Chin. Ann. Math. Ser. B
38(5), 2017, 1037–1046
DOI: 10.1007/s11401-017-1020-x

Chinese Annals of Mathematics, Series B

Boundedness of Solutions for Duffing Equation with Low Regularity in Time*

Xiaoping YUAN[1]

Abstract It is shown that all solutions are bounded for Duffing equation $\ddot{x} + x^{2n+1} + \sum\limits_{j=0}^{2n} P_j(t)x^j = 0$, provided that for each $n+1 \le j \le 2n$, $P_j \in C^\gamma(\mathbb{T}^1)$ with $\gamma > 1 - \frac{1}{n}$ and for each j with $0 \le j \le n$, $P_j \in L(\mathbb{T}^1)$ where $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$.

Keywords Duffing equation, Boundedness of solutions, Lagrange stability, Moser twist theorem, Quasi-periodic solution
2000 MR Subject Classification 34D20, 37J40, 70K43

1 Introduction

In 1962, Moser [6] proposed to study the boundedness of all solutions (Lagrange stability) for Duffing equation

$$\ddot{x} + \beta x^3 + \alpha x = P(t), \quad P \in C(\mathbb{T}^1),\ \mathbb{T}^1 := \mathbb{R}/\mathbb{Z},$$

where $\beta > 0$, $\alpha \in \mathbb{R}$ are constants.

In 1976, Morris [5] proved the boundedness of all solutions for

$$\ddot{x} + 2x^3 = P(t).$$

Subsequently, Morris' boundedness results was, by Dieckerhoff-Zehnder [1] in 1987, extended to a wider class of systems

$$\ddot{x} + x^{2n+1} + \sum_{j=0}^{2n} P_j(t)x^j = 0, \quad n \ge 1, \tag{1.1}$$

where

$$P_j \in C^\nu, \quad \nu \ge 1 + \frac{4}{n} + [\log_2^n] \to \infty \quad \text{as } n \to \infty.$$

Then they remarked that:

"It is not clear whether the boundedness phenomenon is related to the smoothness in the t-variable or whether this requirement is a shortcoming of our proof."

In 1989 and 1992, Liu [3–4] proved the boundedness for

$$\ddot{x} + x^{2n+1} + a(t)x + P(t) = 0, \quad a \in C^0(\mathbb{T}^1),\ P \in C^0(\mathbb{T}^1).$$

Manuscript received June 20, 2017; Revised July 7, 2017.
[1]School of Mathematical Sciences, Fudan University, Shanghai 200433, China.
E-mail: xpyuan@fudan.edu.cn
*Project supported by the National Natural Science Foundation of China (No. 11421061).

In 1991, Laederich-Levi [2] relaxed the smoothness requirement of $P_j(t)$ $(j = 0, 1, \cdots, 2n)$ for (1.1) to

$$P_j \in C^{5+\varepsilon}(\mathbb{T}^1), \quad \varepsilon > 0.$$

In his PhD thesis (1995), the present author further relaxed the requirement to C^2 (see [12–14]).

In the present paper, we will relax the smoothness requirement to Hölder continuity. More exactly, we have the following theorem

Theorem 1.1 *For arbitrary given constant $\gamma \in (1-\frac{1}{n}, 1)$, assume $P_j \in C^\gamma(\mathbb{T}^1)$ for $n+1 \le j \le 2n$ and $P_j \in L(\mathbb{T}^1)$ for $0 \le j \le n$. Then every solution $x(t)$ of the equation* (1.1)*,*

$$\ddot{x} + x^{2n+1} + \sum_{j=0}^{2n} P_j(t)x^j = 0, \quad n \ge 1,$$

is bounded, i.e. it exists for all $t \in \mathbb{R}$ and $\sup\limits_{t\in\mathbb{R}}(|x(t)| + |\dot{x}(t)|) < C < \infty$, where the constant $C = C(x(0), \dot{x}(0))$ depends the initial data $(x(0), \dot{x}(0))$.

Remark 1.1 In [11], it is proved that there is a continuous periodic function $p(t)$ such that the Duffing equation $\frac{d^2x}{dt^2} + x^{2n+1} + p(t)x^l = 0$ with $p(t) \in C^0(\mathbb{T}^1)$, $n \ge 2$, $2n+1 > l \ge n+2$ possesses an unbounded solution, which shows that the Hölder continuity of the coefficients P_j's is necessary for the boundedness of solutions. In this sense, the result is almost sharp.

