Quasiperiodic dynamics in Hamiltonian $1\frac{1}{2}$ degree of freedom systems far from integrability

Shui-Nee Chow, Martijn van Noort*, Yingfei Yi†
Center for Dynamical Systems and Nonlinear Studies
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332, USA

January 21, 2005

Abstract

The subject of this paper is two-quasiperiodicity in a large class of one-and-a-half degree of freedom Hamiltonian systems. The main result is that such systems have invariant tori for any internal frequency that is of constant type and sufficiently large, relative to the forcing frequency. An explicit bound on the minimum value of the internal frequency is presented. The systems under consideration are not required to be small perturbations of integrable ones.

keywords: KAM theory, Aubry-Mather theory, nonintegrable Hamiltonian systems
mathematics subject classification (2000): Primary: 37J40 - Secondary 70H08, 70H11, 70H07, 37E40

1 Introduction and main result

This paper deals with a class of real-analytic one-and-a-half degree of freedom Hamiltonian systems, that is, Hamiltonian systems of one degree of freedom depending periodically on time. Without loss of generality the period in time is set to 1. We further assume that the systems in our class can be written in action-angle variables $(x, y) \in (\mathbb{R}/\mathbb{Z}) \times \mathbb{R}$. The equations of motion of such a system are given by

$$\begin{aligned} \dot{x} &= \frac{\partial}{\partial y} H(x, y, t) \\ \dot{y} &= -\frac{\partial}{\partial x} H(x, y, t) \\ \dot{t} &= 1, \end{aligned}$$

*corresponding author: **fax USA 404-894-4409, email: martijn@math.gatech.edu**
†The last author was partially supported by NSF grant DMS0204119

where $t \in (\mathbb{R}/\mathbb{Z})$, and $H(x, y, t)$ is the (non-autonomous) Hamilton function. The Hamiltonian can be expanded as

$$H(x, y, t) = \omega y + \frac{1}{2} m y^2 + H_g(y, t) + H_R(x, y, t), \tag{1}$$

with $\omega \in \mathbb{R}$, $m \in \mathbb{R}$, $\operatorname{av}(H_g, t) = O(y^3)$ and $\operatorname{av}(H_R, x) = 0$. Here $\operatorname{av}(f, s)$ denotes the average of a function f with respect to the angle s:

$$\operatorname{av}(f, s) := \int_0^1 f \, \mathrm{d}s.$$

We assume that $m \neq 0$, while $\omega \neq 0$ is used as a parameter. Without loss of generality we restrict to $\omega > 0$ and $m > 0$. This can always be achieved by sign changes in t and y. A final assumption, to be made more precise below in (3), is that $\frac{\partial^2}{\partial y^2} H_g$ is sufficiently small. This defines a class of one-parameter families of Hamiltonian vector fields of the form

$$X = X(x, y, t; \omega) = (\omega + my + g(y, t)) \frac{\partial}{\partial x} + \frac{\partial}{\partial t} + R(x, y, t), \tag{2}$$

$g = \frac{\partial}{\partial y} H_g$ and $R = \frac{\partial}{\partial y} H_R \frac{\partial}{\partial x} - \frac{\partial}{\partial x} H_R \frac{\partial}{\partial y}$. Hence $\operatorname{av}(g, t) = O(y^2)$ as $y \to 0$ and $\operatorname{av}(R, x) = 0$. It is sometimes convenient to consider the Poincaré map of this system, defined to be the first return map on the section $t = 0$.

The family X is a perturbation of $X^{(0)}$, defined by $X = X^{(0)} + R$. Each member of the unperturbed family is independent of x and hence integrable. It has an invariant torus $y = 0$ with frequency vector $(\omega, 1)$, that is, frequency ω in x, and frequency 1 in t. It is well known that this torus will persist under sufficiently small perturbations if the frequency vector satisfies a Diophantine condition. However, in the present setting the perturbation R is an arbitrary Hamiltonian vector field, possibly large, and our purpose is to study the persistence of the invariant torus $y = 0$ in this more general case. Our main result is that this torus will persist if $\frac{\partial}{\partial y} g$ is sufficiently small compared to m, while ω is of constant type and larger than some ω_0. An explicit formula for ω_0, depending on g, m and R, will be presented.

Arnol'd [Arn63] has studied the Hamiltonian (1) in the (equivalent) context of adiabatic theory, in a more qualitative setup, proving the existence of invariant tori for ω "sufficiently large" (but without providing estimates on the size of ω). The same problem also arises in the question of boundedness of solutions of (1), going back to a result by Littlewood [Lit66]. The literature on this is mainly concerned with two special cases, and both are discussed as examples in the present paper. In the first case H_R is periodic in (x, t) and independent of y, and one proves the existence of invariant tori for large Diophantine ω, that is, for large y. In the second case x is a variable in $\mathbb{R}$, and H_R is a superquadratic polynomial in x and independent of y. As an example, one can think of a periodically forced Duffing equation. After passing to action angle coordinates this reduces to a system of the form (1), and one can prove the existence of invariant tori surrounding the circle $x = y = 0$ at large distance. In both cases the existence of invariant tori implies boundedness of all solutions, with the tori acting as barriers. Studies in this area include Dieckerhoff and Zehnder [DZ87], Moser [Mos89a, Mos89b], Liu [Liu89], Levi [Lev90a, Lev91], You [You90], Laederich and Levi [LL91], Levi and Zehnder [LZ95]. Gallavotti et al. [GGM99] discuss a problem from celestial mechanics with incommensurate frequencies, but with a small perturbation.

We remark that by a trivial localization our method may be used to discuss the existence of a (Cantor) family of invariant tori of a single system, rather than a single torus for a (Cantor) family of systems. Indeed, consider for example the Hamiltonian

$$H(x, y, t) = H^{(0)}(y, t) + H^{(1)}(x, y, t),$$

where $\mathrm{av}(H^{(1)}, x) = 0$ and the function $H^{(1)}$ is bounded. The Hamiltonian $H^{(0)}$ has invariant tori $y = y_0$ for any constant y_0, with frequency $\Omega(y_0) = \mathrm{av}(\frac{\partial}{\partial y} H^{(0)}(y_0, t), t)$ in x. Assume that the frequency map $\Omega : \mathbb{R} \longrightarrow \mathbb{R}$ is surjective, then for every x-frequency ω there exists a (possibly non-unique) torus $y = y_0$ of $H^{(0)}$ with $\Omega(y_0) = \omega$. Our result implies that there exists an invariant torus near $y = y_0$, with x-frequency $\omega = \Omega(y_0)$, if $\Omega'(y_0) \neq 0$ and $\omega \geq \omega_0$ is of constant type. Indeed, this fits in the setting sketched above if we take ω as parameter and translate the unperturbed torus to the origin by $y \mapsto y - y_0$.

More can be said in case the map Ω is surjective and $\frac{\partial^2}{\partial y^2} H^{(0)}(y, t) > 0$ for all t and all $|y| > \delta$, for some $\delta > 0$. The phase space can be divided into three invariant connected components. In two of these, corresponding to $\omega > \omega_0$ and $\omega < -\omega_0$, the Poincaré map P is a monotone twist map, and has an invariant circle with rotation number ω for each $|\omega| \geq \omega_0$ of constant type - corresponding to a quasiperiodic invariant torus of the Hamiltonian with frequencies $\{\omega, 1\}$. In fact, one can show that in these regions the measure of invariant tori converges exponentially fast to full measure as $\omega \to \pm\infty$, see for example [Pös82, DG96, BHS96, JV97, BNS]. The existence of invariant tori in these region also implies the existence of Aubry-Mather sets. Indeed, for any irrational number ω in these regions, there are two invariant circles of P with rotation numbers $\omega_1 < \omega_2$ and $\omega \in (\omega_1, \omega_2)$ such that the restriction of P to the annulus bounded by these circles is a monotone twist map that preserves area and boundary components. According to the well-known Aubry-Mather theory ([ALD83, Mat82, Mos86]), the map P admits an Aubry-Mather set with rotation number ω which gives rise to a 1-parameter family of Aubry-Mather solutions

$$x(t) = U(t, \omega t + \theta) \ (\mathrm{mod}\ 1), \ y(t) = V(t, \omega t + \theta)$$

of the Hamiltonian system, where $U : \mathbb{S}^1 \times \mathbb{R} \longrightarrow \mathbb{R}$ is monotone in the second argument, and both $U(t, \theta) - \theta$ and $V(t, \theta)$ are 1-periodic in t and θ but in general are not necessarily continuous.

The dynamics in the "region of interest", given by $-\omega_0 < \omega < \omega_0$, is bounded but not restricted otherwise. We note that if P does not satisfy the twist condition everywhere in this region, even the existence of *any* Aubry-Mather set is in question (compare [Jia99], where Aubry-Mather sets are shown to exist for all irrational rotation numbers).

1.1 Main result

To state our main result we will first introduce some notation. The vector field X is real-analytic and hence can be extended analytically to a neighborhood of the unperturbed invariant torus $\{(x,y,t)\} = (\mathbb{R}/\mathbb{Z})\times\{0\}\times(\mathbb{R}/\mathbb{Z})$ in $(\mathbb{C}/\mathbb{Z})\times\mathbb{C}\times(\mathbb{C}/\mathbb{Z})$, defined as follows. Let

$$\mathcal{D} = \mathcal{D}(\eta,\rho) := ((\mathbb{R}/\mathbb{Z}) \oplus [-i\eta, i\eta]) \times B(\rho) \times (\mathbb{R}/\mathbb{Z}) \subset (\mathbb{C}/\mathbb{Z}) \times \mathbb{C} \times (\mathbb{C}/\mathbb{Z}),$$

where $B(\rho)$ is the closed disc of radius $\rho > 0$ and center 0 in $\mathbb{C}$, and $\eta > 0$. For $d > 0$ and any set $\mathcal{E} \subset (\mathbb{C}/\mathbb{Z}) \times \mathbb{C} \times (\mathbb{C}/\mathbb{Z})$ we denote by $\mathcal{E} + d$ the neighborhood of radius d around $\mathcal{E}$, that is

$$\mathcal{E} + d = \{z \in (\mathbb{C}/\mathbb{Z}) \times \mathbb{C} \times (\mathbb{C}/\mathbb{Z}) : \text{ there exists } x \in \mathcal{E} \text{ such that } |z - x| \le d\}.$$

For η, ρ and d sufficiently small X can be considered as an analytic vector field on $\mathcal{D}_0 = \mathcal{D} + 2d$.

