/d$. Here P is the positive root abscissa. Use Proposition 3.1 with $a_0 = P^{-s}$, so $r = (RP)^{-s}$, and call its finite plane-wave sum $C_{d,R}$, including its prefactor r. Orthonormality gives

$$\|C_{d,R}(t)\|_s^2 = \sum_{v\in\mathcal{V}} r^2\langle e + Rv\rangle^{2s}|x_v(2|g|rt)|^2, \qquad e = (1,0). \tag{4.3}$$

For each fixed tree the coefficients $r^2\langle e + Rv\rangle^{2s}$ tend to $|v/P|^{2s}$. The summed main weight $\sum_v |v/P|^{2s}$ over any single primary generation is at most one, since each split of a weight w replaces it by $w(\zeta^s + (1-\zeta)^s) \le w$. The primary amplitudes are constant within a generation and their squares sum over the generations to at most two by (2.7). Thus $\sup_{t\ge0}\|C_{d,R}(t)\|_s \le 3$ for all sufficiently large R at each d. Its initial squared norm is at most $2c_s^d + 1/d + o_{R\to\infty}(1)$, by (2.8).

We now apply Proposition 4.2 with the old pair U, V fixed and with a sequence indexed by the depth d. Write c for the gauge of V and $W_{d,R}$ for the solution of Proposition 3.1 at this depth. For each d choose $R = R(d)$ large enough that

$$\begin{aligned} |c_R - c| + \sup_{t\ge0}\|W_{d,R}(t) - V(t)\|_{s_0} &\le 2^{-d},\\ \sup_{t\ge0}\|W_{d,R}(t) - V(t) - C_{d,R}(t)\|_s &\le 2^{-d},\\ \|C_{d,R}(0)\|_s &\le (2c_s^d + 1/d)^{1/2} + 2^{-d}. \end{aligned}$$

The first two choices use the fixed-tree limits, and the third uses the leaf estimate above. At time zero the last two inequalities give $\|W_{d,R}(0) - V(0)\|_s \le (2c_s^d + 1/d)^{1/2} + 2^{1-d} \to 0$. Thus the global weak comparison and the strong initial comparison hold on the same sequence. The strong norm of the new contribution can still be of order one at late times.

Let $b_{d,R}$ be the correction given by Proposition 4.2 and put $\widetilde{U} \coloneqq W_{d,R} + b_{d,R}$. Subtracting the old solution gives

$$\widetilde{U} - U - C_{d,R} = (W_{d,R} - V - C_{d,R}) + [b_{d,R} - (U - V)].$$

The first term is the error for a single tree, made small by choosing R after d. The second tends uniformly to zero by Proposition 4.2, since the strong initial cost tends to zero with depth and the weak error is arbitrarily small for each fixed tree. We conclude that

$$\sup_{t\ge0}\|\widetilde{U}(t) - U(t) - C_{d,R}(t)\|_s \longrightarrow 0 \tag{4.4}$$

along the choices. The terms in this sum are compared in their specified gauges. Multiplication by any of these scalar phases preserves each Fourier modulus. We also have $\sup_{t\ge0}\|\widetilde{U}(t) - U(t)\|_{s_0} \to 0$, and $\widetilde{U}$ is exponentially asymptotic to V_R. The initial distances in H^s tend to zero. All these conclusions hold with any additional lower requirements on $R(d)$, since for each fixed tree the limits in Proposition 3.1 hold as $R \to \infty$.

