# Growth rates of orbits in non-periodic twist maps and a theorem by Neishtadt

Markus Kunze[1] & Rafael Ortega[2]

[1] Universität Köln, Institut für Mathematik, Weyertal 86-90, D - 50931 Köln, Germany

[2] Departamento de Matemática Aplicada, Universidad de Granada, E-18071 Granada, Spain

**Key words:** non-periodic twist map, orbit growth, Hamiltonian normal form, exponential averaging

2010 *Mathematics Subject Classification.* 34C29, 37E40, 37J25, 37J40, 70H11, 70H14, 70K40

**Abstract**

We consider non-periodic holomorphic twist maps of the form

$$\theta_1 = \theta + \frac{1}{r^\alpha}(\gamma + F_1(\theta, r)), \quad r_1 = r + r^{1-\alpha} F_2(\theta, r),$$

for $\alpha \in ]0, 1[$ and $\gamma \in \mathbb{R} \setminus \{0\}$. Under appropriate assumptions on $F_1$, $F_2$ and a primitive $\mathfrak{h}$ of $r_1\, d\theta_1 - r\, d\theta$ it is shown that $r_n = \mathcal{O}((\log n)^{1/\alpha})$, if $(\theta_n, r_n)_{n\in\mathbb{N}_0}$ is a forward complete real orbit of the map.

## 1 Introduction

Over the last years we have examined the dynamics of twist maps with non-periodic angles [4, 5, 6, 7]. Motivated by the Fermi-Ulam ping-pong model, and also by the Littlewood boundedness problem, we have obtained results on the role of the bounded orbits in the general dynamics [8] and also on the improbability of escaping orbits [9, 14, 15]. However, the first result for this class of maps is older and due to Neishtadt [11]. He studied the ping-pong model in the analytic case and proved that for any orbit the velocity $v_n$ after the impact $n$ must satisfy

$$v_n = \mathcal{O}(\log n), \quad n \to \infty. \tag{1.1}$$

In this paper we consider more general holomorphic maps $f : (\theta, r) \mapsto (\theta_1, r_1)$ of the form

$$\theta_1 = \theta + \frac{1}{r^\alpha}(\gamma + F_1(\theta, r)), \quad r_1 = r + r^{1-\alpha} F_2(\theta, r), \tag{1.2}$$

where $\alpha \in ]0,1[$ and $\gamma \in \mathbb{R} \setminus \{0\}$. They should be viewed as perturbations of

$$\theta_1 = \theta + \frac{\gamma}{r^\alpha}, \quad r_1 = r.$$

The latter map is well-defined on $\mathbb{R}\times]0,\infty[$ and has a holomorphic extension to the complex domain $\mathbb{C} \times \{r \in \mathbb{C} : \operatorname{Re} r > 0\}$. Moreover, it is symplectic, due to $r_1\, d\theta_1 - r\, d\theta = d\mathfrak{h}_0$ for $\mathfrak{h}_0(\theta, r) = -\frac{\alpha\gamma}{1-\alpha}\, r^{1-\alpha}$. We will investigate the dynamics of (1.2), being defined on a set of the type

$$\Omega = \mathbb{R}_\delta \times \{r \in \mathbb{C} : \operatorname{Re} r > \underline{r}, |\operatorname{Im} r| < \eta |r|\}$$

for some $\delta, \underline{r} > 0$, $\eta \in ]0,1[$, with $\mathbb{R}_\delta = \{\theta \in \mathbb{C} : |\operatorname{Im}\theta| < \delta\}$ denoting the open strip in the complex plane about $\mathbb{R}$ of width $\delta$. Our main assumptions are:

(i) the smallness of the holomorphic functions $F_j$ on $\Omega$ (supposed to map reals into reals), in the sense that $F_j(\theta, r) = \mathcal{O}(r^{-\alpha})$, uniformly in $\theta \in \mathbb{R}_\delta$, for $j = 1, 2$;

(ii) $\mathfrak{h}(\theta, r) = \mathfrak{h}_0(\theta, r) + \mathcal{O}(r^{1-2\alpha})$ uniformly in $\theta \in \mathbb{R}_\delta$, where $r_1\, d\theta_1 - r\, d\theta = d\mathfrak{h}$ holds for (1.2).

Under these hypotheses we are going to show (Theorem 3.1) that there exists a constant $C > 0$ such that if $(\theta_n, r_n)_{n\in\mathbb{N}_0}$ is a forward complete real orbit of (1.2), then there is $n_0 \in \mathbb{N}$ so that

$$r_n \le C(\log n)^{1/\alpha}, \quad n \ge n_0.$$

For the proof, we apply a rescaling $\xi = \varepsilon^{1/\alpha} r$ to put $f$ from (1.2) into the form

$$\psi_\varepsilon : \quad \theta_1 = \theta + \varepsilon R_1(\theta, \xi, \varepsilon), \quad \xi_1 = \xi + \varepsilon R_2(\theta, \xi, \varepsilon), \tag{1.3}$$

where $R_1(\theta, \xi, \varepsilon) = \frac{1}{\xi^\alpha}(\gamma + F_1(\theta, \frac{\xi}{\varepsilon^{1/\alpha}}))$ and $R_2(\theta, \xi, \varepsilon) = \xi^{1-\alpha} F_2(\theta, \frac{\xi}{\varepsilon^{1/\alpha}})$. It turns out that the family of maps $\{\psi_\varepsilon\}$ can be defined on a common domain $G_\rho$, where $G = \mathbb{R}\times]1,2[$ and

$$G_\rho = \{x = (q,p) \in \mathbb{C}^2 : |\operatorname{Im} q| < \rho,\ \operatorname{dist}(p, I) < \rho\}.$$

This leads us to study (see Section 2, also for more discussion of the subtleties) general maps $P_\varepsilon : G_\rho \to \mathbb{C}^2$ given by

$$P_\varepsilon : \quad x_1 = x + \varepsilon l(x, \varepsilon), \quad x_1 = (q_1, p_1), \quad x = (q, p),$$

where $l$ belongs to a certain class of maps $\mathcal{M}_{1,\rho,\sigma}$ that has to be carefully set up in order to account for singularities of $l$ or $\frac{\partial l}{\partial \varepsilon}$ at $\varepsilon = 0$; recall the definition of $R_1$, $R_2$ in terms of $F_1$, $F_2$ above. Inspired by [7], we call the family of maps $\{P_\varepsilon\}$ E-symplectic, if $p_1\, dq_1 - p\, dq = dh(\cdot, \varepsilon)$ for a function $h \in \mathcal{M}_{1,\rho,\sigma}$ such that, as $\varepsilon \to 0$,

$$h(q,p,\varepsilon) = \varepsilon\, \mathfrak{m}(q,p) + \mathcal{O}(\varepsilon^2), \quad \frac{\partial h}{\partial \varepsilon}(q,p,\varepsilon) = \mathfrak{m}(q,p) + \mathcal{O}(\varepsilon),$$

uniformly in $(q,p) \in G_\rho$ for a bounded function $\mathfrak{m} : G_\rho \to \mathbb{C}$. It turns out that all these conditions can be verified for (1.3) after rescaling $\mathfrak{h}$ from (ii) to $h$. Furthermore, it is possible to construct a function $E = E(x)$ satisfying $J\nabla E(x) = l(x, 0)$, where $J$ denotes the standard symplectic matrix; in fact $E(\theta, \xi) = E(\xi) = \frac{\gamma}{1-\alpha}\, \xi^{1-\alpha}$ for the maps from (1.3). The function $E$

should be thought of as an approximate first integral (adiabatic invariant) for the family $\{P_\varepsilon\}$. This means that the variation of $E$ along the orbit remains small for an exponentially large time. More precisely, Theorem 2.5 ensures that if

$$(x_n)_{0\le n\le N} = (P_\varepsilon^n(x_0))_{0\le n\le N}$$

is a real forward orbit piece of $P_\varepsilon$ so that $x_n \in G$ for all $0 \le n \le N$, then

$$|E(x_n) - E(x_0)| \le \hat{C}\varepsilon, \quad 0 \le n \le \min\{N, N_\varepsilon\}, \quad N_\varepsilon = [e^{\hat{D}/\varepsilon}], \tag{1.4}$$

for constants $\hat{C}, \hat{D} > 0$ and if $\varepsilon > 0$ is small enough (all independent of the orbit). Going back to the original variables $(\theta, r)$, it follows that

$$|s_n - s_m| \le C s_m^\beta, \quad m \le n \le m + [e^{s_m^\delta}],$$

where $s_n \sim r_n^{1-\alpha}$ (up to a multiplicative constant), $\beta = \frac{1-2\alpha}{1-\alpha} < 1$ and $\delta = \frac{\alpha}{1-\alpha} > 0$. Then to complete the proof of Theorem 3.1 we need to show that $\limsup_{m\to\infty} \frac{s_m}{(\log m)^{1/\delta}} \le C_1$. This is accomplished in a clean way by using Lemma 3.4, which is related to upper and lower solutions to the difference equation $x_{n+1} = x_n + C x_n^\beta$.

The proof of Theorem 2.5, which implies the adiabatic bound (1.4), is given in an appendix (Section 7), and it is based on realizing $P_\varepsilon$ as the Poincaré map of a periodic Hamiltonian system, as outlined in [11]. Nevertheless we include the necessary details, since what we precisely need is a version that allows for a non-smooth dependence on $\varepsilon$. Furthermore, the fact that in $G = \mathbb{R}\times]1, 2[$ the first coordinate can be unbounded poses some technical challenges; this is accounted for by introducing assumptions on the primitive of the 1-form that are not explicit in [11]. This leads to the introduction of suitable function classes $\mathcal{H}_{\rho,\sigma}$ and $\widetilde{\mathcal{H}}_{\rho,\sigma}$ for the relevant Hamiltonians $H = H(x, t, \varepsilon)$ in the Hamiltonian normal form theorem (Section 6).

Section 4 concerns the application to the ping-pong map. Note that $\alpha = 1/2$ in this case, but in the notation of the main theorem (Theorem 3.1) as mentioned above $r_n = E_n = v_n^2/2$ corresponds to energy, not velocity, and hence we recover (1.1). An important issue here is how to extend the ping-pong map to the analytic setup and how to change variables appropriately. In fact it will turn out that even in this case our result is slightly stronger as what is mentioned in [11], since it comes with some uniformity, in the sense that it provides the estimate

$$\limsup_{n\to\infty} \frac{v_n}{\log n} \le C_0$$

for a constant $C_0 > 0$ that is independent of the chosen orbit.

It remains an open question, if Theorem 3.1 yields an optimal bound. In Section 5 we consider the examples

$$\theta_1 = \theta + \sqrt{\frac{2}{r_1}}, \quad r_1 = r + \frac{2\theta}{(1+\theta^2)^N},$$

for $N \in \mathbb{N}$. By direct analysis, it can be shown that for $(\theta_0, r_0) \in \mathbb{R}^2$ such that $\theta_0 > 0$ and $r_0 > 0$ the forward complete orbit $(\theta_n, r_n)_{n\in\mathbb{N}_0}$ does exist, and moreover

$$\sup_{n\in\mathbb{N}_0} r_n < \infty \quad \text{if} \quad N \ge 2,$$

whereas

$$0 < \liminf_{n\to\infty} \frac{r_n}{(\log n)^2} \le \limsup_{n\to\infty} \frac{r_n}{(\log n)^2} < \infty \quad \text{if} \quad N = 1.$$

In addition, it will turn out that Theorem 3.1 is applicable for $N \ge 2$ only, but not for $N = 1$. This indicates that: (i) some assumption of Theorem 3.1 could maybe be relaxed (to cover $N = 1$); (ii) in some examples maybe unbounded orbits do exist. However, given the advanced technical machinery that is used to establish Theorem 3.1, both questions seem to be difficult to answer.

# 2 E-symplectic families of maps

An important observation in [11] is the existence of adiabatic invariants for families of analytic canonical maps close to the identity. Given a convex domain $G \subset \mathbb{R}^N \times \mathbb{R}^N$ and a family of symplectic maps

$$P_\varepsilon : G \to \mathbb{R}^N \times \mathbb{R}^N, \quad x_1 = x + \varepsilon l(x, \varepsilon),$$

it is possible to construct a function $E = E(x)$ satisfying

$$J\nabla E(x) = l(x, 0), \tag{2.1}$$

where $J = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix}$. For small $\varepsilon$ the iteration $x_{n+1} = P_\varepsilon(x_n)$ can be interpreted as a numerical integration method for the Hamiltonian system $\dot{x} = J\nabla E(x)$. This fact suggests that $E(x)$ should be an adiabatic invariant for $P_\varepsilon$, meaning that

$$|E(P_\varepsilon^n(x)) - E(x)| \le C\varepsilon, \quad 0 \le n \le N_\varepsilon, \tag{2.2}$$

where $N_\varepsilon$ is of the order $e^{D/\varepsilon}$; the constants $C, D > 0$ should only depend upon an appropriate norm of $l$. In essence this is discussed in Remark 5 and Proposition 3 of [11]. Additional details can be found in [2], in particular in the case of bounded domains.

However, the previous statements must be taken with some caution in the case where the underlying domain is unbounded. As a counter-example we consider the family of translations

$$x_1 = x + \varepsilon Jv + \varepsilon^2 v,$$

defined on the whole space $G = \mathbb{R}^N \times \mathbb{R}^N$. Here $v \ne 0$ is a fixed vector and $E(x) = \langle x, v\rangle$ satisfies (2.1), since $l(x, \varepsilon) = Jv + \varepsilon v$. Due to $P_\varepsilon^n(x) = x + n\varepsilon l(x, \varepsilon)$ we obtain

$$|E(P_\varepsilon^n(x)) - E(x)| = \varepsilon^2 n|v|^2.$$

Therefore (2.2) does hold only for $n \le N_\varepsilon = \mathcal{O}(1/\varepsilon)$ many steps.

To overcome this inherent difficulty, Benettin and Giorgilli in [2] considered an unbounded domain $G$ and a family of maps derived from a symplectic integration algorithm for a Newtonian system of the type $\ddot{q} = -\nabla V(q)$. Then they impose some growth conditions on $V(q)$ as $|q| \to \infty$. We will follow a different approach and assume that our family $\{P_\varepsilon\}$ satisfies a condition inspired by the notion of an exact symplectic map (called E-symplectic), as it was understood in our previous work [7]. Furthermore, to simplify matters, we will restrict ourselves to the case of direct interest to us for applications. Throughout we will take

$$N = 1 \quad \text{and} \quad G = \mathbb{R} \times I,$$

where $I \subset \mathbb{R}$ is an open and bounded interval. Our goal will be to understand the dynamics of a map on the plane $(\theta, r) \mapsto (\theta_1, r_1)$ when $r \to \infty$. For this reason our family of maps $\{P_\varepsilon\}$, $P_\varepsilon : (q, p) \mapsto (q_1, p_1)$, will be obtained after a rescaling $q = \theta$, $p = \varepsilon r$ with $q \in \mathbb{R}$ and $p \in ]1, 2[$. This procedure will lead to functions $l(x, \varepsilon)$ that are analytic in $x$, but not necessarily smooth in $\varepsilon$; a prototype can be the function $l(x, \varepsilon) = h(x/\varepsilon^2)$, where $h$ is real analytic in $[1, \infty[$ and $h(\zeta) \to 0$ as $\zeta \to \infty$. Then $l$ is continuous as a function of the two variables $(x, \varepsilon)$, but the partial derivatives $\partial_\varepsilon^k l$ do not always exist at $\varepsilon = 0$.

The following definitions are motivated by the previous discussions. In general, for the norms on $\mathbb{C}^d$ and $\mathbb{C}^{d_1 \times d_2}$ we will take $|x| = \max_{1 \le i \le d} |x_i|$ and $|A| = \max_{1 \le i \le d_1, 1 \le j \le d_2} |a_{ij}|$, respectively. Note that for $A \in \mathbb{C}^{d \times d}$, $x \in \mathbb{C}^d$, $A_1 \in \mathbb{C}^{d_1 \times d}$ and $A_2 \in \mathbb{C}^{d \times d_2}$ this implies

$$|Ax| \le d|A||x|, \quad |A_1 A_2| \le d|A_1||A_2|.$$

The points in $G = \mathbb{R} \times I$ will be denoted by $x = (q, p)$. For $\rho > 0$ we will write

$$G_\rho = \{x = (q, p) \in \mathbb{C}^2 : |\operatorname{Im} q| < \rho, \ \operatorname{dist}(p, I) < \rho\}.$$

Given $\varphi : G_\rho \to \mathbb{C}$ holomorphic, let

$$\|\varphi\|_\rho = \sup\{|\varphi(x)| : x \in G_\rho\}.$$

If $0 < r < \rho$, then by the Cauchy integral formula one has

$$\|D\varphi\|_r \le \frac{1}{\rho - r} \|\varphi\|_\rho,$$

where $D\varphi$ is the Jacobian.

**Definition 2.1 (The classes $\mathcal{M}_{\rho,\sigma}$ and $\mathcal{M}_{1,\rho,\sigma}$)** *Let $\rho > 0$ and $\sigma \in ]0, 1[$.*

*(i) The class $\mathcal{M}_{\rho,\sigma}$ consists of those continuous maps $l : G_\rho \times [0, \sigma] \to \mathbb{C}^2$, $l = l(x, \varepsilon)$, which satisfy:*

*(a) $l$ maps real into reals; and*

*(b) for every $\varepsilon \in [0, \sigma]$ the map $l(\cdot, \varepsilon)$ is holomorphic on $G_\rho$ and*

$$\|l\|_{\rho,\sigma} = \sup\{\|l(\cdot, \varepsilon)\|_\rho : \varepsilon \in [0, \sigma]\} < \infty.$$

*(ii) The class $\mathcal{M}_{1,\rho,\sigma}$ consists of those continuous maps $l : G_\rho \times [0, \sigma] \to \mathbb{C}^2$, $l = l(x, \varepsilon)$, satisfying*

*(a) $l$ maps real into reals;*

*(b) $l$ is $C^\infty$ in $G_\rho \times ]0, \sigma]$;*

(c) *for every* $\varepsilon \in [0,\sigma]$ *the map* $l(\cdot,\varepsilon)$ *is holomorphic on* $G_\rho$*;*

(d) *one has*

$$\|l\|_{1,\rho,\sigma} = \|l\|_{\rho,\sigma} + \sup\left\{\left\|\frac{\partial l}{\partial\varepsilon}(\cdot,\varepsilon)\right\|_\rho : \varepsilon \in ]0,\sigma]\right\} < \infty.$$

**Remark 2.2** Note that, for a map $l \in \mathcal{M}_{\rho,\sigma}$ or $l \in \mathcal{M}_{1,\rho,\sigma}$, all the derivatives $\partial_x^\alpha \partial_\varepsilon^k l(\cdot,\varepsilon) : G_\rho \to \mathbb{C}^2$ for $\varepsilon \in ]0,\sigma]$ are holomorphic, where $\alpha \in \mathbb{N}_0^2$ and $k \in \mathbb{N}_0$. Similarly, all the $\partial_x^\alpha l : G_\rho \times [0,\sigma] \to \mathbb{C}^2$ are continuous functions of both variables. This follows from the Cauchy integral formula and the continuity of $l$. Furthermore, the derivatives can be interchanged: $\partial_x^\alpha \partial_\varepsilon^k l(\cdot,\varepsilon) = \partial_\varepsilon^k \partial_x^\alpha l(\cdot,\varepsilon)$.

**Definition 2.3** *Suppose that* $l \in \mathcal{M}_{1,\rho,\sigma}$*, and for* $\varepsilon \in [0,\sigma]$ *consider the family of maps* $P_\varepsilon : G_\rho \to \mathbb{C}^2$ *given by*

$$P_\varepsilon : \quad x_1 = x + \varepsilon l(x,\varepsilon), \quad x_1 = (q_1,p_1), \quad x = (q,p). \tag{2.3}$$

*We say that the family* $\{P_\varepsilon\}$ *is E-symplectic, if there is a function* $h \in \mathcal{M}_{1,\rho,\sigma}$ *such that*

$$p_1\, dq_1 - p\, dq = dh(\cdot,\varepsilon) \tag{2.4}$$

*and there exists a bounded function* $\mathfrak{m} : G_\rho \to \mathbb{C}$ *satisfying*

$$h(q,p,\varepsilon) = \varepsilon\,\mathfrak{m}(q,p) + \mathcal{O}(\varepsilon^2) \quad as \quad \varepsilon \to 0 \tag{2.5}$$

*and*

$$\frac{\partial h}{\partial\varepsilon}(q,p,\varepsilon) = \mathfrak{m}(q,p) + \mathcal{O}(\varepsilon) \quad as \quad \varepsilon \to 0 \tag{2.6}$$

*uniformly in* $(q,p) \in G_\rho$*.*

**Remark 2.4** (a) $\mathfrak{m}$ is holomorphic in $G_\rho$. To see this, note that $\frac{\partial h}{\partial\varepsilon}(\cdot,\varepsilon)$ is holomorphic for $\varepsilon > 0$ by Remark 2.2. Since $\mathfrak{m}$ is the uniform limit of $\int_0^1 \frac{\partial h}{\partial\varepsilon}(q,p,t\varepsilon)\, dt$ as $\varepsilon \to 0$, it is holomorphic itself.
(b) $\mathfrak{m}$ satisfies

$$\frac{\partial\mathfrak{m}}{\partial q}(q,p) = p\,\frac{\partial l_1}{\partial q}(q,p,0) + l_2(q,p,0), \quad \frac{\partial\mathfrak{m}}{\partial p}(q,p) = p\,\frac{\partial l_1}{\partial p}(q,p,0), \tag{2.7}$$

where $l = (l_1,l_2)$. For, we observe from (2.5) that $\varepsilon^{-1}h \to \mathfrak{m}$ uniformly on $G_\rho$. Therefore also the derivatives converge, uniformly on compact subsets of $G_\rho$. From (2.4),

$$\varepsilon^{-1}\frac{\partial h}{\partial q} = l_2 + p\,\frac{\partial l_1}{\partial q} + \varepsilon\, l_2\,\frac{\partial l_1}{\partial q}, \quad \varepsilon^{-1}\frac{\partial h}{\partial p} = p\,\frac{\partial l_1}{\partial p} + \varepsilon\, l_2\,\frac{\partial l_1}{\partial p}.$$

Thus it remains to pass to the limit $\varepsilon \to 0$ and use Remark 2.2. Relation (2.7) can also be stated as

$$\nabla\mathfrak{m}(x) = p\,\nabla l_1(x,0) + \begin{pmatrix} l_2(x,0) \\ 0 \end{pmatrix}, \quad x = (q,p). \tag{2.8}$$

(c) One has
$$\frac{\partial l_1}{\partial q}(q,p,0) + \frac{\partial l_2}{\partial p}(q,p,0) = 0, \tag{2.9}$$
as follows from $\frac{\partial^2 \mathfrak{m}}{\partial q \partial p} = \frac{\partial^2 \mathfrak{m}}{\partial p \partial q}$. Relation (2.9) implies that the Jacobian matrix $Dl(x,0)$ is Hamiltonian, i.e., it satisfies $Dl(x,0)^* J + JDl(x,0) = 0$, or equivalently, $JDl(x,0)$ is symmetric. Since $G_\rho$ is simply connected, we conclude that there is a holomorphic function $E : G_\rho \to \mathbb{C}$ such that $J\nabla E = l(\cdot,0)$, i.e., (2.1) holds. Actually, (2.8) shows that we can take
$$E(x) = l_1(x,0)p - \mathfrak{m}(x), \quad x = (q,p). \tag{2.10}$$

(d) The relation $J\nabla E = l(\cdot,0)$ yields
$$dE = \frac{\partial E}{\partial q}\, dq + \frac{\partial E}{\partial p}\, dp = -l_2\, dq + l_1\, dp.$$
Hence $E(x) = E(x_0) + \int_\gamma (-l_2\, dq + l_1\, dp)$ for every path $\gamma$ that connects a fixed $x_0 \in G$ to $x$. This observation makes the connection to the formula for $E$ given in [11] below (2.7).

(e) Condition (2.6) does not follow from (2.5), as the example
$$h(q,p,\varepsilon) = \varepsilon\, \mathfrak{m}(q,p) + \varepsilon^2 \sin\left(\frac{1}{\varepsilon}\right)$$
shows.

The key result of this section is the following theorem. It should be compared to [11, (2.7), p. 135 and Prop. 3, p. 136].

**Theorem 2.5** *Suppose that $l \in \mathcal{M}_{1,\rho,\sigma}$, and for $\varepsilon \in [0,\sigma]$ consider the family of maps $P_\varepsilon : G_\rho \to \mathbb{C}^2$ given by*
$$P_\varepsilon : \quad x_1 = x + \varepsilon l(x,\varepsilon). \tag{2.11}$$
*Let the family $\{P_\varepsilon\}$ be E-symplectic. Then there exist $\hat\sigma \in ]0,\sigma]$ and constants $\hat C, \hat D > 0$ (depending upon $\rho$, $\sigma$, $\|l\|_{1,\rho,\sigma}$, the interval $I$, $\|h\|_{1,\rho,\sigma}$ and $\sup_{\varepsilon\in]0,\sigma]} \|\varepsilon^{-1}(\frac{\partial h}{\partial \varepsilon}(\cdot,\varepsilon) - \mathfrak{m})\|_\rho$) such that if*
$$(x_n)_{0\le n\le N} = (P_\varepsilon^n(x_0))_{0\le n\le N}$$
*is a real forward orbit piece of $P_\varepsilon$ so that $x_n \in G$ for all $0 \le n \le N$, then*
$$|E(x_n) - E(x_0)| \le \hat C \varepsilon, \quad 0 \le n \le \min\{N, N_\varepsilon\}, \quad N_\varepsilon = [e^{\hat D/\varepsilon}]. \tag{2.12}$$

We postpone the proof to Section 7.

# 3 Main result

To motivate our main result let $\alpha \in ]0,1[$ and $\gamma \in \mathbb{R} \setminus \{0\}$. Consider the map $(\theta, r) \mapsto (\theta_1, r_1)$ given by

$$\theta_1 = \theta + \frac{\gamma}{r^\alpha}, \quad r_1 = r.$$

It is well-defined on $\mathbb{R} \times ]0, \infty[$ and has a holomorphic extension to the complex domain $\mathbb{C} \times \{r \in \mathbb{C} : \operatorname{Re} r > 0\}$. Moreover, the map is symplectic, since it satisfies

$$r_1\, d\theta_1 - r\, d\theta = d\mathfrak{h}_0$$

for

$$\mathfrak{h}_0(\theta, r) = -\frac{\alpha\gamma}{1-\alpha}\, r^{1-\alpha}. \tag{3.1}$$

We will consider perturbations of this map on a sub-domain of $\mathbb{C}^2$ of the type

$$\Omega = \mathbb{R}_\delta \times \{r \in \mathbb{C} : \operatorname{Re} r > \underline{\mathrm{r}}, |\operatorname{Im} r| < \eta |r|\}, \tag{3.2}$$

where $\delta, \underline{\mathrm{r}} > 0$, $\eta \in ]0,1[$, and $\mathbb{R}_\delta = \{\theta \in \mathbb{C} : |\operatorname{Im} \theta| < \delta\}$ denotes the open strip in the complex plane about $\mathbb{R}$ of width $\delta$.