2 Action-Angle Variable

Replacing x by Ax in (1.1), we get

$$A\ddot{x} + A^{2n+1}x^{2n+1} + \sum_{j=0}^{2n} P_j(t)x^jA^j = 0, \tag{2.1}$$

where A is a constant large enough. That is,

$$\ddot{x} + A^{2n}x^{2n+1} + \sum_{j=0}^{2n} P_j(t)x^jA^{j-1} = 0. \tag{2.2}$$

Let

$$y = A^{-n}\dot{x} \quad \text{or} \quad \dot{x} = A^n y.$$

Then

$$\dot{y} = A^{-n}\ddot{x} = A^{-n}\Big(- A^{2n}x^{2n+1} - \sum_{j=0}^{2n} P_j(t)x^jA^{j-1}\Big) = -A^n x^{2n+1} - \sum_{j=0}^{2n} P_j(t)x^jA^{j-n-1}.$$

Thus,

$$\dot{x} = \frac{\partial H}{\partial y}, \quad \dot{y} = -\frac{\partial H}{\partial x}, \tag{2.3}$$

where

$$H = A^n\Big(\frac{1}{2}y^2 + \frac{1}{2(n+1)}x^{2(n+1)}\Big) + \sum_{j=0}^{2n} \frac{P_j(t)}{j+1}x^{j+1}A^{j-n-1}. \tag{2.4}$$

Let $\mathbb{T}_s^1 = \{t \in \mathbb{C}/\mathbb{Z} : |\mathrm{Im}\, t| < s\}$ for any $s > 0$. Consider an auxiliary Hamiltonian system

$$\dot{x} = \frac{\partial H_0}{\partial y}, \quad \dot{y} = -\frac{\partial H_0}{\partial x}, \quad H_0 = \frac{1}{2}y^2 + \frac{1}{2(n+1)}x^{2(n+1)}. \tag{2.5}$$

Let $(x_0(t), y_0(t))$ be the solution to (2.5) with initial $(x_0(0), y_0(0)) = (1, 0)$. Then this solution is clearly periodic. Let T_0 be its minimal positive period. By energy conservation, we have

$$(n+1)y_0^2(t) + x_0^{2n+2}(t) \equiv 1, \quad t \in \mathbb{R}, \tag{2.6}$$

by which, we construct the following symplectic transformation

$$\Psi_0 : \quad \begin{cases} x = c^\alpha I^\alpha x_0(\theta T_0), \\ y = c^\beta I^\beta y_0(\theta T_0), \end{cases}$$

where $\alpha = \frac{1}{n+2}$, $\beta = 1-\alpha = \frac{n+1}{n+2}$, $c = \frac{1}{\alpha T_0}$ and where $(I, \theta) \in \mathbb{R}^+ \times \mathbb{T}^1$ is action-angle variables. By calculation, $\det \frac{\partial\,(x,y)}{\partial(\theta,I)} = 1$. Thus the transformation is indeed symplectic. Clearly $\Psi_0(I, \theta)$ is analytic in $(I, \theta) \in \mathbb{R}^+ \times \mathbb{T}_{s_0}^1$ with some constant $s_0 > 0$.

Under Ψ_0, (2.3) is changed

$$\dot{\theta} = \frac{\partial H}{\partial I}, \quad \dot{I} = -\frac{\partial H}{\partial \theta}, \tag{2.7}$$

where $H = H_0(I) + R(I, \theta, t)$ with

$$H_0(I) = d \cdot A^n \cdot I^{2\beta} = d \cdot A^n \cdot I^{\frac{2(n+1)}{n+2}}, \quad d = \frac{c^{2\beta}}{2(n+1)} \tag{2.8}$$

and

$$R(I, \theta, t) = \sum_{j=0}^{2n} \frac{P_j(t)}{j+1}(c^{\frac{1}{n+1}}\, x_0(\theta T_0))^{j+1} A^{j-n-1} I^{\frac{j+1}{n+2}}. \tag{2.9}$$

Clearly, $R(I, \theta, t) = O(A^{n-1})$ for $A \to \infty$ and fixed I in some compact intervals.

3 Approximation Lemma

First, we cite an approximation lemma (see [9–10] for the detail). We start by recalling some definitions and setting some new notations. Assume that X is a Banach space with the norm $\|\cdot\|_X$. First recall that $C^\mu(\mathbb{R}^n; X)$ for $0 < \mu < 1$ denotes the space of bounded Hölder continuous functions $f : \mathbb{R}^n \to X$ with the form

$$\|f\|_{C^\mu, X} = \sup_{0<|x-y|<1} \frac{\|f(x) - f(y)\|_X}{|x-y|^\mu} + \sup_{x \in \mathbb{R}^n} \|f(x)\|_X.$$

If $\mu = 0$ then $\|f\|_{C^\mu, X}$ denotes the sup-norm. For $\ell = k+\mu$ with $k \in \mathbb{N}$ and $0 \le \mu < 1$, we denote by $C^\ell(\mathbb{R}^n; X)$ the space of functions $f : \mathbb{R}^n \to X$ with Hölder continuous partial derivatives, i.e., $\partial^\alpha f \in C^\mu(\mathbb{R}^n; X_\alpha)$ for all multi-indices $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}^n$ with the assumption that $|\alpha| := |\alpha_1| + \cdots + |\alpha_n| \le k$ and X_α is the Banach space of bounded operators $T : \prod^{|\alpha|}(\mathbb{R}^n) \to X$ with the norm $\|T\|_{X_\alpha} = \sup\{\|T(u_1, u_2, \cdots, u_{|\alpha|})\|_X : \|u_i\| = 1,\ 1 \le i \le |\alpha|\}$. We define the norm $\|f\|_{C^\ell} = \sup\limits_{|\alpha| \le \ell} \|\partial^\alpha f\|_{C^\mu, X_\alpha}$.