If f is a function defined on $\mathcal{E}$, then we denote by $||f||_{\mathcal{E}}$ the supremum norm of f on $\mathcal{E}$. If f is vector valued, then $||f||_{\mathcal{E}}$ is the supremum norm of the maximum of the components of f.

We say that a frequency ω is of constant type with parameter γ if

$$|\omega - p/q| \ge \gamma |q|^{-2} \text{ for all } p/q \in \mathbb{Q}.$$

Our main result is as follows.

Main Theorem. *Let X be the vector field given by (2). Choose $\gamma \in (0,1)$, $\eta > 6\gamma$, $\rho = (3m)^{-1}\eta$ and $d > 0$ sufficiently small such that X is analytic on $\mathcal{D}_0 = \mathcal{D} + 2d$, where $\mathcal{D} = \mathcal{D}(\eta,\rho)$. Assume that*

$$||\frac{\partial}{\partial y} g||_{\mathcal{D}+d/2} \le \frac{1}{4} m. \tag{3}$$

Then X has an invariant torus with frequency vector $(\omega, 1)$ if ω is of constant type with parameter γ, and $\omega \ge \omega_0 = \omega_0(m, ||g||_{\mathcal{D}_0}, ||R||_{\mathcal{D}_0}, \gamma, \eta, d)$, defined below. The y coordinate of every point on this torus satisfies

$$|y| < \frac{11}{19} d + \rho. \tag{4}$$

The threshold frequency ω_0 is given by

$$\begin{aligned} \omega_0 &= \alpha_0 \max\left\{1, \frac{1}{\log 2} W(b \log 2)\right\} \qquad (5)\\ b &= \frac{2+3m}{c\gamma^2} \max\left\{12, \frac{7^8}{108(\eta - 6\gamma)^4}\right\} \max\{md, 2||R||_{\mathcal{D}_0}\}\\ \alpha_0 &= \frac{6a}{d}\left(||R||_{\mathcal{D}_0} + 2C \max\left\{1, \frac{2||R||_{\mathcal{D}_0}}{md}\right\}\right)\\ C &= \frac{2}{3}\max\{m(2d+\rho) + ||g||_{\mathcal{D}_0} + ||R||_{\mathcal{D}_0}, 1\}, \end{aligned}$$

where $a = \frac{1}{4} + \eta + 2d$, W is the lambert W function, that is, the inverse of $W \mapsto We^W$, and $c = 5.04$ is the constant appearing in Herman's version of the twist theorem [Her83, Her86].

We note that the function $W(x)$ is well-defined and strictly increasing for $x > 0$. This implies that the result of the theorem also holds if we replace the condition $\omega \geq \omega_0$ by the stronger condition

$$\omega \geq \alpha_0 \max\{1, \log_2 b\}.$$

This theorem can be applied for any system of the form (2), but it is most useful in case R is large. Indeed, for small R any standard KAM theorem can be used. For large R the theorem states that X has an invariant torus with frequency vector $(\omega, 1)$ if ω is of constant type, and

$$\omega \gg \frac{||R||^2_{\mathcal{D}_0} \log\left((1+m)||R||_{\mathcal{D}_0}\right)}{m}.$$

Here we use that $W(x) \approx \log x$ for large x.

Remark 1.1 : In the case discussed above, of a single system with a Cantor family of invariant tori, there exist Aubry-Mather sets for all intermediate rotation numbers As far as we know it is an open question whether a similar result holds in the current context, that is, whether the family X has Aubry-Mather sets of rotation number ω for all parameter values $\omega \geq \omega_0$.

Remark 1.2 : We restrict to frequencies ω of constant type because the twist theorem in the version of Herman [Her83, Her86] is used in our proof. Other versions of the twist theorem, e.g. by Moser [Mos62] and Rüssmann [Rüs70, Rüs83], allow a larger set of Diophantine frequencies, but require a smaller perturbation, which in the present setting translates into a larger value of ω_0. Thus, in this alternative approach, one "loses" the lower frequencies, and "gains" higher frequencies. In this case the set of frequencies for which an invariant torus exists can be shown to have positive measure, converging to full measure exponentially fast as $\omega \to +\infty$, cf. [Pös82, DG96, BHS96, JV97, BNS].

Remark 1.3 : Condition (3) on g can be eliminated by applying a symplectic coordinate change $(H, t) \mapsto (K(H, t), \tau(H, t))$, where the Hamiltonian H is considered as a variable conjugate to t. We refer to Arnol'd [Arn63] for more details on the transformation. It can be interpreted as a reparametrization of the independent variable t, that removes the time-dependence of the function g. Thus the new Hamiltonian K is of the form

$$K(x, y, \tau) = \omega y + \frac{1}{2} m y^2 + K_g(y) + K_R(x, y, \tau),$$

with $\mathrm{av}(K_R, x) = 0$ and $K_g = O(y^3)$. Thus condition (3) is satisfied for y sufficiently small.

1.2 Method

The proof of the main theorem consists of two steps. First the vector field X is averaged with respect to x. Indeed, if ω is large, then x is a fast angle, and X is conjugate up to a small perturbation to its average with respect to x, that is, to $X^{(0)}$. In this sense, X is nearly integrable.

We use Neishtadt's averaging procedure, cf. [Nei84], to transform X to a vector field $\tilde{X}$, defined on $\mathcal{D}$, of the form

$$\tilde{X} = \tilde{X}(x, y, t; \omega) = \tilde{X}^{(0)} + \tilde{R} = (\omega + my + \tilde{g}(y, t)) \frac{\partial}{\partial x} + \frac{\partial}{\partial t} + \tilde{R}(x, y, t), \tag{6}$$

where $\mathrm{av}(\tilde{g}, t) = O(y)$ and $\tilde{R}$ is exponentially small as $\omega \to +\infty$. Thus $\tilde{X}$ is an exponentially small perturbation of the integrable vector field $\tilde{X}^{(0)} := \tilde{X} - \tilde{R}$, which has an invariant torus $y = 0$ with frequency vector $(\omega, 1)$. The vector fields X and $\tilde{X}$ are conjugate by a symplectic, orientation and t preserving transformation that is analytic in (x, y, t) and piecewise analytic in ω.

We observe that for small ω this averaging procedure does not decrease the perturbation, which is the reason that our theorem is not an improvement of standard results in this setting. Indeed, as usual in Neishtadt's procedure (and we will make this explicit in section 3), for small ω only very few, possibly only one, averaging steps will be performed. Thus in this case the perturbation will not decrease much, and hence no benefit can be expected from it.

In the second step of the proof we consider the Poincaré map $\tilde{P}$ of the averaged vector field $\tilde{X}$, defined to be the first return map on the section $t = 0$. This map is of the form

$$\tilde{P} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + \omega + my + q(y) \\ y \end{pmatrix} + Q(x, y),$$

where $q(0) = 0$, $|q'| < \frac{1}{2}m$ and Q is exponentially small as $\omega \to +\infty$. Thus $\tilde{P}$ is a twist map, and its quasiperiodic dynamics can be studied using a twist theorem. An invariant circle of $\tilde{P}$ of rotation number ω corresponds to an invariant torus of X of frequency vector $(\omega, 1)$.

Remark 1.4 : The standard proof of the KAM theorem uses averaging with respect to both angles. In the present context this procedure will diverge if the perturbation R is large, and is thus not applicable.

Remark 1.5 : The dynamics on the invariant torus $y = 0$ of the integrable vector field $\tilde{X}^{(0)}$ is in general not close to parallel flow, due to the fact that $\tilde{g}(0, t)$ may be large. By using the Poincaré map, we resolve this problem, since the average contribution of the term $\tilde{g}(0, t)$ vanishes over one period in t. Alternatively, one can use a KAM theorem after removing this term by either a reparametrization of time $t \mapsto \tau(t)$ or a symplectic translation $x \mapsto x + f(t)$, leaving all other variables invariant.

The rest of the paper is organized as follows. We first give two examples in the next section to illustrate the application of our main result. The averaging part of the proof is described in section 3, followed by the construction of the twist map in section 4. The proof of the main theorem is completed in section 5. Appendix A deals with the construction of action angle coordinates in one of the examples.

Acknowledgements

The authors thank the referee for valuable comments and suggestions.