The forward orbit of U is relatively compact in H^s because its restriction to a bounded interval is compact and its late tail approaches the compact range of V. Let $P_{\ge N}$ retain the Fourier coefficients at $|k| \ge N$. Then $\sup_{t\ge0}\|P_{\ge N}U(t)\|_s \to 0$ as $N \to \infty$. Indeed, cover the compact orbit closure by finitely many H^s balls of radius ε and take N large enough that the tail of each center has norm below ε. The projection has norm at most one, giving 2ε for every point of the orbit. Each frequency in $C_{d,R}$ has length at least $R - 1$. Thus its cross term with U is bounded by

$$|\langle U(t), C_{d,R}(t)\rangle_{H^s}| \le \|P_{\ge R-1}U(t)\|_s\|C_{d,R}(t)\|_s \le 3\|P_{\ge R-1}U(t)\|_s.$$

Here $\langle\cdot,\cdot\rangle_{H^s}$ is the inner product associated with our Fourier norm. The other error in expanding the squared norm is bounded by the left side of (4.4) times $2\sup_t\|U(t)\|_s+6+o(1)$. It follows that

$$\sup_{t\geq 0}\Big|\|\widetilde{U}(t)\|_s^2-\|U(t)\|_s^2-\|C_{d,R}(t)\|_s^2\Big|\longrightarrow 0. \tag{4.5}$$

This estimate includes the period during which the transfer is taking place.

Set $K\coloneqq 1+RP$, $D\coloneqq K^s$, and consider the window $[D^{3/4},D^{5/4}]$. With the finite tree fixed, rt tends uniformly to zero on $0\leq t\leq D^{3/4}$ as $R\to\infty$. The early value of (4.3) is therefore bounded by $2c_s^d+1/d+o_{R\to\infty}(1)$. On $t\geq D^{5/4}$ the slow time tends uniformly to infinity. The terminal amplitudes of Lemma 2.5 then imply

$$\sup_{t\geq D^{5/4}}\big|\|C_{d,R}(t)\|_s^2-1\big|\leq 1/d+o_{R\to\infty}(1).$$

Only the main root and the finitely many axial auxiliary amplitudes survive. A bound on the time of capture itself comes from the finite system. Put

$$t_*\coloneqq\frac{C_{3/4}d^{3/2}(K-1)^s}{2|g|}. \tag{4.6}$$

By this time the model root has reached $3/4$ and stays there or above, by Proposition 2.6. The old field U has uniformly small normalized Fourier coefficient at $(K,0)$ by compactness. The root coefficient of $C_{d,R}$ has normalized modulus $r\langle K\rangle^s H(2|g|rt)$, where $r\langle K\rangle^s\to 1$. Together with (4.4) this gives $\langle K\rangle^s|\widehat{\widetilde{U}}((K,0),t)|\geq 2/3$ for every $t\geq t_*$, if the choice is sufficiently far along the sequence.

We give the simultaneous choices explicitly. At stage $j\geq 1$ use the preceding procedure with $U=U_{j-1}$, $V=V_{j-1}$. Denote the new solutions by U_j,V_j, the root index by K_j, the time by $D_j\coloneqq K_j^s$, and the plane sum by C_j. Fix any $\delta>0$. Use the convergence in depth and, at each depth, the freedom to enlarge R, to meet the following conditions.

(1) The initial distance $\|U_j(0)-U_{j-1}(0)\|_s$ is less than $\delta 2^{-j}$. The error in (4.5) is at most 2^{-j}. The early squared norm of C_j and the late error from one in its squared norm are at most 2^{-j}.

(2) Require $G(D_j)\geq j$ at every stage. For $j>1$, also require $K_j>K_{j-1}$ and $D_j^{3/4}>D_{j-1}^{5/4}+1$. Take $\ell(K_j)\geq C_{3/4}d^{3/2}/(2|g|)$, which makes (4.6) at most $K_j^s\ell(K_j)$.

(3) At every earlier recorded frequency $(K_i,0)$, $i<j$, the change in normalized Fourier modulus, uniformly for all $t\geq 0$, is at most $2^{-j}/6$. Indeed the modulus difference is at most $\langle K_i\rangle^{s-s_0}\sup_t\|U_j(t)-U_{j-1}(t)\|_{s_0}$ and only finitely many indices have been specified. At the new frequency impose the $2/3$ bound above.

(4) Keep the successive changes of the periodic targets in uniform H^1 below $|g_0|2^{-j-2}$, and the actions in the ball of Lemma 2.2. At each fixed tree this follows by taking R large, since the total h^1 norm of the finitely many new coordinates is $O_d(Rr)\to 0$, the old action change is $O_d(r^2)$, and their angle speeds are unchanged. The common-period integral (2.2) then keeps $|g|\geq|g_0|/2$ throughout the construction.

