ISSN 1560-3547, Regular and Chaotic Dynamics, 2023, Vol. 28, Nos. 4–5, pp. 395–424. © Pleiades Publishing, Ltd., 2023, corrected publication 2023.
ISSN 1560-3547, Regular and Chaotic Dynamics, 2023. © Pleiades Publishing, Ltd., 2023.

Complex Arnol'd – Liouville Maps

Luca Biasco[1*] and Luigi Chierchia[1]**

[1] *Dipartimento di Matematica e Fisica,*
Università degli Studi Roma Tre,
Largo San Leonardo Murialdo 1, 00146 Roma, Italy

Received February 27, 2023; revised May 05, 2023; accepted June 20, 2023

Abstract—We discuss the holomorphic properties of the complex continuation of the classical Arnol'd – Liouville action-angle variables for real analytic 1 degree-of-freedom Hamiltonian systems depending on external parameters in suitable Generic Standard Form, with particular regard to the behaviour near separatrices. In particular, we show that near separatrices the actions, regarded as functions of the energy, have a special universal representation in terms of affine functions of the logarithm with coefficients analytic functions. Then, we study the analyticity radii of the action-angle variables in arbitrary neighborhoods of separatrices and describe their behaviour in terms of a (suitably rescaled) distance from separatrices. Finally, we investigate the convexity of the energy functions (defined as the inverse of the action functions) near separatrices, and prove that, in particular cases (in the outer regions outside the main separatrix, and in the case the potential is close to a cosine), the convexity is strictly defined, while in general it can be shown that inside separatrices there are inflection points.

MSC2010 numbers: 37J05, 37J35, 37J40, 70H05, 70H08, 70H15

DOI: 10.1134/S1560354723520064

Keywords: Hamiltonian systems, action-angle variables, Arnol'd – Liouville integrable systems, complex extensions of symplectic variables, KAM theory

1. INTRODUCTION

Vladimir Igorevič Arnol'd, in his didactic masterpiece [2] where most of us learned modern mechanics, explained, in precise mathematical terms, how to construct action-angle variables for an integrable Hamiltonian system with compact energy levels (for a different construction, see, also, [15]). In this paper we consider real-analytic, one-degree-of-freedom Hamiltonian systems depending on adiabatic invariants and discuss the fine holomorphic properties of the complex Arnold – Liouville map, especially near their singularities.

Analytic properties of the action-angle map are difficult to find in the mathematical literature despite their interest *per se* and, especially, in view of their relevance in modern perturbation theory; see, however, A. Neishtadt's remarkable thesis [20] (in Russian). Indeed, in real-analytic theories (such as averaging theory, KAM theory, Nekhoroshev-like theorems, *etc.*[1)]), it is necessary to control complex domains, whose characteristics appear explicitly in iterative perturbative constructions. This is even more relevant if one needs to have holomorphic information *arbitrarily* close to the singularities of the action-angle variables (separatrices).

For example, in real-analytic models of Arnol'd diffusion, starting with Arnol'd' pioneering 1964 paper [1], one often considers perturbations of the simple pendulum[2)], while the analysis

*E-mail: luca.biasco@uniroma3.it
**E-mail: luigi.chierchia@uniroma3.it

1) Compare [3] and references therein for general information.

2) For references on Arnol'd diffusion, besides [1]; compare also [5, 10, 12, 13, 16, 18, 23, 24] among many other interesting results.

presented here allows for a real-analytic generic class of one-degrees-of-freedom Hamiltonians, which is a significant generalization of pendulum-like models. Furthermore, in connection with the formidable problem of Arnol'd diffusion in generic real-analytic *a-priori* stable systems, the fine complex analytic understanding of the integrable limit appears to be an essential tool.

Another example concerns the detection of primary and secondary Lagrangian tori in generic analytic nearly-integrable systems in phase space regions very close or inside the separatrices arising near simple resonances. In fact, besides the usual KAM primary tori (i. e., the tori which are deformation of integrable ones at a distance $\gg\sqrt{\varepsilon}$ from separatrices), one expects the appearance of more tori very close to separatrices and secondary tori of different homotopy inside separatrices; compare [3, Section 6.3.3-C] and [19]. In this kind of analysis it also essential to have *uniform estimates* over the relevant analyticity parameters, which is one of the main issues addressed in this paper. For more information on this subject, see [17] for lower-dimensional tori and [6–8], and [9], where the authors develop a "singular KAM theory" for real-analytic generic natural systems, which shows, in particular, that the total Liouville measure of Lagrangian KAM tori has a density larger than $1-\varepsilon|\log\varepsilon|^c$. For such results, the analytic tools developed in this paper play an essential role.

The main results of this paper are Theorem 1 and Theorem 2.

In Theorem 1 it is shown that near separatrices the actions, regarded as functions of the energy, have a special universal representation in terms of affine functions of the logarithm with coefficients analytic functions. Such a representation allows one to give several estimates on the (derivatives of the) action functions.

In Theorem 2 we compute the analyticity radii of the action-angle variables in arbitrary neighborhoods of separatrices and see how they behave in terms of a (suitably rescaled) distance from separatrices.

In Section 5 we investigate the convexity of the energy functions (defined as the inverse of the action functions) near separatrices, showing that, in particular cases (in the outer regions outside the main separatrix, and in the case when the potential is "close" to a cosine), the convexity is strictly defined, while in general it can be shown that inside separatrices there are inflection points.

For completeness we include in Appendix A a rather standard hyperbolic normal form near hyperbolic equilibria. Appendix B contains the proofs of two simple lemmata.

Finally, we point out that some results of this paper are not new. Indeed, similar analytic estimates were obtained by different methods by A. I. Neishtadt in [20, Chapter 3, Section 7], which, however, does not contain explicit estimates of constants related to analyticity properties as discussed here.

2. 1D HAMILTONIANS IN STANDARD FORM

Consider a 1-degree-of-freedom real-analytic Hamiltonian system possibly parameterized by external parameters with the Hamiltonian function

$$\mathrm{H}(\mathrm{p},\mathrm{q}_1)=\mathrm{h}(\mathrm{p})+\varepsilon\mathrm{f}(\mathrm{p},\mathrm{q}_1),\qquad(\mathrm{p},\mathrm{q}_1)\in\mathrm{D}\times\mathbb{T}^1,\tag{2.1}$$

where $n\geqslant 1$, $\mathrm{D}\subset\mathbb{R}^n$ is a bounded domain, $\mathbb{T}^n:=\mathbb{R}^n/(2\pi\mathbb{Z}^n)$, $\varepsilon\geqslant 0$ a "perturbation parameter", $\hat{\mathrm{p}}=(\mathrm{p}_2,\dots,\mathrm{p}_n)$ are external "adiabatic invariants", and h and f are real-analytic functions.