2 Examples

2.1 Oscillator with periodic potential

The first example is the periodically forced oscillator with equation of motion

$$\ddot{x} + \frac{\partial V}{\partial x}(x,t) = 0 \ , \ \dot{t} = 1, \tag{7}$$

where V is real-analytic and 1-periodic in x and t. Choose $\omega > 0$ and define y by $my = \dot{x} - \omega$, where $m > 0$ is a constant to be determined later. Then we obtain the system

$$\begin{aligned} \dot{x} &= \omega + my \\ \dot{y} &= -m^{-1}\frac{\partial V}{\partial x}(x,t) \\ \dot{t} &= 1, \end{aligned}$$

with Hamilton function

$$H(x,y,t;\omega) = \omega y + \frac{1}{2}my^2 + m^{-1}V(x,t).$$

Choose positive constants γ, ρ, η and d as in the main theorem. Assume that $||V||_{\mathcal{D}_0} \gg 1$, then a brief calculation shows that

$$\omega_0 \sim ||V||_{D_0}^{1/2} \log ||V||_{D_0},$$

if we take $m = ||V||_{D_0}^{1/2}$. Thus, if $\omega \geq \omega_0$ and ω is of constant type (with parameter γ), then the forced oscillator (7) has an invariant torus of the form $\dot{x} = f(x,t)$, with frequency ω in x, and

$$|f(x,t) - \omega| \leq \left(\frac{11}{19}d + \rho\right) ||V||_{D_0}^{1/2} \text{ for all } (x,t) \in \mathbb{S}^1 \times \mathbb{S}^1.$$

As a specific example one can think of a forced pendulum, where

$$V(x,t) = (\alpha + \beta \cos t)\cos x, \tag{8}$$

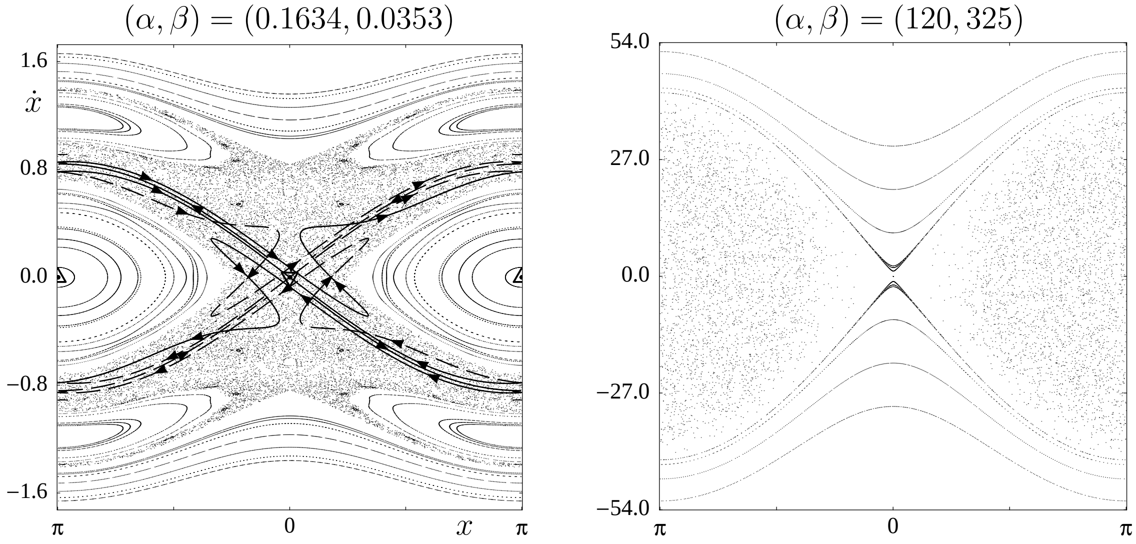

Figure 1: Poincaré map of the forced pendulum (8) on the cilindrical section $t = 0$, $(x, \dot{x}) \in \mathbb{S}^1 \times \mathbb{R}$, at two typical parameter points. The invariant tori of rotational type correspond to invariant circles of this map that wind around the cilinder, constituting a set of positive measure in two unbounded regions in phase space, above and below a central "region of interest".

for some parameters α and β sufficiently large. Here $x = \pi$ corresponds to the lower equilibrium, where the pendulum is hanging down, and $x = 0$ to the upper equilibrium, where it is standing up. The main theorem states that this pendulum has invariant tori corresponding to rotational motion for any frequency ω in x of constant type that satisfies $\omega \geq \omega_0$, where $\omega_0 \sim (\alpha + \beta)\log(\alpha + \beta)$, see figure 1, also cf. [BHN+].

Remark 2.1 : An alternative method is to construct a twist map for (7) directly, see e.g. [Lev90b]. A brief calculation (implicitly present in the quoted paper) shows that in this setting the threshold frequency satisfies $\omega_0 \sim ||V||_{\mathcal{D}_0}$. Thus for large potentials the preliminary averaging procedure of our approach significantly improves the estimate.

Remark 2.2 : An oscillator with "large" potential function V occurs naturally in adiabatic theory. Indeed, consider an oscillator with a potential function that varies slowly and periodically in time. The corresponding equation of motion can be written as

$$\ddot{x} + \frac{\partial V}{\partial x}(x, \lambda) = 0 \ , \ \dot{\lambda} = \varepsilon, \tag{9}$$

where $\varepsilon > 0$ is a small parameter, V is periodic in both arguments, and depends on a slowly varying parameter λ. Assume that $V \sim 1$, and introduce $y = \dot{x} - \omega$. This gives

$$\dot{x} \ = \ \omega + y$$

$$
\begin{aligned}
\dot{y} &= -\frac{\partial V}{\partial x}(x,\lambda) \\
\dot{\lambda} &= \varepsilon.
\end{aligned}
$$

Rescaling this system by a factor ε, and applying the main theorem shows the existence of tori $\dot{x} = f(x,\lambda)$ of (9) for any frequency $\omega \geq \omega_0$, such that ω/ε is of constant type. Here $\omega_0 \sim -\log\varepsilon$, and $|f(x,\lambda) - \omega| \leq \frac{11}{19}d + \rho$.

Remark 2.3 : The forced oscillator (7) has Aubry-Mather sets for any frequency ω, including frequencies lower than ω_0, cf. [Jia99, Mos86].

2.2 Oscillator with polynomial potential

Our second example is also a periodically forced oscillator, with equation of motion

$$
\ddot{x} + \frac{\partial V}{\partial x}(x,t) = 0 \ , \ \dot{t} = 1,
$$

where $x \in \mathbb{R}$ and $V(x,t)$ is a superquadratic polynomial potential of the form

$$
V(x,t) = \sum_{k=1}^{2n} a_k(t)x^k,
$$

where $n \geq 2$, all a_k are 1-periodic and real-analytic in t, and $a_{2n} > 0$. If V is independent of t, then outside a neighborhood of the circle $(x,\dot{x}) = 0$, $t \in \mathbb{S}^1$ the phase space is filled with invariant tori, surrounding this circle. Let us consider the existence of such tori in case V depends on time. We refer to [LZ95] and appendix A for more details on the constructions below.

For $x^2 + \dot{x}^2$ sufficiently large one can construct action-angle variables $(\theta, I) \in \mathbb{S}^1 \times \mathbb{R}$ in the usual way, that is, $I = I(h,t)$ is the area enclosed by the curve $H(x,y,t) = h$ for a fixed t, and θ is the conjugate angle. Then the system takes the form

$$
\begin{aligned}
\dot{\theta} &= \Omega(I,t) + \frac{\partial K_1}{\partial I}(\theta,I,t) \\
\dot{I} &= -\frac{\partial K_1}{\partial \theta}(\theta,I,t) \\
\dot{t} &= 1,
\end{aligned}
$$

corresponding to a Hamiltonian $K(\theta,I,t) = K_0(I,t) + K_1(\theta,I,t)$, with $\Omega = \frac{\partial K_0}{\partial I}$. Here $K_0 \sim I^{2n/(n+1)}$ and $K_1 \sim I$ for large I.

The time dependence can be removed from the leading term by a symplectic transformation of time and the Hamiltonian:

$$
(t,K) \mapsto (\tau,\bar{K}),
$$

where the Hamiltonian is considered as a variable conjugate to time. This leads to the following system (where the dot now denotes the derivative with respect to τ):

$$\begin{aligned}\dot{\theta} &= \bar{\Omega}(I) + \frac{\partial \bar{K}_1}{\partial I}(\theta, I, \tau) \\ \dot{I} &= -\frac{\partial \bar{K}_1}{\partial \theta}(\theta, I, \tau) \\ \dot{\tau} &= 1,\end{aligned}$$

with Hamiltonian $\bar{K} = \bar{K}_0 + \bar{K}_1$ and $\bar{\Omega} = \frac{\partial}{\partial I}\bar{K}_0$. A computation shows that the Hamiltonian is given by

$$\begin{aligned}\bar{K}_0(I) &= c_n \bar{a} I^{2n/(n+1)} + O(I^{(2n-1)/(n+1)}), \\ \bar{K}_1(\theta, I, \tau) &= f_n(\theta, I, \tau)\frac{\bar{a} a'_{2n}}{a_{2n}^{(n+2)/(n+1)}} I + O(I^{n/(n+1)}), \\ \bar{a} &= \mathrm{av}(a_{2n}(t)^{1/(n+1)}, t),\end{aligned}$$

where $c_n \in [0.03, 0.12]$, $|f_n| \leq 0.4$. The unperturbed Hamiltonian $\bar{K}_0$ has a family of invariant tori $I =$ constant, with frequency vector $(\Omega(I), 1)$.

Application of the main theorem now shows that for sufficiently large ω of constant type there exists an invariant torus of the form $I = f(\theta, \tau)$ with frequency ω in θ and frequency 1 in τ. This corresponds to an invariant torus in the original coordinates (x, y, t) of the same frequencies. Indeed, since $\bar{\Omega} \to +\infty$ as $I \to +\infty$ there exists a (unique) I_0 such that

$$\bar{\Omega}(I_0) = \omega.$$

We now apply a localization and rescaling of the action variable, which puts the unperturbed invariant torus with frequency $(\omega, 1)$ at the origin. Define a new action variable J by $I = I_0 + I_0^{(n+3)/(2n+2)} J$. Although this is not an action variable in a strict sense, that is, the transformation $(\theta, I) \mapsto (\theta, J)$ is not symplectic (there is a time rescaling involved), it does take Hamiltonian systems to Hamiltonian systems. In the new variables (θ, J, τ) the components of (2) satisfy

$$\omega \sim I_0^{(n-1)/(n+1)}\ ,\ g \sim 1\ ,\ \text{and}\ m \sim R \sim I_0^{(n-1)/(2n+2)}.$$

Thus ω grows faster than the other components, and therefore the conditions of the main theorem are satisfied.