Theorem 3.1 (Jackson-Moser-Zehnder) *Let $f \in C^\ell(\mathbb{R}^n; X)$ for some $\ell > 0$ with finite C^ℓ norm over $\mathbb{R}^n$. Let ϕ be a radical-symmetric, C^∞ function, having as support the closure of the unit ball centered at the origin, where ϕ is completely flat and takes value 1, and let $K = \widehat{\phi}$ be its Fourier transform. For all $\sigma > 0$ define*

$$f_\sigma(x) := K_\sigma * f = \frac{1}{\sigma^n} \int_{\mathbb{T}^n} K\Big(\frac{x-y}{\sigma}\Big) f(y) \mathrm{d}y.$$

Then there exists a constant $C \geq 1$ depending only on ℓ and n such that the following holds: For any $\sigma > 0$, the function $f_\sigma(x)$ is a real-analytic function from $\mathbb{C}^n$ to X such that if Δ_σ^n denotes the n-dimensional complex strip of width σ,

$$\Delta_\sigma^n := \{x \in \mathbb{C}^n \big| |\mathrm{Im} x_j| \leq \sigma,\ 1 \leq j \leq n\},$$

then for $\forall \alpha \in \mathbb{N}^n$ with $|\alpha| \leq \ell$ one has

$$\sup_{x \in \Delta_\sigma^n} ||\partial^\alpha f_\sigma(x) - \sum_{|\beta| \leq \ell - |\alpha|} \frac{\partial^{\beta+\alpha} f(\mathrm{Re} x)}{\beta!} (\sqrt{-1} \mathrm{Im} x)^\beta ||_{X_\alpha} \leq C \|f\|_{C^\ell} \sigma^{\ell - |\alpha|}, \tag{3.1}$$

and for all $0 \leq s \leq \sigma$,

$$\sup_{x \in \Delta_s^n} \|\partial^\alpha f_\sigma(x) - \partial^\alpha f_s(x)\|_{X_\alpha} \leq C \|f\|_{C^\ell} \sigma^{\ell - |\alpha|}. \tag{3.2}$$

The function f_σ preserves periodicity (i.e., if f is T-periodic in any of its variable x_j, so is f_σ).

By this theorem, for each $P_j \in C^\gamma(\mathbb{T}^1)$, $j = n+1, n+2, \cdots, 2n$, and any $\varepsilon > 0$, there is a real analytic function[1] $P_{j,\varepsilon}(t)$ from $\mathbb{T}_\varepsilon^1$ to $\mathbb{C}$ such that

$$\sup_{t \in \mathbb{T}^1} |P_{j,\varepsilon}(t) - P_j(t)| \leq C\, \varepsilon^\gamma\, \|P_j\|_{C^\gamma} \tag{3.3}$$

and

$$\sup_{t \in \mathbb{T}_\varepsilon^1} |P_{j,\varepsilon}(t)| \leq C \|P_j\|_{C^\gamma}. \tag{3.4}$$

Write

$$R(I, \theta, t) = R_\varepsilon(I, \theta, t) + R^\varepsilon(I, \theta, t), \tag{3.5}$$

where

$$R_\varepsilon(I, \theta, t) = \sum_{j=n+1}^{2n} \frac{1}{j+1} A^{j-n-1} I^{\frac{j+1}{n+2}} c^{\frac{j}{n+2}} x_0^{j+1}(\theta T_0) P_{j,\varepsilon}(t), \tag{3.6}$$

$$\begin{aligned} R^\varepsilon(I, \theta, t) = &\sum_{j=0}^{n} \frac{1}{j+1} A^{j-n-1} I^{\frac{j+1}{n+2}} c^{\frac{j}{n+2}} x_0^{j+1}(\theta T_0) P_j(t) \\ &+ \sum_{j=n+1}^{2n} \frac{1}{j+1} A^{j-n-1} I^{\frac{j+1}{n+2}} c^{\frac{j}{n+2}} x_0^{j+1}(\theta T_0)(P_j(t) - P_{j,\varepsilon}(t)). \end{aligned} \tag{3.7}$$

[1]A complex value function $f(t)$ of complex variable t in some domain in $\mathbb{C}$ is called real analytic if it is analytic in the domain and is real for real argument t.

Now let us restrict I to some compact intervals, $[1,4]$, say. Let $A^{-1} < \varepsilon_0$.

For a sufficiently small $\varepsilon_0 > 0$, letting

$$\varepsilon = \Big(\frac{\varepsilon_0}{A^{n-1}}\Big)^{\frac{1}{\gamma}}, \tag{3.8}$$

by Theorem 3.1, we have the following facts:

(i) $R^\varepsilon(I,\theta,t)$ is real analytic in $(I,\theta) \in [1,4] \times \mathbb{T}^1_{s_0}$ for fixed $t \in \mathbb{T}^1$ and $R^\varepsilon(I,\theta,\cdot) \in L^1(\mathbb{T}^1)$ for fixed $(I,\theta) \in [1,4] \times \mathbb{T}^1_{s_0}$, and

$$\sup_{(I,\theta,t)\in[1,4]\times\mathbb{T}^1_{s_0}\times\mathbb{T}^1} |R^\varepsilon(I,\theta,t)| \le C\varepsilon_0, \tag{3.9}$$

where C is a constant[2] depending on only $\|P_j\|_{C^\gamma}$.