**Theorem 3.1** *Consider the map $f : (\theta, r) \mapsto (\theta_1, r_1)$ given by*

$$\theta_1 = \theta + \frac{1}{r^\alpha}(\gamma + F_1(\theta, r)), \quad r_1 = r + r^{1-\alpha} F_2(\theta, r), \tag{3.3}$$

*under the following hypotheses:*

*(a) $F_1$ and $F_2$ are holomorphic in $\Omega$ from (3.2).*

*(b) If $(\theta, r) \in \Omega \cap \mathbb{R}^2$, then $F_1(\theta, r), F_2(\theta, r) \in \mathbb{R}$.*

*(c) $F_j(\theta, r) = \mathcal{O}(r^{-\alpha})$, uniformly in $\theta \in \mathbb{R}_\delta$ and for $j = 1, 2$.*

*(d) There is a holomorphic function $\mathfrak{h} : \Omega \to \mathbb{C}$ that maps reals into reals and such that $r_1\, d\theta_1 - r\, d\theta = d\mathfrak{h}$ as well as*

$$\mathfrak{h}(\theta, r) = \mathfrak{h}_0(\theta, r) + \mathcal{O}(r^{1-2\alpha}), \tag{3.4}$$

*uniformly in $\theta \in \mathbb{R}_\delta$, where $\mathfrak{h}_0$ is defined in (3.1).*

*Then there exists a constant $C > 0$ such that if $(\theta_n, r_n)_{n \in \mathbb{N}_0}$ is a forward complete real orbit of $f$, then there is $n_0 \in \mathbb{N}$ so that*

$$r_n \le C (\log n)^{1/\alpha}, \quad n \ge n_0.$$

**Remark 3.2** (a) The dependence of $C$ with respect to the parameters will be discussed along the proof; $C$ will be obtained from a sequence of constants $C_1, \dots C_{13}$.
(b) If the functions $F_1$ and $F_2$ are $2\pi$-periodic in $\theta$, then $f$ can be defined on a cylinder and the conclusion can be improved to $r_n = \mathcal{O}(1)$ as $n \to \infty$ for each complete real orbit. This is a consequence of the Small Twist Theorem, see [17, Chapter III]. In fact, after the rescaling $\rho = \varepsilon r$ with $\rho \in [1, 2]$, the map $f$ has an expansion of the form

$$\theta_1 = \theta + \varepsilon^\alpha \frac{\gamma}{\rho^\alpha} + \mathcal{O}(\varepsilon^{2\alpha}), \quad \rho_1 = \rho + \mathcal{O}(\varepsilon^{2\alpha}),$$

as $\varepsilon \to 0$, uniformly in $\theta \in \mathbb{R}_\delta$. Taking a sequence $\varepsilon_n \to 0$, we find corresponding invariant curves $r = \psi_n(\theta)$ such that $1/\varepsilon_n \le \psi_n(\theta) \le 2/\varepsilon_n$ for $\theta \in \mathbb{R}$. These curves are closed in the cylinder and act as barriers for all real orbits, preventing them to escape. The same conclusion is valid if then dependence on $\theta$ is quasiperiodic and the frequencies satisfy a Diophantine condition, cf. [19].
(c) Without any further assumptions, for a map $f$ which satisfies (a)-(d), there are infinitely many forward complete real orbits such that $r_n = \mathcal{O}(1)$ along the orbit. This is a consequence of the results in [5]. To establish the claim, we first observe that (c) yields for $r \in \mathbb{R}$ the bound

$$\frac{\partial F_1}{\partial r}(\theta, r) = \mathcal{O}(r^{-(1+\alpha)}), \quad r \to \infty, \tag{3.5}$$

uniformly in $\theta \in \mathbb{R}$; (3.5) follows from the Cauchy formula, see the proof of Theorem 3.1 below. According to [7], the latter estimate is sufficient to guarantee the existence of a generating function $h = h(\theta, \theta_1)$ associated to $f$, i.e., $r = \frac{\partial h}{\partial \theta}$ and $r_1 = -\frac{\partial h}{\partial \theta_1}$ are verified. Actually one can take $h(\theta, \theta_1) = -\mathfrak{h}(\theta, R(\theta, \theta_1))$, where $r = R(\theta, \theta_1)$ is implicitly defined by the first equation in (3.3). Some computations then show that

$$R(\theta, \theta_1) \sim \gamma^{1/\alpha}(\theta_1 - \theta)^{-1/\alpha}, \quad h(\theta, \theta_1) \sim \frac{\alpha\gamma^{1/\alpha}}{1-\alpha}(\theta_1 - \theta)^{-\frac{1-\alpha}{\alpha}},$$

as $\theta_1 - \theta \to 0^+$, where as usual $F(x) \sim G(x)$ as $x \to x_0$ means that $\lim_{x\to x_0} F(x)/G(x) = 1$. Hence we can invoke [5, Thm. 2.5] or [7, Exercise 5.6] to deduce that for each $\hat{r} > 0$ the map $f$ has an orbit $(\theta_n, r_n)_{n\in\mathbb{N}_0}$ such that $r_n \ge \hat{r}$ for all $n \in \mathbb{N}_0$ and furthermore $\sup_n r_n < \infty$.

To prepare for the proof of the theorem, we are going to discuss some aspects of the method of upper and lower solutions for the difference equation

$$x_{n+1} = g(x_n), \tag{3.6}$$

where $g : I \to \mathbb{R}$ is an increasing function that is defined on an interval $I \subset \mathbb{R}$.

A sequence $(\gamma_n)_{0\le n\le N} \subset I$ is called a lower solution of (3.6), if $\gamma_{n+1} \le g(\gamma_n)$ for $n = 0, \dots, N-1$. An upper solution is defined by reversing the previous inequality.

**Lemma 3.3** *Let $(\gamma_n)$ and $(\Gamma_n)$ be a lower solution and an upper solution of (3.6). If $\gamma_0 \le \Gamma_0$, then $\gamma_n \le \Gamma_n$ for all $n$.*

**Proof:** This follows by induction from the monotonicity of $g$. $\square$

Next we will show how to construct lower and upper solutions for an equation that will be important for the proof of Theorem 3.1. Consider

$$x_{n+1} = x_n + Cx_n^\beta,$$

where $C > 0$ and $\beta < 1$. The function $g(x) = x + Cx^\beta$ is increasing in $I = [0, \infty[$, if $\beta \geq 0$, and it is increasing in $I = [(C|\beta|)^{\frac{1}{1-\beta}}, \infty[$, if $\beta < 0$. Inspired by the general solution of the differential equation $\dot{x} = Cx^\beta$, we test sequences of the type

$$\gamma_n = (A + Bn)^{\frac{1}{1-\beta}}, \quad n \geq 0,$$

for some $A, B > 0$; the condition $A \geq (C|\beta|)^{\frac{1}{1-\beta}}$ is also assumed, if $\beta < 0$, to make sure that $\gamma_n \in I$. From the mean value theorem we obtain

$$\gamma_{n+1} - \gamma_n = \frac{B}{1-\beta} (A + B(n + \zeta_n))^{\frac{\beta}{1-\beta}}$$

for some $\zeta_n \in ]0, 1[$.

Let us first look at the case where $\beta \in [0, 1[$. Here we deduce that

$$\frac{B}{1-\beta} \gamma_n^\beta \leq \gamma_{n+1} - \gamma_n \leq \frac{B}{1-\beta} \gamma_{n+1}^\beta. \tag{3.7}$$

Hence $\gamma_n$ will be an upper solution, as soon as $B \geq C(1-\beta)$. To get a lower solution, we observe that

$$\frac{\gamma_{n+1}^{1-\beta}}{\gamma_n^{1-\beta}} = 1 + \frac{B}{A + Bn} \leq 1 + \frac{B}{A}.$$

Therefore we get

$$\gamma_{n+1}^\beta \leq \left(1 + \frac{B}{A}\right)^{\frac{\beta}{1-\beta}} \gamma_n^\beta,$$

and thus, due to (3.7), $\gamma_n$ will be a lower solution, if $B(1 + \frac{B}{A})^{\frac{\beta}{1-\beta}} \leq C(1-\beta)$.

For $\beta < 0$ the inequality (3.7) is reversed. As a consequence, $\gamma_n$ will be a lower solution, if $A \geq (C|\beta|)^{\frac{1}{1-\beta}}$ and $B \leq C(1-\beta)$, and it is an upper solution for $A \geq (C|\beta|)^{\frac{1}{1-\beta}}$ and $B(1 + \frac{B}{A})^{\frac{\beta}{1-\beta}} \geq C(1-\beta)$. Thus to summarize:

(a) If $\beta < 1$ is fixed and $B = \frac{1}{2}C(1-\beta)$, then $\Gamma_n = (A + Bn)^{\frac{1}{1-\beta}}$ will be a lower solution, if $A > 0$ is taken sufficiently large (depending on $C$ and $\beta$); this fact won't be needed in what follows.

(b) If $\beta < 1$ is fixed and $B = 2C(1-\beta)$, then $\Gamma_n = (A + Bn)^{\frac{1}{1-\beta}}$ will be an upper solution, if $A > 0$ is taken sufficiently large (depending on $C$ and $\beta$).

Returning to the general setup, let us now assume that the interval $I$ is of the type $I = ]b, \infty[$ and let $h : \mathbb{N}_0 \to \mathbb{N}_0$ be a given function with the property that

$$h(n) \geq n + 1, \quad n \in \mathbb{N}_0. \tag{3.8}$$

Let $(\gamma_n)_{n \in \mathbb{N}_0} \subset I$ be a sequence such that

$$\gamma_m \leq g(\gamma_n), \quad 0 \leq n \leq m \leq h(n). \tag{3.9}$$

This sequence is a lower solution of (3.6), but it has additional favorable properties; in this case it will be possible to sharpen the conclusion of Lemma 3.3 as follows.

**Lemma 3.4** *Let $(\gamma_n)_{n \in \mathbb{N}_0} \subset I$ and $(\Gamma_n)_{n \in \mathbb{N}_0} \subset I$ be such that:*

*(a)* $(\gamma_n)$ *satisfies (3.9),*

*(b)* $(\Gamma_n)$ *is an upper solution to (3.6),*

*(c)* $\gamma_0 \leq \Gamma_0$,

*(d)* $(\Gamma_n)$ *is increasing and* $\limsup_{n \to \infty} \gamma_n = \infty$.

*Then there is an increasing function $\sigma : \mathbb{N}_0 \to \mathbb{N}$ such that*

$$\gamma_{\sigma(n)} > \Gamma_n \quad \textit{and} \quad \gamma_m \leq \Gamma_n, \quad m \in \{0, \dots, \sigma(n) - 1\}. \tag{3.10}$$

*In addition,*

$$\sigma(n + 1) > h(\sigma(n) - 1), \quad n \in \mathbb{N}_0. \tag{3.11}$$

**Proof:** Define

$$\sigma(n) = \min\{k \in \mathbb{N}_0 : \gamma_k > \Gamma_n\}.$$

It follows from (d) that $\sigma$ is well-defined and monotone increasing. Also, thanks to (c), we have $\sigma(0) \geq 1$ and accordingly $\sigma(n) \geq 1$ for all $n \in \mathbb{N}_0$; in particular, the statement (3.11) makes sense. The relations in (3.10) are obtained directly from the definition of $\sigma$, and we are going to prove (3.11) by contradiction. So assume that for some $n \in \mathbb{N}_0$ we have $\sigma(n+1) \leq h(\sigma(n) - 1)$. Then

$$h(\sigma(n) - 1) \geq \sigma(n + 1) \geq \sigma(n) > \sigma(n) - 1$$

shows that we can use (3.9), with $n$ replaced by $\sigma(n) - 1$ and $m$ replaced by $\sigma(n+1)$, to deduce that $\gamma_{\sigma(n+1)} \leq g(\gamma_{\sigma(n)-1})$. Since $g$ is increasing and due to (b), this would yield

$$\gamma_{\sigma(n+1)} \leq g(\gamma_{\sigma(n)-1}) \leq g(\Gamma_n) \leq \Gamma_{n+1},$$

which is impossible by (3.10). □

**Remark 3.5** The previous proof is still valid, if the sequence $(\gamma_n)$ does not lie in $I$, but satisfies a modified version of (3.9). Assume that there a numbers $b^* > b_* > b$ such that $\Gamma_0 \geq b^*$ and

$$\gamma_{n+1} \geq b^* \quad \Longrightarrow \quad \gamma_n \geq b_*. \tag{3.12}$$

Then $\gamma_n$ is required to have the property that

$$\gamma_n \geq b_* \quad \Longrightarrow \quad b \leq \gamma_m \leq g(\gamma_n), \quad 0 \leq n \leq m \leq h(n). \tag{3.13}$$

**Proof of Theorem 3.1 :** Step 1: Some estimates. We will show that, after restricting the size of $\Omega$, the functions $F_1$ and $F_2$ will satisfy some additional estimates. From (c) we know that there are numbers $C_j > 0$ such that

$$|F_j(\theta, r)| \leq C_j r^{-\alpha}, \quad (\theta, r) \in \Omega, \quad j = 1, 2. \tag{3.14}$$

We consider the smaller region

$$\Omega_* = \mathbb{R}_\delta \times \Big\{ r \in \mathbb{C} : \operatorname{Re} r > 2\underline{r}, |\operatorname{Im} r| < \frac{\eta}{2} |r| \Big\}$$

and claim that there are constants $C_j^{(1)} > 0$ for $j = 1, 2$ such that

$$\Big| \frac{\partial F_j}{\partial r}(\theta, r) \Big| \leq C_j^{(1)} r^{-(\alpha+1)}, \quad (\theta, r) \in \Omega_*, \quad j = 1, 2, \tag{3.15}$$

where $C_j^{(1)}$ only depends upon $\underline{r}$, $\eta$ and $C_j$. To prove this we first use an elementary geometric argument to find a constant $\kappa \in ]0, 1[$, depending upon $\underline{r}$ and $\eta$, such that if $(\theta, r) \in \Omega_*$, then all points $(\theta, \rho)$ with $\theta \in \mathbb{R}_\delta$ and $|\rho - r| \leq \kappa |r|$ will belong to $\Omega$. Now it is possible to use the Cauchy formula

$$\frac{\partial F_j}{\partial r}(\theta, r) = \frac{1}{2\pi i} \int_\gamma \frac{F_j(\theta, \rho)}{(\rho - r)^2} \, d\rho,$$

where $\gamma$ is a circle with center $r$ and radius $\kappa |r|$. Then (3.14) leads, after a short computation, to (3.15). The same kind of arguments in conjunction with (d) yields the following bounds for $\mathfrak{h}$:

$$|\mathfrak{h}(\theta, r) - \mathfrak{h}_0(\theta, r)| \leq C_3 \, r^{1-2\alpha}, \quad (\theta, r) \in \Omega, \tag{3.16}$$

$$\Big| \frac{\partial \mathfrak{h}}{\partial r}(\theta, r) - \frac{\partial \mathfrak{h}_0}{\partial r}(\theta, r) \Big| \leq C_3^{(1)} r^{-2\alpha}, \quad (\theta, r) \in \Omega_*, \tag{3.17}$$

where $C_3 > 0$ and $C_3^{(1)} > 0$ are suitable constants. From now on the domain $\Omega$ will be replaced by $\Omega_*$. To simplify notation, we will assume that already $\Omega$ is a domain on which the estimates (3.14), (3.15), (3.16) and (3.17) are verified.

Step 2: Rescaling. Under the transformation $\xi = \varepsilon^{1/\alpha} r$ the map $f$ becomes

$$\psi_\varepsilon : \quad \theta_1 = \theta + \varepsilon R_1(\theta, \xi, \varepsilon), \quad \xi_1 = \xi + \varepsilon R_2(\theta, \xi, \varepsilon),$$

where

$$R_1(\theta, \xi, \varepsilon) = \frac{1}{\xi^\alpha} \Big( \gamma + F_1\Big(\theta, \frac{\xi}{\varepsilon^{1/\alpha}}\Big)\Big), \quad R_2(\theta, \xi, \varepsilon) = \xi^{1-\alpha} F_2\Big(\theta, \frac{\xi}{\varepsilon^{1/\alpha}}\Big).$$

According to (a), $\psi_\varepsilon$ is defined on

$$\Sigma_\varepsilon = \mathbb{R}_\delta \times \{\xi \in \mathbb{C} : |\mathrm{Im}\,\xi| < \eta|\xi|, \mathrm{Re}\,\xi > \varepsilon^{1/\alpha}\underline{\mathrm{r}}\}.$$

We intend to apply Theorem 2.5 to the family of maps $\{\psi_\varepsilon\}$, and the first task will be to determine a common domain. Let us fix $I = ]1,2[$ and define $G = \mathbb{R} \times I$. A generic point in $G$ will be denoted by $x = (\theta, \xi)$ and we also recall that $|x| = \max\{|\theta|, |\xi|\}$ will be taken as the norm on $\mathbb{C}^2$. Elementary geometric considerations show that it is possible to select $\rho \in ]0, \min\{1/2, \delta\}[$ and $\sigma > 0$ such that $G_\rho \subset \Sigma_\varepsilon$ for $\varepsilon \in [0, \sigma]$. The next step is to show that $l = (R_1, R_2)$ belongs to $\mathcal{M}_{1,\rho,\sigma}$. Note that we are extending this map to $\varepsilon = 0$ by letting

$$R_1(\theta, \xi, 0) = \frac{\gamma}{\xi^\alpha}, \quad R_2(\theta, \xi, 0) = 0.$$

The functions $R_i(\cdot,\cdot,0)$ are obviously continuous on $G_\rho$. We are going to show that

$$R_1(\theta, \xi, \varepsilon) \to \frac{\gamma}{\xi^\alpha}, \quad R_2(\theta, \xi, \varepsilon) \to 0, \tag{3.18}$$

as $\varepsilon \to 0$, uniformly in $G_\rho$. This implies that the extension of $R_i$ to $G_\rho \times [0, \sigma]$ is continuous, and hence the same holds for $l$. The limits in (3.18) are a consequence of (3.14) and the bounds $1/2 \leq 1 - \rho \leq |\xi| \leq 2 + \rho \leq 5/2$ for $(\theta, \xi) \in G_\rho$.

Now that we know that $l$ is continuous in $G_\rho \times [0, \sigma]$, the conditions (a), (b) and (c) from the definition of the class $\mathcal{M}_{1,\rho,\sigma}$ follow directly from the assumptions on $F_1$ and $F_2$. To establish (d), we first consider $\|l\|_{\rho,\sigma}$. Here

$$\|R_1(\cdot,\cdot,\varepsilon)\|_\rho \leq 2^\alpha(|\gamma| + 2^\alpha\, C_1\, \varepsilon), \quad \|R_2(\cdot,\cdot,\varepsilon)\|_\rho \leq m_\alpha\, C_2\, \varepsilon, \tag{3.19}$$

for $m_\alpha = \max\{(\frac{5}{2})^{1-2\alpha}, 2^{2\alpha-1}\}$, is derived from $1/2 \leq |\xi| \leq 5/2$ and (3.14), so that $\|l\|_{\rho,\sigma} < \infty$. For the derivatives w.r. to $\varepsilon$, we have

$$\frac{\partial R_1}{\partial \varepsilon}(\theta, \xi, \varepsilon) = -\frac{1}{\alpha}\,\frac{\xi^{1-\alpha}}{\varepsilon^{1+1/\alpha}}\,\frac{\partial F_1}{\partial r}\Big(\theta, \frac{\xi}{\varepsilon^{1/\alpha}}\Big), \quad \frac{\partial R_2}{\partial \varepsilon}(\theta, \xi, \varepsilon) = -\frac{1}{\alpha}\,\frac{\xi^{2-\alpha}}{\varepsilon^{1+1/\alpha}}\,\frac{\partial F_2}{\partial r}\Big(\theta, \frac{\xi}{\varepsilon^{1/\alpha}}\Big),$$

for $\varepsilon \in ]0, \sigma]$. Using (3.15), we deduce that

$$\Big\|\frac{\partial R_1}{\partial \varepsilon}(\cdot,\cdot,\varepsilon)\Big\|_\rho \leq \frac{2^{2\alpha}}{\alpha}\, C_1^{(1)}, \quad \Big\|\frac{\partial R_2}{\partial \varepsilon}(\cdot,\cdot,\varepsilon)\Big\|_\rho \leq \frac{m_\alpha}{\alpha}\, C_2^{(1)},$$

so that $\|l\|_{1,\rho,\sigma} < \infty$ and therefore $l \in \mathcal{M}_{1,\rho,\sigma}$.

Step 3: The symplectic condition. We apply assumption (d) to observe that

$$h(\theta, \xi, \varepsilon) = \varepsilon^{1/\alpha}\, \mathfrak{h}\Big(\theta, \frac{\xi}{\varepsilon^{1/\alpha}}\Big) \tag{3.20}$$

is a potential for $\psi_\varepsilon$ on $G_\rho$, i.e., $\xi_1\, d\theta_1 - \xi\, d\theta = dh(\cdot, \varepsilon)$ is satisfied. Moreover, from (3.16), we obtain that

$$h(\theta, \xi, \varepsilon) = \varepsilon\, \mathfrak{m}(\theta, \xi) + \mathcal{O}(\varepsilon^2) \quad \text{as} \quad \varepsilon \to 0, \tag{3.21}$$

uniformly in $(\theta,\xi)\in G_\rho$, where $\mathfrak{m}(\theta,\xi)=\mathfrak{h}_0(\theta,\xi)$; for this note that $\mathfrak{m}$ is homogeneous in $\xi$ of degree $1-\alpha$. The function $\mathfrak{m}$ is bounded on $G_\rho$. Next, condition (2.6) follows from (3.17), and it should be observed that the bound on $\|\varepsilon^{-1}(\frac{\partial h}{\partial\varepsilon}-\mathfrak{m})\|_\rho$ then only depends upon $C_3$, $C_3^{(1)}$ and $\alpha$. It remains to prove that $h\in\mathcal{M}_{1,\rho,\sigma}$ in order to conclude that the family $\{\psi_\varepsilon\}$ is E-symplectic; note that $h$ is extended to $\varepsilon=0$ by $h(\theta,\xi,0)=0$. From (3.21) we get the continuity of $h$ on $G_\rho\times[0,\sigma]$. Next we are going to show that condition (d) in the definition of $\mathcal{M}_{1,\rho,\sigma}$ (see Definition 2.1(ii)) also holds. The definition of $\mathfrak{h}_0$ and (3.16), (3.17) imply that

$$\mathfrak{h}(\theta,r)=\mathcal{O}(r^{1-\alpha})\quad\text{and}\quad\frac{\partial\mathfrak{h}}{\partial r}(\theta,r)=\mathcal{O}(r^{-\alpha})$$

uniformly in $\theta\in\mathbb{R}_\delta$. Thus using these estimates, we obtain a uniform (in $\varepsilon\in]0,\sigma]$) bound on $\|h(\cdot,\varepsilon)\|_\rho$ and $\|\frac{\partial h}{\partial\varepsilon}(\cdot,\varepsilon)\|_\rho$.

Step 4: Application of Theorem 2.5. Since

$$l(x,0)=\begin{pmatrix}R_1(\theta,\xi,0)\\R_2(\theta,\xi,0)\end{pmatrix}=\begin{pmatrix}\gamma\xi^{-\alpha}\\0\end{pmatrix},$$

we deduce from (2.1) that

$$E(\theta,\xi)=\frac{\gamma}{1-\alpha}\,\xi^{1-\alpha},$$

and thus in fact $E=E(\xi)$. Then Theorem 2.5 yields the existence of $\hat\sigma\in]0,\sigma]$ and constants $\hat C,\hat D>0$ such that if $\varepsilon\in[0,\hat\sigma]$ and

$$(x_n)_{0\le n\le N}=(\theta_n,\xi_n)_{0\le n\le N}$$

is a real forward orbit piece of $\psi_\varepsilon$ so that $1<\xi_n<2$ for all $0\le n\le N$, then

$$|E(\xi_n)-E(\xi_0)|\le\hat C\varepsilon,\quad 0\le n\le\min\{N,N_\varepsilon\},\quad N_\varepsilon=[e^{\hat D/\varepsilon}].\tag{3.22}$$

Step 5: Going back to the original system. First we fix two numbers $1<a<3/2<b<2$. Let $(\theta_n,r_n)_{n_1\le n\le n_2}$ be a real forward orbit piece of $f$ such that $\varepsilon=(2r_{n_1}/3)^{-\alpha}<\hat\sigma$, where $\hat\sigma$ is from the previous step; by decreasing $\hat\sigma$ further, we may assume that in addition

$$\hat\sigma\le\frac{\gamma}{(1-\alpha)\hat C}\,\min\Big\{\Big(\frac{3}{2}\Big)^{1-\alpha}-a^{1-\alpha},b^{1-\alpha}-\Big(\frac{3}{2}\Big)^{1-\alpha}\Big\}\tag{3.23}$$

as well as

$$\hat\sigma\le\min\Big\{\frac{2-b}{m_\alpha C_2},\frac{a-1}{m_\alpha C_2}\Big\}\tag{3.24}$$

are verified. Then we have $\xi_{n_1}=\varepsilon^{1/\alpha}r_{n_1}=3/2\in]1,2[$, where in general we let $\xi_n=\varepsilon^{1/\alpha}r_n$. Denote by

$$N_\omega=\max\{m\ge n_1:1<\xi_n<2\text{ for }n\text{ with }n_1\le n\le m\}$$

the longest time such that along the orbit $(\theta_n,\xi_n)_{n_1\le n\le n_2}$ of $\psi_\varepsilon$ it holds that $1<\xi_n<2$. From Step 4 and (3.22) it follows that

$$\Big|\xi_n^{1-\alpha}-\Big(\frac{3}{2}\Big)^{1-\alpha}\Big|\le\frac{1-\alpha}{\gamma}\,\hat C\,\varepsilon,\quad n_1\le n\le\min\{N_\omega,n_1+N_\varepsilon\},\quad N_\varepsilon=[e^{\hat D/\varepsilon}].$$

Thus for $n_1 \le n \le \min\{N_\omega, n_1 + N_\varepsilon\}$ we deduce from $\varepsilon < \hat{\sigma}$ and (3.23) that

$$a^{1-\alpha} < \left(\frac{3}{2}\right)^{1-\alpha} - \frac{1-\alpha}{\gamma}\,\hat{C}\,\varepsilon < \xi_n^{1-\alpha} \le \left(\frac{3}{2}\right)^{1-\alpha} + \frac{1-\alpha}{\gamma}\,\hat{C}\,\varepsilon < b^{1-\alpha}. \tag{3.25}$$

We claim that $N_\omega \ge n_1 + N_\varepsilon$. Otherwise (3.25) would be applicable $n = N_\omega$ to imply that $\xi_{N_\omega} \in [a, b]$. We do also know that $\xi_{N_\omega+1} \notin ]1, 2[$. But then

$$\xi_{N_\omega+1} = \xi_{N_\omega} + \varepsilon R_2(\theta_{N_\omega}, \xi_{N_\omega}, \varepsilon)$$

together with (3.19) and $\varepsilon \le 1$ would lead to $|\xi_{N_\omega+1} - \xi_{N_\omega}| \le m_\alpha\, C_2\, \varepsilon^2 \le m_\alpha\, C_2\, \varepsilon$. This in turn would yield $\xi_{N_\omega+1} \in ]1, 2[$ by (3.24), which is impossible. This completes the argument for $N_\omega \ge n_1 + N_\varepsilon$, and the previous discussion can be summarized as follows: If $r_{n_1} > (3/2)\,\hat{\sigma}^{-1/\alpha}$, then

$$|r_n^{1-\alpha} - r_{n_1}^{1-\alpha}| \le C_4\, r_{n_1}^{1-2\alpha}, \quad n_1 \le n \le n_1 + [e^{C_5\, r_{n_1}^\alpha}], \tag{3.26}$$

where $C_4 = (\frac{2}{3})^{1-2\alpha}(\frac{1-\alpha}{\gamma})\hat{C}$ and $C_5 = (2/3)^\alpha \hat{D}$.