If we assume that a_{2n} and a'_{2n} are large compared to the other coefficients $a_1, \ldots, a_{2n-1}$, then a more precise estimate is possible, and we find that invariant tori exist for $I \geq I_0$, where

$$I_0 \sim \left(\frac{\max_t |a'_{2n}(t)|}{\min_t a_{2n}(t)}\right)^{(n+1)/(n-1)}.$$

3 An averaging theorem

This section describes the first step of the proof of the main result. Starting with the vector field X given by (2) we perform a sequence of transformations to obtain a vector field $\tilde{X}$ as in (6), which is of similar form but with a perturbation $\tilde{R}$ that is exponentially small as $\omega \to +\infty$. This procedure is due to Neishtadt [Nei84, BRS96, SV01], and in our case it leads to the following theorem.

Theorem 3.1 *Choose $\eta > 0$, $\rho > 0$ and $d > 0$. Let $\mathcal{D} = \mathcal{D}(\eta, \rho)$, $\mathcal{D}_0 = \mathcal{D} + 2d$ and define α_0 as before. Let X be given by (2), and be defined analytically on $\mathcal{D}_0$.*

If condition (3) holds, then for $\omega \geq \alpha_0$ there exists a symplectic transformation $\Psi : \mathcal{D} \longrightarrow \mathcal{D}_0$, such that $\tilde{X} = \Psi_*^{-1} X$ is a Hamiltonian vector field defined on $\mathcal{D}$ by (6), where

$$\begin{aligned} ||\tilde{g}||_{\mathcal{D}} &< ||g||_{\mathcal{D}+d/2} + ||R||_{\mathcal{D}_0} \\ ||\frac{\partial}{\partial y}\tilde{g}||_{\mathcal{D}} &< \frac{m}{2} \\ av(\tilde{g}, t) &= O(y) \\ ||\tilde{R}||_{\mathcal{D}} &\leq \frac{\alpha_0}{\omega} 2^{-\omega/\alpha_0} \min\left\{ ||R||_{\mathcal{D}_0}, \frac{md}{2} \right\} \end{aligned}$$

The transformation Ψ preserves orientation and t, and is analytic on $\mathcal{D}$, piecewise analytic in ω, and near identity for large ω in the sense that

$$||\Psi - Id||_{\mathcal{D}} \leq \frac{\alpha_0 d}{19\omega}\left(10 + \frac{\alpha_0}{\omega}\right).$$

The remainder of this section deals with the proof of this theorem, which consists of three parts. First the vector field is averaged once with respect to x. This will reduce the perturbation sufficiently to ensure that the twist condition is satisfied throughout. Then Neishtadt's procedure is used to further decrease the perturbation R, by repeatedly averaging with respect to x. In general these averaging transformations change the frequency of the torus $y = 0$ of the unperturbed system. Therefore the final transformation is a translation in y to restore the frequency to the value ω. We first introduce some notation and give an overview of the proof, and subsequently deal with each of the three parts in turn.

Outline of the proof

Choose $\omega \geq \alpha_0$. We will perform $N + 1$ averaging steps, where $N \geq 1$ is such that $\omega \in [N\alpha_0, (N + 1)\alpha_0)$. Define $\sigma = d/N$, then $\omega\sigma \geq \alpha_0 d$. Let

$$\mathcal{D}_j = \mathcal{D} + \frac{1}{2}d + \left(1 - \frac{j-1}{N}\right) d\ ,\ j = 1, \ldots, N + 1.$$

Then $\mathcal{D} + d/2 = \mathcal{D}_{N+1} \subset \mathcal{D}_N \subset \ldots \subset \mathcal{D}_1 = \mathcal{D} + 3d/2 \subset \mathcal{D}_0 = \mathcal{D} + 2d$. Define $X_0 = X$, and for $j = 1, \ldots, N+1$, let X_j be a vector field defined on $\mathcal{D}_j$, of the form

$$X_j = (\omega + my + g_j(y,t)) \frac{\partial}{\partial x} + \frac{\partial}{\partial t} + R_j(x,y,t).$$

We construct analytic, symplectic averaging transformations $\Psi_j : \mathcal{D}_{j+1} \longrightarrow \mathcal{D}_j$, invertible on their image, such that $\Psi_{j*}^{-1} X_j = X_{j+1}$, for $j = 0, \ldots, N$. The transformation Ψ_j equals the flow over time 1 of a Hamiltonian analytic vector field Y_j, and hence

$$X_{j+1} = \Psi_{j*}^{-1} X_j = \exp(\mathrm{ad} Y_j)(X_j),$$

where $\mathrm{ad}Y(X) = [Y, X]$ denotes the Lie bracket of vector fields Y and X. The vector field Y_j satisfies the homological equation

$$\left[\omega \frac{\partial}{\partial x}, Y_j\right] = R_j - \mathrm{av}(R_j, x). \tag{10}$$

Combining this with the above expression for X_{j+1} we find that

$$\begin{aligned} g_{j+1} &= g_j + \mathrm{av}(R_j, x) \\ R_{j+1} &= \sum_{k=1}^{+\infty} \frac{1}{k!} (\mathrm{ad}Y_j)^k \left((my + g_j) \frac{\partial}{\partial x} + \frac{\partial}{\partial t} + \frac{k}{k+1} R_j + \frac{1}{k+1} \mathrm{av}(R_j, x) \right). \end{aligned}$$

We observe that $g_1 = g$, but generically $g_j \neq g$ for $j > 1$. Another difference between the first averaging step and the subsequent steps is that the transformation Ψ_0 maps its domain $\mathcal{D}_1$ into a neighborhood $\mathcal{D}_0$ of $\mathcal{D}_1$ of fixed radius $d/2$, whereas the subsequent transformations map their domains into a neighborhood of radius d/N which decreases as ω increases. This leads to an improved estimate on the first averaging step, compared to the estimate in case of an ω dependent radius.

The final step is a (symplectic) translation $T : \mathcal{D} \longrightarrow \mathcal{D} + d/2$, defined by $T(x,y,t) = (x, y + \Delta y, t)$ for some constant Δy, depending on ω but not on the phase variables. We define $\tilde{X} = T_*^{-1} X_{N+1}$, and write X_{N+1} as $X_{N+1} = X_{N+1}^{(0)} + R_{N+1}$. The translation is necessary because the frequency of the invariant torus $y = 0$ of the unperturbed vector field $X_{N+1}^{(0)}$ may have shifted away from ω due to the averaging transformations. After translation the torus $y = 0$ of the unperturbed vector field $\tilde{X}^{(0)}$ has frequency ω. This means that Δy has to satisfy the equation

$$m\Delta y + \mathrm{av}(g_{N+1}(\Delta y, t), t) = 0. \tag{11}$$

We note that this equation has a unique solution if $|\frac{\partial}{\partial y} g_{N+1}(\Delta y, t)|$ is sufficiently small compared to m, that is, if the twist condition is satisfied. To facilitate the estimates for this translation we introduce a second sequence of sets

$$\mathcal{D}'_j = \mathcal{D} + \left(1 - \frac{j-1}{N}\right) d \ , \ j = 1, \ldots, N+1,$$

with $\mathcal{D}'_1 = \mathcal{D} + d$, $\mathcal{D}'_{N+1} = \mathcal{D}$, $\mathcal{D}'_{j+1} \subset \mathcal{D}'_j$ and $\mathcal{D}'_j \subset D_j$.

The first averaging step

We use that

$$||\frac{1}{k!}(\mathrm{ad}Y)^k(X)||_{\mathcal{E}} \leq \frac{2}{3}\left(\frac{3}{\delta}||Y||_{\mathcal{E}+\delta}\right)^k ||X||_{\mathcal{E}+\delta}$$

for arbitrary analytic vector fields X and Y on $\mathcal{E}+\delta$, where $\mathcal{E} \subset (\mathbb{C}/\mathbb{Z}) \times \mathbb{C} \times (\mathbb{C}/\mathbb{Z})$ and $\delta > 0$, cf. [Fas90]. This leads to the following bounds:

$$\begin{aligned}
||g_1 - g||_{\mathcal{D}_1} &= 0 \\
||R_1||_{\mathcal{D}_1} &\leq \frac{2}{3}\sum_{k=1}^{+\infty}\left(\frac{6}{d}||Y_0||_{\mathcal{D}_0}\right)^k (\max\{2md + m\rho + ||g||_{\mathcal{D}_0} + ||R||_{\mathcal{D}_0}, 1\})\,.
\end{aligned}$$

Let $a = \frac{1}{4} + \eta + 2d$ as before, then for $\omega \geq \alpha_0$ we obtain that

$$\frac{6}{d}||Y_0||_{\mathcal{D}_0} \leq \frac{6a}{\omega d}||R||_{\mathcal{D}_0} < 1,$$

hence the series for R_1 converges, and

$$\begin{aligned}
||R_1||_{\mathcal{D}_1} &\leq \frac{6a}{\omega d}||R||_{\mathcal{D}_0}\left(1 - \frac{6a}{\omega d}||R||_{\mathcal{D}_0}\right)^{-1} C \\
&\leq \frac{\alpha_0}{2\omega}\min\left\{||R||_{\mathcal{D}_0}, \frac{dm}{2}\right\}.
\end{aligned}$$