There is no uniformity in the finite-tree constants required by these choices. More formally one can first impose all the lower requirements on $R(d)$ involving the given geometry, including the conditions on the clock. The input errors for Proposition 4.2 still tend to zero as $d\to\infty$. Taking one sufficiently large depth in this sequence imposes the remaining strong and weak error tolerances.

The initial fields converge in H^s to a datum u_0 with $\|u_0-U_0(0)\|_s\leq\delta$. Global well-posedness and persistence of regularity for the defocusing flow [3, 36] give convergence of the physical

trajectories in H^s on every compact time interval. Equivalently, one can use the H^s algebra estimate for local existence, uniqueness and continuous dependence at this regularity and the global defocusing theory to continue the limit. The masses converge, so the convergences on bounded intervals also hold in the respective mass gauges. For fixed i and $t \geq t_i$, where t_i is (4.6) for its tree, condition (3) and convergence at the fixed frequency K_i give

$$\langle K_i\rangle^s |\widehat{u}((K_i,0),t)| \geq \frac{2}{3} - \sum_{j=i+1}^{\infty} \frac{2^{-j}}{6} \geq \frac{1}{2}.$$

In particular, for any $m \geq 1$ the squared norm is at least $m/4$ at every time $t \geq \max_{1\leq i\leq m} t_i$, by summing over those m distinct frequencies. This already proves the divergence in Theorem 1.1. The more precise time dependence follows by also accounting for what happens during a transfer.

Finally sum (4.5) from stage one to stage n. The sum of the absolute errors is at most one. At any fixed t the windows are disjoint, so at most one term in $\sum_{j\leq n} \|C_j(t)\|_s^2$ is in a transition window. That term is between zero and nine. At stage j compare the squared norm to the indicator $\mathbf{1}_{\{D_j\leq t\}}$, which is one if $D_j \leq t$ and zero otherwise. Before the window both are within 2^{-j} of zero, and after the window their difference from each other is at most 2^{-j}. Inside the one possible active window the difference is at most nine. Summing these errors gives ten as a uniform bound for the difference between the sum of the squared carrier norms and the counting function. Adding the error of at most one from (4.5) gives

$$\Big| \|U_n(t)\|_s^2 - \#\{j \leq n : D_j \leq t\} \Big| \leq 11 + \sup_{v\geq 0} \|U_0(v)\|_s^2.$$

The right-hand side is finite by periodicity and independent of n, t. For each t, the left-hand norm converges by continuous dependence, and the count stabilizes since $D_j \to \infty$. This proves (1.2) uniformly in t. On $[D_j, D_{j+1})$ it gives $\|u(t)\|_s \leq (j + O(1))^{1/2}$ and $G(t) \geq j$. Hence $\|u(t)\|_s/G(t) \to 0$. The counting function tends to infinity, and the initial norm can be made arbitrarily small by the choices of U_0 and δ. This proves Theorems 1.1 and 1.2.

4.3. **The limiting axial solution and the height of a transfer.** The periodic axial solutions constructed above have a limit in every lower Sobolev norm if the weak tolerances are chosen summably. One can also measure the speed at which the two-dimensional solution approaches this limit. Let α be as in (3.2) and put $k_s := 3\alpha/4$, so the error scale in that section is $\rho = R^{-k_s}$. We use almost periodicity in the following sense. An H^σ-valued function is almost periodic if it is a uniform limit on $\mathbb{R}$ of finite sums $\sum_m f_m e^{i\omega_m t}$, with $f_m \in H^\sigma$ and $\omega_m \in \mathbb{R}$.