Step 6: Conclusion. In terms of $s_n = C_5^{\frac{1-\alpha}{\alpha}}\, r_n^{1-\alpha}$ and writing $m = n_1$, (3.26) reads as follows: If $s_m > C_6$, then

$$|s_n - s_m| \le C_7\, s_m^\beta, \quad m \le n \le m + [e^{s_m^\delta}], \tag{3.27}$$

where $\beta = \frac{1-2\alpha}{1-\alpha} < 1$, $\delta = \frac{\alpha}{1-\alpha} > 0$, $C_6 = (3/2)^{1-\alpha}\,(C_5/\hat{\sigma})^{1/\delta}$ and $C_7 = C_4 C_5$. We need to prove that $\limsup_{m\to\infty} \frac{s_m}{(\log m)^{1/\delta}} \le C_*$ for an appropriate constant $C_* > 0$ that will be independent of the initial condition $s_0$. Suppose now that $\limsup_{n\to\infty} s_n = \infty$, or equivalently $\limsup_{n\to\infty} r_n = \infty$.

We intend to adapt (3.27) to the framework described in Lemma 3.4 and Remark 3.5. The function $g(x) = x + C_7\, x^\beta$ is increasing in $I = ]b, \infty[$, where we take $b = 0$ for $\beta \ge 0$ and $b = (C_7|\beta|)^{\frac{1}{1-\beta}}$ for $\beta < 0$. In addition,

$$h(n) = n + [e^{s_n^\delta}]$$

satisfies (3.8), since all the $s_n$ are positive. The numbers $b_*$ and $b^*$ are defined as follows. First we take

$$b_* = \max\{C_6 + 1, 2b, (2C_7)^{\frac{1}{1-\beta}}\}.$$

To obtain $b^*$ so that (3.12) is satisfied for $\gamma_n = s_n$ we need an estimate of the type $s_n \ge C_9\, s_{n+1} - 1$, which is valid for all $n \in \mathbb{N}$. By the definition of the map and by (3.14):

$$r_{n+1} = r_n + r_n^{1-\alpha} F_2(\theta_n, r_n) \le r_n + C_2\, r_n^{1-2\alpha},$$

which translates into

$$\begin{aligned} s_{n+1} &\le C_5^{\frac{1-\alpha}{\alpha}} (r_n + C_2\, r_n^{1-2\alpha})^{1-\alpha} \le C_5^{\frac{1-\alpha}{\alpha}} (r_n^{1-\alpha} + C_2\, r_n^{(1-2\alpha)(1-\alpha)}) \\ &= s_n + C_8\, s_n^{1-2\alpha} \le (C_8 + 1)(s_n + 1) \end{aligned}$$

for a suitable constant $C_8 > 0$. Expressed differently, we have the bound

$$s_n \ge C_9\, s_{n+1} - 1 \quad \text{for} \quad C_9 = (C_8 + 1)^{-1}. \tag{3.28}$$

Therefore an appropriate choice for $b^* > b_* > b$ is $b^* = (C_8 + 1)(b_* + 1)$. Finally we take $\Gamma_n = (A + Bn)^{\frac{1}{1-\beta}}$ with $B = 2C_7(1-\beta)$ and $A^{\frac{1}{1-\beta}} \geq \max\{s_0, b^*\}$. From the discussion prior to this proof we know that $(\Gamma_n)$ will be an upper solution to $x_{n+1} = x_n + C_7\, x_n^\beta$, if $A$ is fixed to be sufficiently large (depending on $s_0$, $\beta$, $C_6$, $C_7$). Clearly $\Gamma_0 \geq \max\{s_0, b^*\}$ and $(\Gamma_n)$ is increasing. Lastly, $(s_n)$ satisfies (3.13), the latter due to (3.27): if $s_m \geq b_*$, then $s_m > C_6$ and (3.27) applies. It follows that $s_m - C_7 s_m^\beta \leq s_n \leq g(s_m)$ for $0 \leq m \leq n \leq h(m)$. The lower bound also yields

$$s_n \geq s_m(1 - C_7 s_m^{\beta-1}) \geq \frac{1}{2}\, s_m \geq b$$

for such $n$. Hence Lemma 3.4 provides us with an increasing function $\sigma : \mathbb{N}_0 \to \mathbb{N}$ such that $s_{\sigma(n)} > \Gamma_n$, $s_m \leq \Gamma_n$ for $m \in \{0, \ldots, \sigma(n) - 1\}$ and furthermore

$$\sigma(n+1) > \sigma(n) - 1 + [e^{s^\delta_{\sigma(n)-1}}], \quad n \in \mathbb{N}_0. \tag{3.29}$$

Thus from (3.29), (3.28) and $s_{\sigma(n)} > \Gamma_n$ we deduce

$$\sigma(n+1) > \sigma(n) + e^{s^\delta_{\sigma(n)-1}} - 2 \geq \sigma(n) + e^{(C_9\, s_{\sigma(n)} - 1)^\delta} - 2 \geq \sigma(n) + e^{(C_9 \Gamma_n - 1)^\delta} - 2.$$

After some straightforward manipulations using the definition of $\Gamma_n$ and $\frac{\delta}{1-\beta} = 1$, this yields

$$\sigma(n+1) \geq \sigma(n) + C_{10}\, e^{c_{11} n} - 2$$

for constants $C_{10}, c_{11} > 0$ depending upon $C_9$, $\delta$, $A$ and $B$. Therefore

$$\sigma(n) = \sigma(0) + \sum_{k=0}^{n-1} (\sigma(k+1) - \sigma(k)) \geq 1 + C_{10} \sum_{k=0}^{n-1} e^{c_{11} k} - 2n \geq C_{10}\, \frac{e^{c_{11} n} - 1}{e^{c_{11}} - 1} - 2n. \tag{3.30}$$

Thus $\sigma$ has at least exponential growth, which means that its 'inverse' will remain below a logarithm. More precisely, let

$$\psi(m) = \min\{n \in \mathbb{N}_0 : m < \sigma(n)\}.$$

Then $m \leq \sigma(\psi(m)) - 1$ and hence $s_m \leq \Gamma_{\psi(m)}$. In addition, from (3.30) it follows that $\psi(m) \leq \log(m + C_{12}) + C_{13}$ for suitable constants $C_{12}, C_{13} > 0$. This in turn leads to

$$s_m \leq (A + B\psi(m))^{\frac{1}{1-\beta}} \leq (A + B \log(m + C_{12}) + BC_{13})^{1/\delta}$$

and therefore also

$$\limsup_{m \to \infty} \frac{s_m}{(\log m)^{1/\delta}} \leq C_*$$

for $C_* = B^{1/\delta}$, which completes the proof. Note that $C_*$ is independent of the initial condition $s_0$, but in general the $n_0$ from the statement of Theorem 3.1 will depend on $s_0$. □

# 4 Application to the ping-pong map

The Fermi-Ulam ping-pong map (see [5]) for the forcing function $p$ is usually expressed in terms of the variables time and velocity at the impacts with one of the rackets. Assuming that this racket is fixed, the equations for the map $(t_0, v_0) \mapsto (t_1, v_1)$ are

$$t_1 = \tilde{t} + \frac{p(\tilde{t})}{v_1}, \quad v_1 = v_0 - 2\dot{p}(\tilde{t}),$$

where $\tilde{t} = \tilde{t}(t_0, v_0)$ denotes the hitting time to the other racket, which is obtained from the relation $(\tilde{t} - t_0)v_0 = p(\tilde{t})$. A computation shows that $v_1\, dt_1 \wedge dv_1 = v_0\, dt_0 \wedge dv_0$, and this formula suggests the energy $E = \frac{1}{2}v^2$ to be used as the conjugate variable of time. In this way we obtain the symplectic map $\Psi : (t_0, E_0) \mapsto (t_1, E_1)$,

$$t_1 = \tilde{t} + \frac{p(\tilde{t})}{\sqrt{2}(\sqrt{E_0} - \sqrt{2}\,\dot{p}(\tilde{t}))}, \quad E_1 = (\sqrt{E_0} - \sqrt{2}\,\dot{p}(\tilde{t}))^2,$$

where $\tilde{t} = \tilde{t}(t_0, v_0)$ is implicitly defined by means of

$$\tilde{t} = t_0 + \frac{p(\tilde{t})}{\sqrt{2E_0}}.$$

The real domain of the map $\Psi$ contains a half-plane of the type $t_0 \in \mathbb{R}$, $E > R_*$ (see [5]).

As an application of Theorem 3.1 we will obtain the following result, which is an upper bound for the velocities in the analytic case; also see [11, Example 5].

**Theorem 4.1** *Let $\delta > 0$ and $p : \mathbb{R}_\delta \to \mathbb{C}$ be holomorphic and such that $p$ maps reals into reals, $|p(z)| \le C$ for $z \in \mathbb{R}_\delta$, and $0 < a \le p(t) \le b$ for $t \in \mathbb{R}$. Then there exist constants $C_*, E_* > 0$, depending only upon the parameters, such that if $(t_n, E_n)_{n\in\mathbb{N}_0}$ is a forward complete real orbit of $\Psi$ with $\liminf_{n\to\infty} E_n \ge E_*$, then there is $n_0 \in \mathbb{N}$ so that*

$$|E_n| \le C_*(\log n)^2, \quad n \ge n_0,$$

*and for the velocities $v_n = \sqrt{2E_n}$ this means $|v_n| \le \sqrt{2C_*}\log n$ for $n \ge n_0$.*

**Remark 4.2** According to [11], only logarithmic accelerations are possible. This means that the estimate $v_n = \mathcal{O}(\log n)$ as $n \to \infty$ holds for any forward complete orbit. The above statement is slightly stronger, since the estimate has some uniformity: note that

$$\limsup_{n\to\infty} \frac{v_n}{\log n} \le \sqrt{2C_*},$$

and $C_*$ is independent of the chosen orbit.

The idea of the proof is to use Theorem 3.1, not in the coordinates $(t, E)$, but in $(w, W)$ given by $w = \int_0^t \frac{ds}{p(s)^2}$ and $W = p(t)^2 E$. Thus we need to verify that the map $(w_0, W_0) \mapsto (w_1, W_1)$ satisfies the assumptions of Theorem 3.1. This will be accomplished in three steps. First we are going to show that $\Psi$ has a well-defined holomorphic extension. In the second step we will prove that the map $(w_0, W_0) \mapsto (w_1, W_1)$ is exact symplectic, and the function $\mathfrak{h} = \mathfrak{h}(w_0, W_0)$ satisfying

$$W_1\, dw_1 - W_0\, dw_0 = d\mathfrak{h}$$

will be computed. Finally, after applying Theorem 3.1 to this new map, we will go back to the original to obtain the conclusion. Incidentally, we would like to mention that the quantity $W^{1/2}$ appears in [11, Example 5], where it is considered as an adiabatic invariant.

## 4.1 The complexified map

We start with two lemmas on holomorphic functions.

**Lemma 4.3** *Let $g : \mathbb{R}_\delta \to \mathbb{C}$ be holomorphic and such that $\operatorname{Re} g'(z) > 0$ for $z \in \mathbb{R}_\delta$. Then $g$ is one-to-one.*

**Proof:** This is a particular case of [13, Prop. 1.10]. □

**Remark 4.4** Under the assumptions of Lemma 4.3, as $g$ is non-constant and holomorphic, the image $g(\mathbb{R}_\delta) \subset \mathbb{C}$ is open. Thus $g^{-1} : g(\mathbb{R}_\delta) \to \mathbb{R}_\delta$ is well-defined and holomorphic by the inverse function theorem.

**Lemma 4.5** *Let $g : \mathbb{R}_\delta \to \mathbb{C}$ be holomorphic such that $g$ maps reals into reals and there exists $\alpha > 0$ so that $\operatorname{Re} g'(z) > \alpha$ for $z \in \mathbb{R}_\delta$. Then $\mathbb{R}_{\alpha\delta} \subset g(\mathbb{R}_\delta)$.*

**Proof:** Fix $w = a + ib \in \mathbb{R}_{\alpha\delta}$, i.e., we have $|b| < \alpha\delta$. In particular, we can choose $\sigma \in ]0, \delta[$ such that $|b| < \alpha\sigma$ holds. To find a solution $z \in \mathbb{R}_\delta$ of $g(z) = w$, note first that $g(\mathbb{R}) = \mathbb{R}$ by assumption. Hence there is $x \in \mathbb{R}$ satisfying $g(x) = a$. We consider the functions $f_1(z) = g(z) - a$ and $f_2(z) = -ib$ and our intention is to apply Rouché's Theorem on the rectangular region bounded by

$$\Gamma = \{\xi : \operatorname{Re}\xi \in [x - \Delta, x + \Delta], \operatorname{Im}\xi = \pm\sigma\} \cup \{\xi : \operatorname{Re}\xi = x \pm \Delta, \operatorname{Im}\xi \in [-\sigma, \sigma]\},$$

where $\Delta > 0$ will be taken to be large enough (see below). Once we have established that

$$|f_1(\xi)| > |f_2(\xi)|, \quad \xi \in \Gamma, \tag{4.1}$$

the proof will be complete, since then $f_1(z) = g(z) - a$ and $f_1(z) + f_2(z) = g(z) - w$ will have the same numbers of zeros inside of $\Gamma$; this number is one, by the choice of $x$ and as $g$ is one-to-one by Lemma 4.3. To check (4.1) on the horizontal parts of $\Gamma$ take $\xi = t \pm i\sigma$, where $t \in [x - \Delta, x + \Delta]$. Then $g(x) \in \mathbb{R}$ yields

$$|f_1(\xi)| = |g(t \pm i\sigma) - g(x)| \geq |\operatorname{Im} g(t \pm i\sigma)|.$$

As $g'$ is real on $\mathbb{R}$, one has, using the hypothesis,

$$|\operatorname{Im} g(t \pm i\sigma)| = \left|\operatorname{Im} \int_x^{t \pm i\sigma} g'(z)\, dz\right| = \left|\operatorname{Im} \int_t^{t \pm i\sigma} g'(z)\, dz\right| = \left|\operatorname{Re} \int_0^{\sigma} g'(t \pm is)\, ds\right| \geq \alpha|\sigma|.$$

It follows that $|f_1(\xi)| \geq \alpha|\sigma| > |b| = |f_2(\xi)|$. It remains to verify (4.1) on the vertical parts of $\Gamma$. For, take $\xi = (x \pm \Delta) + is$, where $s \in [-\sigma, \sigma]$. Define $K = \max\{|g(x+iv)| : v \in [-\sigma, \sigma]\}$. Now observe that, for $t \in \mathbb{R}$,

$$
\begin{aligned}
|\operatorname{Re} g(t+is)| &= \left|\operatorname{Re} g(x+is) + \operatorname{Re} \int_{x+is}^{t+is} g'(z)\,dz\right| \geq \left|\operatorname{Re} \int_x^t g'(u+is)\,du\right| - K \\
&\geq \alpha|t-x| - K
\end{aligned}
$$

due to the hypothesis. As a consequence,

$$
|f_1(\xi)| = |g(x \pm \Delta + is) - a| \geq \alpha\Delta - K - |a| > |b| = |f_2(\xi)|,
$$

provided that we fix $\Delta > \alpha^{-1}(|a| + |b| + K)$. □

Now we return to the ping-pong map and define $\varphi(z) = p(z)^2$ for $z \in \mathbb{R}_\delta$. From the assumptions on $p$ we may assume that $a \leq b \leq C$. Then

$$
\delta_1 = \frac{a^2\delta}{4C^2}
$$

satisfies $\delta_1 \leq \frac{\delta}{4}$. The expression $\operatorname{Arg} z \in ]-\pi, \pi]$ for $z \in \mathbb{C} \setminus \{0\}$ will denote the argument of $z$.

**Lemma 4.6** *The function $\varphi$ satisfies*

$$
\operatorname{Re}\varphi(z) \geq \frac{1}{2}\,a^2 \quad \textit{and} \quad |\varphi(z)| \leq 2b^2, \quad z \in \mathbb{R}_{\delta_1},
$$

*and moreover*

$$
|\operatorname{Arg}\varphi(z)| < \frac{\pi}{4}, \quad z \in \mathbb{R}_{\delta_1/2}.
$$

**Proof:** From the Cauchy integral formula we deduce that

$$
|\varphi'(z)| \leq \frac{2C^2}{\delta} \quad \text{for} \quad z \in \mathbb{R}_{\delta/2}.
$$

Take $z = t + is \in \mathbb{R}_{\delta/2}$. Then $\varphi(z) = \varphi(t) + \int_t^{t+is} \varphi'(\xi)\,d\xi$ implies that

$$
\operatorname{Re}\varphi(z) \geq a^2 + \operatorname{Re} \int_t^{t+is} \varphi'(\xi)\,d\xi \geq a^2 - \frac{2C^2}{\delta}\,|s|
$$

as well as

$$
|\varphi(z)| \leq b^2 + \left|\int_t^{t+is} \varphi'(\xi)\,d\xi\right| \leq b^2 + \frac{2C^2}{\delta}\,|s|
$$

and

$$
|\operatorname{Im}\varphi(z)| \leq \frac{2C^2}{\delta}\,|s|.
$$

Thus if $z \in \mathbb{R}_{\delta_1}$, then $\operatorname{Re}\varphi(z) \geq a^2/2$ and $|\varphi(z)| \leq b^2 + 2\frac{C^2}{\delta}\frac{a^2\delta}{4C^2} \leq 3b^2/2$. In addition, if $z \in \mathbb{R}_{\delta_1/2}$, then $|\operatorname{Im}\varphi(z)| \leq a^2/4 < \operatorname{Re}\varphi(z)$, which yields the claim on the argument. □

**Lemma 4.7** *The function $\tau : \mathbb{R}_{\delta_1} \to \mathbb{C}$, $\tau(z) = \int_0^z \frac{d\zeta}{p(\zeta)^2} = \int_0^z \frac{d\zeta}{\varphi(\zeta)}$, is holomorphic, one-to-one with holomorphic inverse, and satisfies*

$$\mathbb{R}_{\sigma(\Delta)} \subset \tau(\mathbb{R}_\Delta) \quad \textit{for} \quad \Delta \in ]0, \delta_1], \tag{4.2}$$

*where $\sigma(\Delta) = \frac{a^2}{2C^4}\Delta$.*

**Proof:** By Lemma 4.6 the function $\varphi(z)$ does not vanish on the simply connected domain $\mathbb{R}_{\delta_1}$, and hence $\frac{1}{\varphi(z)}$ has a holomorphic primitive $\tau(z)$. Lemma 4.3 in conjunction with Lemma 4.6 implies that $\tau$ is one-to-one. Also $\tau$ maps reals into reals and satisfies

$$\operatorname{Re}\tau'(z) = \operatorname{Re}\left(\frac{1}{\varphi(z)}\right) = \frac{1}{|\varphi(z)|^2}\operatorname{Re}\varphi(z) \geq \frac{a^2}{2C^4}.$$

Thus Lemma 4.5 applies with $\alpha = \frac{a^2}{2C^4}$ to prove that (4.2) holds. Concerning the fact that $\tau^{-1} : \tau(\mathbb{R}_{\delta_1}) \to \mathbb{C}$ is holomorphic, cf. Remark 4.4. □

To extend the ping-pong map $\Psi$ as a holomorphic map, we take the complex square root to be $\sqrt{z} = |z|^{1/2}\exp((i/2)\operatorname{Arg} z)$, where the complex plane is cut along $]-\infty, 0]$. In particular, $\sqrt{z}$ is holomorphic on $\mathbb{C}\setminus]-\infty, 0]$ and extends the positive square root. Note that $\sqrt{\varphi(z)} = p(z)$ holds for all $z \in \mathbb{R}_{\delta_1/2}$. To establish this identity, it is sufficient to adapt the proof of Lemma 4.6 to the function $p$ to conclude that $|\operatorname{Arg} p(z)| < \frac{\pi}{4}$ for $z \in \mathbb{R}_{\delta_1/2}$.

**Lemma 4.8** *Let $\underline{e} = \frac{8C^2}{\delta^2}$. Then for every $z \in \mathbb{R}_{\delta/4}$ and $E \in \mathbb{C}\setminus]-\infty, 0]$ such that $|E| > \underline{e}$ the equation*

$$\tilde{z} = z + \frac{p(\tilde{z})}{\sqrt{2E}} \tag{4.3}$$

*has a unique solution $\tilde{z} = \tilde{z}(z, E)$ lying in $\mathbb{R}_{\delta/2}$. Moreover, $(z, E) \mapsto \tilde{z}(z, E)$ is holomorphic as a function of two variables. In addition, $\Delta \in ]0, \delta/4]$ and $z \in \mathbb{R}_\Delta$ implies that $\tilde{z}(z, E) \in \mathbb{R}_{2\Delta}$.*

**Proof:** Consider the function $g(\tilde{z}, E) = \tilde{z} - \frac{p(\tilde{z})}{\sqrt{2E}}$, so that we need to solve $g(\tilde{z}, E) = z$. For $\tilde{z} \in \mathbb{R}_{\delta/2}$ one has

$$|p'(\tilde{z})| \leq \frac{2C}{\delta} \tag{4.4}$$

by the Cauchy integral formula. It follows that

$$\operatorname{Re}\frac{\partial g}{\partial \tilde{z}}(\tilde{z}, E) = 1 - \operatorname{Re}\frac{p'(\tilde{z})}{\sqrt{2E}} \geq 1 - \frac{|p'(\tilde{z})|}{\sqrt{2}|E|^{1/2}} \geq 1 - \frac{2C}{\sqrt{2}\delta\underline{e}^{1/2}} = \frac{1}{2}$$

for $\tilde{z} \in \mathbb{R}_{\delta/2}$. Thus from Lemma 4.3 we infer that $g(\tilde{z}, E) = z$ can have at most one solution $\tilde{z} \in \mathbb{R}_{\delta/2}$. Next let $\Delta \in ]0, \delta/4]$. Then $\operatorname{Re}\frac{\partial g}{\partial \tilde{z}}(\tilde{z}, E) \geq 1/2$ for $\tilde{z} \in \mathbb{R}_{2\Delta}$. Therefore we can invoke Lemma 4.5 with $\alpha = 1/2$ to obtain $\mathbb{R}_\Delta \subset g(\cdot, E)(\mathbb{R}_{2\Delta})$. Finally we can apply the implicit function theorem to deduce that $\tilde{z} = \tilde{z}(z, E)$ is holomorphic on the domain $\mathbb{R}_{\delta/4} \times \{E \in \mathbb{C}\setminus]-\infty, 0] : |E| > \underline{e}\}$. □

Now we are in a position to define the holomorphic extension of the ping-pong map,

$$\Psi : \mathcal{U}_0 \subset \mathbb{C}^2 \to \mathbb{C}^2, \quad (z, E) \mapsto (z_1, E_1),$$

given by

$$z_1 = \tilde{z} + \frac{p(\tilde{z})}{\sqrt{2}(\sqrt{E} - \sqrt{2}\, p'(\tilde{z}))}, \quad E_1 = (\sqrt{E} - \sqrt{2}\, p'(\tilde{z}))^2, \tag{4.5}$$

where $\tilde{z} = \tilde{z}(z, E)$ is from Lemma 4.8 and

$$\mathcal{U}_0 = \{(z, E) \in \mathbb{R}_{\delta/4} \times (\mathbb{C} \setminus ]-\infty, 0]) : |E| > \underline{e}\}.$$

To see that this map is well-defined, we first observe that, according to Lemma 4.8, $\tilde{z} \in \mathbb{R}_{\delta/2}$. Thus both $p$ and $p'$ can be evaluated at $\tilde{z}$. Moreover, using (4.4) it follows that the denominator in the equation defining $z_1$ never vanishes: we have

$$|E|^{1/2} > \underline{e}^{1/2} = \frac{2\sqrt{2}\, C}{\delta} \geq \sqrt{2}\, |p'(\tilde{z})|.$$

## 4.2 The change of variables and the new map

The map $\Psi$ is exact symplectic on the domain

$$\widehat{\mathcal{U}_0} = \{(z, E) \in \mathcal{U}_0 : z \in \mathbb{R}_{\delta_1/2}\}.$$

More precisely, it satisfies the identity

$$E_1\, dz_1 - E\, dz = dh \tag{4.6}$$

for

$$h(z, E) = -\frac{1}{2}\, p(\tilde{z})^2 \left( \frac{1}{z_1 - \tilde{z}} + \frac{1}{\tilde{z} - z} \right). \tag{4.7}$$

The new restriction on the size of $|\mathrm{Im}\, z|$ guarantees that $h$ is holomorphic on $\widehat{\mathcal{U}_0}$. In fact, both denominators $z_1 - \tilde{z}$ and $\tilde{z} - z$ do not vanish. This is a consequence of the definitions of $z_1$ and $\tilde{z}$, together with the inequality

$$|p(\tilde{z})| \geq \frac{1}{\sqrt{2}}\, a > 0, \quad z \in \mathbb{R}_{\delta_1/2},$$

which in turn follows from Lemmas 4.6 and 4.8.

The generating function for the ping-pong map was computed in [5]. This computation, together with the relationship between the function $h$ and the generating function (see [7]), imply that (4.6), (4.7) holds. Note that all computations in [5] were done on the real domain of $\Psi$, but once again we rely on the uniqueness of holomorphic extensions.

Later we will need to reformulate (4.6), (4.7) in the new variables $(w, W)$, where $w = \tau(z)$ and $W = p(z)^2 E$. This can be achieved from general principles, without any further computation. For this reason we include a short digression into general maps.