Moreover, the flow Ψ_0 over time 1 of Y_0 maps $\mathcal{D}_1 = \mathcal{D} + 3d/2$ into $\mathcal{D}_0 = \mathcal{D} + 2d$, and

$$||\Psi_0 - Id||_{\mathcal{D}_1} \leq ||Y_0||_{\mathcal{D}_0} \leq \frac{\alpha_0 d}{38\omega}.$$

In the same way one shows that

$$||R_1||_{\mathcal{D}_1'} \leq \frac{\alpha_0 md}{8\omega}.$$

Neishtadt's procedure

A short proof of Neishtadt's theorem in the current context is included for the convenience of the reader. Analogous to the first averaging step, we have the following inequalities for $j = 1, .., N$:

$$\begin{aligned}
||g_{j+1} - g_j||_{\mathcal{D}_{j+1}} &\leq ||R_j||_{\mathcal{D}_j} \\
||R_{j+1}||_{\mathcal{D}_{j+1}} &\leq \frac{2}{3}\sum_{k=1}^{+\infty}\left(\frac{3}{\sigma}||Y_j||_{\mathcal{D}_j}\right)^k \left(\max\left\{\frac{3}{2}md + m\rho + ||g_j||_{\mathcal{D}_j} + ||R_j||_{\mathcal{D}_j}, 1\right\}\right) \\
\frac{3}{\sigma}||Y_j||_{\mathcal{D}_j} &\leq \frac{6a}{\omega\sigma}||R_j||_{\mathcal{D}_j}.
\end{aligned}$$

We claim that

$$\begin{aligned}
||g_j - g||_{\mathcal{D}_j} &\leq 2(1-2^{1-j})||R_1||_{\mathcal{D}_1} \\
||\frac{\partial}{\partial y}g_j - \frac{\partial}{\partial y}g||_{\mathcal{D}_j} &\leq \frac{1}{4}(1-2^{1-j})m \\
||R_j||_{\mathcal{D}_j} &\leq \frac{1}{2}||R_{j-1}||_{\mathcal{D}_{j-1}}
\end{aligned}$$

for $j = 1, \ldots, N+1$. Using condition (3) this obviously implies that

$$\begin{aligned}
||R_j||_{\mathcal{D}_j} &\leq 2^{1-j}||R_1||_{\mathcal{D}_1} \text{ , and} \\
||g_j - g||_{\mathcal{D}_j} + ||R_j||_{\mathcal{D}_j} &< 2||R_1||_{\mathcal{D}_1} \\
||\frac{\partial}{\partial y}g_j||_{\mathcal{D}+d/2} &< \frac{m}{2}.
\end{aligned}$$

The claim is proved by induction. For $j = 1$, it is clearly true, using $g_1 = g$ and the bound on $||R_1||_{\mathcal{D}_1}$ obtained above. For $j > 1$, the first inequality is easy to show. Furthermore,

$$\begin{aligned}
||\frac{\partial}{\partial y}g_j - \frac{\partial}{\partial y}g||_{\mathcal{D}_j} &\leq ||\frac{\partial}{\partial y}g_{j-1} - \frac{\partial}{\partial y}g||_{\mathcal{D}_{j-1}} + ||\frac{\partial}{\partial y}R_{j-1}||_{\mathcal{D}_j} \\
&\leq \frac{1}{4}(1-2^{2-j})m + \frac{1}{\sigma}2^{1-j}||R_1||_{\mathcal{D}_1} \\
&\leq \frac{1}{4}(1-2^{2-j})m + \frac{\alpha_0 md}{4\omega\sigma}2^{1-j},
\end{aligned}$$

from which the desired estimate follows. We observe that the last inequality holds since R_1 is considerably smaller than R_0, that is, the special first averaging step is necessary to preserve the twist condition. Using the induction hypothesis on R_{j-1} we observe that

$$\frac{3}{\sigma}||Y_{j-1}||_{\mathcal{D}_{j-1}} \leq \frac{6a}{\omega\sigma}||R_{j-1}||_{\mathcal{D}_{j-1}} < 1,$$

hence Ψ_{j-1} maps $\mathcal{D}_j$ into $\mathcal{D}_{j-1}$, and

$$||\Psi_{j-1} - Id||_{\mathcal{D}_j} \leq ||Y_{j-1}||_{\mathcal{D}_{j-1}} \leq \frac{\alpha_0^2 d}{38\omega^2}2^{2-j}.$$

Moreover, the expansion for R_j converges, and

$$||R_j||_{\mathcal{D}_j} \leq \frac{6a}{\omega\sigma}||R_{j-1}||_{\mathcal{D}_{j-1}}\left(1 - \frac{6a}{\omega\sigma}||R||_{\mathcal{D}_0}\right)^{-1} C \leq \frac{1}{2}||R_{j-1}||_{\mathcal{D}_{j-1}}.$$

Similarly one shows that

$$||g_j - g||_{\mathcal{D}'_j} \leq 2(1-2^{1-j})||R_1||_{\mathcal{D}'_1} \leq \frac{\alpha_0 md}{4\omega}.$$

Translation

Define $\bar{g}_{N+1} = \text{av}(g_{N+1}, t)$, and $f(\Delta y) = m\Delta y + \bar{g}_{N+1}(\Delta y)$. The translation $T : (x, y, t) \mapsto (x, y + \Delta y, t)$ is defined implicitly by (11), that is, by $f(\Delta y) = 0$. By a straightforward argument one shows that

$$f\left(\frac{\alpha_0 d}{2\omega}\right) > 0 \text{ and } f\left(-\frac{\alpha_0 d}{2\omega}\right) < 0.$$

Hence the equation $f(\Delta y) = 0$ has a solution with

$$|\Delta y| \leq \frac{\alpha_0 d}{2\omega} \leq \frac{d}{2}\ , \text{ and thus } T(\mathcal{D}) \subset \mathcal{D} + \frac{d}{2}.$$

Completion of the proof

For $\omega \in [N\alpha_0, (N+1)\alpha_0)$, let the transformation $\Psi : \mathcal{D} \mapsto \mathcal{D} + 2d$ be defined by

$$\Psi = \Psi_0 \circ \Psi_1 \circ \ldots \circ \Psi_N \circ T.$$

Then $\Psi_*^{-1} X = \tilde{X}$, where the vector fields X and $\tilde{X}$ are given by (2) and (6), respectively. The translation T is chosen such that $\text{av}(\tilde{g}, t) = O(y)$. We have that

$$||\Psi - Id||_{\mathcal{D}} \leq \sum_{j=0}^{N} ||\Psi_j - Id||_{\mathcal{D}_{j+1}} + ||T - Id||_{\mathcal{D}} < \frac{\alpha_0 d}{38\omega}\left(1 + \frac{2\alpha_0}{\omega}\right) + \frac{\alpha_0 d}{2\omega}.$$

The estimates from Neishtadt's procedure show that the components of $\tilde{X}$ satisfy

$$\begin{aligned} ||\tilde{g}||_{\mathcal{D}} &< ||g||_{\mathcal{D}+d/2} + ||R||_{\mathcal{D}_0} \\ ||\frac{\partial}{\partial y}\tilde{g}||_{\mathcal{D}} &< \frac{1}{2}m \\ ||\tilde{R}||_{\mathcal{D}} &\leq 2^{-N-1}\frac{\alpha_0}{\omega}\min\left\{||R||_{\mathcal{D}_0}, \frac{md}{2}\right\}. \end{aligned}$$

Since $\omega < (N+1)\alpha_0$, the conclusion of the theorem follows from this.

4 Construction of the twist map

In this section we first construct the Poincaré map $\tilde{P}$ of the averaged vector field $\tilde{X}$, and subsequently transform it to the normal form used in KAM theory. Let $\tilde{P}$ be defined as the return map of $\tilde{X}$ on the section $t = 0$. This map is area preserving, and of the form $\tilde{P} = \tilde{P}^{(0)} + Q$, where

$$\tilde{P}^{(0)} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + \omega + my + q(y) \\ y \end{pmatrix}\ , \ q = \text{av}(\tilde{g}, t),$$

is the return map of the unperturbed vector field

$$\tilde{X}^{(0)} = (\omega + my + \tilde{g}(y,t))\,\frac{\partial}{\partial x} + \frac{\partial}{\partial t},$$

while

$$Q(x,y) = \begin{pmatrix} Q_1(x,y) \\ Q_2(x,y) \end{pmatrix}$$

is due to the perturbation $\tilde{R}$, and is exponentially small as $\omega \to +\infty$. We define

$$\bar{\mathcal{D}}(\eta,\rho) = ((\mathbb{R}/\mathbb{Z}) \oplus [-i\eta, i\eta]) \times B(\rho) \subset (\mathbb{C}/\mathbb{Z}) \times \mathbb{C}$$

to be the projection of $\mathcal{D}(\eta,\rho)$ onto the first two coordinates. From theorem 3.1 it follows that $q(0) = 0$ and $||q'||_{\bar{\mathcal{D}}(\eta,\rho)} < \frac{1}{2}m$. According to the following lemma, the image of a set $\bar{\mathcal{D}}(\eta_1,\rho_1)$ under the Poincaré map lies inside $\bar{\mathcal{D}}(\eta,\rho)$ if η_1 and ρ_1 are sufficiently small compared to η and ρ.