An important example is when $n\geqslant 2$ and H is the "secular" Hamiltonian of a nearly-integrable system $(\mathtt{H},\mathrm{D}\times\mathbb{T}^n)$ with

$$\mathtt{H}=h(y)+\varepsilon f(y,x)\tag{2.2}$$

real-analytic, after averaging (for small ε) around a simple resonance

$$\mathcal{R}_k:=\{y\in\mathrm{D}:\partial_y h(y)\cdot k=0\}\tag{2.3}$$

for some nonvanishing $k\in\mathbb{Z}^n$, and after a suitable linear symplectic change of variables such that $\mathrm{q}_1=k\cdot x$ is the resonant angle and such that $\partial_{\mathrm{p}_1}\mathrm{h}|_{\mathrm{p}_1=0}=0$. Here, "simple resonance" means that, in the fixed neighborhood of $\mathcal{R}_k$ where averaging is performed, there are no other independent resonant relations $\partial_y h(y)\cdot\ell=0$ for some ℓ independent of k (and of not too high order); "secular" means that H is obtained disregarding the high-order perturbation obtained after averaging.

Remark 1. As just seen in the above example, the case when $\partial_{\mathrm{p}_1}\mathrm{h}(\mathrm{p}^0) = 0$ for some point $\mathrm{p}^0 \in \mathrm{D}$ appears naturally in perturbation theory.

On the other hand, the case when $\partial_{\mathrm{p}_1}\mathrm{h}(\mathrm{p}^0) \neq 0$ is trivial: indeed, by the implicit function theorem, for values of the energy E close to $\mathrm{h}(\mathrm{p}^0)$, p close to p^0 and ε small enough, there exists a function $v(E, \hat{\mathrm{p}}, \mathrm{q}_1)$ such that $\mathrm{h}(v(E, \hat{\mathrm{p}}, \mathrm{q}_1), \hat{\mathrm{p}}, \mathrm{q}_1) = E$; therefore, one can define new action variables

$$I_1 := \int_0^{2\pi} v(E, \hat{\mathrm{p}}, \mathrm{q}_1) d\mathrm{q}_1, \qquad \hat{I} := \hat{\mathrm{p}},$$

which, by the classical Arnol'd – Liouville construction, can be completed into a symplectic transformation $\phi : (I, \varphi) \to (\mathrm{p}, \mathrm{q})$ such that $\mathrm{H} \circ \phi = h(I)$ with h real-analytic; see [2, Chapter 10] for details.

Next, we show that, in general, the Hamiltonian H in (2.1), in a neighborhood of a critical point p^0 of h, can be symplectically put into a "standard form", which generalizes the features of the standard pendulum and it is particularly suited to study its (complex) Arnol'd – Liouville action-angle variables. The precise quantitative definition of "Hamiltonian in standard form" is given in the following two definitions.

Definition 1. A $C^2(\mathbb{T}, \mathbb{R})$ Morse function F with distinct critical values is called β**-Morse**, with $\beta > 0$, if

$$\min_{\theta \in \mathbb{T}} \big(|F'(\theta)| + |F''(\theta)| \big) \geqslant \beta, \quad \min_{i \neq j} |F(\theta_i) - F(\theta_j)| \geqslant \beta, \tag{2.4}$$

where $\theta_i \in \mathbb{T}$ are the critical points of F.

To formulate the next definition we need some notation: Given $D \subseteq \mathbb{R}^m$, and $r > 0$, we denote by D_r the complex neighborhood of D given by

$$D_r := \bigcup_{z \in D} \{y \in \mathbb{C}^m \text{ s.t. } |y - z| < r\},$$

and, for $s > 0$, by $\mathbb{T}^m_s$ the complex neighborhood of width $2s$ of $\mathbb{T}^m$ given by

$$\mathbb{T}^m_s := \{x = (x_1, \ldots, x_m) \in \mathbb{C}^m : \quad |\operatorname{Im} x_j| < s\}/(2\pi\mathbb{Z}^m). \tag{2.5}$$

Definition 2. Let $\hat{D} \subseteq \mathbb{R}^{n-1}$ be a bounded domain, $\mathtt{R} > 0$ and $D := (-\mathtt{R}, \mathtt{R}) \times \hat{D}$. We say that the real-analytic Hamiltonian $\mathtt{H}$ is in generic standard form (in short, "standard form") with respect to standard symplectic variables $(p_1, q_1) \in (-\mathtt{R}, \mathtt{R}) \times \mathbb{T}$ and "external actions" $\hat{p} = (p_2, \ldots, p_n) \in \hat{D}$ if $\mathtt{H}$ has the form

$$\mathtt{H}(p, q_1) = \big(1 + \nu(p, q_1)\big) p_1^2 + \mathtt{G}(\hat{p}, q_1), \tag{2.6}$$

where $p = (p_1, \hat{p}) = (p_1, \ldots, p_n)$ and:

- ν and $\mathtt{G}$ are real-analytic functions defined on, respectively, $D_{\mathtt{r}} \times \mathbb{T}_{\mathtt{s}}$ and $\hat{D}_{\mathtt{r}} \times \mathbb{T}_{\mathtt{s}}$ for some $0 < \mathtt{r} \leqslant \mathtt{R}$ and $\mathtt{s} > 0$;
- $\mathtt{G}$ has zero average and there exists a function $\bar{\mathtt{G}}$ (the "reference potential") depending only on q_1 such that, for some $\beta > 0$,

$$\bar{\mathtt{G}} \ \text{ is } \ \beta\text{–Morse}, \qquad \langle \bar{\mathtt{G}} \rangle = 0; \tag{2.7}$$

- the following estimates hold:

$$\begin{cases} \sup\limits_{\mathbb{T}^1_{\mathtt{s}}} |\bar{\mathtt{G}}| \leqslant \varepsilon, \\ \sup\limits_{\hat{D}_{\mathtt{r}} \times \mathbb{T}^1_{\mathtt{s}}} |\mathtt{G} - \bar{\mathtt{G}}| \leqslant \varepsilon\mu, \quad \text{for some} \quad 0 < \varepsilon \leqslant \mathtt{r}^2/2^{16}, \quad 0 \leqslant \mu < 1, \\ \sup\limits_{D_{\mathtt{r}} \times \mathbb{T}^1_{\mathtt{s}}} |\nu| \leqslant \mu. \end{cases} \tag{2.8}$$

We shall call $(\hat{D}, \mathtt{R}, \mathtt{r}, \mathtt{s}, \beta, \varepsilon, \mu)$ *the analyticity characteristics of* $\mathtt{H}$ *with respect to the unperturbed potential* $\bar{\mathtt{G}}$.