Consider the space $\mathbb{C}^2$ endowed with the 1-form

$$\sigma = p\,dq,$$

where $q, p \in \mathbb{C}$ are the coordinates of a point. Assume that $D, \mathcal{D} \subset \mathbb{C}^2$ are two domains with sub-domains $D_1 \subset D$ and $\mathcal{D}_1 \subset \mathcal{D}$. Let $\chi : \mathcal{D} \to D$ be a holomorphic diffeomorphism such that $\chi(\mathcal{D}_1) \subset D_1$ and

$$\chi^*\sigma = \sigma + dm$$

for some holomorphic function $m : \mathcal{D} \to \mathbb{C}$. In addition, let $T : D_1 \to \mathbb{C}^2$ be a holomorphic map with $T(D_1) \subset D$ and

$$T^*\sigma = \sigma + dh$$

for some holomorphic function $h : D_1 \to \mathbb{C}$. Then $\tilde{T} = \chi^{-1} \circ T \circ \chi : \mathcal{D}_1 \to \mathcal{D}$ is well-defined and a short calculation reveals that

$$\tilde{T}^*\sigma = \sigma + d\mathfrak{h} \tag{4.8}$$

for

$$\mathfrak{h} = h \circ \chi + m - m \circ \tilde{T}. \tag{4.9}$$

In fact, the standard properties of pullbacks of differential forms yield

$$\begin{aligned} d(h \circ \chi) &= \chi^*(dh) = \chi^*(T^*\sigma - \sigma) = (T \circ \chi)^*\sigma - \chi^*\sigma = (\chi \circ \tilde{T})^*\sigma - \chi^*\sigma \\ &= \tilde{T}^*(\chi^*\sigma) - \chi^*\sigma = \tilde{T}^*(\sigma + dm) - \sigma - dm = \tilde{T}^*\sigma - \sigma + d(m \circ \tilde{T} - m), \end{aligned}$$

which proves (4.8).

Now we go back to the ping-pong and introduce the full change of variables $\Gamma : (z, E) \mapsto (w, W)$.

**Lemma 4.9** *The map*

$$\Gamma : \mathbb{R}_{\delta_1} \times \mathbb{C} \to \mathbb{C}^2, \quad (z, E) \mapsto (w, W),$$

*where $w = \tau(z)$ and $W = p(z)^2 E$, is a holomorphic diffeomorphism between $\mathbb{R}_{\delta_1} \times \mathbb{C}$ and $\Gamma(\mathbb{R}_{\delta_1} \times \mathbb{C})$ that verifies $\Gamma^*\sigma = \sigma$. Moreover, if $\Delta \in ]0, \delta_1]$, then $\mathbb{R}_{\sigma(\Delta)} \times \mathbb{C} \subset \Gamma(\mathbb{R}_\Delta \times \mathbb{C})$.*

**Proof:** According to Lemma 4.7, $\Gamma$ is holomorphic and satisfies $\mathbb{R}_{\sigma(\Delta)} \times \mathbb{C} \subset \Gamma(\mathbb{R}_\Delta \times \mathbb{C})$ for $\Delta \in ]0, \delta_1]$, due to (4.2) and the fact that its inverse $\Gamma^{-1}$ is given by $(w, W) \mapsto (z, p(z)^{-2}W)$, where $z = \tau^{-1}(w)$. This inverse is also holomorphic. To prove that $\Gamma^*\sigma = \sigma$, observe that $dw = \frac{1}{p(z)^2}\,dz$, and hence $W\,dw = E\,dz$ as desired. □

To adjust our situation to the general framework as outlined above, we define

$$\mathcal{D} = \Gamma(\mathbb{R}_{\delta_1} \times \mathbb{C}), \quad D = \mathbb{R}_{\delta_1} \times \mathbb{C}, \quad \chi = \Gamma^{-1}, \quad T = \Psi.$$

Furthermore, we take $m = 0$ and $h$ from (4.7). To introduce $\mathcal{D}_1$ and $D_1$, let $\Delta = \frac{\delta_1}{4}$ and take $\rho \geq 4$ so large that $2\Delta + \frac{\delta}{2\sqrt{\rho}} < \delta_1$. Then we define

$$\begin{aligned} \mathcal{D}_1 &= \Big\{(w, W) \in \mathbb{R}_{\sigma(\Delta)} \times \mathbb{C} : |\mathrm{Arg}\, W| < \frac{\pi}{4}, |W| > \rho\, C^2 \underline{e}\Big\}, \\ D_1 &= \{(z, E) \in \mathbb{R}_\Delta \times \mathbb{C} : E \in \mathbb{C} \setminus ]-\infty, 0], |E| > \rho \underline{e}\}, \end{aligned}$$

and check all the conditions set out before. First of all, $D_1 \subset D$ and $\mathcal{D}_1 \subset \mathcal{D}$ are immediate, using Lemma 4.9 for the latter. Also $\chi : \mathcal{D} \to D$ is a holomorphic diffeomorphism by definition, and moreover $\chi(\mathcal{D}_1) \subset D_1$. For, let $(z, E) = \Gamma^{-1}(w, W) \in \chi(\mathcal{D}_1)$. Then $w \in \mathbb{R}_{\sigma(\Delta)}$ implies $z \in \mathbb{R}_\Delta$ by Lemma 4.9. Furthermore, since $|p(z)| \le C$,

$$|E| = \frac{|W|}{|p(z)|^2} > \frac{\rho\, C^2 \underline{e}}{C^2} = \rho \underline{e}$$

and due to Lemma 4.6,

$$|\operatorname{Arg} E| \le |\operatorname{Arg} W| + \left|\operatorname{Arg} \frac{1}{\varphi(z)}\right| < \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2},$$

so that in particular $E \not\in ]-\infty, 0]$ and accordingly $(z, E) \in D_1$. Next, to see that $T(D_1) \subset D$, we have to make sure that $(z_1, E_1) = T(z, E)$ has $|\operatorname{Im} z_1| < \delta_1$ for $(z, E) \in D_1$. As $z \in \mathbb{R}_\Delta$ and $|E| > \rho\underline{e} > \underline{e}$, we have $\tilde{z} \in \mathbb{R}_{2\Delta}$ by Lemma 4.8. In addition, (4.4) yields $|p'(\tilde{z})| \le \frac{2C}{\delta}$ and thus by the definition of $E_1$:

$$|E_1|^{1/2} = |\sqrt{E} - \sqrt{2}\, p'(\tilde{z})| \ge |E|^{1/2} - \frac{2\sqrt{2}\, C}{\delta} \ge \frac{1}{2}\, |E|^{1/2};$$

recall that $\underline{e} = \frac{8C^2}{\delta^2}$ and $\rho \ge 4$. Thus by the definition of $z_1$:

$$|\operatorname{Im} z_1| \le |\operatorname{Im} \tilde{z}| + \frac{|p(\tilde{z})|}{\sqrt{2}\, |E_1|^{1/2}} \le 2\Delta + \frac{\sqrt{2}\, C}{|E|^{1/2}} \le 2\Delta + \frac{\delta}{2\sqrt{\rho}} < \delta_1,$$

which completes the argument for $T(D_1) \subset D$. Since $D_1 \subset \widehat{\mathcal{U}}_0$, $h$ from (4.7) is well-defined and holomorphic on $D_1$ and we have $T^*\sigma = \sigma + dh$ due to (4.6).

Now that we have verified the conditions of the general argument, we can conclude from (4.8) and (4.9) that $\Phi = \Gamma \circ \Psi \circ \Gamma^{-1} : \mathcal{D}_1 \to \mathbb{C}^2$, $(w, W) \mapsto (w_1, W_1)$, is well-defined and satisfies

$$W_1\, dw_1 - W\, dw = d\mathfrak{h}, \quad \mathfrak{h} = h \circ \Gamma^{-1}. \tag{4.10}$$

## 4.3 Application of the main theorem and proof of Theorem 4.1

To summarize, so far we have established that the map $\Phi : \mathcal{D}_1 \to \mathbb{C}^2$ is well-defined and holomorphic. We are going to apply Theorem 3.1 with $f = \Phi$, $\underline{r} = \rho\, C^2 \underline{e}$, $\eta = \frac{1}{2}$, $\alpha = \frac{1}{2}$ and $\gamma = \sqrt{2}$. For the width of the strip we take $\delta$ to be $\sigma(\Delta)$, and in order to write $\Phi$ in the required form (3.3), we introduce $F_1 = F$ and $F_2 = G$ for

$$F(w, W) = \sqrt{W}(w_1 - w) - \sqrt{2}, \quad G(w, W) = \frac{W_1 - W}{\sqrt{W}}. \tag{4.11}$$

We also define $\tilde{F}(w, W) = w_1 - w - \frac{\sqrt{2}}{\sqrt{W}}$, so that $F(w, W) = \sqrt{W}\, \tilde{F}(w, W)$. Then the assumptions (a) and (b) of Theorem 3.1 are satisfied. For, recall from (3.2) that $\Phi$ needs to be defined on

$$\Omega = \{(w, W) \in \mathbb{R}_\delta \times \mathbb{C} : \operatorname{Re} W > \underline{r}, |\operatorname{Im} W| < \eta |W|\},$$

which is the case due to $\Omega \subset \mathcal{D}_1$.

To derive the needed bounds $|F| = \mathcal{O}(|W|^{-1/2})$ and $|G| = \mathcal{O}(|W|^{-1/2})$ as required by (c), we need to make some preliminary observations. From Lemma 4.8 we know that $\tilde{z} \in \mathbb{R}_{\delta/2}$, so that $|p'(\tilde{z})| \le 2C/\delta$ by (4.4). In general, if $\xi \in \mathbb{C}$ satisfies $\operatorname{Re}\xi > 0$, then $\operatorname{Re}\sqrt{\xi} > \frac{1}{\sqrt{2}}|\xi|^{1/2}$. Therefore if $|E| > 2\underline{e}$ and $\operatorname{Re} E > 0$, then

$$\operatorname{Re}(\sqrt{E} - \sqrt{2}\,p'(\tilde{z})) > \frac{1}{\sqrt{2}}\sqrt{2\underline{e}} - \sqrt{2}\,\frac{2C}{\delta} = 0,$$

and hence $E_1 \in \mathbb{C}\backslash]-\infty, 0]$. As a consequence, $\sqrt{E_1}$ can be understood as a single-valued expression and we can write the first equation in (4.5) as

$$z_1 = \tilde{z} + \frac{p(\tilde{z})}{\sqrt{2E_1}}. \tag{4.12}$$

Lastly, if even $z \in \mathbb{R}_\Delta$, $|E| > \rho\underline{e}$ and $\operatorname{Re} E > 0$ holds, then one also has

$$|\sqrt{E} - \sqrt{2}\,p'(\tilde{z})| \ge \sqrt{|E|} - \sqrt{2}\,\frac{2C}{\delta} \ge \frac{1}{2}\sqrt{|E|},$$

and hence

$$|E_1| \ge \frac{1}{4}\,|E|. \tag{4.13}$$

We also have $z \in \mathbb{R}_\Delta \subset \mathbb{R}_{\delta_1}$, and therefore $|\varphi(z)| \le 2b^2$ by Lemma 4.6. It follows that $|E| = |\frac{W}{\varphi(z)}| \ge \frac{1}{2b^2}|W|$, and thus $|E_1| \ge \frac{1}{8b^2}|W|$ due to (4.13). Then from the definitions of $\tilde{z}$ and $z_1$, cf. (4.3) and (4.12),

$$|z - \tilde{z}| \le \frac{C}{\sqrt{2}|E|^{1/2}} \le Cb\,|W|^{-1/2}, \quad |z_1 - \tilde{z}| \le \frac{C}{\sqrt{2}|E_1|^{1/2}} \le 2Cb\,|W|^{-1/2}. \tag{4.14}$$

In particular, $|z - z_1| \le 3Cb\,|W|^{-1/2}$. To bound $\tilde{F}$, we first observe that

$$z_1 - z = \tilde{z} + \frac{p(\tilde{z})}{\sqrt{2E_1}} - z = \Big[\frac{1}{\sqrt{2E}} + \frac{1}{\sqrt{2E_1}}\Big]\,p(\tilde{z}). \tag{4.15}$$

Since $|\operatorname{Arg} W| < \frac{\pi}{4}$ and $|\operatorname{Arg}\frac{1}{\varphi(z)}| < \frac{\pi}{4}$, we have $\sqrt{W} = \sqrt{\varphi(z)}\sqrt{E} = p(z)\sqrt{E}$. Now the expression for $\tilde{F}$ is split up according to

$$\tilde{F}(w, W) = w_1 - w - \frac{\sqrt{2}}{\sqrt{W}} = w_1 - w - \frac{\sqrt{2}}{\sqrt{E}}\,\frac{1}{p(z)} = \tilde{F}_1 + \tilde{F}_2,$$

where

$$\begin{aligned} \tilde{F}_1 &= w_1 - w - \Big[\frac{1}{\sqrt{2E}} + \frac{1}{\sqrt{2E_1}}\Big]\,\frac{1}{p(z)} \\ &= \tau(z_1) - \tau(z) - \frac{z_1 - z}{p(z)p(\tilde{z})} \end{aligned}$$

by (4.15), and

$$\tilde{F}_2 = \Big[\frac{1}{\sqrt{2E_1}} - \frac{1}{\sqrt{2E}}\Big]\,\frac{1}{p(z)}.$$

For the first term,

$$|\tilde{F}_1| = \Big|\int_z^{z_1} \Big(\frac{1}{p(\zeta)^2} - \frac{1}{p(z)p(\tilde{z})}\Big)\,d\zeta\Big|. \tag{4.16}$$

From geometric considerations we deduce that

$$\begin{aligned} |\zeta - z| &\leq |z_1 - z| \leq 3Cb\,|W|^{-1/2}, \\ |\zeta - \tilde{z}| &\leq \max\{|\tilde{z} - z|, |\tilde{z} - z_1|\} \leq 2Cb\,|W|^{-1/2}, \end{aligned}$$

for any point $\zeta$ on the segment $[z, z_1]$. The upper and lower bounds for $p$ provided by Lemma 4.6 together with the estimate for $|p'|$ allow us to find a constant $K_1 > 0$ such that for each $\zeta \in [z, z_1]$:

$$\Big|\frac{1}{p(\zeta)^2} - \frac{1}{p(z)p(\tilde{z})}\Big| \leq K_1 \max\{|\zeta - z|, |\zeta - \tilde{z}|\} \leq 3CK_1 b\,|W|^{-1/2}.$$

As a consequence, (4.16) yields

$$|\tilde{F}_1| \leq 3CK_1 b\,|W|^{-1/2}\,|z_1 - z| \leq 9C^2K_1b^2\,|W|^{-1}.$$

To bound $\tilde{F}_2$, we note that by (4.5) and (4.4)

$$|\sqrt{E_1} - \sqrt{E}| \leq \sqrt{2}\,|p'(\tilde{z})| \leq \frac{2\sqrt{2}\,C}{\delta}.$$

Therefore, due to (4.13),

$$|\tilde{F}_2| = \frac{|\sqrt{E_1} - \sqrt{E}|}{\sqrt{2}\sqrt{|E_1|}\sqrt{|E|}}\,\frac{1}{|p(z)|} \leq \frac{2C}{\delta\sqrt{|E_1|}\sqrt{|W|}} \leq \frac{4C|p(z)|}{\delta|W|} \leq \frac{4C^2}{\delta}\,|W|^{-1},$$

and thus we deduce that altogether

$$|F(w,W)| = \sqrt{|W|}\,|\tilde{F}(w,W)| \leq \sqrt{|W|}\,(|\tilde{F}_1| + |\tilde{F}_2|) \leq C_1|W|^{-1/2}$$

holds for an appropriate constant $C_1 > 0$ depending only upon $\delta$, $C$, $a$, $b$. Concerning the bound on $|G|$, according to [9, (5.10)] one has

$$W_1 - W = \frac{1}{2}\,\varphi(\tilde{z})\int_0^1 (1-\lambda)\Big[\varphi''((1-\lambda)\tilde{z} + \lambda z) - \varphi''((1-\lambda)\tilde{z} + \lambda z_1)\Big]\,d\lambda\,, \tag{4.17}$$

where once again $\varphi(z) = p(z)^2$. In the paper just mentioned this relation was used for a real-valued $p$, but as all functions involved are holomorphic, it extends to the complex-valued case due to the uniqueness theorem in complex analysis. Now $\Delta \leq \delta_1/4$ and $\delta_1 \leq \delta/4$ yields $|\mathrm{Im}\,((1-\lambda)\tilde{z} + \lambda z)| \leq |\mathrm{Im}\,\tilde{z}| + |\mathrm{Im}\,z| < 2\Delta + \Delta = 3\Delta < \delta/2$ and similarly $|\mathrm{Im}\,((1-\lambda)\tilde{z} + \lambda z_1)| \leq |\mathrm{Im}\,\tilde{z}| + |\mathrm{Im}\,z_1| < 2\Delta + \delta_1 < \delta/2$. Owing to the Cauchy integral formula one has $|\varphi'''(z)| \leq (2/\delta)^3C^2$ for $z \in \mathbb{R}_{\delta/2}$. Therefore (4.17) implies that

$$|W_1 - W| \leq 4\,C^4\delta^{-3}|z_1 - z| \leq 12\,C^5 b\,\delta^{-3}\,|W|^{-1/2},$$

and hence

$$|G(w,W)| \le 12\, C^5 b\, \delta^{-3}\, |W|^{-1},$$

which is in fact better than $G = \mathcal{O}(|W|^{-1/2})$ what we would have needed in assumption (c) of Theorem 3.1.

Lastly we are going to verify the hypothesis (d) of Theorem 3.1, the function $\mathfrak{h}$ being given by (4.10) with $h$ from (4.7). We also note that $\mathfrak{h}_0(w,W) = -\sqrt{2W}$ for our choice of parameters and we need to establish that $|\mathfrak{h} - \mathfrak{h}_0| = \mathcal{O}(1)$. To simplify the estimates, it is convenient to express $h$ and $\mathfrak{h}$ is a different way, which is based on the definition of the maps $\Psi$ and $\Phi$. More precisely, using the various definitions we write

$$\begin{aligned} h(z,E) &= -\frac{1}{2}\, p(\tilde{z})^2 \Big( \frac{1}{z_1 - \tilde{z}} + \frac{1}{\tilde{z} - z} \Big) \\ &= -\frac{1}{2}\, p(\tilde{z})^2 \Big( \frac{\sqrt{2E_1}}{p(\tilde{z})} + \frac{\sqrt{2E}}{p(\tilde{z})} \Big) \\ &= -\frac{1}{2}\, p(\tilde{z}) \Big( \sqrt{2}\, (\sqrt{E} - \sqrt{2}\, p'(\tilde{z})) + \sqrt{2E} \Big) \\ &= -\sqrt{2E}\, p(\tilde{z}) + p(\tilde{z}) p'(\tilde{z}) \end{aligned}$$

and

$$\mathfrak{h}(w,W) = -\sqrt{2W}\, \frac{p(\tilde{z})}{p(z)} + p(\tilde{z}) p'(\tilde{z}),$$

where $(z,E) = \Gamma^{-1}(w,W)$. As a consequence,

$$\mathfrak{h}(w,W) - \mathfrak{h}_0(w,W) = \sqrt{2W} \Big( 1 - \frac{p(\tilde{z})}{p(z)} \Big) + p(\tilde{z}) p'(\tilde{z}).$$

Similarly as before, the lower bound on $|p(z)|$ and the upper bound on $|p'|$ together with (4.14) lead to

$$\Big| 1 - \frac{p(\tilde{z})}{p(z)} \Big| \le K_2\, |z - \tilde{z}| \le C K_2\, b\, |W|^{-1/2},$$

which proves that

$$|\mathfrak{h}(w,W) - \mathfrak{h}_0(w,W)| \le C_2.$$

Let us now fix $E_* > \frac{1}{a^2}\, \underline{r}$, where $\underline{r}$ appears in the definition of the domain $\Omega$. Suppose that $(t_n, E_n)_{n \in \mathbb{N}_0}$ is a forward complete real orbit of $\Psi$ with $\liminf_{n \to \infty} E_n \ge E_*$. By assumption there exists $N \in \mathbb{N}_0$ such that $E_n > \frac{1}{a^2}\, \underline{r}$ for $n \ge N$. Then $W_n = p(t_n)^2 E_n \ge a^2 E_n > \underline{r}$ for $n \ge N$ shows that $(w_n, W_n)_{n \ge N}$ is a forward complete real orbit for $\Phi$, and hence Theorem 3.1 is applicable. □

# 5 An example

Consider the map $f : (\theta, r) \mapsto (\theta_1, r_1)$ defined as

$$\theta_1 = \theta + \sqrt{\frac{2}{r_1}}, \quad r_1 = r - q(\theta), \tag{5.1}$$

where $q$ is a given function. This map is symplectic, because it can be expressed in the form $r = \frac{\partial g}{\partial \theta}$, $r_1 = -\frac{\partial g}{\partial \theta_1}$, for the generating function

$$g(\theta, \theta_1) = \frac{2}{\theta_1 - \theta} + Q(\theta)$$

with $Q$ denoting a primitive of $q$. This is possibly the simplest family of maps in the framework of Section 3. We will analyze the dynamics for the particular case where

$$q(\theta) = -\frac{2\theta}{(1+\theta^2)^N}$$

for an integer $N \geq 1$.

Assuming that $(\theta_0, r_0) \in \mathbb{R}^2$ is such that $\theta_0 > 0$ and $r_0 > 0$, we observe that a forward complete orbit $(\theta_n, r_n)_{n\in\mathbb{N}_0}$ can be produced; the sequences $\theta_n$ and $r_n$ are positive and increasing. We are going to prove that always

$$\sup_{n\in\mathbb{N}_0} r_n < \infty \quad \text{if} \quad N \geq 2,$$

whereas

$$0 < \liminf_{n\to\infty} \frac{r_n}{(\log n)^2} \leq \limsup_{n\to\infty} \frac{r_n}{(\log n)^2} < \infty \quad \text{if} \quad N = 1.$$

As we will see, Theorem 3.1 is applicable for $N \geq 2$, but not for $N = 1$. Therefore the predictions of the theorem in this particular case are far from being optimal. So the main open questions in this respect are: Is the growth rate $r_n = \mathcal{O}((\log n)^2)$ best possible in Theorem 3.1? On the contrary, are all orbits bounded?

## 5.1 Applicability of Theorem 3.1

First note that the map (5.1) can be written in the form (3.3), with $\alpha = 1/2$, $\gamma = \sqrt{2}$,

$$F_1(\theta, r) = \sqrt{2}\left(\frac{1}{\sqrt{1 - \frac{q(\theta)}{r}}} - 1\right), \quad F_2(\theta, r) = -\frac{1}{\sqrt{r}}\, q(\theta), \tag{5.2}$$

where in the following the complex square root will be understood as before.

The function $q$ is bounded and holomorphic on any strip $\mathbb{R}_\delta$ with $\delta < 1$; we fix $\delta = 1/2$ for definiteness. It follows that $|q(\theta)| \leq C$ for $\theta \in \mathbb{R}_{1/2}$, where $C = 4^{N+1}$. To prove this, consider $\theta \in \mathbb{R}_{1/2}$ and $|\theta| \geq 2$ first. Here we have

$$|q(\theta)| \leq \frac{2|\theta|}{(|\theta|^2 - 1)^N} \leq \frac{2|\theta|}{(|\theta|^2/2)^N} \leq \frac{1}{2^{N-2}}.$$

If $\theta \in \mathbb{R}_{1/2}$ and $|\theta| \leq 2$, then $|1 + \theta^2| = |\theta + i||\theta - i| \geq 1/4$ yields

$$|q(\theta)| = \frac{2|\theta|}{|1+\theta^2|^N} \leq 2 \cdot 4^N |\theta| \leq 4^{N+1}.$$

We now proceed as in the previous section to find an appropriate domain of holomorphy $\Omega \subset \mathbb{C}^2$. Clearly the hypotheses (a)-(c) of Theorem 3.1 are satisfied. The validity of (d) is more delicate. As has been used before, in general the primitive of the form $r_1\, d\theta_1 - r\, d\theta$ is computed from the generating function $g$ via $\mathfrak{h}(\theta, r) = -g(\theta, \theta_1(\theta, r))$. Thus for the map from (5.1) we get

$$\mathfrak{h}(\theta, r) = -\sqrt{2(r - q(\theta))} - Q(\theta).$$

We also note that

$$\mathfrak{h}_0(\theta, r) = -\sqrt{2r},$$

cf. (3.1), and (3.4) says that we should have $|\mathfrak{h}(\theta, r) - \mathfrak{h}_0(\theta, r)|$ bounded, uniformly in $(\theta, r) \in \Omega$, in order that Theorem 3.1 is applicable. For $N \geq 2$ the primitive $Q(\theta) = \frac{1}{(N-1)(1+\theta^2)^{N-1}}$ is bounded, which shows that Theorem 3.1 can be used. On the other hand, for $N = 1$ the primitive is $Q(\theta) = -\log(1+\theta^2)$, and the best estimate one can get is $|\mathfrak{h}(\theta, r) - \mathfrak{h}_0(\theta, r)| = \mathcal{O}(1)$ for each $\theta \in \mathbb{R}_{1/2}$, but it is not uniform.

## 5.2 The real dynamics

We start with a useful notion of equivalence for sequences.

**Definition 5.1** *Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be two sequences of eventually positive numbers. We say that $(a_n)$ is equivalent to $(b_n)$, if there exist constants $C > c > 0$ and $n_0 \in \mathbb{N}$ such that $ca_n \leq b_n \leq Ca_n$ is verified for $n \geq n_0$; this will be written as $(a_n) \simeq (b_n)$.*

**Lemma 5.2** *Let $(\rho_n)_{n\in\mathbb{N}}$ be an increasing sequence of positive numbers such that*

$$(\rho_n) \simeq \Big(\sum_{k=2}^{n-1} \frac{1}{R_k^\sigma}\Big) \tag{5.3}$$

*for some $\sigma \geq 1$, where $R_k = \sum_{j=1}^{k-1} \frac{1}{\rho_j^{1/2}}$. If $\sigma > 1$, then $(\rho_n)$ is bounded. If $\sigma = 1$, then $(\rho_n) \simeq ((\log n)^2)$.*

The proof is given in the next subsection.

Going back to the map $f$ from (5.1), we consider $(\theta_0, r_0) \in \mathbb{R}^2$ such that $\theta_0 > 0$ and $r_0 > 0$. Let $(\theta_n, r_n)_{n\in\mathbb{N}_0}$ denote the resulting forward complete orbit; we already noted above that $\theta_n$ and $r_n$ are positive and increasing. As a first step we are going to show that $\lim_{n\to\infty} \theta_n = \infty$. Otherwise we would have $0 < \theta_0 \leq \theta_n \leq C$ for $n \in \mathbb{N}$. Then $r_{n+1} - r_n = -q(\theta_n) = |q(\theta_n)|$ implies that

$$0 < m \leq r_{n+1} - r_n \leq M < \infty, \quad n \in \mathbb{N},$$

where $m = \inf_{n\in\mathbb{N}} |q(\theta_n)|$ and $M = \sup_{n\in\mathbb{N}} |q(\theta_n)|$. Then $(r_n) \simeq (n)$ and consequently

$$\theta_n = \theta_0 + \sum_{j=1}^{n} \sqrt{\frac{2}{r_j}} \to \infty, \tag{5.4}$$

which is a contradiction.