Lemma 4.1 *Choose $\eta > 0$ and $\rho > 0$, and let $\mathcal{D} = \mathcal{D}(\eta,\rho)$. Let $\tilde{P}$ be the Poincaré map of the vector field $\tilde{X}$ given by (6), and assume that $\tilde{g}$ satisfies the conclusion of theorem 3.1. Let ρ_1 and η_1 satisfy*

$$\begin{aligned} \rho_1 &\le \rho - ||\tilde{R}||_{\mathcal{D}} \\ \eta_1 &\le \eta - \frac{3}{2}m\rho_1 - \left(\frac{3}{4}m + 1\right)||\tilde{R}||_{\mathcal{D}}. \end{aligned}$$

If $\rho_1 > 0$ and $\eta_1 > 0$, then

$$\begin{aligned} \tilde{P}(\bar{\mathcal{D}}(\eta_1,\rho_1)) &\subset \bar{\mathcal{D}}(\eta,\rho) \\ ||Q_1||_{\bar{\mathcal{D}}(\eta_1,\rho_1)} &\le \left(1 + \frac{3}{4}m\right)||\tilde{R}||_{\mathcal{D}} \\ ||Q_2||_{\bar{\mathcal{D}}(\eta_1,\rho_1)} &\le ||\tilde{R}||_{\mathcal{D}}. \end{aligned}$$

Proof: The proof of this lemma consists of rather straightforward estimates. Take $(x_0,y_0) \in \bar{D}(\eta_1,\rho_1)$ and let $(x(t),y(t),t)$ be the integral curve of $\tilde{X}$ that satisfies $x(0) = x_0$ and $y(0) = y_0$. Let $T > 0$ be such that $(x(t),y(t),t) \in \mathcal{D}$ for all $t \in [0,T]$. Then, for some $y_*(t)$ on the line segment from y_0 to $y(t)$,

$$\begin{aligned} |y(t) - y_0| &\le t||\tilde{R}||_{\mathcal{D}} \\ x(t) &= x_0 + \omega t + my_0 t + \int_0^t \tilde{g}(y_0,t)\,\mathrm{d}t + p(t)\ ,\ \text{where} \\ |p(t)| &\le m\int_0^t |y(t) - y_0|\,\mathrm{d}t + \int_0^t \left|\frac{\partial}{\partial y}\tilde{g}(y_*(t),t)(y(t)-y_0)\right|\mathrm{d}t + \mathrm{t}||\tilde{\mathrm{R}}||_{\mathcal{D}} \\ &\le \left(\frac{3}{4}mt^2 + t\right)||\tilde{R}||_{\mathcal{D}}, \end{aligned}$$

for $t \in [0,T]$. Since ωt and $\int_0^t \tilde{g}(0,t)\,\mathrm{d}t$ are real valued, and

$$\left|\int_0^t \tilde{g}(y_0,t) - \tilde{g}(0,t)\,\mathrm{d}t\right| \leq \frac{1}{2}m|y_0|t,$$

we have that

$$|\mathrm{Im}(x(t)-x_0)| \leq \frac{3}{2}m|y_0|t + |p(t)|.$$

Together with the bound on $y(t)$ and the conditions on η_1 and ρ_1 this shows that we can take $T=1$. The bounds on Q_1 and Q_2 follow from the above estimates.

□

We now introduce a coordinate transformation that simplifies the dynamics of the Poincaré map in the angle coordinate. Define $\Phi : (\mathbb{C}/\mathbb{Z}) \times \mathbb{C} \mapsto (\mathbb{C}/\mathbb{Z}) \times \mathbb{C}$ by

$$\Phi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \theta \\ r \end{pmatrix} = \begin{pmatrix} x \\ my + q(y) + Q_1(x,y) \end{pmatrix}.$$

Let $\delta > 0$ be the unique positive solution of

$$\eta = 6\gamma + 7\delta + \frac{11}{96} \cdot \frac{4+3m}{2+3m} c\gamma^2\delta^4,$$

where γ, c, m and ρ are as in the main theorem. The next theorem shows that Φ can be inverted and that $F = \Phi_*^{-1}\tilde{P}$ is a return map of the required form, defined on the set $\bar{\mathcal{D}}(\delta, \gamma+\delta)$, with $\gamma > 0$ and $\delta > 0$.

Theorem 4.2 *Let $\tilde{X}$ be the averaged vector field (6) defined on $\mathcal{D} = \mathcal{D}(\eta,\rho)$, with Poincaré map $\tilde{P}$. Assume ω is sufficiently large such that*

$$\eta \;\geq\; 6\gamma + 7\delta + \left(11 + \frac{33}{4}m\right) ||\tilde{R}||_{\mathcal{D}} \tag{12}$$

$$\rho \;\geq\; \frac{2}{m}(\gamma+\delta) + \left(\frac{4}{m} + 4\right) ||\tilde{R}||_{\mathcal{D}}. \tag{13}$$

Let $\eta_2 = \delta$ and

$$\rho_2 = \frac{2}{m}(\gamma+\delta) + \left(\frac{2}{m} + \frac{3}{2}\right) ||\tilde{R}||_{\mathcal{D}},$$

then Φ restricted to $\bar{\mathcal{D}}(\eta_2,\rho_2)$ is one-to-one. Moreover, there exists $U \subset \bar{\mathcal{D}}(\eta_2,\rho_2)$ such that

$$\Phi|_U : U \longrightarrow \bar{\mathcal{D}}(\delta, \gamma+\delta)$$

is invertible. Thus the map $F = (\Phi|_U)_*\tilde{P}$ is well-defined on $\bar{\mathcal{D}}(\delta,\gamma+\delta)$. It is given by

$$F : \begin{pmatrix} \theta \\ r \end{pmatrix} \mapsto \begin{pmatrix} \theta + \omega + r \\ r + f(\theta, r) \end{pmatrix},$$

where

$$||f||_{\bar{\mathcal{D}}(\delta,\gamma+\delta)} \leq (2+3m)||\tilde{R}||_{\mathcal{D}}.$$

In view of the estimate on $||\tilde{R}||_{\mathcal{D}}$ in theorem 3.1, the last terms in (12) and (13) can be made arbitrarily small by increasing ω, and hence these two conditions are always satisfied for sufficiently large ω.

Invariant circles of F obviously correspond to invariant circles of $\tilde{P}$, of the same rotation number. The transformation Φ is not area-preserving in general, and neither is F. However, F does preserve the coordinate dependent area form $(\partial y/\partial r)\mathrm{d}\theta\wedge\mathrm{d}r$.
Proof: Define η_1 and ρ_1 as in lemma 4.1, then

$$\tilde{P}(\bar{\mathcal{D}}(\eta_2,\rho_2)) \subset \bar{\mathcal{D}}(\eta_1,\rho_1) \text{ and } \tilde{P}(\bar{\mathcal{D}}(\eta_1,\rho_1)) \subset \bar{\mathcal{D}}(\eta,\rho).$$

The transformation Φ is one to one since $\frac{\partial}{\partial y}r$ is nonzero:

$$\begin{aligned} ||\frac{\partial}{\partial y}r||_{\bar{\mathcal{D}}(\eta_2,\rho_2)} &\geq m - ||q'||_{\bar{\mathcal{D}}(\eta_2,\rho_2)} - ||\frac{\partial}{\partial y}Q_1(x,y)||_{\bar{\mathcal{D}}(\eta_2,\rho_2)} \\ &> \frac{1}{2}m - \frac{1}{\rho_1-\rho_2}||Q_1||_{\bar{\mathcal{D}}(\eta_1,\rho_1)} \geq 0. \end{aligned}$$

To show that $(\Phi|_U)^{-1}$ is well-defined, consider the equation $r = my + \varepsilon q(y) + \varepsilon Q_1(x,y)$. First let $\varepsilon = 0$. If $|r| \leq \gamma+\delta$ and $|\mathrm{Im}(x)| \leq \delta$, then $|y| \leq \rho_2$. Let $\varepsilon_0 > 0$ be the largest number such that this is true for all $\varepsilon \in [0,\varepsilon_0]$. Then by continuity there are $x, y, r = my + \varepsilon_0 q(y) + \varepsilon_0 Q_1(x,y)$ such that $|\mathrm{Im}(x)| \leq \delta$, $|y| = \rho_2$ and $|r| \leq \gamma+\delta$. Hence

$$\gamma+\delta \geq (m - \frac{1}{2}m\varepsilon_0)\rho_2 - \varepsilon_0(1+\frac{3}{4}m)||\tilde{R}||_{\mathcal{D}},$$

which implies $\varepsilon_0 \geq 1$, by the definition of ρ_2.

Let $(\theta,r) \in \bar{\mathcal{D}}(\delta,\gamma+\delta)$, $(x,y) = (\Phi|_U)^{-1}(\theta,r)$ and $(x_1,y_1) = \tilde{P}(x,y)$. Then the function f is given by

$$\begin{aligned} f(\theta,r) &= my_1 + q(y_1) + Q_1(x_1,y_1) - my - q(y) - Q_1(x,y) \\ &= mQ_2(x,y) + \frac{\partial}{\partial y}q(y_*)Q_2(x,y) + Q_1(x_1,y_1) - Q_1(x,y), \end{aligned}$$

for some y_* between y and y_1. Hence

$$||f||_{\bar{\mathcal{D}}(\delta,\delta+\rho)} \leq \frac{3}{2}m||Q_2||_{\bar{\mathcal{D}}(\eta_1,\rho_1)} + 2||Q_1||_{\bar{\mathcal{D}}(\eta_1,\rho_1)}.$$

□

5 Proof of the main result

Let $F : \bar{\mathcal{D}}(\delta,\gamma+\delta) \longrightarrow (\mathbb{C}/\mathbb{Z}) \times \mathbb{C}$ be the twist map defined in the previous section. By the coordinate transformations Φ and Ψ, see theorems 3.1 and 4.2, this map is conjugate to the Poincaré map of the original vector field X. Thus, an invariant circle of F with rotation number ω corresponds to an invariant torus of X with frequency vector $(\omega, 1)$. According to the twist theorem, cf. [Her83, Her86], the map F has an invariant circle with rotation number ω if

$$||f||_{\bar{\mathcal{D}}(\delta,\gamma+\delta)} \leq \frac{c\gamma^2\delta^4}{24}\ , \ c = 5.04,$$

and ω is of constant type with parameter γ, that is,

$$|\omega - p/q| \geq \gamma|q|^{-2} \text{ for all } p/q \in \mathbb{Q}.$$

The invariant circle lies in the annulus $(\mathbb{R}/\mathbb{Z}) \times [-\gamma, \gamma]$. Using the results of the previous two sections, the smallness condition on f can be replaced by

$$(2+3m)||\tilde{R}||_{\mathcal{D}} \leq \frac{c\gamma^2\delta^4}{24}, \tag{14}$$

if the assumptions of theorem 3.1 and conditions (12) and (13) hold. According to theorem 3.1, equation (14) is satisfied for $\omega \geq \omega_0$. Using this, a tedious but straightforward calculation shows that conditions (12) and (13) are satisfied as well. Thus the existence of an invariant circle is proven.