Corollary 4.3. *The data in Theorem 1.2 can be chosen so that there exists an axial solution v, almost periodic in time in H^σ for each $1 < \sigma < s$, satisfying*

$$\lim_{t\to+\infty} (1+t)^b \|u(t) - v(t)\|_{H^\sigma} = 0 \qquad \left(0 \leq b < \frac{\min(s-\sigma, k_s)}{s}\right).$$

At every t, $\langle K_j\rangle^s |\widehat{v}((K_j,0),t)| \geq 1/2$ for every j. In particular $v(t) \in \bigcap_{1<\sigma<s} H^\sigma \setminus H^s$.

Proof. Fix a finite tree and one pair (σ, b) as in the statement. In the notation of Section 3, on the terminal interval $rt \geq T$ the difference of the old Birkhoff variables from their terminal values at the same time is $O(\rho e^{-\sigma_*(rt-T)})$ in any fixed sequence norm on F. Here $\sigma_* > 0$ denotes the time weight used for this tree. On F^c the difference has Z norm $O(re^{-\sigma_*(rt-T)})$, by subtracting rY_∞ in (3.11). The transverse part obeys the same bound. On $0 \leq rt \leq T$ the backward estimate gives the respective bounds $O(\rho)$ and $O(r)$. In h^σ the new axial coordinates are $O(R^\sigma r)$, tending to zero. The compact set of old angle values thus has a neighborhood

where the differential of the inverse chart $h^\sigma \to H^\sigma$ is uniformly bounded. Applying it to the difference of the actual and terminal coordinates at the same time gives, for R large,

$$\|W_R(t) - V_R(t)\|_{H^\sigma} \le C(\rho + R^{\sigma - s}) \exp[-\sigma_*(rt - T)_+], \qquad t \ge 0.$$

The constants here may depend on the tree. Writing $\tau = rt$ and using $0 < r \le 1$, we have $(1+t)^b \le r^{-b}(1+\tau)^b$. The supremum of $(1+\tau)^b e^{-\sigma_*(\tau - T)_+}$ over $\tau \ge 0$ is finite. Therefore the weighted error is at most $C(R^{sb - k_s} + R^{sb + \sigma - s})$. Both exponents are negative precisely under the two strict upper bounds on b in the statement.

Use a countable dense collection of pairs (σ, b) strictly inside the allowed range, with σ, b rational. In the construction of stage j, impose estimates for its first j pairs, taking $s_0 < s$ above their spatial exponents if necessary. For each pair the preceding bound can be made arbitrarily small. On the history correction, the convergence of $b_n - (U - V)$ in the exponential strong norm in the proof of Proposition 4.2 and on the fixed initial interval gives convergence in the polynomially weighted H^σ norm. It follows that the successive differences of $U_j - V_j$ are summably small in that norm, each field in its own mass gauge. Indeed $(U_j - V_j) - (U - V) = (W_R - V_R) + (b_n - (U - V))$. For each fixed pair, the weighted differences are bounded, for example, by 2^{-j} at all stages after it enters the list. For any n beyond that stage, the weighted distance from the limit to $U_n - V_n$ is at most $\sum_{j>n} 2^{-j}$. Let $t \to \infty$ first, using the exponential decay of this fixed $U_n - V_n$, and then let $n \to \infty$. This proves the asserted vanishing of the weighted limit.

We also require summable uniform H^σ differences of the targets V_j on the whole real line. The locked old angles and the small h^σ norm of the added coordinates give this by Lemma 2.2 and the bounded chart differential. Each V_j is approximable uniformly in H^σ on its period by its time Fourier means (convolution with the Fejér kernels). The uniform limit of these periodic targets is therefore almost periodic in the sense just defined. Local well-posedness identifies it in its limiting gauge as a solution independent of y. At each stage U_j, V_j have the same mass. Their limits consequently have the same mass, and removing the common gauge preserves the norm of the difference. It preserves almost periodicity too, since multiplication by e^{ict} preserves uniform approximation by trigonometric polynomials in time. The strict inequalities on σ, b allow each real pair in the statement to be dominated by one of the rational pairs. To see why the bound for the axial limit holds even at early times, fix i, $j \ge i$, and $t \in \mathbb{R}$. Denote the period of V_j in its mass gauge by T_j^{ax}. For all sufficiently large integers m, the time $t + mT_j^{\mathrm{ax}}$ is past the capture time of K_i. At these times U_j satisfies the lower bound $1/2$ at that index and $U_j - V_j \to 0$ in H^s. Periodicity gives the same lower bound for $V_j(t)$. Let $j \to \infty$, using continuity of the single Fourier coefficient in H^σ. The limit $v(t)$ has normalized modulus at least $1/2$ at every K_i, and hence the sum defining $\|v(t)\|_{H^s}^2$ diverges. □