In terms of $R_n = \sum_{j=1}^{n-1} \frac{1}{\sqrt{r_j}}$ the relation (5.4) can be written as $\theta_n = \theta_0 + \sqrt{2}\, R_{n+1}$, which implies that also $\lim_{n\to\infty} R_n = \infty$ holds, and furthermore $(\theta_n) \simeq (R_{n+1})$. Since $R_{n+1} - R_n \to 0$, we deduce that $(\theta_n) \simeq (R_n)$ is verified. Due to $\lim_{n\to\infty} \theta_n = \infty$ and

$$|q(\theta)|\, \theta^{2N-1} = \frac{2\theta^{2N}}{(1+\theta^2)^N},$$

there are constants $K > k > 0$ such that

$$\frac{k}{\theta_n^{2N-1}} \le |q(\theta_n)| \le \frac{K}{\theta_n^{2N-1}}, \quad n \ge 0.$$

If follows that for suitable $n_0 \in \mathbb{N}$ and constants $K^* > k^* > 0$ one has

$$\frac{k^*}{R_n^{2N-1}} \le |q(\theta_n)| \le \frac{K^*}{R_n^{2N-1}}, \quad n \ge n_0.$$

As a consequence, for $n \ge n_0$ we obtain from $r_n = r_{n_0-1} + \sum_{j=n_0}^{n} |q(\theta_j)|$ that

$$r_{n_0-1} + k^* \sum_{j=n_0}^{n} \frac{1}{R_j^{2N-1}} \le r_n \le r_{n_0-1} + K^* \sum_{j=n_0}^{n} \frac{1}{R_j^{2N-1}}.$$

From this it is easily deduced that $(r_n) \sim (\sum_{j=2}^{n-1} \frac{1}{R_j^{2N-1}})$, and hence Lemma 5.2 applies with $\sigma = 2N-1$. Its conclusion is that $(r_n)$ is bounded, if $N \ge 2$, whereas $(r_n) \simeq ((\log n)^2)$, if $N = 1$.

## 5.3 Some auxiliary results

**Lemma 5.3** *For some $a, b > 0$ consider the differential equation $y' = a e^{-by^{1/2}}$, $y > 0$. Then every solution satisfies*

$$\lim_{x\to\infty} \frac{y(x)}{(\ln x)^2} = \frac{1}{b^2}. \tag{5.5}$$

**Proof:** First we observe that $y$ is increasing and $y'(x) < a$. Thus $y$ is well-defined for $x \to \infty$ and satisfies $y(x) \to \infty$ as $x \to \infty$, since the equation has no equilibrium. By separation of variables,

$$\left(y(x)^{1/2} - \frac{1}{b}\right) e^{by(x)^{1/2}} = \frac{ab}{2}\, x + C,$$

where $C$ is a constant. Taking any $b_1 < b < b_2$ we deduce that, for large $x$,

$$e^{b_1 y(x)^{1/2}} < \frac{ab}{2}\, x < e^{b_2 y(x)^{1/2}},$$

which yields the claim upon taking the logarithm. □

**Lemma 5.4** *Let $\rho : [1,\infty[ \to [1,\infty[$ be continuous. Furthermore, suppose that there are constants $0 < \gamma < \Gamma$ and $x_0 > 1$ so that*

$$\gamma\,\rho(x) \le \int_1^x \frac{dy}{R(y)^\sigma} \le \Gamma\,\rho(x), \quad x \ge x_0, \tag{5.6}$$

*for some $\sigma \ge 1$, where $R(y) = \int_1^y \frac{d\xi}{\rho(\xi)^{1/2}}$ for $y \ge 1$. If $\sigma > 1$, then $\rho$ is bounded. If $\sigma = 1$, then there are constants $0 < c < C$ such that*

$$c \le \frac{\rho(x)}{(\log x)^2} \le C \tag{5.7}$$

*for $x$ sufficiently large.*

**Proof:** Let $\phi(x) = \int_1^x \frac{dy}{R(y)^\sigma}$. Then $\phi' = R^{-\sigma}$ and $\phi'' = -\sigma R^{-(\sigma+1)} R'$ together with $R' = \rho^{-1/2}$ leads to the differential equation $\phi'' = -\sigma(\phi')^{\frac{\sigma+1}{\sigma}}\rho^{-1/2}$. Then (5.6) yields

$$\gamma\sigma^2\,\frac{\phi'(x)^{\frac{2(\sigma+1)}{\sigma}}}{\phi''(x)^2} \le \phi(x) \le \Gamma\sigma^2\,\frac{\phi'(x)^{\frac{2(\sigma+1)}{\sigma}}}{\phi''(x)^2}, \quad x \ge x_0.$$

Owing to $\phi'' < 0$, this can be rewritten as

$$-\Gamma^{1/2}\sigma\,\frac{\phi'(x)}{\phi(x)^{1/2}} \le \frac{\phi''(x)}{\phi'(x)^{\frac{1}{\sigma}}} \le -\gamma^{1/2}\sigma\,\frac{\phi'(x)}{\phi(x)^{1/2}}, \quad x \ge x_0. \tag{5.8}$$

First we consider the case where $\sigma = 1$. Here we obtain

$$\frac{d}{dx}\Big(\log\phi'(x) + 2\Gamma^{1/2}\phi(x)^{1/2}\Big) \ge 0,$$
$$\frac{d}{dx}\Big(\log\phi'(x) + 2\gamma^{1/2}\phi(x)^{1/2}\Big) \le 0,$$

for $x \ge x_0$. Upon integration and exponentiation one gets

$$\phi'(x)\,e^{2\gamma^{1/2}\phi(x)^{1/2}} \le e^{b_0}, \quad \phi'(x)\,e^{2\Gamma^{1/2}\phi(x)^{1/2}} \ge e^{B_0},$$

for $x \ge x_0$, where $b_0 = \log\phi'(x_0) + 2\gamma^{1/2}\phi(x_0)^{1/2}$ and $B_0 = \log\phi'(x_0) + 2\Gamma^{1/2}\phi(x_0)^{1/2}$. Therefore $\phi$ is a lower solution of $y' = ae^{-by^{1/2}}$ for $\underline{a} = e^{b_0}$, $\underline{b} = 2\gamma^{1/2}$ and an upper solution for $\overline{a} = e^{B_0}$, $\overline{b} = 2\Gamma^{1/2}$. Let $\underline{y}$ and $\overline{y}$ denote the corresponding solutions with common initial values $\underline{y}(x_0) = \overline{y}(x_0) = \phi(x_0)$. Then

$$\underline{y}(x) \le \phi(x) \le \overline{y}(x), \quad x \ge x_0.$$

According to Lemma 5.3 one has $\lim_{x\to\infty}\frac{\underline{y}(x)}{(\ln x)^2} = \frac{1}{4\gamma}$ and $\lim_{x\to\infty}\frac{\overline{y}(x)}{(\ln x)^2} = \frac{1}{4\Gamma}$. Recalling (5.6), this leads to (5.7), where we can take for instance $c = \frac{1}{8\gamma\Gamma}$ and $C = \frac{1}{2\gamma\Gamma}$.

In the second case $\sigma > 1$, (5.8) can be expressed as

$$-\Gamma^{1/2}\sigma\,\frac{\phi'(x)}{\phi(x)^{1/2}} \le \frac{\phi''(x)}{\phi'(x)^{\frac{1}{\sigma}}} \le -\gamma^{1/2}\sigma\,\frac{\phi'(x)}{\phi(x)^{1/2}}.$$

Upon integration of the inequality on the right-hand side, it is found that

$$\frac{d}{dx}\Big(\frac{\sigma}{\sigma-1}\,\phi'(x)^{\frac{\sigma-1}{\sigma}}+2\gamma^{1/2}\sigma\,\phi(x)^{1/2}\Big)\le 0.$$

Therefore it follows from $\phi'\ge 0$ that

$$2\gamma^{1/2}\sigma\,\phi(x)^{1/2}\le\frac{\sigma}{\sigma-1}\,\phi'(x)^{\frac{\sigma-1}{\sigma}}+2\gamma^{1/2}\sigma\,\phi(x)^{1/2}\le\frac{\sigma}{\sigma-1}\,\phi'(x_0)^{\frac{\sigma-1}{\sigma}}+2\gamma^{1/2}\sigma\,\phi(x_0)^{1/2},$$

which shows that $\phi$ is bounded. Since $1\le\rho(x)\le\gamma^{-1}\phi(x)$, also $\rho$ is bounded. □

**Proof of Lemma 5.2 :** First we consider the case where $\sigma=1$.

Step 1: $\rho_n\to\infty$ as $n\to\infty$. Otherwise we would have $\rho_n\to\rho_\infty\in]0,\infty[$ as $n\to\infty$. But then $(R_n)\simeq(n)$, and consequently the series $\sum\frac{1}{R_n}$ is divergent. However, this contradicts (5.3).

Step 2: $R_n\to\infty$ as $n\to\infty$. Otherwise we would have $R_n\to R_\infty\in]0,\infty[$ as $n\to\infty$. Then (5.3) yields

$$(\rho_n)\simeq\Big(\sum_{k=2}^{n-1}\frac{1}{R_k}\Big)\simeq(n),$$

but this in turn leads to $(R_n)\simeq(\sum_{j=1}^{n-1}\frac{1}{j^{1/2}})$, which is divergent as $n\to\infty$.

Step 3: $(\rho_{n+1})_{n\ge1}\simeq(\rho_n)_{n\ge1}$ and $(R_{n+1})_{n\ge1}\simeq(R_n)_{n\ge1}$. To establish these assertions, we first introduce a convenient notion. A sequence $(a_n)$ will be said to have the bounded difference property (BD property, for short), if $a_n\to\infty$ as $n\to\infty$ and the sequence of progressive differences $(a_{n+1}-a_n)$ is bounded. If $(a_n)$ has the BD property, then $(a_{n+1})_{n\ge1}\simeq(a_n)_{n\ge1}$, since $|a_{n+1}/a_n-1|\le C/|a_n|\le 1/2$ for $n$ large enough. The BD property is not invariant under the equivalence of sequences, but if $(a_n)$ has the BD property and $(b_n)\simeq(a_n)$, then $(b_{n+1})\simeq(b_n)$.

Returning to $(\rho_n)$ and $(R_n)$, owing to (5.3) and Step 1 we know that $\sum_{k=2}^{n-1}\frac{1}{R_k}\to\infty$ as $n\to\infty$. Since the differences are bounded (even converging to zero) by Step 2, $(\sum_{k=2}^{n-1}\frac{1}{R_k})$ has the BD property. Invoking (5.3) once more, it follows that $(\rho_{n+1})\simeq(\rho_n)$. Similarly, $(R_n)$ has the BD property, and thus $(R_{n+1})\simeq(R_n)$.

Step 4: To prove that $(\rho_n)\simeq((\log n)^2)$, we may assume that $\rho_n\ge1$ for $n\in\mathbb{N}$. Then the function $\rho:[1,\infty[\to[1,\infty[$ obtained by piecewise linear interpolation from $\rho(n)=\rho_n$ is continuous, increasing and such that $\lim_{x\to\infty}\rho(x)=\infty$. Let $R(y)=\int_1^y\frac{d\xi}{\rho(\xi)^{1/2}}$ for $y\ge1$. According to Lemma 5.4 it is sufficient to establish the estimate (5.6). If $j\le\xi\le j+1$, then $\rho_j\le\rho(\xi)\le\rho_{j+1}$. Since

$$R(n)=\int_1^n\frac{d\xi}{\rho(\xi)^{1/2}}=\sum_{j=1}^{n-1}\int_j^{j+1}\frac{d\xi}{\rho(\xi)^{1/2}}$$

for $n\in\mathbb{N}$, we deduce that $R_{n+1}-\rho_1^{-1/2}\le R(n)\le R_n$. Hence we may employ Step 3 to obtain $(R(n))\simeq(R_n)$. Finally we observe that if $y\in[j,j+1]$, then $R(j)\le R(y)\le R(j+1)$. For $x\in[N,N+1]$ then

$$\int_1^x\frac{dy}{R(y)}=\sum_{j=1}^{N-1}\int_j^{j+1}\frac{dy}{R(y)}+\int_N^x\frac{dy}{R(y)}$$

yields

$$\sum_{j=1}^{N-1} \frac{1}{R(j+1)} \leq \int_1^x \frac{dy}{R(y)} \leq \sum_{j=1}^{N} \frac{1}{R(j)}.$$

If we now use (5.3) in conjunction with $\rho_N \leq \rho(x) \leq \rho_{N+1}$ and $(\rho_{n+1}) \simeq (\rho_n)$, the relation (5.6) follows easily.

In the case where $\sigma > 1$ we need to prove that $(\rho_n)$ is bounded. Assume on the contrary that we would have $\rho_n \to \infty$ as $n \to \infty$ (recall that the sequence is increasing). This would imply $R_n \to \infty$ as $n \to \infty$, as otherwise $R_n \to R_\infty \in ]0, \infty[$ as $n \to \infty$ for an appropriate $R_\infty$. Then (5.3) yields

$$(\rho_n) \simeq \Big( \sum_{k=2}^{n-1} \frac{1}{R_k^\sigma} \Big) \simeq (n),$$

but this in turn leads to $(R_n) \simeq (\sum_{j=1}^{n-1} \frac{1}{j^{1/2}})$, which is divergent as $n \to \infty$. Thus we are in the same position as after Steps 1 and 2 in the above argument. An inspection of Steps 3 and 4 shows that they can be straightforwardly adapted to the current setting. In other words, we can apply the case $\sigma > 1$ of Lemma 5.4, and hence the function $\rho(x)$ is found to be bounded. Since $\rho_n = \rho(n)$, the sequence $(\rho_n)$ must be bounded which is a contradiction and completes the proof of Lemma 5.2. □

# 6 Appendix I: A Hamiltonian normal form

In this appendix we will give fully detailed proofs of some of the results in [11] and we will discuss the assumptions that are needed for those proofs to work.

**Definition 6.1 (The class $\mathcal{H}_{\rho,\sigma}$)** *For $\rho > 0$ and $\sigma \in ]0, 1[$ let $\mathcal{H}_{\rho,\sigma}$ be the class of continuous functions $H : G_\rho \times \mathbb{R} \times [0, \sigma] \to \mathbb{C}$, $H = H(x, t, \varepsilon)$, satisfying*

*(a) $H$ is $T$-periodic in $t$;*

*(b) $H$ maps reals into reals;*

*(c) for every $t \in \mathbb{R}$ and $\varepsilon \in [0, \sigma]$ the function $H(\cdot, t, \varepsilon)$ is holomorphic on $G_\rho$; and*

*(d) the gradient w.r. to $x$, $\nabla H = \nabla_x H(q, p, t, \varepsilon)$, is a continuous function from $G_\rho \times \mathbb{R} \times [0, \sigma]$ to $\mathbb{C}^2$ such that*

$$\|\nabla H\|_{\rho,\sigma} := \sup \{ \|\nabla H(\cdot, \cdot, t, \varepsilon)\|_\rho : t \in \mathbb{R}, \varepsilon \in [0, \sigma] \} < \infty.$$

**Remark 6.2** Note that for a function $H \in \mathcal{H}_{\rho,\sigma}$ all partial derivatives $\partial_x^\alpha H : G_\rho \times \mathbb{R} \times [0, \sigma] \to \mathbb{C}^d$ w.r. to $x$ are again continuous functions of all three variables, where as usual $\partial_x^\alpha H = \frac{\partial^{|\alpha|}}{\partial_q^{\alpha_1} \partial_p^{\alpha_2}} H$ for a multi-index $\alpha \in \mathbb{N}_0^2$. This is a consequence of the fact that the Cauchy integral formula can be differentiated w.r. to $x$.

**Definition 6.3 (The class $\widetilde{\mathcal{H}}_{\rho,\sigma}$)** *The class $\widetilde{\mathcal{H}}_{\rho,\sigma}$ consists of those $H \in \mathcal{H}_{\rho,\sigma}$ with the additional property that*

$$\int_0^T H(x,t,\varepsilon)\,dt = 0 \tag{6.1}$$

*for $x \in G_\rho$ and $\varepsilon \in [0,\sigma]$.*

Observe that if $H \in \widetilde{\mathcal{H}}_{\rho,\sigma}$, then $t \mapsto \int_0^t H(x,s,\varepsilon)\,ds$ is $T$-periodic.

For $h \in \widetilde{\mathcal{H}}_{\rho,\sigma}$ consider the (time-dependent) implicit Euler transformation $\Phi : (x,t,\varepsilon) \mapsto y$ with inverse $(y,t,\varepsilon) \mapsto x = \Psi(y,t,\varepsilon)$, $x = (q,p)$, $y = (q_1,p_1)$, which is given by

$$q_1 = q - \varepsilon \int_0^t \frac{\partial h}{\partial p_1}(q,p_1,s,\varepsilon)\,ds, \quad p_1 = p + \varepsilon \int_0^t \frac{\partial h}{\partial q}(q,p_1,s,\varepsilon)\,ds. \tag{6.2}$$

Solving the second equation, we obtain $p_1 = p_1(q,p,t,\varepsilon)$, and the first equation then determines $q_1 = q_1(q,p,t,\varepsilon)$. We will show that the map $\Psi$ is well-defined and it is an admissible change of variables, in a sense that is made precise in the following definition.

**Definition 6.4** *Let $0 < \rho_1 \le \rho$ and $0 < \sigma_1 \le \sigma$. A map $\Psi : G_{\rho_1} \times \mathbb{R} \times [0,\sigma_1] \to \mathbb{C}^2$, $x = \Psi(y,t,\varepsilon)$, will be called an admissible change of variables, if it satisfies*

*(a) $\Psi$ maps reals into reals;*

*(b) $\Psi$ is $T$-periodic in $t$ and $\Psi(y,0,\varepsilon) = \Psi(y,T,\varepsilon) = y$;*

*(c) $\Psi$ is continuous;*

*(d) for every $t \in \mathbb{R}$ and $\varepsilon \in [0,\sigma_1]$ the map $\Psi(\cdot,t,\varepsilon)$ is holomorphic in $G_{\rho_1}$, and for every $y \in G_{\rho_1}$ and $\varepsilon \in [0,\sigma_1]$ the map $\Psi(y,\cdot,\varepsilon) \in C^1(\mathbb{R})$;*

*(e) all admissible partial derivatives with regard to $y$ and $t$ are continuous functions of all the arguments $(y,t,\varepsilon)$; and*

*(f) for every $t \in \mathbb{R}$ and $\varepsilon \in [0,\sigma_1]$ the map $\Psi(\cdot,t,\varepsilon)$ is a symplectic diffeomorphism from $G_{\rho_1}$ onto its image.*

**Lemma 6.5** *For $0 < r < \rho$ and $\sigma > 0$ given, let $\sigma_1 = \min\{\frac{\rho-r}{12},\sigma\}$. Then, for each $h \in \widetilde{\mathcal{H}}_{\rho,\sigma}$ with $T\|\nabla h\|_{\rho,\sigma} \le 1$, the equations (6.2) define a map*

$$\Psi : G_r \times \mathbb{R} \times [0,\sigma_1] \to \mathbb{C}^2$$

*that is an admissible change of variables and satisfies $\Psi(G_r,t,\varepsilon) \subset G_\rho$ for every $t \in \mathbb{R}$ and $\varepsilon \in [0,\sigma_1]$. Moreover,*

$$\|\Psi(\cdot,\cdot,\varepsilon) - I\|_r \le \varepsilon T \|\nabla h\|_{\rho,\sigma} \tag{6.3}$$

*for $\varepsilon \in [0,\sigma_1]$.*

**Remarks 6.6** (a) The simple geometry of $G$ implies the following useful fact: if $(q,p),(q_1,p_1) \in G_\rho$, then also $(q,p_1),(q_1,p) \in G_\rho$. For this reason the equations (6.2) are well-defined.
(b) The condition $T\|\nabla h\|_{\rho,\sigma} \le 1$ is just imposed to get a definitive value for $\sigma_1$. When we are going to apply the lemma to an arbitrary $h \in \widetilde{\mathcal{H}}_{\rho,\sigma}$ later, a rescaling argument (in $\varepsilon$) can be used.

**Proof of Lemma 6.5:** To solve the first equation in (6.2) we evenly split the interval $[r,\rho]$ into $r < R_1 < R_2 < \rho$, where $R_1 - r = R_2 - R_1 = \rho - R_2 = \frac{\rho-r}{3}$. Define

$$X = \{q \in \mathbb{C} : |\mathrm{Im}\, q| \le R_2\}$$

as well as $\hat\sigma = \min\{\frac{\rho-r}{6}, \sigma\}$. For $(q_1,p_1) \in G_{R_1}$ and $\varepsilon \in [0,\hat\sigma]$ fixed, let

$$\mathcal{F}(q) = q_1 + \varepsilon \int_0^t \frac{\partial h}{\partial p_1}(q,p_1,s,\varepsilon)\, ds.$$

Then $\mathcal{F} : X \to X$ is a self-map, since $\varepsilon T\|\nabla h\|_{\rho,\sigma} \le \varepsilon \le \hat\sigma < \frac{\rho-r}{3} = R_2 - R_1$. The condition (6.1) in Definition 6.3 allows us to restrict to the time interval $[0,T]$. From the Cauchy integral formula we deduce

$$\|D^2 h\|_{R_2,\sigma} \le \frac{1}{\rho - R_2}\, \|\nabla h\|_{\rho,\sigma} = \frac{3}{\rho - r}\, \|\nabla h\|_{\rho,\sigma}.$$

This estimate applies in particular to the cross-derivative $\frac{\partial^2 h}{\partial q \partial p_1}$ and ensures that $\mathcal{F}$ is a contraction, due to

$$\varepsilon T\|D^2 h\|_{R_2,\sigma} \le \varepsilon T \frac{3}{\rho - r}\, \|\nabla h\|_{\rho,\sigma} \le \frac{3\hat\sigma}{\rho - r} \le \frac{1}{2}. \tag{6.4}$$

The unique fixed point of $\mathcal{F}$ defines a continuous map $q = q(q_1,p_1,t,\varepsilon) : G_{R_1} \times \mathbb{R} \times [0,\hat\sigma] \to X$ Then the definition of $\Psi$ is completed by setting

$$p = p_1 - \varepsilon \int_0^t \frac{\partial h}{\partial q}(q(q_1,p_1,t,\varepsilon),p_1,s,\varepsilon)\, ds.$$

Note that

$$\mathrm{dist}(p,I) \le \mathrm{dist}(p_1,I) + \varepsilon T\|\nabla h\|_{\rho,\sigma} \le R_1 + \varepsilon \le R_1 + \hat\sigma < R_2,$$

and hence $\Psi$ is defined on $G_{R_1} \times \mathbb{R} \times [0,\hat\sigma]$ and takes values in $\overline{G_{R_2}} \subset G_\rho$. The bound

$$|\Psi(q_1,p_1,t,\varepsilon) - (q_1,p_1)| \le \varepsilon T\|\nabla h\|_{\rho,\sigma} \tag{6.5}$$

is a direct consequence of the definition of $\Psi$, and in particular (6.5) implies (6.3), since $r < R_1$.

To prove the smoothness of $\Psi$, we observe that $q$ is defined implicitly by the equation $F = 0$, where

$$F(q,q_1,p_1,t,\varepsilon) = q - q_1 - \varepsilon \int_0^t \frac{\partial h}{\partial p_1}(q,p_1,s,\varepsilon)\, ds.$$

The transversality condition

$$\frac{\partial F}{\partial q} = 1 - \varepsilon \int_0^t \frac{\partial^2 h}{\partial q \partial p_1}(q,p_1,s,\varepsilon)\, ds \ne 0$$

is satisfied, due to (6.4). Hence the implicit function theorem applies to yield that $q$ (and hence $p$) verifies all the smoothness requirements for an admissible change of variables.

It remains to establish that $\Psi(\cdot, t, \varepsilon)$ is a symplectic diffeomorphism from $G_r$ onto its image, for $t \in [0, T]$ and $\varepsilon \in [0, \sigma_1]$. Using (6.5), which is valid for $y = (q_1, p_1) \in G_{R_1}$, we deduce that

$$|D\Psi(q_1, p_1, t, \varepsilon) - I| \le \frac{\varepsilon}{R_1 - r} T\|\nabla h\|_{\rho,\sigma} = \frac{3\varepsilon}{\rho - r} T\|\nabla h\|_{\rho,\sigma} \le \frac{3\sigma_1}{\rho - r} \le \frac{1}{4}, \tag{6.6}$$

where $D\Psi = D_y\Psi = D_{(q_1,p_1)}\Psi$ is the Jacobian. This will allow us to interpret $\Psi(\cdot, t, \varepsilon)$ as a Lipschitz continuous perturbation of the identity. Indeed, if we define $\Gamma = \Psi - I$, then owing to the convexity of $G_r$ and from (6.6) we obtain the bound

$$\begin{aligned} |\Gamma(y, t, \varepsilon) - \Gamma(\tilde{y}, t, \varepsilon)| &= \Big| \int_0^1 \frac{d}{ds} [\Gamma(sy + (1 - s)\tilde{y}, t, \varepsilon)]\, ds \Big| \\ &\le 2 \cdot \frac{1}{4} |y - \tilde{y}| = \frac{1}{2} |y - \tilde{y}| \end{aligned}$$

for $y, \tilde{y} \in G_r$. Hence the Lipschitz constant of $\Gamma(\cdot, t, \varepsilon)$ is $\le 1/2$. This in turn implies that $\Psi(\cdot, t, \varepsilon)$ is one-to-one on $G_r$. According to (6.6), i.e., $|D\Psi(q_1, p_1, t, \varepsilon) - I| \le 1/4$, the matrix $D\Psi$ has an inverse. Thus the inverse function theorem can be applied at each fixed $y \in G_r$ to deduce that $\Psi(\cdot, t, \varepsilon)$ is a diffeomorphism from $G_r$ onto the open set $\Psi(G_r, t, \varepsilon) \subset G_\rho$. This diffeomorphism is symplectic, because it has been obtained from the equations (6.2), which can be derived from the generating function

$$S(q, p_1, t, \varepsilon) = qp_1 - \varepsilon \int_0^t h(q, p_1, s, \varepsilon)\, ds. \tag{6.7}$$

This completes the proof of the lemma. □

**Corollary 6.7** *Under the assumptions of Lemma 6.5, let* $0 < \hat{r} < r < \rho$ *and denote by* $\Psi : y \mapsto x$ *the map that is induced by (6.2). Let* $\sigma_2 = \min\{\sigma_1, \frac{r-\hat{r}}{2}\} = \min\{\frac{\rho-r}{12}, \frac{r-\hat{r}}{2}, \sigma\}$*. If* $t \in \mathbb{R}$ *and* $\varepsilon \in [0, \sigma_2]$*, then* $\Psi(G_r, t, \varepsilon) \supset G_{\hat{r}}$*.*

For the proof, the following result will be helpful, which is [16, Prop. I.3, p. 50].