A straightforward estimate on the transformation Φ, cf. theorem 4.2, shows that in (x, y) coordinates the invariant circle lies in a strip $(\mathbb{R}/\mathbb{Z}) \times [-y_0, y_0]$, with

$$y_0 = \frac{2}{m}\left(\gamma + \frac{4+3m}{96(2+3m)}c\gamma^2\delta^4\right) < \rho.$$

Since none of the estimates change if $\tilde{P}$ were defined on an arbitrary section $t = t_0$ (instead of $t = 0$), the y coordinate of the corresponding torus of $\tilde{X}$ is also bounded by y_0. Together with the estimate in theorem 3.1 on $||\Psi - Id||_{\mathcal{D}}$ this implies that in the original coordinates y satisfies (4). That concludes the proof.

Remark 5.1 : Actually the cited twist theorem requires a bound on the C^4 norm of f. Using a Cauchy inequality this can be replaced by the above condition. Moreover, the theorem only implies the existence of a translated curve. However, in the present Hamiltonian context, this curve has to be invariant.

A Action angle coordinates for the oscillator with polynomial potential

Recall that the oscillator with polynomial potential is given by the real-analytic Hamiltonian

$$H(x, y, t) = \frac{1}{2} y^2 + V(x, t), \tag{15}$$

corresponding to the equations

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -V_x(x, t), \\ \dot{t} &= 1, \end{aligned}$$

where $V(x, t) = \sum_{k=1}^{2n} a_k(t) x^k$, $n \geq 2$ is 1-periodic in t, and $a_{2n} > 0$. We will construct action-angle coordinates for this system, such that the action is close to constant along solutions far away from the origin $x = y = 0$. Subsequently the resulting system will be averaged once in t to remove the time dependence from the leading terms.

Fixed t

We first take t to be a fixed parameter. Choose h_0 to be strictly larger than any critical value of $x \mapsto V(x, t)$. Then $H(x, y, t) = h$ corresponds to a simple closed curve for any $h \geq h_0$. We construct action-angle coordinates (θ, I) on $A = \{(x, y) : H(x, y, t) \geq h_0\}$, using a generating function $S = S(x, I, t)$:

$$\begin{aligned} y &= \frac{\partial}{\partial x} S(x, I, t), \\ \theta &= \frac{\partial}{\partial I} S(x, I, t). \end{aligned}$$

For given h, let $x_{\max}$ and $x_{\min}$ be the maximal and minimal x coordinates on the curve $H = h$, i.e., $V(x_{\max}, t) = V(x_{\min}, t) = h$. Define the action $I = I(h, t)$ to be the area enclosed by this curve (see figure 2):

$$\begin{aligned} I(h, t) &= \oint_{H=h} y \, \mathrm{d}x \\ &= 2\sqrt{2} \int_{x_{\min}}^{x_{\max}} \sqrt{h - \sum_{k=1}^{2n} a_k(t) x^k} \, \mathrm{d}x \end{aligned}$$

where y was solved from (15). We write the integral as

$$\int_{x_{\min}}^{x_{\max}} \ldots = \int_{x_{\min}}^{0} \ldots + \int_{0}^{x_{\max}} \ldots =: I_- + I_+,$$

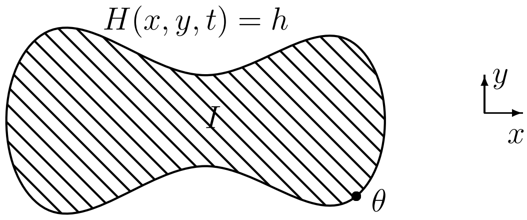

Figure 2: Definition of action-angle coordinates. The action I is the area enclosed by $H = h$, for fixed t, while the angle θ determines a point on the boundary $H = h$.

and use the substitution $x = ux_{\min}$ to obtain

$$\begin{aligned} I_- &= \sqrt{a_{2n}(t)}|x_{\min}|^{n+1}\int_0^1 \sqrt{1-u^{2n}} + O(x_{\min}^{-1})\,\mathrm{d}x \\ &= b_n\sqrt{a_{2n}(t)}|x_{\min}|^{n+1} + O(x_{\min}^n) \\ &= b_n a_{2n}(t)^{-1/(2n)}h^{(n+1)/(2n)} + O(\sqrt{h}), \end{aligned}$$

where b_n is an increasing sequence, $b_2 \approx 0.87$ and $\lim_{n\to\infty} b_n = 1$, and we used that $h^{(n+1)/(2n)} = a_{2n}^{1/2+1/(2n)}|x_{\min}|^{n+1} + O(x_{\min}^n)$. In fact,

$$b_n = \int_0^1 \sqrt{1-u^{2n}}\,\mathrm{d}u = \frac{1}{2}\sqrt{\pi}\frac{\Gamma(1+\frac{1}{2n})}{\Gamma(\frac{3}{2}+\frac{1}{2n})}.$$

Consequently, we find that

$$I(h,t) = 4\sqrt{2}b_n a_{2n}(t)^{-1/(2n)}h^{(n+1)/(2n)} + O(\sqrt{h}), \tag{16}$$

with inverse

$$\begin{aligned} K_0(I,t) &= c_n a_{2n}(t)^{1/(n+1)}I^{2n/(n+1)} + O(I^{(2n-1)/(n+1)}), \\ c_n &= \left(4\sqrt{2}b_n\right)^{-2n/(n+1)}, \end{aligned} \tag{17}$$

that is, $K_0(I(h,t)) = h$. The function K_0 is the Hamiltonian in action-angle coordinates for fixed t. The sequence c_n is decreasing, with $c_2 \approx 0.12$ and $\lim_{n\to\infty} c_n = \frac{1}{32}$.

Now define $S(x, I, t)$ to be the area enclosed by $H = K_0(I,t)$ between $x_{\min}$ and x:

$$S(x,I,t) = \int_{x_{\min},H=K_0(I,t)}^{x} y(x,H,t)\,\mathrm{d}x. \tag{18}$$

Then $y = \frac{\partial}{\partial x}S$ and we define $\theta = \frac{\partial}{\partial I}S$. Observe that S is defined modulo I, and hence θ is an angle.

Varying t

Now we let t vary. In the same action-angle coordinates, the Hamiltonian is given by

$$K(\theta, I, t) = K_0(I, t) + K_1(x(\theta, I, t), I, t),$$

where $K_1 = \frac{\mathrm{d}}{\mathrm{d}t} S$, see for example [Arn89]. We note that $\frac{\partial}{\partial t} S$ is unique.

Differentiating the integral formula (18) for S with respect to the boundary $x_{\min}$ gives zero, and thus we obtain, using $K_0(I,t) - V(x,t) = \sum_{k=1}^{2n} a_k(t)(x_{\min}^k - x^k)$ and substituting $x = x_{\min} u$ as before,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} S &= \int_{x_{\min}}^{x} \frac{\partial}{\partial t} y + \frac{\partial}{\partial h} y \frac{\partial}{\partial t} K_0(I,t)\,\mathrm{d}\xi \\
&= \frac{1}{2}\sqrt{2} \int_{x_{\min}}^{x} (K_0(I,t) - V(\xi,t))^{-1/2} \left(\frac{\partial}{\partial t} K_0(I,t) - \frac{\partial}{\partial t} V(\xi,t)\right) \mathrm{d}\xi \\
&= \frac{1}{2}\sqrt{2} |x_{\min}|^{n+1} \int_{u}^{1} \left(a_{2n}(1 - s^{2n})\right)^{-1/2} a_{2n}'(t) \left(\frac{1}{n+1} - s^{2n}\right) \mathrm{d}s + O(x_{\min}^n) \\
&= \frac{\tilde{d}_n a_{2n}'(t)}{8 b_n a_{2n}(t)} I + O(I^{n/(n+1)}),
\end{aligned}$$

where

$$\tilde{d}_n = \int_{x/x_{\min}}^{1} \frac{(n+1)^{-1} - s^{2n}}{\sqrt{1 - s^{2n}}} \mathrm{d}s,$$

hence (using $x_{\max} \approx -x_{\min}$, and we can set them equal modulo the remainder in the formula above),

$$|\tilde{d}_n| \leq d_n := \int_{-1}^{1} \frac{1}{\sqrt{1 - s^{2n}}} \mathrm{d}s.$$

Here d_n is a decreasing sequence, $d_2 \approx 2.62$, and $\lim_{n \to +\infty} d_n = 2$. We conclude that

$$\begin{aligned}
K_1 = \frac{\mathrm{d}}{\mathrm{d}t} S &= f_n \frac{a_{2n}'(t)}{a_{2n}(t)} I + O(I^{n/(n+1)}) \text{ , where} \\
|f_n| &\leq \frac{d_n}{8 b_n}
\end{aligned}$$

Averaging

The differential equations corresponding to the Hamiltonian K are

$$\begin{aligned}
\dot{\theta} &= \Omega(I,t) + F(\theta, I, t) \\
\dot{I} &= G(\theta, I, t) \\
\dot{t} &= 1,
\end{aligned}$$

where $F = \frac{\partial}{\partial I} K_1$, $G = -\frac{\partial}{\partial \theta} K_1$, and

$$\Omega = \frac{\partial}{\partial I} K_0 = \frac{2n}{n+1} c_n a_{2n}(t)^{1/(n+1)} I^{(n-1)/(n+1)} + O(I^{(n-2)/(n+1)}).$$

Thus $\frac{\partial}{\partial I}\Omega > 0$ for I large enough and $\Omega \to +\infty$ as $I \to +\infty$.