(ii) $R_\varepsilon(I,\theta,t)$ is real analytic in $(I,\theta,t) \in [1,4] \times \mathbb{T}^1_{s_0} \times \mathbb{T}^1_\varepsilon$ and

$$\sup_{(I,\theta,t)\in[1,4]\times\mathbb{T}^1_{s_0}\times\mathbb{T}^1_\varepsilon} |R_\varepsilon(I,\theta,t)| \le CA^{n-1}, \tag{3.10}$$

where C is a constant depending on only $\|P_j\|_{C^\gamma}$. Therefore, we have

$$H(I,\theta,t) = H_0(I) + R_\varepsilon(I,\theta,t) + R^\varepsilon(I,\theta,t). \tag{3.11}$$

4 Symplectic Transformations

We will look for a series of symplectic transformations $\Psi_1, \cdots, \Psi_N$ such that $H^{(N)} = H \circ \Psi_1 \circ \cdots \circ \Psi_N = H_0^N + O(\varepsilon_0)$, where $H_0^N(\mu) \approx A^n \mu^{\frac{2(n+1)}{n+2}}$ such that Moser's twist theorem works for $H^{(N)}$.

To this end, let $\Psi_1 : (\mu,\phi) \to (I,\theta)$ is implicitly defined by

$$\Psi_1 : \begin{cases} I = \mu + \dfrac{\partial S_1}{\partial \theta}, \\ \phi = \theta + \dfrac{\partial S_1}{\partial \mu} \end{cases}$$

with $S_1 = S_1(\mu,\theta,t)$ to be specified latter. If Ψ_1 is well-defined, then it is symplectic, since

$$\mathrm{d}I \wedge \mathrm{d}\theta = \Big(1 + \frac{\partial^2 S_1}{\partial\mu\partial\theta}\Big)\mathrm{d}\mu \wedge \mathrm{d}\theta = \mathrm{d}\mu \wedge \mathrm{d}\phi.$$

The transformed Hamiltonian function $H^{(1)}(\mu,\phi,t) = H \circ \Psi_1(\mu,\phi,t)$. We express temporarily in the variable (μ,θ) instead of (μ,ϕ):

$$H^{(1)}(\mu,\theta,t) = H\Big(\mu + \frac{\partial S_1}{\partial\theta}, \theta, t\Big) + \frac{\partial S_1}{\partial t}. \tag{4.1}$$

By Taylor's formula and (3.11)

$$\begin{aligned} H^{(1)}(\mu,\theta,t) &= H_0\Big(\mu + \frac{\partial S_1}{\partial\theta}, \theta, t\Big) + R_\varepsilon\Big(\mu + \frac{\partial S_1}{\partial\theta}, \theta, t\Big) + R^\varepsilon \circ \Psi_1(\mu,\phi,t) + \frac{\partial S_1}{\partial t} \\ &= H_0(\mu) + \partial_\mu H_0(\mu)\frac{\partial S_1}{\partial\theta} + R_\varepsilon(\mu,\theta,t) + R^1_\varepsilon(\mu,\theta,t) + R^\varepsilon \circ \Psi_1(\mu,\phi,t), \end{aligned} \tag{4.2}$$

[2] Denote by C a universal constant which may be different in different place.

where

$$R_\varepsilon^1(\mu,\theta,t)=\int_0^1(1-\tau)\partial_\mu^2 H_0\Big(\mu+\frac{\partial S_1}{\partial\theta}\tau,\theta,t\Big)\Big(\frac{\partial S_1}{\partial\theta}\Big)^2\mathrm{d}\tau$$
$$+\int_0^1\partial_\mu R_\varepsilon\Big(\mu+\frac{\partial S_1}{\partial\theta}\tau,\theta,t\Big)\frac{\partial S_1}{\partial\theta}\mathrm{d}\tau+\frac{\partial S_1}{\partial t}. \tag{4.3}$$

Let

$$\partial_\mu H_0\cdot\frac{\partial S_1}{\partial\theta}+R_\varepsilon(\mu,\theta,t)=[R_\varepsilon](\mu,t),\quad [R_\varepsilon](\mu,t)=\int_0^1 R_\varepsilon(\mu,\theta,t)\mathrm{d}\theta. \tag{4.4}$$

Then

$$H^{(1)}(\mu,\theta,t)=H_0(\mu)+[R_\varepsilon](\mu,t)+R_\varepsilon^1(\mu,\theta,t)+R^\varepsilon\circ\Psi_1(\mu,\phi,t)$$
$$=H_0^1(\mu,t)+R_\varepsilon^1(\mu,\theta,t)+R^\varepsilon\circ\Psi_1(\mu,\phi,t), \tag{4.5}$$

where

$$H_0^1(\mu,t)=H_0(\mu)+[R_\varepsilon](\mu,t). \tag{4.6}$$

We are now in position to solve (4.4). Actually,

$$S_1(\mu,\theta,t)=\int_0^\theta\frac{[R_\varepsilon](\mu,t)-R_\varepsilon(\mu,\theta,t)}{\partial_\mu H_0(\mu)}\mathrm{d}\theta. \tag{4.7}$$