Remark 2. (i) The Hamiltonian in standard form $\mathtt{H}$ retains the basic features of the standard pendulum or, more precisely, of a natural system with a generic periodic potential, having, in particular all equilibria on the $p_1 = 0$ axis in the (p_1, q_1)-phase space.

(ii) If $\mathtt{H}$ is in standard form, then the parameters β and ε satisfy the relation[3)]

$$\varepsilon/\beta \geqslant 1/2. \tag{2.9}$$

Furthermore, one can always fix $\kappa \geqslant 4$ such that:

$$1/\kappa \leqslant \mathtt{s} \leqslant 1, \qquad 1 \leqslant \mathtt{R}/\mathtt{r} \leqslant \kappa, \qquad 1/2 \leqslant \varepsilon/\beta \leqslant \kappa. \tag{2.10}$$

Such a parameter κ rules the scaling properties of these Hamiltonians and is the only constant (besides the dimension n) on which the various constants depend.

(iii) The critical points of a Morse function on $\mathbb{T}$ (i. e., a function which has only nondegenerate critical points), by compactness, cannot accumulate, hence, they are in a finite, even number (alternately, a relative maximum and a relative minimum). For β-Morse functions one can easily estimate the number of critical points:

If G is a β-Morse function, then the number $2N$ of its critical points does not exceed $\pi\sqrt{2\max_{\mathbb{R}}|G''|/\beta}$.

Proof. If θ_i and θ_j are different critical points of G, then, by Taylor expansion at order two and by (2.4), one has $\beta \leqslant |G(\theta_i) - G(\theta_j)| \leqslant \frac{1}{2}(\max_{\mathbb{R}}|G''|)|\theta_i - \theta_j|^2$, which implies that

$$\min_{i \neq j} |\theta_i - \theta_j| \geqslant \sqrt{2\beta/\max_{\mathbb{R}}|G''|}, \tag{2.11}$$

from which the claim follows at once. □

(iv) Of course, the constant $1/2^{16}$ appearing in the definition is quite arbitrary (as long as it is $\ll 1$).

(v) Hamiltonians in standard form have been investigated in [6, 8] and [9].

Proposition 1. *Let* H *in* (2.1) *be a real-analytic function and assume that at* $\mathrm{p}^0 \in \mathrm{D}$ $\mathrm{p}_1 \to \mathrm{h}$ *has a nondegenerate critical point*[4)]. *Assume also that* $\mathrm{q}_1 \to \mathrm{f}(\mathrm{p}^0, \mathrm{q}_1)$ *is a Morse function with distinct critical values. Then, for ε small enough,* H *is symplectically conjugated to a Hamiltonian in standard form in an* (ε*-independent*) *neighborhood of* $\{\mathrm{p}^0\} \times \mathbb{T}^n$.

Proof. Assume that h and f have holomorphic extension on, respectively, $\mathrm{D_r}$ and $\mathrm{D_r} \times \mathbb{T}^1_{\mathrm{s}}$ for some $\mathrm{r}, \mathrm{s} > 0$, and that $|\mathrm{h}|$ and $|\mathrm{f}|$ are uniformly bounded on their complex domains by some constant $\mathrm{M} > 0$. Let us consider H as a 1-degree-of-freedom Hamiltonian in action-angle variables $(\mathrm{p}_1, \mathrm{q}_1)$, depending on parameters $\hat{\mathrm{p}} = (\hat{\mathrm{p}}_2, \ldots, \hat{\mathrm{p}}_n)$.

By assumption, there exist $\delta, \beta > 0$ such that $|\partial^2_{\mathrm{p}_1^2}\mathrm{h}(\mathrm{p}^0)| = \delta$, and $F(\theta) = \mathrm{f}(\mathrm{p}^0, \theta)$ verifies (2.4). By the implicit function theorem[5)], for $c = c(n) > 1$ large enough, setting

$$\rho := \delta \mathrm{r}^3/c\mathrm{M} \leqslant \mathrm{r}/4,$$

there exists a function $u(\hat{\mathrm{p}})$ holomorphic in $|\hat{\mathrm{p}} - \hat{\mathrm{p}}^0| < \rho$, such that $\mathrm{p}_1^0 = u(\hat{\mathrm{p}}^0)$,

$$\partial_{\mathrm{p}_1}\mathrm{h}(u(\hat{\mathrm{p}}), \hat{\mathrm{p}}) = 0 \qquad \text{and} \qquad |u(\hat{\mathrm{p}}) - \mathrm{p}_1^0| \leqslant \rho, \quad \forall\, |\hat{\mathrm{p}} - \hat{\mathrm{p}}^0| < \rho. \tag{2.12}$$

---

3) By (2.4) and (2.8), $\beta \leqslant |\bar{\mathtt{G}}(\theta_i) - \bar{\mathtt{G}}(\theta_j)| \leqslant 2\max_{\mathbb{T}}|\bar{\mathtt{G}}| \leqslant 2\varepsilon$.

4) Explicitly: $\partial_{\mathrm{p}_1}\mathrm{h}(\mathrm{p}^0) = 0$ and $\partial^2_{\mathrm{p}_1^2}\mathrm{h}(\mathrm{p}^0) \neq 0$.

5) See, e. g., [11, Appendix A].