We apply a symplectic transformation of time and the Hamiltonian to average out the time dependence of K_0. Define $\bar{a} = \mathrm{av}(a_{2n}(t)^{1/(n+1)}, t)$, and

$$\bar{K}_0(I) = \mathrm{av}(K_0(I,t),t) = c_n \bar{a} I^{2n/(n+1)} + O(I^{(2n-1)/(n+1)}), \tag{19}$$

then the transformation is of the form $(t, K) \mapsto (\tau, \bar{K})$, determined by a generating function $S(t, \bar{K})$:

$$K = \bar{K} + \frac{\partial}{\partial t} S(t, \bar{K}) \ , \ \tau = t + \frac{\partial}{\partial \bar{K}} S(t, \bar{K}).$$

We want the transformation to map K_0 to $\bar{K}_0$, that is, we want that

$$K_0(I,t) = \bar{K}_0(I) + \frac{\partial}{\partial t} S(t, \bar{K}_0(I)),$$

which is equivalent to

$$\frac{\partial}{\partial t} S(t,k) = \tilde{K}_0\left(\bar{K}_0^{-1}(k), t\right) \ , \ \tilde{K}_0 = K_0 - \bar{K}_0. \tag{20}$$

A short calculation reveals that the transformation is of the form $\tau(t,K) = \tau_0(t) + O(K^{-1/(2n)}))$ and $\bar{K} = (\tau_0'(t))^{-1} K + O(K^{(2n-1)/(2n)})$, where

$$\tau_0(t) = \frac{1}{\bar{a}} \int_0^t a(s)^{1/(n+1)}\, \mathrm{d}s.$$

We observe that $\tau_0(0) = 0$, $\tau_0(t+1) = \tau_0(t) + 1$, and $\tau_0'(t) > 0$ (since $a_{2n}(t) > 0$). The Hamiltonian $\bar{K}$ is of the form $\bar{K}(\theta, I, \tau) = \bar{K}_0(I) + \bar{K}_1(\theta, I, \tau)$, where $\bar{K}_0$ is as in (19), and

$$\bar{K}_1(\theta, I, \tau) = f_n \frac{\bar{a} a_{2n}'(\tau)}{a_{2n}(\tau)^{(n+2)/(n+1)}} I + O(I^{n/(n+1)}).$$

The corresponding system is

$$\begin{aligned} \dot{\theta} &= \bar{\Omega}(I) + \frac{\partial}{\partial I} \bar{K}_1 \\ \dot{I} &= -\frac{\partial}{\partial \theta} \bar{K}_1 \\ \dot{\tau} &= 1, \end{aligned}$$

where $\bar{\Omega} = \frac{\partial}{\partial I} \bar{K}_0$ and the dot now denotes differentiation with respect to τ. Indeed, $\bar{K}_1$ is given by

$$\begin{aligned} \bar{K}_1 + \frac{\partial}{\partial t} S(t, \bar{K}_0 + \bar{K}_1) - \frac{\partial}{\partial t} S(t, \bar{K}_0) &= K_1 \text{ , that is,} \\ \left(1 + \frac{\partial^2}{\partial k \partial t} S(t, \bar{K}_0 + \sigma \bar{K}_1)\right) \bar{K}_1 &= K_1, \end{aligned}$$

for some $\sigma \in (0,1)$. The formula for $\bar{K}_1$ now follows from the fact that

$$1 + \frac{\partial^2}{\partial k \partial t} S(t,k) = \frac{1}{\bar{a}} a^{1/(n+1)} + O(k^{-1/(2n)}).$$

References

[ALD83] S. Aubry and P. Y. Le Daeron. The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states. *Phys. D*, 8(3):381–422, 1983.

[Arn63] V. I. Arnol'd. Small denominators and problems of stability of motion in classical and celestial mechanics. *Uspehi Mat. Nauk*, 18(6 (114)):91–192, 1963. translated in Russian Math. Surveys.

[Arn89] V.I. Arnol'd. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, second edition, 1989.

[BHN$^+$] H.W. Broer, I. Hoveijn, M. van Noort, C. Simó, and G. Vegter. The parametrically forced pendulum: a case study in $1\frac{1}{2}$ degree of freedom. Submitted.

[BHS96] H.W. Broer, G.B. Huitema, and M.B. Sevryuk. *Quasi-periodic motions in families of dynamical systems. Order amidst chaos*, volume 1645 of *Lecture Notes in Mathematics*. Springer-Verlag, 1996.

[BNS] H.W. Broer, M. van Noort, and C. Simó. Existence and measure of 2-quasiperiodicity in Hamiltonian one-and-a-half degree of freedom systems. To appear in the proceedings of Equadiff 2003.

[BRS96] H.W. Broer, R. Roussarie, and C. Simó. Invariant circles in the Bogdanov-Takens bifurcation for diffeomorphisms. *Ergodic Theory Dynam. Systems*, 16(6):1147 – 1172, 1996.

[DG96] A. Delshams and P. Gutiérrez. Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system. *J. Diff. Eq.*, 131(2):277 – 303, 1996.

[DZ87] R. Dieckerhoff and E. Zehnder. Boundedness of solutions via the twist-theorem. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 14(1):79–95, 1987.

[Fas90] F. Fassò. Lie series method for vector fields and Hamiltonian perturbation theory. *Z. Angew. Math. Phys.*, 41(6):843 – 864, 1990.

[GGM99] G. Gallavotti, G. Gentile, and V. Mastropietro. A field theory approach to Lindstedt series for hyperbolic tori in three time scales problems. *J. Math. Phys.*, 40(12):6430–6472, 1999.

[Her83] M.R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau, vol. 1*, volume 103-104 of *Astérisque*. Société Mathématique de France, 1983.

[Her86] M.R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau, vol. 2*, volume 144 of *Astérisque*. Société Mathématique de France, 1986.

[Jia99] M. Jiang. On a theorem by Mather and Aubry-Mather sets for planar Hamiltonian systems. *Sci. China Ser. A*, 42(11):1121–1128, 1999.

[JV97] À. Jorba and J. Villanueva. On the persistence of lower-dimensional invariant tori under quasi-periodic perturbations. *J. Nonlinear Sci.*, 7(5):427–473, 1997.

[Lev90a] M. Levi. KAM theory for particles in periodic potentials. *Ergod. Th. Dynam. Sys.*, 10(4):777 – 785, 1990.

[Lev90b] M. Levi. A period-adding phenomenon. *SIAM J. Appl. Math.*, 50(4):943 – 955, 1990.

[Lev91] M. Levi. Quasiperiodic motions in superquadratic time-periodic potentials. *Comm. Math. Phys.*, 143(1):43 – 83, 1991.

[Lit66] J. E. Littlewood. Unbounded solutions of an equation $\ddot{y} + g(y) = p(t)$, with $p(t)$ periodic and bounded, and $g(y)/y \to \infty$ as $y \to \pm\infty$. *J. London Math. Soc.*, 41:491–507, 1966.

[Liu89] B. Liu. Boundedness for solutions of nonlinear Hill's equations with periodic forcing terms via Moser's twist theorem. *J. Differential Equations*, 79(2):304–315, 1989.

[LL91] S. Laederich and M. Levi. Invariant curves and time-dependent potentials. *Ergod. Th. Dynam. Sys.*, 11(2):365–378, 1991.

[LZ95] M. Levi and E. Zehnder. Boundedness of solutions for quasiperiodic potentials. *SIAM J. Math. Anal.*, 26(5):1233 – 1256, 1995.

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

[Mos62] J. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. II*, pages 1 – 20, 1962.

[Mos86] J. Moser. Recent developments in the theory of Hamiltonian systems. *SIAM Rev.*, 28(4):459–485, 1986.

[Mos89a] J. Moser. Minimal foliations on a torus. In M. Giaquinta, editor, *Topics in calculus of variations (Montecatini Terme, 1987)*, volume 1365 of *Lecture Notes in Mathematics*, pages 62–99. Springer-Verlag, 1989.

[Mos89b] J. Moser. Quasi-periodic solutions of nonlinear elliptic partial differential equations. *Bol. Soc. Brasil. Mat. (N.S.)*, 20(1):29–45, 1989.

[Nei84] A.I. Neishtadt. The separation of motions in systems with rapidly rotating phase. *J. Appl. Math. and Mech.*, 48(2):133 – 139, 1984.

[Pös82] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.*, 35(5):653 – 696, 1982.

[Rüs70] H. Rüssmann. Kleine Nenner. I. Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, pages 67 – 105, 1970.

[Rüs83] H. Rüssmann. On the existence of invariant curves of twist mappings of an annulus. In J. Palis Jr., editor, *Geometric Dynamics*, volume 1007 of *Lecture Notes in Mathematics*, pages 677 – 719. Springer-Verlag, 1983.

[SV01] C. Simó and C. Valls. A formal approximation of the splitting of separatrices in the classical Arnold's example of diffusion with two equal parameters. *Nonlinearity*, 14(6):1707–1760, 2001.

[You90] J. You. Invariant tori and Lagrange stability of pendulum-type equations. *J. Diff. Eq.*, 85:54 – 65, 1990.