By (3.8) and (3.10), S_1 is well-defined in $(\mu,\theta,t)\in[1,4]\times\mathbb{T}_{s_0}\times\mathbb{T}_\varepsilon$, and analytic in the domain

$$\sup_{(\mu,\theta,t)\in[1,4]\times\mathbb{T}_{s_0}\times\mathbb{T}_\varepsilon}|S_1(\mu,\theta,t)|\le CA^{-1}. \tag{4.8}$$

Thus, by the implicit function theorem, $\Psi_1(\mu,\phi,t):[1+O(A^{-1}),4-O(A^{-1})]\times\mathbb{T}^1_{s_0/2}\times\mathbb{T}^1_\varepsilon\to[1,4]\times\mathbb{T}^1_{s_0}\times\mathbb{T}^1_\varepsilon$.

(1) Estimate of $H_0^1(\mu,t)$.

By (3.10) and (4.9), we have that $H_0^1(\mu,t)$ is analytic in $[1,4]\times\mathbb{T}_\varepsilon$, and

$$CA^n\ge|\partial_\mu^2 H_0^1(\mu,t)|\ge\frac{A^n}{C},\quad t\in\mathbb{T}_{\frac{\varepsilon}{2}}, \tag{4.9}$$

and by Cauchy's estimate

$$\sup_{(\mu,t)\in[1,4]\times\mathbb{T}^1_{\frac{\varepsilon}{2}}}|\partial_t H_0^1(\mu,t)|\le\sup_{(\mu,t)\in[1,4]\times\mathbb{T}^1_{\frac{\varepsilon}{2}}}|\partial_t[R_\varepsilon](\mu,t)|\le\sup_{(\mu,\theta,t)\in[1,4]\times\mathbb{T}^1_{s_0}\times\mathbb{T}^1_{\frac{\varepsilon}{2}}}|\partial_t R_\varepsilon(\mu,\theta,t)|$$
$$\le\frac{2}{\varepsilon}\sup_{(\mu,\theta,t)\in[1,4]\times\mathbb{T}^1_{s_0}\times\mathbb{T}^1_\varepsilon}|R_\varepsilon(\mu,\theta,t)|\le\frac{2}{\varepsilon}CA^{n-1}\lesssim C\varepsilon_0^{-\frac{1}{\gamma}}A^{\frac{n-1}{\gamma}}A^{n-1}$$
$$\le C\varepsilon_0^{-\frac{1}{\gamma}}A^{(n-1)(1+\frac{1}{\gamma})}. \tag{4.10}$$

(2) Estimate of $R_\varepsilon^1(\mu,\theta,t)$.

By (4.8) and Cauchy' estimate,

$$\sup_{(\mu,\theta,t)\in[1,4]\times\mathbb{T}_{s_0}\times\mathbb{T}_{\frac{\varepsilon}{2}}}|\partial_t S_1(\mu,\theta,t)|\le\frac{2CA^{-1}}{\varepsilon}\le CA^{-1}\Big(\frac{\varepsilon_0}{A^{n-1}}\Big)^{-\frac{1}{\gamma}}\le C\varepsilon_0^{-\frac{1}{\gamma}}A^{-1+\frac{n-1}{\gamma}}$$
$$=C\varepsilon_0^{-\frac{1}{\gamma}}A^{n-1-\varpi}, \tag{4.11}$$

where

$$\varpi := n - \frac{n-1}{\gamma} = \frac{n}{\gamma}\Big(\gamma - \Big(1 - \frac{1}{n}\Big)\Big). \tag{4.12}$$

By the assumption $\gamma \in \big(1 - \frac{1}{n}, 1\big)$,

$$0 < \varpi \leq 1.$$

By (3.10) and noting $H_0(\mu) = dA^{n-1}\mu^{\frac{2n+2}{n+2}}$, we have

$$\sup_{(\mu,\theta,t)\in\mathcal{D}_1} |R^1_\varepsilon(\mu,\theta,t)| \leq CA^nA^{-2} + CA^{n-1}A^{-1} + C\varepsilon_0^{-\frac{1}{\gamma}}A^{-1+\frac{n-1}{\gamma}} \leq C\,\varepsilon_0^{-\frac{1}{\gamma}}\,A^{n-1-\varpi}, \tag{4.13}$$

where

$$\mathcal{D}_1 = [1 + O(A^{-1}), 4 - O(A^{-1})] \times \mathbb{T}^1_{s_0/2} \times \mathbb{T}^1_{\frac{\varepsilon}{2}}.$$