**Lemma 6.8** *Let* $X, Y$ *be Banach spaces and suppose that* $U \subset Y$ *is open. If* $\Psi : U \to \Psi(U) \subset X$ *is a homeomorphism,* $\Psi^{-1}$ *is Lipschitz continuous with constant* $\mathrm{Lip}(\Psi^{-1}) < \lambda$*, and* $\overline{B_r(y)} \subset U$*, then*

$$\Psi(\overline{B_r(y)}) \supset \overline{B_{r/\lambda}(\Psi(y))}.$$

**Proof of Corollary 6.7:** We are going to apply Lemma 6.8 with $U = G_r$, $\Psi = \Psi(\cdot, t, \varepsilon)$ and $\lambda = 2$. Inspecting the proof of Lemma 6.5, we have shown in (6.6) that $|D\Psi(y) - I| \le 1/4$ for $y \in G_{R_1} \supset G_r$, where we write $\Psi(y) = \Psi(y, t, \varepsilon)$. This yields

$$|D\Psi(y)^{-1}| = \Big| \sum_{j=0}^{\infty} (-1)^j (D\Psi(y) - I)^j \Big| \le \sum_{j=0}^{\infty} 2^j \, |D\Psi(y) - I|^j \le 2.$$

Next observe that if $y \in G_{\hat{r}}$, then $\overline{B_{r-\hat{r}}(y)} \subset G_r$, as a consequence of the geometry of $G$ and the choice of the norm. For $y \in G_{\hat{r}}$ we also have $|y - \Psi(y)| \le \varepsilon$ by (6.3), which means that $y \in \overline{B_\varepsilon(\Psi(y))}$. Owing to Lemma 6.8 we obtain

$$y \in \overline{B_\varepsilon(\Psi(y))} \subset \overline{B_{(r-\hat{r})/2}(\Psi(y))} \subset \Psi(\overline{B_{r-\hat{r}}(y)}) \subset \Psi(G_r),$$

as claimed. □

**Lemma 6.9** *For $0 < r < \rho$ and $\sigma > 0$ given, let $H \in \mathcal{H}_{\rho,\sigma}$ and $h \in \widetilde{\mathcal{H}}_{\rho,\sigma}$ be such that $T\|\nabla h\|_{\rho,\sigma} \le 1$. Defining $\sigma_1 = \min\{\frac{\rho-r}{12}, \sigma\}$ as before, we consider the admissible change of variables $x = \Psi(y,t,\varepsilon)$ for $(y,t,\varepsilon) \in G_r \times \mathbb{R} \times [0,\sigma_1]$ according to Lemma 6.5. Then, for every $\varepsilon \in [0,\sigma_1]$, the $T$-periodic Hamiltonian system*

$$\dot{x} = \varepsilon J \nabla_x H(x,t,\varepsilon) \tag{6.8}$$

*is transformed (pulled back via $\Psi$) into*

$$\dot{y} = \varepsilon J \nabla_y K(y,t,\varepsilon), \tag{6.9}$$

*where*

$$K(y,t,\varepsilon) = H(\Psi(y,t,\varepsilon),t,\varepsilon) - h(q(y,t,\varepsilon),p_1,t,\varepsilon); \tag{6.10}$$

*recall that $y = (q_1,p_1)$, $\Psi = (\Psi_1,\Psi_2)$ with $\Psi_1 = q$ and $\Psi_2 = p$. Moreover, $K \in \mathcal{H}_{r,\sigma_1}$ and*

$$\|\nabla K\|_{r,\sigma_1} \le 3\,\|\nabla H\|_{\rho,\sigma} + \frac{5}{2}\,\|\nabla h\|_{\rho,\sigma}. \tag{6.11}$$

**Proof:** Given a Hamiltonian system $\dot{x} = J\nabla_x H(x,t)$ and a change of variables $x = \Psi(y,t)$ that is induced by a generating function of the type $S = S(q,p_1,t)$, the pull-back of the system is $\dot{y} = J\nabla_y K(y,t)$, where

$$K(y,t) = H(\Psi(y,t),t) + \frac{\partial S}{\partial t}(q(y,t),p_1,t).$$

This is part of the classical theory of non-autonomous Hamiltonian systems, cf. [1]. It was known early on, see [12, pp. 13-16] for an elegant exposition. In our case a generating function $S$ of $\Psi$ is given in (6.7), and the formula (6.10) follows.

To show that $K \in \mathcal{H}_{r,\sigma_1}$ we differentiate (6.10) to obtain

$$\nabla K = (D\Psi)^* \nabla H - \frac{\partial h}{\partial q}\,\nabla q - \frac{\partial h}{\partial p_1}\,(0,1)^*.$$

From (6.6) we know that $|D\Psi(y) - I| \le 1/4$ for $y \in G_{R_1} \supset G_r$, which in turn yields $|D\Psi(y)| \le 5/4$. In particular, also $|\nabla q| \le 5/4$, dropping the arguments. Therefore

$$|\nabla K| \le 2|D\Psi||\nabla H| + \left|\frac{\partial h}{\partial q}\right||\nabla q| + \left|\frac{\partial h}{\partial p_1}\right| \le \frac{5}{2}\,|\nabla H| + \frac{5}{4}\left|\frac{\partial h}{\partial q}\right| + \left|\frac{\partial h}{\partial p_1}\right|$$

leads to (6.11). □

Given $H \in \mathcal{H}_{\rho,\sigma}$ we define the function

$$\bar{H}(x,\varepsilon) = \frac{1}{T}\int_0^T H(x,t,\varepsilon)\,dt \quad \text{and} \quad \tilde{H} = H - \bar{H}.$$

Then $\bar{H} \in \mathcal{H}_{\rho,\sigma}$ is autonomous and $\tilde{H} \in \widetilde{\mathcal{H}}_{\rho,\sigma}$. Moreover, we have the bounds

$$\|\nabla\bar{H}\|_{\rho,\sigma} \le \|\nabla H\|_{\rho,\sigma} \quad \text{and} \quad \|\nabla\tilde{H}\|_{\rho,\sigma} \le 2\|\nabla H\|_{\rho,\sigma}. \tag{6.12}$$

**Lemma 6.10** *For $0 < r < \rho$ and $\sigma > 0$ given, let $H \in \mathcal{H}_{\rho,\sigma}$ be such that $T\|\nabla\tilde{H}\|_{\rho,\sigma} \le 1$. We apply Lemma 6.9 with $h = \tilde{H}$. Then the admissible change of variables $\Psi : y \mapsto x$ and the new Hamiltonian function $K$ satisfy*

$$\begin{aligned} \|\Psi(\cdot,\cdot,\varepsilon) - I\|_r &\le \varepsilon T\|\nabla\tilde{H}\|_{\rho,\sigma}, & (6.13)\\ \|\bar{K} - \bar{H}\|_{r,\sigma_1} &\le 2\varepsilon T\,\|\nabla\tilde{H}\|_{\rho,\sigma}\,(\|\nabla\bar{H}\|_{\rho,\sigma} + \|\nabla\tilde{H}\|_{\rho,\sigma}), & (6.14)\\ \|\tilde{K}\|_{r,\sigma_1} &\le 4\varepsilon T\,\|\nabla\tilde{H}\|_{\rho,\sigma}\,(\|\nabla\bar{H}\|_{\rho,\sigma} + \|\nabla\tilde{H}\|_{\rho,\sigma}), & (6.15)\end{aligned}$$

*for $\varepsilon \in [0,\sigma_1]$.*

**Proof:** The first estimate (6.13) is a direct consequence of (6.3). To derive (6.14), we rewrite (6.10) in the form

$$K(y,t,\varepsilon)-\bar{H}(y,\varepsilon) = \bar{H}(\Psi(y,t,\varepsilon),\varepsilon)-\bar{H}(y,\varepsilon)+\tilde{H}(\Psi(y,t,\varepsilon),t,\varepsilon)-\tilde{H}(q(y,t,\varepsilon),p_1,t,\varepsilon). \tag{6.16}$$

Since $q = \Psi_1$ is just a coordinate of $\Psi$,

$$|\Psi(y,t,\varepsilon) - (q(y,t,\varepsilon),p_1)| = |p(y,t,\varepsilon) - p_1| \le |\Psi(y,t,\varepsilon) - y|,$$

and hence it follows from (6.16) and (6.13) that

$$\begin{aligned} \|K - \bar{H}\|_{r,\sigma_1} &\le 2\|\nabla\bar{H}\|_{\rho,\sigma}\,\|\Psi - I\|_{r,\sigma_1} + 2\|\nabla\tilde{H}\|_{\rho,\sigma}\,\|\Psi - I\|_{r,\sigma_1}\\ &\le 2\varepsilon T\|\nabla\tilde{H}\|_{\rho,\sigma}(\|\nabla\bar{H}\|_{\rho,\sigma} + \|\nabla\tilde{H}\|_{\rho,\sigma}). \qquad (6.17)\end{aligned}$$

Observing that

$$\begin{aligned} \|\bar{K} - \bar{H}\|_{r,\sigma_1} &= \sup\left\{\left|\frac{1}{T}\int_0^T [K(y,t,\varepsilon) - \bar{H}(y,\varepsilon)]\,dt\right| : y \in G_r, \varepsilon \in [0,\sigma_1]\right\}\\ &\le \sup\{|K(y,t,\varepsilon) - \bar{H}(y,\varepsilon)| : y \in G_r, t \in \mathbb{R}, \varepsilon \in [0,\sigma_1]\}\\ &= \|K - \bar{H}\|_{r,\sigma_1}, \qquad (6.18)\end{aligned}$$

(6.14) is a consequence of (6.17). Concerning (6.15), it suffices to write $\tilde{K} = (K-\bar{H})+(\bar{H}-\bar{K})$ and to use (6.18) as well as (6.17). □

For the next result we are going to apply Lemma 6.10 $N$ times.

**Lemma 6.11** *Let $0 < r < \rho$ and $\sigma > 0$ be given. For every integer $N \geq 1$ and $H \in \mathcal{H}_{\rho,\sigma}$ so that $T\|\nabla H\|_{\rho,\sigma} \leq 1/2$ there exists an admissible change of variables $x = \Psi_N(y,t,\varepsilon)$, which is defined on $G_r \times \mathbb{R} \times [0,\sigma_N]$ for*

$$\sigma_N = \min\left\{\frac{\rho - r}{72N}, \sigma\right\},$$

*and which satisfies $\Psi(G_r, t, \varepsilon) \subset G_\rho$ for $t \in \mathbb{R}$ and $\varepsilon \in [0,\sigma_N]$. Furthermore,*

$$\dot{x} = \varepsilon J \nabla_x H(x,t,\varepsilon) \tag{6.19}$$

*is transformed (pulled back via $\Psi_N$) into*

$$\dot{y} = \varepsilon J \nabla_y H_N(y,t,\varepsilon) \tag{6.20}$$

*for $H_N \in \mathcal{H}_{r,\sigma_N}$, and moreover we have*

$$\begin{aligned}
\|\Psi_N(\cdot,\cdot,\varepsilon) - I\|_r &\leq 2\varepsilon, &(6.21)\\
\|\nabla \tilde{H}_N\|_{r,\sigma_N} &\leq \left(\frac{1}{T}\right) 2^{-N}, &(6.22)\\
\|\nabla \bar{H}_N\|_{r,\sigma_N} &\leq \frac{3}{2T}, &(6.23)\\
|\bar{H}_N(y,\varepsilon) - \bar{H}_N(y,0)| &\leq \frac{24}{T}\varepsilon + |\bar{H}(y,\varepsilon) - \bar{H}(y,0)|, &(6.24)
\end{aligned}$$

*for $y \in G_r$ and $\varepsilon \in [0,\sigma_N]$.*

First we are going to state an auxiliary result that will be useful in the proof of this lemma.

**Lemma 6.12** *Let $(b_k)_{0\leq k\leq K}$ and $(c_k)_{0\leq k\leq K}$ for some $K \in \mathbb{N} \cup \{\infty\}$ be sequences of positive numbers such that*

$$b_k \leq \alpha b_{k-1}(b_{k-1} + c_{k-1}) \quad \text{and} \quad c_k \leq b_k + c_{k-1}$$

*for $1 \leq k \leq K$, where $\alpha > 0$ is such that $4\alpha(b_0 + c_0) \leq 1$. Then*

$$b_k \leq \frac{1}{2^k} b_0 \quad \text{and} \quad c_k \leq b_0 + c_0$$

*for $0 \leq k \leq K$.*

**Proof:** We check that $b_k \leq 2^{-k} b_0$ and $c_k \leq b_0 \sum_{j=1}^{k} 2^{-j} + c_0$ by induction. Clearly this holds for $k = 0$. For the induction step, by hypothesis we have

$$b_{k+1} \leq \alpha b_k(b_k + c_k) \leq \alpha b_0 2^{-k}(b_0 2^{-k} + b_0 + c_0) \leq 2\alpha b_0 2^{-k}(b_0 + c_0) \leq b_0 2^{-(k+1)},$$

and hence in particular

$$c_{k+1} \leq b_{k+1} + c_k \leq b_0 2^{-(k+1)} + b_0 \sum_{j=1}^{k} 2^{-j} + c_0 = b_0 \sum_{j=1}^{k+1} 2^{-j} + c_0,$$

which completes the argument. ∎

**Proof of Lemma 6.11:** We introduce a uniform partition of the interval $[r,\rho]$ by

$$\rho_N = r < \rho_{N-1} < \ldots < \rho_1 < \rho_0 = \rho,$$

where $\rho_k - \rho_{k+1} = \frac{\rho-r}{N}$ for $k = 0,\ldots,N-1$. The midpoint of $[\rho_{k+1},\rho_k]$ will be denoted by $r_{k+1}$, so that $\rho_k - r_{k+1} = r_{k+1} - \rho_{k+1} = \frac{\rho-r}{2N}$.

Set $H_0 = H$ and observe that $T\|\nabla\tilde{H}_0\|_{\rho,\sigma} = T\|\nabla\tilde{H}\|_{\rho,\sigma} \le 2T\|\nabla H\|_{\rho,\sigma} \le 1$ by (6.12) and by assumption. Hence we can apply Lemma 6.10 for $r$ replaced by $r_1$ to obtain an admissible change of variables $\Psi^{(1)}$ that is defined on $G_{r_1} \times \mathbb{R} \times [0,\hat{\sigma}_1]$ and takes values in $G_\rho$, where $\hat{\sigma}_1 = \min\{\frac{\rho-r_1}{12},\sigma\} = \min\{\frac{\rho-r}{24N},\sigma\}$. The transformed Hamiltonian is denoted by $H_1 \in \mathcal{H}_{r_1,\hat{\sigma}_1}$, and from (6.13)–(6.15) we have the bounds

$$\begin{aligned}
\|\Psi^{(1)}(\cdot,\cdot,\varepsilon) - I\|_{r_1} &\le \varepsilon T\|\nabla\tilde{H}_0\|_{\rho,\sigma}, &(6.25)\\
\|\bar{H}_1 - \bar{H}_0\|_{r_1,\hat{\sigma}_1} &\le 2\varepsilon T\,\|\nabla\tilde{H}_0\|_{\rho,\sigma}\,(\|\nabla\bar{H}_0\|_{\rho,\sigma} + \|\nabla\tilde{H}_0\|_{\rho,\sigma}), &(6.26)\\
\|\tilde{H}_1\|_{r_1,\hat{\sigma}_1} &\le 4\varepsilon T\,\|\nabla\tilde{H}_0\|_{\rho,\sigma}\,(\|\nabla\bar{H}_0\|_{\rho,\sigma} + \|\nabla\tilde{H}_0\|_{\rho,\sigma}), &(6.27)
\end{aligned}$$

for $\varepsilon \in [0,\hat{\sigma}_1]$. Since $\sigma_N \le \hat{\sigma}_1$, we may replace $\hat{\sigma}_1$ by $\sigma_N$ in all of the above. Next we are going to derive some preliminary estimates on $H_0$ and $H_1$. Let

$$b_0 = \|\nabla\tilde{H}_0\|_{\rho,\sigma} \quad \text{and} \quad c_0 = \|\nabla\bar{H}_0\|_{\rho,\sigma}$$

as well as

$$b_1 = \|\nabla\bar{H}_1 - \nabla\bar{H}_0\|_{\rho_1,\sigma_N} + \|\nabla\tilde{H}_1\|_{\rho_1,\sigma_N} \quad \text{and} \quad c_1 = \|\nabla\bar{H}_1\|_{\rho_1,\sigma_N}.$$

Note that by (6.12),

$$b_0 \le 2\|\nabla H\|_{\rho,\sigma} \le \frac{1}{T} \quad \text{and} \quad c_0 \le \|\nabla H\|_{\rho,\sigma} \le \frac{1}{2T}. \tag{6.28}$$

Furthermore,

$$b_k \le \frac{12N\sigma_N T}{\rho - r}\, b_{k-1}(b_{k-1} + c_{k-1}) \quad \text{and} \quad c_k \le b_k + c_{k-1} \tag{6.29}$$

are verified for $k = 1$. To establish this claim, note that by the Cauchy integral formula, (6.26) and (6.27),

$$\begin{aligned}
b_1 &= \|\nabla\bar{H}_1 - \nabla\bar{H}_0\|_{\rho_1,\sigma_N} + \|\nabla\tilde{H}_1\|_{\rho_1,\sigma_N} \le \frac{1}{r_1 - \rho_1}\Big(\|\bar{H}_1 - \bar{H}_0\|_{r_1,\sigma_N} + \|\tilde{H}_1\|_{r_1,\sigma_N}\Big)\\
&\le \frac{2N}{\rho - r}\,(2\sigma_N T + 4\sigma_N T)\,\|\nabla\tilde{H}_0\|_{\rho,\sigma}\,(\|\nabla\bar{H}_0\|_{\rho,\sigma} + \|\nabla\tilde{H}_0\|_{\rho,\sigma}) = \frac{12N\sigma_N T}{\rho - r}\, b_0(b_0 + c_0).
\end{aligned} \tag{6.30}$$

Concerning the bound on $c_1$, we have

$$c_1 = \|\nabla\bar{H}_1\|_{\rho_1,\sigma_N} \le \|\nabla\bar{H}_1 - \nabla\bar{H}_0\|_{\rho_1,\sigma_N} + \|\nabla\bar{H}_0\|_{\rho_1,\sigma_N} \le b_1 + c_0. \tag{6.31}$$

We are going to prove that this process can be repeated $N$ times, if we consider the sequence of nested domains

$$G_r = G_{\rho_N} \subset G_{r_N} \subset G_{\rho_{N-1}} \subset \ldots \subset G_{r_2} \subset G_{\rho_1} \subset G_{r_1} \subset G_{\rho_0} = G_\rho.$$

We will find a sequence $\Psi^{(k)}$, $k = 1, \ldots, N$, of admissible changes of variables sending the set $G_{r_k} \times \mathbb{R} \times [0, \sigma_N]$ into $G_{\rho_{k-1}}$. These changes of variable $\Psi^{(k)}$ and Hamiltonian functions $H_k \in \mathcal{H}_{r_k,\sigma_N} \subset \mathcal{H}_{\rho_k,\sigma_N}$ will be constructed by finite induction w.r.t. to $k \in \{1, \ldots, N\}$. Suppose that $\Psi^{(1)}, \ldots, \Psi^{(k)}$ and $H_1, \ldots, H_k$ have already been obtained, with the additional property that (6.29) holds, where

$$b_k = \|\nabla \bar{H}_k - \nabla \bar{H}_{k-1}\|_{\rho_k,\sigma_N} + \|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N} \quad \text{and} \quad c_k = \|\nabla \bar{H}_k\|_{\rho_k,\sigma_N}$$

for $k \geq 1$. With $\alpha = \frac{12N\sigma_N T}{\rho - r}$ we note that

$$4\alpha(b_0 + c_0) \leq 1,$$

since by (6.12) and our hypotheses

$$\frac{48N\sigma_N T}{\rho - r} (\|\nabla \tilde{H}_0\|_{\rho,\sigma} + \|\nabla \bar{H}_0\|_{\rho,\sigma}) \leq \frac{144N\sigma_N T}{\rho - r} \|\nabla H\|_{\rho,\sigma} \leq \frac{72N\sigma_N}{\rho - r} \leq 1.$$

Hence Lemma 6.12 applies to yield $b_k \leq 2^{-k} b_0$ and $c_k \leq b_0 + c_0$. In particular, it follows from (6.28) that

$$T\|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N} \leq T b_k \leq T b_0 \leq 1,$$

and Lemma 6.10 is applicable, for $r$ replaced by $r_{k+1}$ and $\sigma$ replaced by $\sigma_N$. The resulting admissible change of variables $\Psi^{(k+1)}$ is defined on $G_{r_{k+1}} \times \mathbb{R} \times [0, \hat{\sigma}_{k+1}]$ and takes values in $G_{\rho_k}$, where $\hat{\sigma}_{k+1} = \min\{\frac{\rho_k - r_{k+1}}{12}, \sigma_N\} = \min\{\frac{\rho - r}{24N}, \sigma_N\} = \sigma_N$. The transformed Hamiltonian is denoted by $H_{k+1} \in \mathcal{H}_{r_{k+1},\sigma_N}$, and from (6.13)–(6.15) we deduce the bounds

$$\|\Psi^{(k+1)}(\cdot, \cdot, \varepsilon) - I\|_{r_{k+1}} \leq \varepsilon T \|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N}, \tag{6.32}$$

$$\|\bar{H}_{k+1} - \bar{H}_k\|_{r_{k+1},\sigma_N} \leq 2\varepsilon T \,\|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N} (\|\nabla \bar{H}_k\|_{\rho_k,\sigma_N} + \|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N}), \tag{6.33}$$

$$\|\tilde{H}_{k+1}\|_{r_{k+1},\sigma_N} \leq 4\varepsilon T \,\|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N} (\|\nabla \bar{H}_k\|_{\rho_k,\sigma_N} + \|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N}), \tag{6.34}$$

for $\varepsilon \in [0, \sigma_N]$. Analogously to (6.30) and (6.31), it follows from the Cauchy integral formula in conjunction with (6.33) and (6.34) that (6.29) holds for $k+1$. Therefore the inductive process to obtain the $\Psi^{(k)}$ and $H_k$ can be completed up to $k = N$.

For the estimate (6.22), note that by (6.29) and (6.28)

$$\|\nabla \tilde{H}_N\|_{r,\sigma_N} = \|\nabla \tilde{H}_N\|_{\rho_N,\sigma_N} \leq b_N \leq 2^{-N} b_0 \leq \frac{1}{T} 2^{-N}.$$

The bound (6.23) is also a consequence of (6.29) and (6.28), since

$$\|\nabla \bar{H}_N\|_{r,\sigma_N} = \|\nabla \bar{H}_N\|_{\rho_N,\sigma_N} = c_N \leq b_0 + c_0 \leq \frac{3}{2T}.$$

The desired admissible change of variables $\Psi_N$ is defined as the composition

$$\Psi_N = \Psi^{(1)} \circ \Psi^{(2)} \circ \ldots \circ \Psi^{(N)},$$

which is defined on $G_r \times \mathbb{R} \times [0, \sigma_N]$ and takes values in $G_\rho$. To obtain (6.21), we are going to use the formula

$$\Psi_N(\cdot,\cdot,\varepsilon) - I = \sum_{k=1}^{N-1} [(\Psi^{(k)} - I) \circ \Psi^{(k+1)} \circ \ldots \circ \Psi^{(N)}](\cdot,\cdot,\varepsilon) + (\Psi^{(N)}(\cdot,\cdot,\varepsilon) - I). \qquad (6.35)$$

For $k \geq 1$ the composition $\Psi^{(k+1)} \circ \ldots \circ \Psi^{(N)}$ maps $G_r$ into $G_{\rho_k} \subset G_{r_k}$ in $y$. Therefore due to (6.32), with $k+1$ replaced by $k$, and using (6.28),

$$\begin{aligned} \|[(\Psi^{(k)} - I) \circ \Psi^{(k+1)} \circ \ldots \circ \Psi^{(N)}](\cdot,\cdot,\varepsilon)\|_r &\leq \|\Psi^{(k)}(\cdot,\cdot,\varepsilon) - I\|_{r_k} \leq \varepsilon T \|\nabla \tilde{H}_{k-1}\|_{\rho_{k-1},\sigma_N} \\ &\leq \varepsilon T b_{k-1} \leq \varepsilon T 2^{-(k-1)} b_0 \leq \varepsilon 2^{-(k-1)}. \end{aligned}$$

Analogously,

$$\|\Psi^{(N)}(\cdot,\cdot,\varepsilon) - I\|_r = \|\Psi^{(N)}(\cdot,\cdot,\varepsilon) - I\|_{\rho_N} \leq \|\Psi^{(N)}(\cdot,\cdot,\varepsilon) - I\|_{r_N} \leq \varepsilon 2^{-(N-1)}.$$

Using (6.35), the foregoing estimates in turn lead to

$$\begin{aligned} \|\Psi_N(\cdot,\cdot,\varepsilon) - I\|_r &\leq \sum_{k=1}^{N-1} \|[(\Psi^{(k)} - I) \circ \Psi^{(k+1)} \circ \ldots \circ \Psi^{(N)}](\cdot,\cdot,\varepsilon)\|_r + \|\Psi^{(N)}(\cdot,\cdot,\varepsilon) - I\|_r \\ &\leq \sum_{k=1}^{N-1} \varepsilon 2^{-(k-1)} + \varepsilon 2^{-(N-1)} \leq 2\varepsilon, \end{aligned}$$

which is (6.21). To prove (6.24), we first note that by (6.33) and (6.28),

$$\begin{aligned} \|\bar{H}_{k+1} - \bar{H}_k\|_{r,\sigma_N} &\leq \|\bar{H}_{k+1} - \bar{H}_k\|_{r_{k+1},\sigma_N} \\ &\leq 2\varepsilon T \, \|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N} \, (\|\nabla \bar{H}_k\|_{\rho_k,\sigma_N} + \|\nabla \tilde{H}_k\|_{\rho_k,\sigma_N}) \\ &\leq 2\varepsilon T b_k (b_k + c_k) \\ &\leq 2\varepsilon T 2^{-k} b_0 (2^{-k} b_0 + b_0 + c_0) \\ &\leq 2^{-k} \left(\frac{6}{T}\right) \varepsilon. \end{aligned}$$

For $y \in G_r$ and $\varepsilon \in [0, \sigma_N]$ it hence follows that

$$\begin{aligned} &|\bar{H}_N(y,\varepsilon) - \bar{H}_N(y,0)| \\ &\quad = \left| \sum_{k=0}^{N-1} (\bar{H}_{k+1}(y,\varepsilon) - \bar{H}_k(y,\varepsilon)) - \sum_{k=0}^{N-1} (\bar{H}_{k+1}(y,0) - \bar{H}_k(y,0)) + (\bar{H}_0(y,\varepsilon) - \bar{H}_0(y,0)) \right| \\ &\quad \leq 2 \sum_{k=0}^{N-1} \|\bar{H}_{k+1} - \bar{H}_k\|_{r,\sigma_N} + |\bar{H}(y,\varepsilon) - \bar{H}(y,0)| \\ &\quad \leq \frac{12}{T} \varepsilon \sum_{k=0}^{N-1} 2^{-k} + |\bar{H}(y,\varepsilon) - \bar{H}(y,0)| \leq \frac{24}{T} \varepsilon + |\bar{H}(y,\varepsilon) - \bar{H}(y,0)|. \end{aligned}$$

This completes the proof of Lemma 6.11. □

Now we are in a position to derive the "Hamiltonian normal form" with exponentially small remainder. For our particular domain $G = \mathbb{R} \times I$, this is essentially the result that is announced in [11, Remark 2, p. 134]. To prepare for the statement, we need to introduce a more relaxed class of transformations, as compared to Definition 6.4.