Under the symplectic change of variables $(p_1, q_1) \to (\mathrm{p}_1, \mathrm{q}_1)$ given by $\mathrm{p}_1 = u(\hat{p}) + p_1$, $\mathrm{q}_1 = q_1$ (with $\hat{\mathrm{p}} = \hat{p}$) the Hamiltonian H becomes

$$H(p, q_1) := h(p) + \varepsilon f(p, q_1),$$

where $h(p) := \mathrm{h}(u(\hat{p}) + p_1, \hat{p})$ and $f(p, q_1) := \mathrm{f}(u(\hat{p}) + p_1, \hat{p}, q_1)$. Note that the functions h and f are holomorphic and uniformly bounded (in modulus) by M on, respectively, $\{|p_1| < \rho\} \times \{|\hat{p} - \hat{\mathrm{p}}^0| < \rho\}$ and $\{|p_1| < \rho\} \times \{|\hat{p} - \hat{\mathrm{p}}^0| < \rho\} \times \mathbb{T}^1_{\mathrm{s}}$. By (2.12) we have that

$$\partial_{p_1} h(0, \hat{p}) = 0 \qquad \forall\, |\hat{p} - \hat{\mathrm{p}}^0| < \rho. \tag{2.13}$$

Moreover, by Cauchy estimates (and taking c large enough) we also have

$$|\partial^2_{p_1^2} h(p)| \geqslant \delta/2, \qquad \forall\ |p_1|,\ |\hat{p} - \hat{\mathrm{p}}^0| < \rho.$$

Now we want to solve the equation $\partial_{p_1} H = 0$. By (2.13) we have

$$\partial_{p_1} H(0, \hat{p}, q_1)|_{\varepsilon=0} = \partial_{p_1} h(0, \hat{p}) = 0.$$

By the implicit function theorem there exists $c_* = c_*(n) \geqslant c > 1$ large enough such that, if $\sqrt{\varepsilon} \leqslant \delta \mathrm{r}^2/c_*\mathrm{M}$, then there exists a function $v(\hat{p}, q_1)$ holomorphic on

$$\{|\hat{p} - \hat{\mathrm{p}}^0| < \rho\} \times \mathbb{T}^1_{\mathrm{s}}$$

with $|v| \leqslant c_*\varepsilon\mathrm{M}/\delta\mathrm{r} \leqslant \rho/4$, satisfying $\partial_{p_1} H(v(\hat{p}, q_1), \hat{p}, q_1) = 0$. Then let us perform the symplectic transformation $p_1 = v(\hat{p}, q_1) + \mathrm{p}_1$, $q_1 = \mathrm{q}_1$ (with $\hat{p} = \hat{\mathrm{p}}$). The new Hamiltonian $H(v(\hat{p}, q_1) + p_1, \hat{p}, q_1)$ is holomorphic on

$$\{|p_1| < \rho/4\} \times \{|\hat{p} - \hat{\mathrm{p}}^0| < \rho\} \times \mathbb{T}^1_{\mathrm{s}},$$

and is given by[6)]

$$\begin{aligned} H(v + p_1, \hat{p}, q_1) \ &= \ H(v, \hat{p}, q_1) + p_1^2 \int_0^1 (1-t) \partial^2_{p_1^2} H(v + t p_1, \hat{p}, q_1)\, dt \\ &= g(\hat{p}) + g_0 \mathtt{H}, \qquad \mathtt{H} := \big(1 + \nu(p, q_1)\big) p_1^2 + \mathtt{G}(\hat{p}, q_1), \end{aligned} \tag{2.14}$$

where $g(\hat{p}) := h(0, \hat{p})$, $g_0 := \frac{1}{2} \partial^2_{\mathrm{p}_1^2} \mathrm{h}(\mathrm{p}^0)$ and

$$\begin{aligned} \nu(p, q_1) &:= \frac{1}{g_0} \int_0^1 (1-t) \Big(\partial^2_{\mathrm{p}_1^2} \mathrm{h}(u + v + t p_1, \hat{p}) - \partial^2_{\mathrm{p}_1^2} \mathrm{h}(\mathrm{p}^0) + \varepsilon \partial^2_{\mathrm{p}_1^2} \mathrm{f}(u + v + t p_1, \hat{p}, q_1) \Big) dt, \\ \mathtt{G}(\hat{p}, q_1) &:= \frac{1}{g_0} \int_0^1 (1-t) \partial^2_{\mathrm{p}_1^2} \mathrm{h}(u + t v, \hat{p}) v^2\, dt \ + \ \varepsilon \mathrm{f}(u + v, \hat{p}, q_1), \\ \bar{\mathtt{G}}(q_1) &:= \frac{\varepsilon}{g_0}\ \mathrm{f}(\mathrm{p}^0, q_1). \end{aligned} \tag{2.15}$$

By (2.12), (2.13), the Cauchy estimates, the facts that

$$\sqrt{\varepsilon} \leqslant \delta \mathrm{r}^2/c_*\mathrm{M}, \quad 4|v| \leqslant \rho = \delta \mathrm{r}^3/c\mathrm{M},$$

and $|u + v + t p_1 - \mathrm{p}_1^0| \leqslant 2\rho \leqslant \mathrm{r}/2$ for every $0 \leqslant t \leqslant 1$, and noting that $|g_0| = \delta/2$ and $\delta \leqslant 2\mathrm{M}/\mathrm{r}^2$ (by the Cauchy estimates), it is easy to see that the Hamiltonian $\mathtt{H}$ in (2.14)–(2.15) is in standard form according to Definition 2 with analyticity characteristics:

$$\hat{D} := \{|\hat{p} - \mathrm{p}^0| < \tfrac{\rho}{8}\},\ \mathtt{R} = \mathtt{r} := \tfrac{\rho}{8},\ \mathtt{s} := \min\{s, 1\},\ \beta := \tfrac{2\varepsilon\beta}{\delta},\ \varepsilon := \tfrac{3\varepsilon\mathrm{M}}{\delta},\ \mu := \tfrac{144}{c},$$

for a suitable constant $c > 144$. In particular, the condition $0 < \varepsilon < \mathtt{r}^2/2^{16}$ in (2.8) is satisfied by taking

$$\varepsilon \leqslant \frac{\delta^3 \mathrm{r}^6}{2^{24} c^2 \mathrm{M}^3}.$$

Finally, taking $\kappa := \max\{4, 1/s\}$, (2.10) holds. □

[6)]For brevity we write u and v instead of $u(\hat{p})$ and $v(\hat{p}, q_1)$, respectively.

3. ANALYTIC PROPERTIES OF ACTIONS AT CRITICAL ENERGIES

In the rest of the paper we shall investigate the complex analytic properties of the action-angle variables for a Hamiltonian in standard form.

In this section we show that near separatrices the actions regarded as functions of the energy E have a quite special "universal" representation (in terms of analytic functions and of logarithms) for energies close to their singular values (namely, the energy of the separatrices).