By (4.8) and the implicit function theorem, there exist $U_1(\mu,\phi,t)$, $V_1(\mu,\phi,t)$ analytic in $\mathcal{D}_1$ such that

$$\sup_{\mathcal{D}_1} |U_1| \leq CA^{-1}, \quad \sup_{\mathcal{D}_1} |V_1| \leq CA^{-1}, \tag{4.14}$$

$$\Psi_1 : \begin{cases} I = \mu + U_1(\mu,\phi,t), \\ \theta = \phi + V_1(\mu,\phi,t), \end{cases} \tag{4.15}$$

and

$$H^1(\mu,\phi,t) = H^1_0(\mu,t) + \widetilde{R}^1_\varepsilon(\mu,\phi,t) + R^\varepsilon \circ \Psi(\mu,\phi,t), \tag{4.16}$$

where

$$\widetilde{R}^1_\varepsilon(\mu,\phi,t) = R^1_\varepsilon(\mu,\phi + V_1(\mu,\phi,t),t) \tag{4.17}$$

and

$$\sup_{\mathcal{D}_1} |\widetilde{R}^1_\varepsilon(\mu,\phi,t)| \leq C\,\varepsilon_0^{-\frac{1}{\gamma}}\,A^{n-1-\varpi}. \tag{4.18}$$

Similarly, let

$$\Psi_2 : \begin{cases} \mu = \lambda + \dfrac{\partial S_2}{\partial \phi}, \\ \widetilde{\phi} = \phi + \dfrac{\partial S_2}{\partial \lambda}, \end{cases} \tag{4.19}$$

where $S_2 = S_2(\lambda,\phi,t)$ is defined by

$$S_2(\lambda,\phi,t) = \int_0^\phi \frac{[\widetilde{R}^1_\varepsilon](\lambda,t) - \widetilde{R}^1_\varepsilon(\lambda,\phi,t)}{\partial_\mu H^1_0(\mu,t)}\mathrm{d}t, \quad [\widetilde{R}^1_\varepsilon](\lambda,t) = \int_0^1 \widetilde{R}^1_\varepsilon(\lambda,\phi,t)\,\mathrm{d}\phi. \tag{4.20}$$

By (4.9) and (4.13),

$$\sup_{\mathcal{D}_1} |S_2(\lambda,\phi,t)| \leq C\,A^{-n}\left(\frac{1}{\varepsilon_0}\right)^{\frac{1}{\gamma}} A^{n-1-\varpi} \leq C\left(\frac{1}{\varepsilon_0}\right)^{\frac{1}{\gamma}} A^{-1-\varpi}. \tag{4.21}$$

It follows from the implicit function theorem that $\Psi_2 : (\lambda, \widetilde{\phi}) \in \mathcal{D}_2 \to \mathcal{D}_1$ is well-defined, where $\mathcal{D}_2 = [1 + O(A^{-1}), 4 - O(A^{-1})] \times \mathbb{T}_{\frac{s_0}{4}} \times \mathbb{T}_{\frac{\varepsilon}{4}}$. By Cauchy estimate,

$$\sup_{\mathcal{D}_2} |\partial_t S_2(\lambda, \phi, t)| \leq \frac{C}{\varepsilon} A^{-1-\varpi} \left(\frac{1}{\varepsilon_0}\right)^{\frac{1}{\gamma}} \leq C\, A^{-\varpi-1} \left(\frac{1}{\varepsilon_0}\right)^{\frac{2}{\gamma}} A^{\frac{n-1}{\gamma}} = C \left(\frac{1}{\varepsilon_0}\right)^{\frac{2}{\gamma}} A^{n-2\varpi-1}. \quad (4.22)$$

Let

$$\begin{aligned} H^{(2)}(\lambda, \phi, t) &:= H^1\Big(\lambda + \frac{\partial S_2}{\partial \phi}, \phi, t\Big) + \frac{\partial S_2}{\partial t} \\ &= H_0^1(\lambda, t) + \partial_\lambda H_0^1(\lambda, t) \frac{\partial S_2}{\partial \phi} + \widetilde{R}_\varepsilon^1(\lambda, \phi, t) \\ &\quad + R_\varepsilon^2(\lambda, \phi, t) + R^\varepsilon \circ \Psi_1 \circ \Psi_2(\lambda, \widetilde{\phi}, t), \end{aligned} \quad (4.23)$$

where

$$\begin{aligned} R_\varepsilon^2(\lambda, \phi, t) &= \int_0^1 (1-\tau) \partial_\lambda^2 H_0^1\Big(\lambda + \frac{\partial S_2}{\partial \phi}\tau, \phi, t\Big) \Big(\frac{\partial S_2}{\partial \phi}\Big)^2 \mathrm{d}\tau \\ &\quad + \int_0^1 \partial_\lambda \widetilde{R}_\varepsilon^1\Big(\lambda + \frac{\partial S_2}{\partial \phi}, \phi, t\Big) \frac{\partial S_2}{\partial \phi} \mathrm{d}\tau + \frac{\partial S_2}{\partial t}. \end{aligned} \quad (4.24)$$

By (4.20), $\partial_\lambda H_0^1(\lambda, t) \frac{\partial S_2}{\partial \phi} + \widetilde{R}_\varepsilon^1(\lambda, \phi, t) = [\widetilde{R}_\varepsilon^1](\lambda, t)$. Let

$$H_0^2(\lambda, t) = H_0^1(\lambda, t) + [\widetilde{R}_\varepsilon^1](\lambda, t). \quad (4.25)$$