**Definition 6.13** *Let $0 < \rho_1 \le \rho$ and $0 < \sigma_1 \le \sigma$. A map $\Psi : G_{\rho_1} \times \mathbb{R} \times [0, \sigma_1] \to \mathbb{C}^2$, $x = \Psi(y, t, \varepsilon)$, will be called a change of variables, if it satisfies*

- *(a) $\Psi$ maps reals into reals;*
- *(b) $\Psi$ is $T$-periodic in $t$ and $\Psi(y, 0, \varepsilon) = \Psi(y, T, \varepsilon) = y$;*
- *(c) for every $\varepsilon \in [0, \sigma_1]$ the map $\Psi(\cdot, \cdot, \varepsilon)$ is $C^1$ in the real sense, and for every $t \in \mathbb{R}$ and $\varepsilon \in [0, \sigma_1]$ the map $\Psi(\cdot, t, \varepsilon)$ is holomorphic in $G_{\rho_1}$; and*
- *(d) for every $t \in \mathbb{R}$ and $\varepsilon \in [0, \sigma_1]$ the map $\Psi(\cdot, t, \varepsilon)$ is a symplectic diffeomorphism from $G_{\rho_1}$ onto its image.*

Note that we are not assuming any property of continuous dependence w.r. to the parameter $\varepsilon$. This is in contrast to the previous notion of an admissible change of variables, introduced in Definition 6.4.

**Theorem 6.14** *For $0 < r < \rho$ and $\sigma > 0$ given, let $H \in \mathcal{H}_{\rho,\sigma}$. Then there exist $C, D > 0$ (depending upon $T, r, \rho, \|\nabla H\|_{\rho,\sigma}$) with the following properties. There is a change of variables $x = \Psi(y, t, \varepsilon)$, which is defined on $G_r \times \mathbb{R} \times [0, \sigma]$ and which satisfies $\Psi(G_r, t, \varepsilon) \subset G_\rho$ for $t \in \mathbb{R}$ and $\varepsilon \in [0, \sigma]$, such that*

$$\dot{x} = \varepsilon J \nabla_x H(x, t, \varepsilon) \tag{6.36}$$

*is transformed (pulled back via $\Psi$) into*

$$\dot{y} = \varepsilon (J \nabla_y \mathcal{N}(y, \varepsilon) + J \nabla_y \mathcal{R}(y, t, \varepsilon)), \tag{6.37}$$

*for functions $\mathcal{N} \in \mathcal{H}_{r,\sigma}$ and $\mathcal{R} \in \widetilde{\mathcal{H}}_{r,\sigma}$. Furthermore,*

$$\begin{aligned}
\|\Psi(\cdot, \cdot, \varepsilon) - I\|_r &\le C\varepsilon, && (6.38)\\
\|\nabla_y \mathcal{N}(\cdot, \varepsilon)\|_r &\le C, && (6.39)\\
\|\nabla_y \mathcal{R}(\cdot, \cdot, \varepsilon)\|_r &\le C e^{-D/\varepsilon}, && (6.40)\\
|\mathcal{N}(y, \varepsilon) - \mathcal{N}(y, 0)| &\le C\varepsilon + \|\bar{H}(\cdot, \varepsilon) - \bar{H}(\cdot, 0)\|_\rho, && (6.41)
\end{aligned}$$

*for $y \in G_r$ and $\varepsilon \in [0, \sigma]$. In addition,*

$$\mathcal{N}(y, 0) = \bar{H}(y, 0). \tag{6.42}$$

**Proof:** We are going to show that

$$C = \max\left\{2\lambda, \frac{2\lambda}{T}, \frac{24\lambda^2}{T}\right\} \quad \text{and} \quad D = \frac{\rho - r}{144\lambda}$$

have the asserted properties, where $\lambda = 2T\|\nabla H\|_{\rho,\sigma}$. The cases $H = 0$ or $\varepsilon = 0$ are trivial, so in particular we may assume that $\lambda > 0$. We rewrite (6.36) as $\dot{x} = \hat{\varepsilon} J \nabla_x \hat{H}(x, t, \hat{\varepsilon})$, where $\hat{\varepsilon} = \lambda\varepsilon \in [0, \hat{\sigma}]$ for $\hat{\sigma} = \lambda\sigma$ and $\hat{H}(x, t, \hat{\varepsilon}) = \lambda^{-1} H(x, t, \lambda^{-1}\hat{\varepsilon})$. It follows that $\hat{H} \in \mathcal{H}_{\rho,\hat{\sigma}}$ satisfies $2T\|\nabla \hat{H}\|_{\rho,\hat{\sigma}} = 2T\lambda^{-1}\|\nabla H\|_{\rho,\sigma} = 1$. Thus we may apply Lemma 6.11 to $\hat{H}$ and with

$$N = \left[\frac{\rho - r}{72\lambda\varepsilon}\right].$$

Hence there exists an admissible change of variables $x = \hat{\Psi}(y, t, \hat{\varepsilon})$, which is defined on $G_r \times \mathbb{R} \times [0, \hat{\sigma}_N]$ for

$$\hat{\sigma}_N = \min\left\{\frac{\rho - r}{72N}, \hat{\sigma}\right\},$$

and which satisfies $\hat{\Psi}(G_r, t, \varepsilon) \subset G_\rho$ for $t \in \mathbb{R}$ and $\varepsilon \in [0, \hat{\sigma}_N]$. Furthermore, $\dot{x} = \hat{\varepsilon} J \nabla_x \hat{H}(x, t, \hat{\varepsilon})$ is transformed into $\dot{y} = \hat{\varepsilon} J \nabla_y K(y, t, \hat{\varepsilon})$ for $K \in \mathcal{H}_{r,\hat{\sigma}_N}$, and in addition we have

$$\|\hat{\Psi}(\cdot, \cdot, \hat{\varepsilon}) - I\|_r \le 2\hat{\varepsilon}, \quad \|\nabla \tilde{K}\|_{r,\hat{\sigma}_N} \le \left(\frac{1}{T}\right) 2^{-N}, \quad \|\nabla \bar{K}\|_{r,\hat{\sigma}_N} \le \frac{3}{2T},$$

$$|\bar{K}(y, \hat{\varepsilon}) - \bar{K}(y, 0)| \le \frac{24}{T}\,\hat{\varepsilon} + |\overline{\hat{H}}(y, \hat{\varepsilon}) - \overline{\hat{H}}(y, 0)|,$$

for $y \in G_r$ and $\hat{\varepsilon} \in ]0, \hat{\sigma}_N]$. Define

$$\Psi(y, t, \varepsilon) = \hat{\Psi}(y, t, \lambda\varepsilon), \quad \mathcal{N}(y, \varepsilon) = \lambda \bar{K}(y, \lambda\varepsilon) \quad \text{and} \quad \mathcal{R}(y, t, \varepsilon) = \lambda \tilde{K}(y, t, \lambda\varepsilon)$$

for $y \in G_r$, $t \in \mathbb{R}$ and $\varepsilon \in ]0, \sigma]$. We also put $\Psi = I$ for $\varepsilon = 0$. If $\varepsilon \in ]0, \sigma]$, then $\hat{\varepsilon} = \lambda\varepsilon \le \lambda\sigma = \hat{\sigma}$ and moreover

$$\hat{\varepsilon} = \frac{1}{N}\,\lambda\varepsilon N \le \frac{1}{N}\,\lambda\varepsilon\left(\frac{\rho - r}{72\lambda\varepsilon}\right) = \frac{\rho - r}{72N},$$

so that $\hat{\varepsilon} \in ]0, \sigma_N]$. Accordingly, the first few claims are straightforwardly verified; this includes (6.38), (6.39) and (6.41). Concerning (6.40), we use the above estimate on $\nabla \tilde{K}$ to get for $\varepsilon \in ]0, \sigma]$

$$\begin{aligned} \|\nabla_y \mathcal{R}(\cdot, \cdot, \varepsilon)\|_r &= \lambda \|\nabla \tilde{K}(\cdot, \cdot, \hat{\varepsilon})\|_r \le \left(\frac{\lambda}{T}\right) 2^{-N} = \left(\frac{2\lambda}{T}\right) 2^{-(N+1)} \le \left(\frac{2\lambda}{T}\right) 2^{-\frac{\rho - r}{72\lambda\varepsilon}} \\ &= \left(\frac{2\lambda}{T}\right) 4^{-\frac{\rho - r}{144\lambda\varepsilon}} \le \left(\frac{2\lambda}{T}\right) e^{-\frac{\rho - r}{144\lambda\varepsilon}}, \end{aligned}$$

which completes the proof of (6.38)–(6.41).

Finally, with regard to (6.42), we observe that in all the previous lemmas we have $\Psi = I$ for $\varepsilon = 0$. Then we can define $\mathcal{N}(y, 0) = \bar{H}(y, 0)$, since $\bar{H}_k(\cdot, 0) = \bar{H}(\cdot, 0)$ for each $k$ throughout the iteration. □

**Corollary 6.15** *Under the assumptions of Theorem 6.14 let $0 < \hat{r} < r < \rho$ and denote by $\Psi : y \mapsto x$ the change of variables that has been constructed there. Let $\sigma_* = \min\{\frac{r - \hat{r}}{12C}, \sigma\}$. If $t \in \mathbb{R}$ and $\varepsilon \in [0, \sigma_*]$, then $\Psi(G_r, t, \varepsilon) \supset G_{\hat{r}}$.*

**Proof:** The argument is similar to the one for Corollary 6.7. □

# 7 Appendix II: Application to maps

Here we will prove Theorem 2.5, following the approach outlined in [11]. First we realize the map $P_\varepsilon$ as the Poincaré map of a periodic Hamiltonian system and then we are going to apply the previous results from Section 6; see [18, p. 13/14] for general information and additional references in a more abstract context.

We start with an auxiliary result on the construction of a Hamiltonian function from an exact symplectic isotopy.

**Lemma 7.1** *Assume that $\Phi : G \times [0,1] \to \mathbb{R}^2$ is $C^\infty$ and that $\Phi(\cdot,t) : G \to G(t) = \Phi(G,t)$ is a diffeomorphism for every $t \in [0,1]$. The inverse map is denoted by $\Psi(\cdot,t)$ and we will also write*

$$X = \Phi(x,t), \quad x = \Psi(X,t), \quad x = (q,p), \quad X = (Q,P), \quad \Phi = (\mathcal{F},\mathcal{G}).$$

*Assume that*

$$P\,dQ - p\,dq = d\eta(\cdot,t) \tag{7.1}$$

*for a $C^\infty$-function $\eta : G \times [0,1] \to \mathbb{R}$. Then*

$$J\nabla h_{\rm aux}(X,t) = \frac{\partial \Phi}{\partial t}(\Psi(X,t),t), \tag{7.2}$$

*where*

$$h_{\rm aux}(X,t) = \frac{\partial \mathcal{F}}{\partial t}(\Psi(X,t),t)\,\mathcal{G}(\Psi(X,t),t) - \frac{\partial \eta}{\partial t}(\Psi(X,t),t) \tag{7.3}$$

*is defined on*

$$\mathcal{D} = \{(X,t) : t \in [0,1], X \in G(t)\}.$$

**Remark 7.2** (a) Note that $G(t) \subset \mathbb{R}^2$ is open and $\mathcal{D}$ is diffeomorphic to $G \times [0,1]$ via the map $(x,t) \mapsto (\Phi(x,t),t)$. Moreover, $X(t) = \Phi(x,t)$ is a solution to $\dot X(t) = J\nabla h_{\rm aux}(X(t),t)$.
(b) Lemma 7.1 remains valid, if $\Phi$ and $\eta$ are $C^1$, and the cross-derivatives

$$\frac{\partial^2 \Phi}{\partial t\,\partial x} = \frac{\partial^2 \Phi}{\partial x\,\partial t}, \quad \frac{\partial^2 \eta}{\partial t\,\partial x} = \frac{\partial^2 \eta}{\partial x\,\partial t},$$

exist, coincide and are continuous functions of $(x,t)$.
(c) If $\Phi(\cdot,t)$, $\Psi(\cdot,t)$ and $\eta(\cdot,t)$ have holomorphic extensions, then also the identity (7.2) can be extended.
(d) We refer to [10, Thm. 6.2.1] for a similar result.

**Proof of Lemma 7.1 :** The identity (7.1) holds in the space of one-forms on $G$. Differentiating w.r. to $t$, we obtain

$$\frac{\partial \mathcal{G}}{\partial t}\,d\mathcal{F} + \mathcal{G}\,d\Big(\frac{\partial \mathcal{F}}{\partial t}\Big) = d\Big(\frac{\partial \eta}{\partial t}\Big).$$

It follows that

$$d\Big(\frac{\partial \mathcal{F}}{\partial t}\,\mathcal{G} - \frac{\partial \eta}{\partial t}\Big) = \frac{\partial \mathcal{F}}{\partial t}\,d\mathcal{G} + d\Big(\frac{\partial \mathcal{F}}{\partial t}\Big)\,\mathcal{G} - d\Big(\frac{\partial \eta}{\partial t}\Big) = \frac{\partial \mathcal{F}}{\partial t}\,d\mathcal{G} - \frac{\partial \mathcal{G}}{\partial t}\,d\mathcal{F} \tag{7.4}$$

on $G$. To pull back this identity under the map $\Psi(\cdot,t) : G(t) \ni X \mapsto x \in G$, denote $\mathfrak{h}(x,t) = \frac{\partial \mathcal{F}}{\partial t}(x,t)\,\mathcal{G}(x,t) - \frac{\partial \eta}{\partial t}(x,t)$. From (7.4) we thus deduce

$$\begin{aligned} dh_{\text{aux}}(\cdot,t) &= d(\mathfrak{h}\circ\Psi) = d(\Psi^*\mathfrak{h}) = \Psi^*(d\mathfrak{h}) = \Psi^*\Big(\frac{\partial \mathcal{F}}{\partial t}\,d\mathcal{G} - \frac{\partial \mathcal{G}}{\partial t}\,d\mathcal{F}\Big) \\ &= \Big(\frac{\partial \mathcal{F}}{\partial t}\circ\Psi\Big)dP - \Big(\frac{\partial \mathcal{G}}{\partial t}\circ\Psi\Big)dQ, \end{aligned}$$

which is equivalent to (7.2). □

**Lemma 7.3** *Let $G = \mathbb{R}\times I \subset \mathbb{R}^2$ for an open and bounded interval $I \subset \mathbb{R}$. Suppose that $l \in \mathcal{M}_{1,\rho,\sigma}$, and for $\varepsilon \in [0,\sigma]$ consider the family of maps $P_\varepsilon : G_\rho \to \mathbb{C}^2$ given by*

$$P_\varepsilon : \quad x_1 = x + \varepsilon l(x,\varepsilon). \tag{7.5}$$

*Let the family $\{P_\varepsilon\}$ be E-symplectic and fix $0 < r < \hat{r} < \rho$. Then there exist $\hat{\sigma} \in ]0,\sigma[$ and a Hamiltonian $H_{\text{aux}} \in \mathcal{H}_{\hat{r},\hat{\sigma}}$ such that for $\varepsilon \in [0,\hat{\sigma}]$ the Poincaré map (time-1-map) of $\dot{x} = \varepsilon J\nabla H_{\text{aux}}(x,t,\varepsilon)$ is $P_\varepsilon$, restricted to $G_r$. Furthermore, there exists a constant $C_{\text{aux}} > 0$ such that*

$$|H_{\text{aux}}(x,t,\varepsilon) - H_{\text{aux}}(x,t,0)| \le C_{\text{aux}}\,\varepsilon \tag{7.6}$$

*for $x \in G_{\hat{r}}$, $t \in [0,1]$ and $\varepsilon \in [0,\hat{\sigma}]$. The constant $C_{\text{aux}}$ will depend upon $\rho$, $\sigma$, $r$, $\hat{r}$, $\|l\|_{1,\rho,\sigma}$, the interval $I$, $\|h\|_{1,\rho,\sigma}$ and $\sup_{\varepsilon\in]0,\sigma]}\|\varepsilon^{-1}(\frac{\partial h}{\partial \varepsilon}(\cdot,\varepsilon) - \mathfrak{m})\|_\rho$ (cf. the notion of E-symplecticity, Definition 2.3).*

**Proof:** Let $\chi : [0,1] \to [0,1]$ be a strictly increasing $C^\infty$-function such that $\chi(0) = 0$, $\chi(1) = 1$ and $\dot{\chi}(0) = \dot{\chi}(1) = 0$. Define

$$\Phi(x,t,\varepsilon) = x + \varepsilon\chi(t)\,l(x,\varepsilon\chi(t)) \tag{7.7}$$

and

$$\eta(x,t,\varepsilon) = h(x,\varepsilon\chi(t)).$$

For fixed $\varepsilon$ we intend to apply the relaxed version of Lemma 7.1, as outlined in Remark 7.2(b), (c). The condition (7.1) holds, due to (2.4) in Definition 2.3.

Our first aim will be to construct the inverse $\Psi$. Define $r_1 = \frac{1}{2}(\rho + \hat{r})$ and fix $\sigma_1 \in ]0,\sigma]$ so that

$$\sigma_1\|l\|_{1,\rho,\sigma} \le \frac{1}{4}(\rho - r_1) = \frac{1}{8}(\rho - \hat{r}). \tag{7.8}$$

We are going to prove that $\Phi(\cdot,t,\varepsilon)$ is a diffeomorphism from $G_{r_1}$ onto its image, if $t \in [0,1]$ and $\varepsilon \in [0,\sigma_1]$. For $\varepsilon = 0$ we have $\Phi(x,t,\varepsilon) = x$, so we can assume that $\varepsilon > 0$. Using the Cauchy integral formula, one gets

$$\varepsilon\Big\|\frac{\partial l}{\partial x}(\cdot,\varepsilon)\Big\|_{r_1} \le \frac{\sigma_1}{\rho - r_1}\,\|l(\cdot,\varepsilon)\|_\rho \le \frac{\sigma_1}{\rho - r_1}\,\|l\|_{1,\rho,\sigma} \le \frac{1}{4}.$$

Hence the matrix

$$M = \frac{\partial \Phi}{\partial x}(x,t,\varepsilon) = I + \varepsilon\chi(t)\,\frac{\partial l}{\partial x}(x,\varepsilon\chi(t)) \tag{7.9}$$

satisfies $|M - I| \leq \frac{1}{4}$. As a consequence, $M$ has an inverse and therefore $\Phi(\cdot, t, \varepsilon)$ is a local diffeomorphism from $G_{r_1}$ onto its image, which is contained in $G_{\frac{3}{4}r_1+\frac{1}{4}\rho}$, the latter by (7.8). If $x_1, x_2 \in G_{r_1}$, then

$$\begin{aligned} |\Phi(x_1,t,\varepsilon) - \Phi(x_2,t,\varepsilon)| &= \Big|x_1 - x_2 + \varepsilon\chi(t)\Big(\int_0^1 \frac{\partial l}{\partial x}(\lambda x_1 + (1-\lambda)x_2, \varepsilon\chi(t))\, d\lambda\Big)(x_1 - x_2)\Big| \\ &\geq |x_1 - x_2| - \frac{1}{2}|x_1 - x_2| \\ &= \frac{1}{2}|x_1 - x_2|; \end{aligned}$$

note that here the convexity of $G$ (and hence $G_{r_1}$) has been used. It follows that $\Phi(\cdot, t, \varepsilon)$ is one-to-one on $G_{r_1}$ and its inverse $\Psi(\cdot, t, \varepsilon)$ has Lipschitz constant 2. Observe that (7.8) also implies that

$$\sigma_1 \|l\|_{1,\rho,\sigma} \leq \frac{1}{4}(\rho - \hat{r}) = \frac{1}{2}(r_1 - \hat{r}).$$

Arguing analogously to Corollary 6.7, it follows that

$$\Phi(G_{r_1}, t, \varepsilon) \supset G_{\hat{r}}$$

for $t \in [0,1]$ and $\varepsilon \in [0, \sigma_1]$. The Hamiltonian function $h_{\mathrm{aux}}$ from Lemma 7.1 will be defined on the domain

$$\mathcal{D} = \{(X, t, \varepsilon) : t \in [0,1], X \in \Phi(G_{r_1}, t, \varepsilon), \varepsilon \in [0, \sigma_1]\} \supset G_{\hat{r}} \times [0,1] \times [0, \sigma_1]. \tag{7.10}$$

Next we choose the number $\hat{\sigma} \in ]0, \sigma_1]$ so that

$$\hat{\sigma}\|l\|_{1,\rho,\sigma} < \hat{r} - r,$$

which in turn implies that

$$\Phi(G_r, t, \varepsilon) \subset G_{\hat{r}} \tag{7.11}$$

for $t \in [0,1]$ and $\varepsilon \in [0, \hat{\sigma}]$, and moreover we have $\Phi(x, t, \varepsilon) = P_{\varepsilon\chi(t)}(x)$ by definition.

From now on we consider $\Phi$ on $G_{r_1} \times [0,1] \times [0, \hat{\sigma}]$ and the inverse $\Psi(\cdot, t, \varepsilon) = \Phi(\cdot, t, \varepsilon)^{-1}$ has domain $\Phi(G_{r_1}, t, \varepsilon)$. Since $\Phi$ is continuous in its three arguments, the same can be said about $\Psi$. In addition, by the inverse function theorem, $\Psi$ is holomorphic in the first variable. Let $\varepsilon \in [0, \hat{\sigma}]$ be fixed. We will prove that $\Phi(\cdot, \cdot, \varepsilon)$ is $C^1$ in $G_{r_1} \times [0,1]$. Moreover, the cross derivatives do exist, they are continuous and coincide. To see this, we can once again restrict our attention to $\varepsilon > 0$. Since $l \in \mathcal{M}_{1,\rho,\sigma}$, the functions $l(\cdot, \varepsilon)$ and $\frac{\partial l}{\partial \varepsilon}(\cdot, \varepsilon)$ are holomorphic. Hence, by Cauchy's integral formula,

$$\Big\|\frac{\partial l}{\partial x}(\cdot, \varepsilon)\Big\|_{r_1} \leq \frac{1}{\rho - r_1}\|l\|_{1,\rho,\sigma}, \tag{7.12}$$

$$\Big\|\frac{\partial^2 l}{\partial x \partial \varepsilon}(\cdot, \varepsilon)\Big\|_{r_1} \leq \frac{1}{\rho - r_1}\|l\|_{1,\rho,\sigma}. \tag{7.13}$$

Note that in (7.12) the case $\varepsilon = 0$ is admissible. By definition, $\Phi(\cdot, \cdot, \varepsilon)$ is $C^\infty$ in $G_\rho \times ]0, 1]$. For $t = 0$, $\Phi(x, 0, \varepsilon) = x$ and $\frac{\partial \Phi}{\partial x}(x, 0, \varepsilon) = I$. From (7.12) and (7.9) we conclude that $\frac{\partial \Phi}{\partial x}(\cdot, \cdot, \varepsilon)$ is continuous in $G_{r_1} \times [0, 1]$. To analyze the derivative w.r. to $t$, we observe that

$$\frac{\partial \Phi}{\partial t}(x, 0, \varepsilon) = \lim_{t \to 0^+} \frac{\Phi(x, t, \varepsilon) - \Phi(x, 0, \varepsilon)}{t} = \varepsilon \lim_{t \to 0^+} \frac{\chi(t)}{t}\, l(x, \varepsilon\chi(t)) = 0,$$

where we used that $\chi(0) = \dot{\chi}(0) = 0$ and $\|l\|_{\rho,\sigma} < \infty$. For $t > 0$,

$$\frac{\partial \Phi}{\partial t}(x, t, \varepsilon) = \varepsilon \dot{\chi}(t) \Big[ l(x, \varepsilon\chi(t)) + \varepsilon\chi(t) \frac{\partial l}{\partial \varepsilon}(x, \varepsilon\chi(t)) \Big]. \tag{7.14}$$

Thus the continuity of $\frac{\partial \Phi}{\partial t}(\cdot, \cdot, \varepsilon)$ is a consequence of $\|l\|_{1,\rho,\sigma} < \infty$. To summarize, so far we have shown that $\Phi(\cdot, \cdot, \varepsilon)$ is $C^1$ in $G_{r_1} \times [0, 1]$. For the cross derivatives, from $\frac{\partial \Phi}{\partial t}(x, 0, \varepsilon) = 0$ we deduce that $\frac{\partial^2 \Phi}{\partial x \partial t}(x, 0, \varepsilon) = 0$. Also, using (7.9) and (7.12),

$$\frac{\partial^2 \Phi}{\partial t \partial x}(x, 0, \varepsilon) = \lim_{t \to 0^+} \frac{\frac{\partial \Phi}{\partial x}(x, t, \varepsilon) - \frac{\partial \Phi}{\partial x}(x, 0, \varepsilon)}{t} = \varepsilon \lim_{t \to 0^+} \frac{\chi(t)}{t} \frac{\partial l}{\partial x}(x, \varepsilon\chi(t)) = 0.$$

Hence the cross derivatives exist at $t = 0$ and they coincide. The continuity of these derivatives is obtained after differentiating (7.9) w.r. to $t$ in $G_{r_1} \times ]0, 1]$; again the bounds (7.12) and (7.13) need to be used here. Both functions $l$ and $h$ belong to the class $\mathcal{M}_{1,\rho,\sigma}$. Thus the previous discussions also apply to the function $\eta(\cdot, \cdot, \varepsilon)$.