Let $\mathtt{H}$ be a Hamiltonian in standard form (Definition 2), let $\bar\theta_0$ be the unique absolute maximum of the reference potential $\bar{\mathtt{G}}$ in $[0, 2\pi)$, and let $2N$ be the number of its critical points (compare Remark 2(iii)). Then, the relative strict nondegenerate maximum and minimum points of $\bar{\mathtt{G}}$, $\bar\theta_i \in [\bar\theta_0, \bar\theta_0 + 2\pi]$, $(0 \leqslant i \leqslant 2N)$ follow in alternating order, $\bar\theta_0 < \bar\theta_1 < \bar\theta_2 < \ldots < \bar\theta_{2N} := \bar\theta_0 + 2\pi$, in particular, $\bar\theta_i$ *are relative maxima/minima points for even/odd* i.
Since $\bar{\mathtt{G}}$ is a β-Morse function, the corresponding critical energies are distinct; let us denote them by $\bar E_i := \bar{\mathtt{G}}(\bar\theta_i)$. Hence, $\bar E_{2N} = \bar E_0$ *is the unique global maximum.*

By the implicit function theorem, for μ small enough, we can continue the $2N$ critical points $\bar\theta_i$ of $\bar{\mathtt{G}}$, obtaining $2N$ critical points $\theta_i = \theta_i(\hat p)$ of $\mathtt{G}(\hat p, \cdot)$ for $\hat p \in \hat D$; the corresponding distinct critical energies become

$$E_j(\hat p) := \mathtt{G}\big(\hat p, \theta_j(\hat p)\big). \tag{3.1}$$

In fact, the following simple lemma, proven in Appendix B.1, based on the implicit function theorem holds.

Lemma 1. *If* [7] $\mu \leqslant 1/(2\kappa)^6$, *then the functions* $\theta_i(\hat p)$ *and* $E_i(\hat p)$ *are real-analytic in* $\hat p \in \hat D_{\mathtt{r}}$ *and*

$$\sup_{\hat p \in \hat D_{\mathtt{r}}} |\theta_i(\hat p) - \bar\theta_i| \leqslant \tfrac{2\varepsilon\mu}{\beta\mathtt{s}}, \qquad \sup_{\hat p \in \hat D_{\mathtt{r}}} |E_i(\hat p) - \bar E_i| \leqslant 3\kappa^3\varepsilon\mu. \tag{3.2}$$

Furthermore, the relative order of $\theta_i(\hat p)$ *and* $E_i(\hat p)$ *is, for every* $\hat p \in \hat D_{\mathtt{r}}$, *the same as that of, respectively,* $\bar\theta_i$ *and* $\bar E_i$.

Now, let $\hat p \in \hat D$ and consider the following phase space of the 1D Hamiltonian system governed by $\mathtt{H}$:

$$\mathcal{M} = \mathcal{M}(\hat p) := \{(p_1, q_1) \in (-\mathtt{R} - \mathtt{r}, \mathtt{R} + \mathtt{r}) \times \mathbb{T} \ \text{ s.t. } \ \mathtt{H}(p_1, \hat p, q_1) < \mathtt{R}^2 + \mathtt{R}\mathtt{r}\}. \tag{3.3}$$

Then, $\mathcal{M}$ decomposes in $2N+1$ open connected components $\mathcal{M}^i = \mathcal{M}^i(\hat p)$, with $0 \leqslant i \leqslant 2N$, plus a zero measure singular set $S = S(\hat p)$ formed by the $2N$ connected separatrices and the $2N$ critical points:

$$\mathcal{M} = \mathcal{M}(\hat p) = \bigcup_{i=0}^{2N} \mathcal{M}^i \ \cup \ S = \bigcup_{i=0}^{2N} \mathcal{M}^i(\hat p) \ \cup \ S(\hat p). \tag{3.4}$$

Although the sets $\mathcal{M}$ and S depend upon the dumb actions $\hat p$, for μ as in Lemma 1, one sees easily that

$$\Big(-\mathtt{R} - \frac{\mathtt{r}}{3}, \mathtt{R} + \frac{\mathtt{r}}{3}\Big) \times \mathbb{T} \subseteq \mathcal{M} \subseteq \Big(-\mathtt{R} - \frac{\mathtt{r}}{2}, \mathtt{R} + \frac{\mathtt{r}}{2}\Big) \times \mathbb{T}. \tag{3.5}$$

For the labeling of the domains $\mathcal{M}^i(\hat p)$ we shall adopt the following conventions:

$\mathcal{M}^0(\hat p)$ is the region below the lowest separatrix and $\mathcal{M}^{2N}(\hat p)$ is the region above the highest separatrix.

For $1 \leqslant i \leqslant 2N-1$ odd the closure of $\mathcal{M}^i(\hat p)$ contains the minimum point $(0, \theta_i)$, while for $2 \leqslant i \leqslant 2N-2$ even the boundary of $\mathcal{M}^i(\hat p)$ is formed by two connected components, the inner one containing the maximum point $(0, \theta_i)$.

[7] κ as in (2.10).

For $0 \leqslant i \leqslant 2N$, define

$$\begin{aligned} &E_-^{(i)}(\hat p) := E_i(\hat p), \quad 0 \leqslant i \leqslant 2N, \quad E_+^{(0)}(\hat p) = E_+^{(2N)}(\hat p) := \mathtt{R}^2 + \mathtt{Rr}, \\ &E_+^{(2j-1)}(\hat p) := \min\{E_{2j-2}(\hat p), E_{2j}(\hat p)\}, \quad 1 \leqslant j \leqslant N, \\ &E_+^{(2j)}(\hat p) := \min\{E_{2j_-}(\hat p), E_{2j_+}(\hat p)\}, \quad 1 \leqslant j < N, \end{aligned} \tag{3.6}$$

with $j_- := \max\{i < j \ \text{ s.t. } \ \bar E_{2i} > \bar E_{2j}\}$, $j_+ := \min\{i > j \ \text{ s.t. } \ \bar E_{2i} > \bar E_{2j}\}$.

Then, for $\hat p \in \hat D$ fixed, and for every $0 \leqslant i \leqslant 2N$, we can define *the action functions*

$$E \in \left(E_-^{(i)}(\hat p), E_+^{(i)}(\hat p)\right) \to I_1^{(i)}(E, \hat p) \tag{3.7}$$

by the standard Arnol'd–Liouville's formula

$$I_1^{(i)}(E, \hat p) := \frac{1}{2\pi} \oint_{\gamma_i} p_1 dq_1, \tag{3.8}$$

where

$$\gamma_i = \gamma_i(E; \hat p) := \mathtt{H}^{-1}(E; \hat p) \cap \mathcal{M}^i(\hat p) \tag{3.9}$$

is the smooth closed curve in the plane (q_1, p_1) with clockwise orientation[8)].