It follows that

$$H^2(\lambda, \phi, t) = H_0^2(\lambda, t) + R_\varepsilon^2(\lambda, \phi, t) + R^\varepsilon \circ \Psi_1 \circ \Psi_2(\lambda, \widetilde{\phi}, t). \quad (4.26)$$

(3) Estimate of $H_0^2(\lambda, t)$.
By (4.9)–(4.10) and (4.13), we have

$$CA^n \geq |\partial_\lambda^2 H_0^2(\lambda, t)| \geq \frac{A^n}{C}, \quad \lambda \in [1, 4],\ t \in \mathbb{T}_{\frac{\varepsilon}{2}}, \quad (4.27)$$

$$\sup_{(\lambda,t) \in [1,4] \times \mathbb{T}^1_{\frac{\varepsilon}{4}}} |\partial_t H_0^2(\lambda, t)| \leq C \varepsilon_0^{-\frac{1}{\gamma}} A^{(n-1)(1+\frac{1}{\gamma})}. \quad (4.28)$$

(4) Estimate of $R_\varepsilon^2(\lambda, \phi, t)$.
By (4.13), (4.21)–(4.22) (4.24) and (4.27), we have

$$\begin{aligned} \sup_{\mathcal{D}_2} |R_\varepsilon^2(\lambda, \phi, t)| &\leq CA^n \Big(\frac{1}{\varepsilon_0}\Big)^{\frac{2}{\gamma}} A^{-2(1+\varpi)} + C\Big(\frac{1}{\varepsilon_0}\Big)^{\frac{2}{\gamma}} A^{n-1-\varpi} A^{-1-\varpi} + C\left(\frac{1}{\varepsilon_0}\right)^{\frac{2}{\gamma}} A^{n-1-2\varpi} \\ &\leq C\left(\frac{1}{\varepsilon_0}\right)^{\frac{2}{\gamma}} A^{n-1-2\varpi}. \end{aligned} \quad (4.29)$$

Take $N \in \mathbb{N}$ with $n - \varpi N \leq -1$. Repeating the above procedure N times, we get a series of symplectic transformations $\Psi_1, \cdots, \Psi_N$ such that

$$H^N(\rho, \xi, t) = H \circ \Psi_1 \circ \cdots \circ \Psi_N = H_0^N(\rho, t) + R_\varepsilon^N(\rho, \xi, t) + R^\varepsilon \circ \Psi_1 \circ \Psi_N(\rho, \xi, t),$$

where $(\rho,\xi,t)\in[1+O(A^{-1}),4-O(A^{-1})]\times\mathbb{T}^1_{\frac{s_0}{2^N}}\times\mathbb{T}^1_{\frac{\varepsilon}{2^N}}$, and

$$\Phi\triangleq\Psi_1\circ\cdots\circ\Psi_N:[1+O(A^{-1}),4-O(A^{-1})]\times\mathbb{T}^1\times\mathbb{T}^1\to[1,4]\times\mathbb{T}^1\times\mathbb{T}^1, \tag{4.30}$$

$$\Phi=\mathrm{id}+O(A^{-1}), \tag{4.31}$$

and $H_0^N(\rho,t)$ satisfies

$$CA^n\ge|\partial_\rho^2H_0^N(\rho,t)|\ge\frac{A^n}{C},\quad \rho\in[2,3],\quad t\in\mathbb{T}, \tag{4.32}$$

$$\sup_{(\rho,t)\in[2,3]\times\mathbb{T}}|\partial_tH_0^N(\rho,t)|\le C\varepsilon_0^{-\frac{1}{\gamma}}A^{(n-1)(1+\frac{1}{\gamma})}, \tag{4.33}$$

and $R_\varepsilon^N(\rho,\xi,t)$ satisfies that for $0\le p+q\le6$,

$$\sup_{(\rho,\xi,t)\in[2,3]\times\mathbb{T}\times\mathbb{T}}|\partial_\rho^p\partial_\xi^qR_\varepsilon^N(\rho,\xi,t)|\le C\,A^{n-\varpi N}\Big(\frac{1}{\varepsilon_0}\Big)^{\frac{N}{\gamma}}\le C\,A^{-1}\Big(\frac{1}{\varepsilon_0}\Big)^{\frac{N}{\gamma}}<C\varepsilon_0, \tag{4.34}$$

where C depends on N and we have assumed that A is large enough such that

$$A^{-1}\Big(\frac{1}{\varepsilon_0}\Big)^{\frac{N}{\gamma}}<\varepsilon_0.$$

Let

$$\mathcal{R}(\rho,\xi,t)=R_\varepsilon^N(\rho,\xi,t)+R^\varepsilon\circ\Psi^1\circ\cdots\circ\Psi^N. \tag{4.35}$$

Then by (3.9), (4.30)–(4.31), (4.34), we have

$$\sup_{(\rho,\xi)\in[2,3]\times\mathbb{T}^1}\int_0^1|\partial_\rho^p\partial_\xi^q\mathcal{R}(\rho,\xi,t)|\,\mathrm{d}t\le C\varepsilon_0,\quad 0\le p+q\le6. \tag{4.36}$$