There is a second way to normalize the trees. A small increase in the initial cost at each leaf allows the Sobolev contribution of a single root to grow exponentially with its depth.

Corollary 4.4. *Fix $1 < s < 2$ and $0 < \kappa < -\frac{1}{2}\log c_s$, with c_s as in the leaf estimate (2.8). There is a single H^s datum, a strictly increasing sequence of depths d_j, and distinct frequencies $(K_j, 0)$, with the following property. Put $B_j := e^{\kappa d_j}$. At that frequency the normalized Fourier modulus $\langle K_j\rangle^s |\widehat{u}((K_j, 0), t)|$ of the solution is at least $B_j/2$ at every time*

$$t \ge \frac{C_{3/4} d_j^{3/2} (K_j - 1)^s}{2|g_j| B_j},$$

where g_j is the coefficient (2.2) of the preceding periodic orbit and $\inf_j |g_j| > 0$.

Proof. Keep $a = 1$ but replace a_0 by $B(d)/P^s$ at depth d, where $B(d) := e^{\kappa d}$. The primary initial cost is now $O(B(d) c_s^{d/2}) = o(1)$. After fixing the geometry, take the positive auxiliary

radii so small that their entire normalized initial cost tends to zero as well and Proposition 2.6 applies. For each fixed depth the weak contribution still tends to zero as $R \to \infty$, and the half-line error in (3.1) can be made arbitrarily small in absolute H^s norm. The hypotheses of Proposition 4.2 therefore hold along a sequence of increasing depths. Its conclusion and compactness of the preceding orbit again give the root lower bound with a fixed relative margin, now $2B(d)/3$. Reserve a total loss of at most $B(d)/6$ in the later weak comparisons at this frequency. The Cauchy argument for the initial data and the calibration of g_j are unchanged. Once depth d_j is retained, $B_j = B(d_j)$ and $r = B_j/(K_j - 1)^s$, which gives the stated time from Proposition 2.6. □

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., vol. 46, American Mathematical Society, Providence, RI (1999).

[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:2609.24930v2 (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*, Institute of Mathematics, University of Zurich, Preprint No. 07-2013 (2013).
[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 and J.-C. Molnar, *On the well-posedness of the defocusing mKdV equation below* L^2, SIAM J. Math. Anal. **49** (2017), 2191–2219.
[38] 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.
[39] T. Kappeler, B. Schaad and P. Topalov, *Scattering-like phenomena of the periodic defocusing NLS equation*, Math. Res. Lett. **24** (2017), 803–826.
[40] S. B. Kuksin, *On turbulence in nonlinear Schrödinger equations*, Geom. Funct. Anal. **7** (1997), 783–822.
[41] 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.
[42] A. Maspero, *Tame majorant analyticity for the Birkhoff map of the defocusing nonlinear Schrödinger equation on the circle*, Nonlinearity **31** (2018), 1981–2030.
[43] 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.
[44] J.-C. Molnar, *New estimates of the nonlinear Fourier transform for the defocusing NLS equation*, Int. Math. Res. Not. (2015), 8309–8352.
[45] J. Molnar, *Features of the nonlinear Fourier transform for the dNLS equation*, Ph.D. thesis, University of Zürich (2016).
[46] 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.
[47] C. Procesi and M. Procesi, *A KAM algorithm for the resonant non-linear Schrödinger equation*, Adv. Math. **272** (2015), 399–470.
[48] G. Staffilani, *Quadratic forms for a 2-D semilinear Schrödinger equation*, Duke Math. J. **86** (1997), 79–107.
[49] 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