Altogether, we see that the relaxed version of Lemma 7.1 can be used to deduce the existence of a function $h_{\text{aux}} = h_{\text{aux}}(X, t, \varepsilon)$, which is defined on $\mathcal{D}$ from (7.10), with the stated properties. In particular, $h_{\text{aux}}(\cdot, t, \varepsilon)$ is well-defined on $G_{\hat{r}}$. Moreover, if $x \in G_r$, then $X(t) = \Phi(x, t, \varepsilon)$ solves

$$\dot{X}(t) = J \nabla h_{\text{aux}}(X(t), t, \varepsilon) \tag{7.15}$$

by Remark 7.2(a), and also $\varepsilon \in [0, \hat{\sigma}]$ yields $X(t) \in G_{\hat{r}}$ for $t \in [0, 1]$ due to (7.11). The Poincaré map of (7.15) is $G_r \ni x \mapsto \Phi(x, 1, \varepsilon) = x + \varepsilon l(x, \varepsilon) = P_\varepsilon(x)$, i.e., the original map restricted to $G_r$.

To express $h_{\text{aux}}$ more explicitly, we recall from the previous computations that

$$\frac{\partial \Phi}{\partial t}(x, t, \varepsilon) = \begin{cases} \varepsilon \dot{\chi}(t) \left[ l(x, \varepsilon\chi(t)) + \varepsilon\chi(t) \frac{\partial l}{\partial \varepsilon}(x, \varepsilon\chi(t)) \right] & : \quad t \in ]0, 1], \varepsilon \in ]0, \hat{\sigma}] \\ 0 & : \quad t = 0 \text{ or } \varepsilon = 0 \end{cases} \tag{7.16}$$

and similarly

$$\frac{\partial \eta}{\partial t}(x, t, \varepsilon) = \begin{cases} \varepsilon \dot{\chi}(t) \frac{\partial h}{\partial \varepsilon}(x, \varepsilon\chi(t)) & : \quad t \in ]0, 1], \varepsilon \in ]0, \hat{\sigma}] \\ 0 & : \quad t = 0 \text{ or } \varepsilon = 0 \end{cases}. \tag{7.17}$$

In the notation of Lemma 7.1 we have

$$\mathcal{F}(x, t, \varepsilon) = q + \varepsilon\chi(t)\, l_1(x, \varepsilon\chi(t)), \quad \mathcal{G}(x, t, \varepsilon) = p + \varepsilon\chi(t)\, l_2(x, \varepsilon\chi(t)),$$

where $x = (q, p)$, and $l = (l_1, l_2)$ are the components. Also observe that by (7.3)

$$h_{\text{aux}}(X, t, \varepsilon) = \frac{\partial \mathcal{F}}{\partial t}(\Psi(X, t, \varepsilon), t, \varepsilon)\, \mathcal{G}(\Psi(X, t, \varepsilon), t, \varepsilon) - \frac{\partial \eta}{\partial t}(\Psi(X, t, \varepsilon), t, \varepsilon).$$

From $\chi(0) = \dot\chi(0) = \dot\chi(1) = 0$ and (7.17) it follows that $h_{\rm aux}(X, t, \varepsilon) = 0$ for $t = 0$ or $t = 1$ or $\varepsilon = 0$. Moreover, if $t \neq 0$ and $\varepsilon > 0$,

$$\begin{aligned} \frac{1}{\varepsilon}\, h_{\rm aux}(X,t,\varepsilon) &= \dot\chi(t)\Big(l_1(x,\varepsilon\chi(t)) + \varepsilon\chi(t)\,\frac{\partial l_1}{\partial\varepsilon}(x,\varepsilon\chi(t))\Big)\Big(p + \varepsilon\chi(t)\, l_2(x,\varepsilon\chi(t))\Big) \\ &\quad - \dot\chi(t)\,\frac{\partial h}{\partial\varepsilon}(x,\varepsilon\chi(t)), \end{aligned} \tag{7.18}$$

and we write $X = (Q, P)$ as well as $x = \Psi(X, t, \varepsilon)$. To pass to the limit $\varepsilon \to 0$ in (7.18), we first recall that $\Psi$ is continuous on $G_{\hat r} \times [0, 1] \times [0, \hat\sigma]$ and $\Psi(X, t, 0) = X$. From (2.6) in Definition 2.3 of an E-symplectic family we know that $\frac{\partial h}{\partial\varepsilon}(x, \varepsilon) \to \mathfrak{m}(x)$ as $\varepsilon \to 0$ uniformly in $x \in G_\rho$. Thus (7.18) yields

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\, h_{\rm aux}(X,t,\varepsilon) = \dot\chi(t)\,[l_1(X,0)P - \mathfrak{m}(X)]$$

and this limit is uniform in $X \in G_{\hat r}$, $t \in [0, 1]$.

Now we define

$$H_{\rm aux}(X,t,\varepsilon) = \begin{cases} \frac{1}{\varepsilon}\, h_{\rm aux}(X,t,\varepsilon) & : \ \varepsilon \in ]0,\hat\sigma] \\ \dot\chi(t)\,[l_1(X,0)P - \mathfrak{m}(X)] & : \ \varepsilon = 0 \end{cases}. \tag{7.19}$$

for $X \in G_{\hat r}$ and $t \in [0, 1]$, and we are going to verify that $H_{\rm aux}$ has the desired properties. From the above discussions we know that $H_{\rm aux}$ is continuous and

$$H_{\rm aux}(X,0,\varepsilon) = H_{\rm aux}(X,1,\varepsilon) = 0. \tag{7.20}$$

As a consequence, $H_{\rm aux}$ can be extended to $G_{\hat r} \times \mathbb{R} \times [0, \hat\sigma]$ in a $T = 1$ periodic fashion. First we need to prove that $H_{\rm aux} \in \mathcal{H}_{\hat r,\hat\sigma}$, cf. Definition 6.1. Here (a)-(c) in this definition are straightforward to check. Concerning (d), for $\varepsilon > 0$ we know from (7.2) that

$$J\nabla H_{\rm aux}(X,t,\varepsilon) = \frac{1}{\varepsilon}\frac{\partial\Phi}{\partial t}(\Psi(X,t,\varepsilon),t,\varepsilon).$$

Thus, by (7.16),

$$\lim_{\varepsilon\to 0} J\nabla H_{\rm aux}(X,t,\varepsilon) = \dot\chi(t)\, l(X,0),$$

and this limit is uniform in $G_{\hat r} \times \mathbb{R}$. On the other hand, the definition of $H_{\rm aux}$ and (2.8) implies that

$$\begin{aligned} J\nabla H_{\rm aux}(X,t,0) &= \dot\chi(t)\, J\Big[P\,\nabla l_1(X,0) - \nabla m(X) + \begin{pmatrix} 0 \\ l_1(X,0) \end{pmatrix}\Big] \\ &= \dot\chi(t)\, J \begin{pmatrix} -l_2(X,0) \\ l_1(X,0) \end{pmatrix} = \dot\chi(t)\, l(X,0). \end{aligned}$$

This shows that $\nabla_X H_{\rm aux}$ is continuous in all of its arguments. Then the bound on $\|\nabla_X H_{\rm aux}\|_{\hat r,\hat\sigma}$ is not difficult to derive from (7.14).

Lastly, we have to establish (7.6). In view of the definition of $H_{\rm aux}$ and (7.20), it suffices to consider $X \in G_{\hat r}$, $t \in ]0, 1]$ and $\varepsilon \in ]0, \hat\sigma]$. From (7.18) we deduce

$$|H_{\rm aux}(X,t,\varepsilon) - H_{\rm aux}(X,t,0)| \le \|\dot\chi\|_\infty\,(R_1 + R_2 + R_3),$$

where

$$\begin{aligned}
R_1 &= |l_1(x,\varepsilon\chi(t))p - l_1(X,0)P|,\\
R_2 &= \Big|\mathfrak{m}(X) - \frac{\partial h}{\partial \varepsilon}(x,\varepsilon\chi(t))\Big|,\\
R_3 &= \varepsilon\chi(t)\,|l_1(x,\varepsilon\chi(t))|\,|l_2(x,\varepsilon\chi(t))| + \varepsilon\chi(t)\left|\frac{\partial l_1}{\partial \varepsilon}(x,\varepsilon\chi(t))\right|\,|p|\\
&\quad + \varepsilon^2\chi(t)^2\left|\frac{\partial l_1}{\partial \varepsilon}(x,\varepsilon\chi(t))\right|\,|l_2(x,\varepsilon\chi(t))|.
\end{aligned}$$

For $R_1$, we observe that by definition of $X = \Phi(x,t,\varepsilon)$, see (7.7),

$$|X - x| = \varepsilon\,\chi(t)\,|l(x,\varepsilon\chi(t))| \le \varepsilon\,\|l\|_{1,\rho,\sigma}.$$

Also note that $x = \Psi(X,t,\varepsilon) \in G_{r_1}$ by construction. Therefore

$$|l_1(x,\varepsilon\chi(t)) - l_1(X,\varepsilon\chi(t))| \le 2\left\|\frac{\partial l_1}{\partial x}\right\|_{r_1,\hat\sigma}|x - X| \le \frac{2\varepsilon}{\rho - r_1}\|l_1\|_{r_1,\hat\sigma}\,\|l\|_{1,\rho,\sigma} \le \frac{2\varepsilon}{\rho - r_1}\|l\|^2_{1,\rho,\sigma}. \tag{7.21}$$

Since $l \in \mathcal{M}_{1,\rho,\sigma}$, also

$$|l(X,\varepsilon\chi(t)) - l(X,0)| \le \|l\|_{1,\rho,\sigma}\,\varepsilon$$

is verified. At this point we need to invoke the geometry of $G = \mathbb{R}\times I$. If $I$ is contained in $[-R,R]$, then $|P| \le R + \hat r \le R + \rho$ as well as $|p| \le R + r_1 \le R + \rho$, due to $X \in G_{\hat r}$ and $x \in G_{r_1}$. Thus altogether, using the foregoing estimates,

$$\begin{aligned}
|R_1| &\le |l_1(x,\varepsilon\chi(t))|\,|p - P| + |l_1(x,\varepsilon\chi(t)) - l_1(X,\varepsilon\chi(t))|\,|P| + |l_1(X,\varepsilon\chi(t)) - l_1(X,0)|\,|P|\\
&\le \varepsilon\,\|l\|^2_{1,\rho,\sigma} + \frac{2(R+\rho)\varepsilon}{\rho - r_1}\|l\|^2_{1,\rho,\sigma} + (R+\rho)\,\|l\|_{1,\rho,\sigma}\,\varepsilon,
\end{aligned}$$

which is acceptable. For $R_2$ we can argue as follows. Since also $h \in \mathcal{M}_{1,\rho,\sigma}$, we obtain as in (7.21) that

$$\left|\frac{\partial h}{\partial \varepsilon}(x,\varepsilon\chi(t)) - \frac{\partial h}{\partial \varepsilon}(X,\varepsilon\chi(t))\right| \le \frac{2\varepsilon}{\rho - r_1}\left\|\frac{\partial h}{\partial \varepsilon}(\cdot,\varepsilon\chi(t))\right\|_{r_1}\|l\|_{1,\rho,\sigma} \le \frac{2\varepsilon}{\rho - r_1}\|h\|_{1,\rho,\sigma}\,\|l\|_{1,\rho,\sigma}\,;$$

observe that $\varepsilon\chi(t) \in ]0,\sigma]$ for $\varepsilon \in ]0,\hat\sigma]$ and $t \in ]0,1]$. If we combine this estimate with (2.6), then $R_2 \le C\varepsilon$ is found. Finally, from $l \in \mathcal{M}_{1,\rho,\sigma}$ and $|p| \le R + \rho$, also $R_3 \le C\varepsilon$ is obtained. This completes the argument for (7.6), and hence the proof of the lemma. □

Now we are in a position to complete the

**Proof of Theorem 2.5:** Let $r_2 = \frac{\rho}{3}$ and $r_1 = \frac{2\rho}{3}$. Then Lemma 7.3 can be applied to $l$ and $0 < r_2 < r_1 < \rho$. We deduce that there exist $\sigma_1 \in ]0,\sigma[$ and a Hamiltonian $H_{\rm aux} \in \mathcal{H}_{r_1,\sigma_1}$ such that for $\varepsilon \in [0,\sigma_1]$ the Poincaré map of

$$\dot x = \varepsilon J\nabla H_{\rm aux}(x,t,\varepsilon) \tag{7.22}$$

is $P_\varepsilon$, restricted to $G_{r_2}$. In addition, one can find a constant $C_{\rm aux} > 0$ so that

$$|H_{\rm aux}(x,t,\varepsilon) - H_{\rm aux}(x,t,0)| \leq C_{\rm aux}\,\varepsilon \tag{7.23}$$

for $x \in G_{r_1}$, $t \in [0,1]$ and $\varepsilon \in [0,\sigma_1]$. The constant $C_{\rm aux}$ depends upon $\rho$, $\sigma$, $\|l\|_{1,\rho,\sigma}$, the interval $I$, $\|h\|_{1,\rho,\sigma}$ and $\sup_{\varepsilon\in]0,\sigma]}\|\varepsilon^{-1}(\frac{\partial h}{\partial\varepsilon}(\cdot,\varepsilon) - \mathfrak{m})\|_\rho$.

Next we are going to invoke Theorem 6.14 to $H_{\rm aux}$ for the parameters $r = r_2$, $\rho = r_1$, $\sigma = \sigma_1$ and $T = 1$. By this result, we can find $C, D > 0$ (depending upon $\rho$ and $\|\nabla H_{\rm aux}\|_{r_1,\sigma_1}$) with the following properties. There is a change of variables $x = \Gamma(y,t,\varepsilon)$, which is defined on $G_{r_2} \times \mathbb{R} \times [0,\sigma_1]$ and which satisfies $\Gamma(G_{r_2},t,\varepsilon) \subset G_{r_1}$ for $t \in \mathbb{R}$ and $\varepsilon \in [0,\sigma_1]$, such that (7.22) is transformed into

$$\dot{y} = \varepsilon(J\nabla_y\mathcal{N}(y,\varepsilon) + J\nabla_y\mathcal{R}(y,t,\varepsilon)), \tag{7.24}$$

for functions $\mathcal{N} \in \mathcal{H}_{r_2,\sigma_1}$ and $\mathcal{R} \in \widetilde{\mathcal{H}}_{r_2,\sigma_1}$. Furthermore,

$$\begin{aligned} \|\nabla_y\mathcal{N}(\cdot,\varepsilon)\|_{r_2} &\leq C, & (7.25)\\ \|\nabla_y\mathcal{R}(\cdot,\cdot,\varepsilon)\|_{r_2} &\leq Ce^{-D/\varepsilon}, & (7.26)\\ |\mathcal{N}(y,\varepsilon) - \mathcal{N}(y,0)| &\leq C\varepsilon + \|\overline{H_{\rm aux}}(\cdot,\varepsilon) - \overline{H_{\rm aux}}(\cdot,0)\|_{r_1}, & (7.27) \end{aligned}$$

for $y \in G_{r_2}$ and $\varepsilon \in [0,\sigma_1]$. In addition,

$$\mathcal{N}(y,0) = \overline{H_{\rm aux}}(y,0) \tag{7.28}$$

is verified. According to the definition of $H_{\rm aux}$ in (7.19) and by (7.18), one sees that it is possible to bound $\|\nabla H_{\rm aux}\|_{r_1,\sigma_1}$ in terms of $\|l\|_{1,\rho,\sigma}$, the interval $I$ and $\|h\|_{1,\rho,\sigma}$.

For later reference we first discuss the connection between $\mathcal{N}$ and the function $E$ from Theorem 2.5, cf. (2.10), and we also consider the variation of $\mathcal{N}$ w.r. to $\varepsilon$. From (7.28), the definition of $H_{\rm aux}(y,0)$ in (7.19) and (2.10),

$$\mathcal{N}(y,0) = \overline{H_{\rm aux}}(y,0) = \int_0^1 \dot{\chi}(t)\,[l_1(y,0)P - \mathfrak{m}(y)]\,dt = l_1(y,0)P - \mathfrak{m}(y) = E(y), \tag{7.29}$$

where $y = (Q,P)$. Using (7.27) and (7.23), we moreover find for $y \in G_{r_2}$ and $\varepsilon \in [0,\sigma_1]$ that

$$|\mathcal{N}(y,\varepsilon) - \mathcal{N}(y,0)| \leq C\varepsilon + \|\overline{H_{\rm aux}}(\cdot,\varepsilon) - \overline{H_{\rm aux}}(\cdot,0)\|_{r_1} \leq C_1\varepsilon, \tag{7.30}$$

where the constant $C_1 = C + C_{\rm aux}$ depends upon $\rho$, $\sigma$, $\|l\|_{1,\rho,\sigma}$, the interval $I$, $\|h\|_{1,\rho,\sigma}$ and $\sup_{\varepsilon\in]0,\sigma]}\|\varepsilon^{-1}(\frac{\partial h}{\partial\varepsilon}(\cdot,\varepsilon) - \mathfrak{m})\|_\rho$; henceforth all constants are allowed to depend upon those parameters.

Now we define $r_3 = \frac{2r_2}{3} = \frac{2\rho}{9}$ and $r_4 = \frac{r_2}{3} = \frac{\rho}{9}$ to obtain $0 < r_4 < r_3 < r_2 < r_1 < \rho$. According to Corollary 6.15 there is $\sigma_2 \in ]0,\sigma_1]$ such that

$$G_{r_3} \subset \Gamma(G_{r_2},t,\varepsilon)$$

for $t \in \mathbb{R}$ and $\varepsilon \in [0,\sigma_2]$; in particular, $\Gamma(\cdot,t,\varepsilon)^{-1} : G_{r_3} \to G_{r_2}$ is well-defined.

Let $\Phi(x,t,\varepsilon)$ denote the solution to (7.22) satisfying $\Phi(x,0,\varepsilon) = x$. Similarly, $\phi(y,t,\varepsilon)$ will be used for the solution to (7.24) so that $\phi(y,0,\varepsilon) = y$. Now we select $\sigma_3 \in ]0,\sigma_2]$ such that

$\Phi(x,t,\varepsilon)$ is well-defined on $G_{r_4}\times[0,1]\times[0,\sigma_2]$ and takes values in $G_{r_3}$. The solutions of the two systems are connected by the formula

$$\phi(y,t,\varepsilon)=\Gamma^{-1}(\Phi(\Gamma(y,0,\varepsilon),t,\varepsilon),t,\varepsilon)=\Gamma^{-1}(\Phi(y,t,\varepsilon),t,\varepsilon)$$

for $y\in G_{r_4}$, $t\in[0,1]$ and $\varepsilon\in[0,\sigma_3]$. Letting $t=1$ and taking into account condition (b) in Definition 6.13, it follows that

$$\phi(y,1,\varepsilon)=\Gamma^{-1}(\Phi(y,1,\varepsilon),1,\varepsilon)=\Phi(y,1,\varepsilon)=P_\varepsilon(y).$$

In other words, $P_\varepsilon$ is also the Poincaré map of (7.24), at least in the domain $G_{r_4}$.

Now we are going to consider the autonomous system

$$\dot{y}=\varepsilon J\nabla_y\mathcal{N}(y,\varepsilon), \tag{7.31}$$

denoting by $\hat{\phi}(y,t,\varepsilon)$ the associated flow. Using (7.25), we deduce that there is $\hat{\sigma}\in]0,\sigma_3]$ with the property that $\hat{\phi}(y,t,\varepsilon)$ is well-defined on $G_{r_4}\times[0,1]\times[0,\hat{\sigma}]$ and moreover

$$\hat{\phi}(G_{r_4}\times[0,1]\times[0,\hat{\sigma}])\subset G_{r_3}.$$

The system (7.31) is Hamiltonian, with Hamiltonian function $\varepsilon\mathcal{N}(\cdot,\varepsilon)$. In particular, if $\hat{P}_\varepsilon=\hat{\phi}(\cdot,y,1)$ denotes the Poincaré map of (7.31), then

$$\mathcal{N}(\hat{P}_\varepsilon(y),\varepsilon)=\mathcal{N}(y,\varepsilon),\quad y\in G_{r_4},\quad \varepsilon\in[0,\hat{\sigma}]. \tag{7.32}$$

To estimate the difference between $\phi$ and $\hat{\phi}$, we first observe that for $\varepsilon\in[0,\hat{\sigma}]$,

$$\|D^2\mathcal{N}(\cdot,\varepsilon)\|_{r_3}\le\frac{1}{r_3-r_2}\,\|\nabla\mathcal{N}(\cdot,\varepsilon)\|_{r_2}\le\frac{1}{r_3-r_2}\,C=C_2,$$

where we have once again resorted to (7.25). If $(y,t,\varepsilon)\in G_{r_4}\times[0,1]\times[0,\hat{\sigma}]$, then the systems (7.24), (7.31) in conjunction with (7.26) yield

$$\begin{aligned}|\phi(y,t,\varepsilon)-\hat{\phi}(y,t,\varepsilon)| &= \varepsilon\Big|\int_0^t[J\nabla_y\mathcal{N}(\phi(y,s,\varepsilon),\varepsilon)+J\nabla_y\mathcal{R}(\phi(y,s,\varepsilon),s,\varepsilon)\\ &\qquad -J\nabla_y\mathcal{N}(\hat{\phi}(y,s,\varepsilon),\varepsilon)]\,ds\Big|\\ &\le C_2\,\varepsilon\int_0^t|\phi(y,s,\varepsilon)-\hat{\phi}(y,s,\varepsilon)|\,ds+C\varepsilon e^{-D/\varepsilon}.\end{aligned}$$

Hence from Gronwall's inequality,

$$|\phi(y,t,\varepsilon)-\hat{\phi}(y,t,\varepsilon)|\le C\varepsilon e^{-D/\varepsilon}\,e^{C_2\varepsilon t}.$$

For the Poincaré maps, i.e., at $t=1$, we deduce

$$|P_\varepsilon(y)-\hat{P}_\varepsilon(y)|\le C_3\,\varepsilon\,e^{-D/\varepsilon},\quad y\in G_{r_4},\quad \varepsilon\in[0,\hat{\sigma}], \tag{7.33}$$

where $C_3=Ce^{C_2\hat{\sigma}}$.

Now we are ready to complete the proof. Let $(x_n)_{0\le n\le N} = (P_\varepsilon^n(x_0))_{0\le n\le N}$ be a real forward orbit piece of $P_\varepsilon$ so that $x_n \in G$ for all $0 \le n \le N$. Since $G \subset G_{r_4}$, all the previous properties can be used along the orbit. From (7.29) and (7.30) we get

$$\begin{aligned} |E(x_n) - E(x_0)| &\le |E(x_n) - \mathcal{N}(x_n,\varepsilon)| + |\mathcal{N}(x_n,\varepsilon) - \mathcal{N}(x_0,\varepsilon)| + |\mathcal{N}(x_0,\varepsilon) - E(x_0)| \\ &= |\mathcal{N}(x_n,0) - \mathcal{N}(x_n,\varepsilon)| + |\mathcal{N}(x_n,\varepsilon) - \mathcal{N}(x_0,\varepsilon)| + |\mathcal{N}(x_0,\varepsilon) - \mathcal{N}(x_0,0)| \\ &\le 2C_1\varepsilon + |\mathcal{N}(x_n,\varepsilon) - \mathcal{N}(x_0,\varepsilon)|. \end{aligned}$$

In addition, (7.32), (7.25) and (7.33) lead to

$$\begin{aligned} |\mathcal{N}(x_n,\varepsilon) - \mathcal{N}(x_0,\varepsilon)| &\le \sum_{j=0}^{n-1} |\mathcal{N}(P_\varepsilon(x_j),\varepsilon) - \mathcal{N}(x_0,\varepsilon)| \\ &= \sum_{j=0}^{n-1} |\mathcal{N}(P_\varepsilon(x_j),\varepsilon) - \mathcal{N}(\hat{P}_\varepsilon(x_0),\varepsilon)| \\ &\le CC_3\, n\, \varepsilon\, e^{-D/\varepsilon}. \end{aligned}$$

Thus the claim follows if we define $\hat{C} = 2C_1 + CC_3$ and $\hat{D} = D$. □

# References


[1] Arnold V.I.: *Mathematical Methods of Classical Mechanics*, 2nd edition, Springer, Berlin-New York 1989

[2] Benettin G. & Giorgilli A.: On the Hamiltonian interpolation of near-to-the-identity symplectic mappings with application to symplectic integration algorithms, *J. Statist. Phys.* **74**, 1117-1143 (1994)

[3] Gorelyshev I.V. & Neishtadt A.I.: On the adiabatic perturbation theory for systems with impacts, *J. Appl. Math. Mech.* **70**, 4-17 (2006)

[4] Kunze M. & Ortega R.: Complete orbits for twist maps on the plane, *Ergodic Theory Dynam. Systems* **28**, 1197-1213 (2008)

[5] Kunze M. & Ortega R.: Complete orbits for twist maps on the plane: the case of small twist, *Ergodic Theory Dynam. Systems* **31**, 1471-1498 (2011)

[6] Kunze M. & Ortega R.: Complete orbits for twist maps on the plane: extensions and applications, *J. Dynamics Differential Equations* **23**, 405-423 (2011)

[7] Kunze M. & Ortega R.: Non-periodic twist maps, in *Stability and Bifurcation Theory for Non-Autonomous Differential Equations*, Eds. Johnson R. & Pera M.P., Lecture Notes in Mathematics Vol. 2065, Springer, Berlin-New York 2013, pp. 265-300

[8] Kunze M. & Ortega R.: Long-time stability estimates for the non-periodic Littlewood boundedness problem, *Proc. London Math. Soc.* **107**, 39-75 (2013)

[9] Kunze M. & Ortega R.: Escaping orbits are rare in the quasi-periodic Fermi-Ulam ping-pong, *Ergodic Theory Dynam. Systems* **40**, 975-991 (2020)

[10] Meyer K. & Offin D.: *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, 3rd edition, Springer, Berlin-New York 2017

[11] Neishtadt A.I.: The separation of motions in systems with rapidly rotating phase, *J. Appl. Math. Mech.* **48**, 133-139 (1985)

[12] Poincaré H.: *Leçons de Mécanique Céleste*, t. I, Gauthier-Villars, Paris 1905

[13] Pommerenke Ch.: *Boundary Behaviour of Conformal Maps*, Springer, Berlin-New York 1992

[14] Schliessauf H.: Escaping orbits are rare in the quasi-periodic Littlewood boundedness problem, *NoDEA Nonlinear Differential Equations Appl.* **26**, Paper No. 24, 21 pp. (2019)

[15] Schliessauf H.: Escaping orbits are also rare in the almost periodic Fermi-Ulam ping-pong, to appear in *Qual. Theory Dyn. Syst.*; arXiv:1908.02529

[16] Shub M.: *Global Stability of Dynamical Systems*, Springer, Berlin-New York 1987

[17] Siegel C.L. & Moser J.K.: *Lectures on Celestial Mechanics*, reprint, Springer, Berlin-New York 1995

[18] Treschev D. & Zubelevich O.: *Introduction to the Perturbation Theory of Hamiltonian Systems*, Springer, Berlin-New York 2010

[19] Zharnitsky V.: Invariant curve theorem for quasiperiodic twist mappings and stability of motion in the Fermi-Ulam problem, *Nonlinearity* **13**, 1123-1136 (2000)