Finally, we denote by $\bar I_1^{(i)}(E)$ the action variables of the "unperturbed" Hamiltonian

$$\bar{\mathtt{H}} := \mathtt{H}|_{\mu=0} := p_1^2 + \bar{\mathtt{G}}(q_1), \tag{3.10}$$

and observe that $I_1^{(i)}(E, \hat p)$ reduces to $\bar I_1^{(i)}(E)$ for $\mu = 0$.

The main result of this section is the following

Theorem 1. *Let* $\mathtt{H}$ *be a Hamiltonian in standard form as in Definition* 2*, let* $\kappa \geqslant 4$ *be such that* (2.10) *holds and let* $2N$ *be the number of critical points of the reference potential* $\bar{\mathtt{G}}$*. Then, there exists a suitable constant* $\mathbf{c} = \mathbf{c}(n, \kappa) \geqslant 2^8 \kappa^3$ *such that, if*

$$\mu \leqslant 1/\mathbf{c}^2, \tag{3.11}$$

then, for all $0 \leqslant i \leqslant 2N$ *and* $\hat I \in \hat D$*, the action functions in* (3.7) *verify the following properties.*

(i) **(Universal behavior at critical energies)** *There exist functions* $\phi_-^i(z, \hat I)$, $\psi_-^i(z, \hat I)$ *for* $0 \leqslant i \leqslant 2N$*, and functions* $\phi_+^i(z, \hat I)$, $\psi_+^i(z, \hat I)$ *for* $0 < i < 2N$*, which are real-analytic in* $\{z \in \mathbb{C} : |z| < 1/\mathbf{c}\} \times \hat D_{\mathtt{r}/2}$ *and satisfy*

$$I_1^i\left(E_\mp^i(\hat I) \pm \varepsilon z,\ \hat I\right) = \phi_\mp^i(z, \hat I) + \psi_\mp^i(z, \hat I)\ z \log z, \quad \forall\ 0 < z < 1/\mathbf{c},\ \hat I \in \hat D. \tag{3.12}$$

On $\{z \in \mathbb{C} : |z| < 1/\mathbf{c}\} \times \hat D_{\mathtt{r}/2}$ *the functions* $\phi_\pm^i(z, \hat I)$, $\psi_\pm^i(z, \hat I)$ *satisfy:*

$$\begin{aligned} &\sup_{|z|<1/\mathbf{c},\, \hat I \in \hat D_{\mathtt{r}/2}} \left(|\phi_\pm^i| + |\psi_\pm^i|\right) \leqslant \mathbf{c}\sqrt{\varepsilon}, \\ &\sup_{|z|<1/\mathbf{c},\, \hat I \in \hat D_{\mathtt{r}/4}} \left(|\partial_{\hat I} \phi_\pm^i| + |\partial_{\hat I} \psi_\pm^i|\right) \leqslant \mathbf{c}\mu_{\mathrm{o}}, \qquad \mu_{\mathrm{o}} := \tfrac{\sqrt{\varepsilon}}{\mathtt{r}}\mu \overset{(2.8)}{\leqslant} 2^{-8}\mu. \end{aligned} \tag{3.13}$$

Moreover,

$$|\phi_\pm^i - \bar\phi_\pm^i|,\ |\psi_\pm^i - \bar\psi_\pm^i| \ \leqslant\ \mathbf{c}\sqrt{\varepsilon}\mu, \tag{3.14}$$

where $\bar\phi_\pm^i := \phi_\pm^i|_{\mu=0}$ *and* $\bar\psi_\pm^i := \psi_\pm^i|_{\mu=0}$.

[8)] For the external curves ($i = 0, 2N$) the orientation is to the right in $\mathcal{M}^{2N}(\hat p)$, to the left in $\mathcal{M}^0(\hat p)$.

(ii) **(Limiting critical values)** *The following bounds at the limiting critical energy values hold:*

$$
\begin{aligned}
&|\psi_+^i(0,\hat I)| \geqslant \sqrt{\varepsilon}/\mathbf{c}, \qquad 0<i<2N, \quad \forall\, \hat I\in \hat D_{\mathtt{r}/2},\\
&|\psi_-^{2j}(0,\hat I)| \geqslant \sqrt{\varepsilon}/\mathbf{c}, \qquad 0\leqslant j\leqslant N, \quad \forall\, \hat I\in \hat D_{\mathtt{r}/2},\\
&\psi_+^i(0,\hat I)>0, \qquad 0<i<2N, \quad \forall\, \hat I\in \hat D,\\
&\psi_-^{2j}(0,\hat I)<0, \qquad 0\leqslant j\leqslant N, \quad \forall\, \hat I\in \hat D,
\end{aligned}
\tag{3.15}
$$

while, in the case of relative minimal critical energies, one has $\forall\, \hat I\in\hat D$, $0<z<1/\mathbf{c}$,

$$
\phi_-^{2j-1}(0,\hat I)=0, \qquad \psi_-^{2j-1}(z,\hat I)=0, \qquad \forall\, 1\leqslant j\leqslant N. \tag{3.16}
$$

(iii) **(Estimates on derivatives of actions on real domains)** *The derivatives of the action functions on real domains satisfy the following estimates:*

$$
\inf_{(E_-^i,E_+^i)} \partial_E I_1^i \geqslant \frac{1}{\mathbf{c}\sqrt{\varepsilon}}, \qquad \forall\, \hat I\in\hat D,\ \forall\, 0<i<2N; \tag{3.17}
$$

$$
\min\left\{\partial_E I_1^{2N},\ \partial_E I_1^0\right\} \geqslant \frac{1}{\mathbf{c}\sqrt{E+\varepsilon}}, \quad \forall\, E>E_{2N},\ \forall\, \hat I\in\hat D. \tag{3.18}
$$