Now,

$$H^N(p,\xi,t)=H_0^N(\rho,t)+\mathcal{R}(\rho,\xi,t). \tag{4.37}$$

5 Proof of Theorem

For H^N the Hamiltonian equation is

$$\begin{cases}\dot\rho=-\dfrac{\partial H^N}{\partial\xi}=-\dfrac{\partial\mathcal{R}(\rho,\xi,t)}{\partial\xi}=O(\varepsilon_0),\\[2mm]\dot\xi=\dfrac{\partial H^N}{\partial\rho}=\dfrac{\partial H_0^N(\rho,t)}{\partial\rho}+\dfrac{\partial\mathcal{R}(\rho,\xi,t)}{\partial\xi}=\dfrac{\partial H_0^N(\rho,t)}{\partial\rho}+O(\varepsilon_0).\end{cases} \tag{5.1}$$

Note

$$H_0^N=d\cdot A^n\cdot\rho^{\frac{2(n+1)}{n+2}}+O(A^{n-1}).$$

By using Picard iteration and Gronwall's inequality and noting (4.36), we get that the time-1 map of (5.1) is of the form

$$\mathcal{P}:\begin{cases}\rho_1=\rho(t)|_{t=1}=\rho_0+F(\rho_0,\xi_0),\\ \xi_1=\xi(t)|_{t=1}=\xi_0+\alpha(\rho_0)+G(\rho_0,\xi_0),\end{cases}\qquad(\rho_0,\xi_0)\in[2,3]\times\mathbb{T}^1$$

with

$$\alpha(\rho_0) = \int_0^1 \frac{\partial H_0^N(\rho_0, t)}{\partial \rho}\,\mathrm{d}t, \quad |\partial_{\rho_0}\alpha(\rho_0)| \geq C\,A^n > 0$$

and

$$|\partial_{\rho_0}^p \partial_{\xi_0}^q F| \leq C\varepsilon_0, \quad |\partial_{\rho_0}^p \partial_{\xi_0}^q G| \leq C\varepsilon_0, \quad p + q \leq 5.$$

Since (5.1) is Hamiltonian, the map P is symplectic. By Moser's twist theorem at pp. 50–54 of [7] (also see [8]), $\mathcal{P}$ has an invariant curve Γ in the annulus $[2,3] \times \mathbb{T}^1$. Since A can be arbitrarily large, it follows that the time-1 map of the original system has an invariant curve Γ_A in the annulus $[2\,A + C, 3\,A - C] \times \mathbb{T}^1$ with C being a constant independent of A. Choosing a sequence $A = A_k \to \infty$ as $k \to \infty$, we have that there are countable many invariant curves Γ_{A_k}, clustering at ∞. Therefore any solution of the original system is bounded. This completes the proof of theorem.

Remark 5.1 Any solutions starting from the invariant curves Γ_{A_k} $(k = 1, 2, \cdots)$ are quasi-periodic with frequencies $(1, \omega_k)$ in time t, where $(1, \omega_k)$ satisfies Diophantine conditions and $\omega > C\,A_k^n$.

References

[1] Dieckerhoff, R. and Zehnder, E., Boundedness of solutions via twist theorem, *Ann. Scula. Norm. Sup. Pisa*, **14**, 1987, 79–95.

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

[3] Liu, B., Boundedness for solutions of nonlinear periodic differential equations via Moser's twist theorem, *Acta Math Sinica* (*N.S.*), **8**, 1992, 91–98.

[4] Liu, B., Boundedness for solutions of nonlinear Hill's equations with periodic forcing terms via Moser's twist theorem, *J. Diff. Eqs*, **79**, 1989, 304–315.

[5] Morris, G. R., A case of boundedness in Littlewood's problem on oscillatory differential equations, *Bull. Austral. Math. Soc.*, **14**, 1976, 71–93.

[6] Moser, J., On invariant curves of area-preserving mappings of annulus, *Nachr. Akad. Wiss. Gottingen Math.-Phys.*, **2**, 1962, 1–20.

[7] Moser, J., Stable and Random Motion in Dynamical Systems, Ann. of Math. Studies, Princeton University Press, Princeton, 1973.

[8] Russman, H., Uber invariante Kurven differenzierbarer Abbildungen eines Kreisringes, *Nachr. Akad. Wiss. Gottingen, Math. Phys.*, **2**, 1970, 67–105.

[9] Salamon, D. and Zehnder, E., KAM theory in configuration space, *Commun. Math. Helv.*, **64**(1), 1989, 84–132.

[10] Salamon, D., The Kolmogorov-Arnold-Moser theorem, *Mathematical Physics Electronic Journal*, **10**(3), 2004, 1–37.

[11] Wang, Y., Unboundedness in a Duffing equation with polynomial potentials, *J. Diff. Eqs.*, **160**(2), 2000, 467–479.

[12] Yuan, X., Invariant tori of Duffing-type equations, *Advances in Math* (*China*), **24**, 1995, 375–376.

[13] Yuan, X., Invariant tori of Duffing-type equations, *J. Diff. Eqs.*, **142**(2), 1998, 231–262.

[14] Yuan, X., Lagrange stability for Duffing-type equations, *J. Diff. Eqs.*, **160**(1), 2000, 94–117.