(iv) **(Estimates on derivatives of actions on complex domains and perturbative bounds)** *For $\lambda>0$ satisfying*

$$
\mathbf{c}\mu\leqslant\lambda\leqslant 1/\mathbf{c}, \tag{3.19}
$$

define the following complex energy-domains:

$$
\mathcal{E}_\lambda^i := \begin{cases}
\{E\in\mathbb{C} : \bar E_-^i-\varepsilon/\mathbf{c} \ < \operatorname{Re}E<\bar E_+^i-\lambda\varepsilon,\ |\operatorname{Im}E|<\varepsilon/\mathbf{c}\},\ i \text{ odd},\\
\{E\in\mathbb{C} : \bar E_-^i+\lambda\varepsilon< \operatorname{Re}E<\bar E_+^i-\lambda\varepsilon,\ |\operatorname{Im}E|<\varepsilon/\mathbf{c}\}, \quad 0,2N\neq i \text{ even},\\
\{E\in\mathbb{C} : \bar E_-^i+\lambda\varepsilon< \operatorname{Re}E<\bar E_+^i,\ |\operatorname{Im}E|<\varepsilon/\mathbf{c}\}, \qquad i=0,2N.
\end{cases}
\tag{3.20}
$$

Then, for $0\leqslant i\leqslant 2N$, the functions I_1^i and $\bar I_1^i$ are holomorphic on the domains $\mathcal{E}_\lambda^i\times\hat D_{\mathtt{r}}$ and satisfy the following estimates:

$$
\sup_{\mathcal{E}_\lambda^i\times\hat D_{\mathtt{r}/4}} |\partial_{\hat I} I_1^i| \leqslant \mathbf{c}^2\,\mu_{\mathrm{o}}, \quad \sup_{\mathcal{E}_\lambda^i}\left|\partial_E\bar I_1^i\right| \leqslant \mathbf{c}^2\,\frac{|\log\lambda|}{\sqrt{\varepsilon}}, \quad \sup_{\mathcal{E}_\lambda^i\times\hat D_{\mathtt{r}/2}}\left|\partial_E I_1^i-\partial_E\bar I_1^i\right| \leqslant \frac{\mathbf{c}^2\mu}{\lambda\sqrt{\varepsilon}}. \tag{3.21}
$$

Remark 3. (i) Statements similar to (3.12) have been also discussed in [20, 21, Lemma 7.2], and [4, Eq. (5.8)]. Analyticity at elliptic equilibria (see Eq. (3.16) above) was proven also in [20, Lemma 7.1].

(ii) Condition (3.11) implies the hypothesis of Lemma 1.

In the rest of the paper we shall use the following

Notation 1. *Given $m,M\geqslant 0$, we say that $m\lessdot M$ if there exists a constant $c=c(n,\kappa)\geqslant 1$ such that $m\leqslant cM$. We shall also say that a function f is of order M, $f=O(M)$, if $|f|\lessdot M$ uniformly on its domain of definition.*

Proof (*of Theorem* 1). For definiteness we consider the case of $i=2j+1$ odd and, in particular, the case with $E_{2j}(\hat p)<E_{2j+2}(\hat p)$. The other cases can be treated in the same way with the obvious changes.

Recalling (3.6), we note that

$$
E_+(\hat p) := E_+^{(2j+1)}(\hat p) = E_{2j}(\hat p), \qquad \bar E_+ := \bar E_+^{(2j+1)} = \bar E_{2j}. \tag{3.22}
$$

For every fixed $\hat{p} \in \hat{D}$ we denote by

$$E \in \Big(E_{2j+1}(\hat{p}), E_{2j}(\hat{p})\Big) \to \Theta_{\star}(E,\hat{p}), \text{ resp.,} \quad E \in \Big(E_{2j+1}(\hat{p}), E_{2j+2}(\hat{p})\Big) \to \Theta^{\star}(E,\hat{p}), \tag{3.23}$$

the (real-analytic) inverse of $\mathtt{G}(\hat{p}, q_1)$ on the interval $\Big(\theta_{2j}(\hat{p}),\theta_{2j+1}(\hat{p})\Big)$, respectively, $\Big(\theta_{2j+1}(\hat{p}),\theta_{2j+2}(\hat{p})\Big)$. As usual, a bar above functions means the limit $\mu = 0$, namely, $E \in (\bar{E}_{2j+1}, \bar{E}_{2j}) \to \bar{\Theta}_{\star}(E)$, respectively, $E \in (\bar{E}_{2j+1}, \bar{E}_{2j+2}) \to \bar{\Theta}^{\star}(E)$, will denote the (real-analytic) inverse of $\bar{\mathtt{G}}(q_1)$ on the interval $(\bar{\theta}_{2j}, \bar{\theta}_{2j+1})$, respectively $(\bar{\theta}_{2j+1}, \bar{\theta}_{2j+2})$. Then, the action function $\bar{I}(E) = \bar{I}_1^{(2j+1)}(E)$ of $\bar{\mathtt{H}}$ in (3.10) can be written as

$$\bar{I}(E) := \frac{1}{\pi}\int_{\bar{\Theta}_{\star}(E)}^{\bar{\Theta}^{\star}(E)} \sqrt{E - \bar{\mathtt{G}}(\theta)}\, d\theta, \qquad E \in (\bar{E}_{2j+1}, \bar{E}_{2j}), \tag{3.24}$$

so that

$$\partial_E \bar{I}(E) := \frac{1}{2\pi}\int_{\bar{\Theta}_{\star}(E)}^{\bar{\Theta}^{\star}(E)} \frac{d\theta}{\sqrt{E - \bar{\mathtt{G}}(\theta)}}. \tag{3.25}$$

We split the proof in four steps.

Step 1: *Explicit expression for the action functions*

In this first step we will obtain an expression analogous to (3.24) for $I(E) = I_1^{(2j+1)}(E)$ in (3.8), see formula (3.41) below; estimates (3.17) will then follow easily.

Let us consider the equation

$$p_1 = \frac{z}{\sqrt{1 + \nu(p, q_1)}}. \tag{3.26}$$

Note that by (2.8) and (3.11) we have

$$\mathrm{Re}\Big(1 + \nu(p, q_1)\Big) \geqslant \frac{1}{2}, \qquad \forall\ (p, q_1) \in D_{\mathtt{r}} \times \mathbb{T}^1_{\mathtt{s}}, \tag{3.27}$$

and, therefore, $\sqrt{1 + \nu(p, q_1)}$ is well defined on $D_{\mathtt{r}} \times \mathbb{T}^1_{\mathtt{s}}$.

Lemma 2. *There exists a unique real-analytic function* $\tilde{\mathcal{P}} : (-\mathtt{R}, \mathtt{R})_{\mathtt{r}/4} \times \mathbb{T}_{\mathtt{s}} \times \hat{D}_{\mathtt{r}} \to \mathbb{C}$ *satisfying the bound*

$$|\tilde{\mathcal{P}}|_{\dagger} := \sup_{(-\mathtt{R},\mathtt{R})_{\mathtt{r}/4} \times \mathbb{T}_{\mathtt{s}} \times \hat{D}_{\mathtt{r}}} |\tilde{\mathcal{P}}| \leqslant 2\mu\mathtt{R} \leqslant \frac{\mathtt{r}}{8}, \tag{3.28}$$

and such that

$$p_1 = \mathcal{P}(z, q_1, \hat{p}) := z + \tilde{\mathcal{P}}(z, q_1, \hat{p}) \tag{3.29}$$

solves (3.26)*:*

$$\mathcal{P}(z, q_1, \hat{p}) = \frac{z}{\sqrt{1 + \nu(\mathcal{P}(z, q_1, \hat{p}), \hat{p}, q_1)}}. \tag{3.30}$$

Moreover,

$$\mathcal{P} : (-\mathtt{R}, \mathtt{R})_{\mathtt{r}/4} \times \mathbb{T}_{\mathtt{s}} \times \hat{D}_{\mathtt{r}} \to (-\mathtt{R}, \mathtt{R})_{\mathtt{r}/2}. \tag{3.31}$$

Proof. We first note that, if $\tilde{\mathcal{P}}$ satisfies the first inequality in (3.28), then, by (2.10) and (3.11), it follows that it also satisfies the second one. Therefore, if $z \in (-\mathtt{R}, \mathtt{R})_{\mathtt{r}/4}$, then $z + \tilde{\mathcal{P}} \in (-\mathtt{R}, \mathtt{R})_{\mathtt{r}/2}$ and (3.31) holds. Let $\mathtt{B}$ denote the closed ball of functions $\tilde{\mathcal{P}}$ satisfying (3.28) and let $\tilde{\mathcal{P}} = \tilde{\mathcal{P}}(z, q_1, \hat{p})$ be the solution of the fixed point equation

$$\tilde{\mathcal{P}} = \Phi(\tilde{\mathcal{P}}) := \Big(\Big(1 + \nu(z + \tilde{\mathcal{P}}, \hat{p}, q_1)\Big)^{-\frac{1}{2}} - 1\Big)z. \tag{3.32}$$

By (2.8), (2.10) and (3.27), it follows that

$$|\Phi(\tilde{\mathcal{P}})|_\dagger \leqslant \mu(\mathtt{R}+\mathtt{r}/4) \leqslant 2\mu\mathtt{R},$$

and, therefore, $\Phi(\mathtt{B}) \subseteq \mathtt{B}$. In fact, Φ is a contraction: Omitting for brevity to write $\hat{p}, q_1$ and setting $\theta(t) = (1-t)\tilde{\mathcal{P}}' + t\tilde{\mathcal{P}}$, we get

$$\nu(z+\tilde{\mathcal{P}}) - \nu(z+\tilde{\mathcal{P}}') = \left(\tilde{\mathcal{P}} - \tilde{\mathcal{P}}'\right)\int_0^1 \partial_{p_1}\nu(z+\theta(t))dt.$$

Since $|\theta(t)|_\dagger \leqslant 2\mu\mathtt{R} \leqslant \mathtt{r}/8$ and $z+\theta(t) \in (-\mathtt{R},\mathtt{R})_{\mathtt{r}/2}$ for every $0 \leqslant t \leqslant 1$, by (2.8) and the Cauchy estimates we get $|\partial_{p_1}\nu\big(z+\theta(t)\big)|_\dagger \leqslant 2\mu/\mathtt{r}$ for any $0 \leqslant t \leqslant 1$. Then, by (2.10) and (3.11),

$$|\Phi(\tilde{\mathcal{P}}) - \Phi(\tilde{\mathcal{P}}')|_\dagger \leqslant 2\Big|\big(\nu(z+\tilde{\mathcal{P}},\hat{p},q_1) - \nu(z+\tilde{\mathcal{P}}',\hat{p},q_1)\big)z\Big|_\dagger \leqslant \frac{8\mu\mathtt{R}}{\mathtt{r}}|\tilde{\mathcal{P}} - \tilde{\mathcal{P}}'|_\dagger \leqslant \frac{1}{2}|\tilde{\mathcal{P}} - \tilde{\mathcal{P}}'|_\dagger,$$

and (3.32) is solved by the standard contraction lemma.

Thus, for real values of $\hat{p}$, q_1, E such that $0 \leqslant E - \mathtt{G}(\hat{p},q_1) \leqslant \mathtt{R}+\mathtt{r}/4$, we have that

$$p_1 = \mathcal{P}\Big(\pm\sqrt{E-\mathtt{G}(\hat{p},q_1)}, q_1, \hat{p}\Big) \quad \text{solves} \quad \mathtt{H}(p_1,\hat{p},q_1) = E, \tag{3.33}$$

where the sign depends on whether $\pm p_1 \geqslant 0$. By (3.28), (3.29) and the Cauchy estimates, for $z \in (-\mathtt{R},\mathtt{R})$,

$$\partial_z\mathcal{P} \geqslant \frac{1}{2}, \tag{3.34}$$

so that for real $q_1, \hat{p}$, the real function $z \in (-\mathtt{R},\mathtt{R}) \mapsto \mathcal{P}(z,q_1,\hat{p})$ is increasing. Note also that $\mathcal{P}(0,q_1,\hat{p}) = 0$.

Define the analytic function

$$\nu_\sharp(z,\theta,\hat{p}) := \frac{1}{2\sqrt{1+\nu\big(\mathcal{P}(z,\theta,\hat{p}),\hat{p},\theta\big)}} + \frac{1}{2\sqrt{1+\nu\big(\mathcal{P}(-z,\theta,\hat{p}),\hat{p},\theta\big)}} - 1. \tag{3.35}$$

Notice that $\nu_\sharp$ is *even* in z and that[9], by (3.31) and (2.8),

$$\sup_{z\in(-\mathtt{R},\mathtt{R})_{\mathtt{r}/4}} |\nu_\sharp(z,\hat{p},\theta)|_{\hat{D},\mathtt{r},\mathtt{s}} \leqslant \sup_{D_\mathtt{r}\times\mathbb{T}^1_\mathtt{s}} |\nu| \leqslant \mu. \tag{3.36}$$

Then, by (3.36), (2.10) and the Cauchy estimates we have

$$\sup_{z\in(-\mathtt{R},\mathtt{R})_{\mathtt{r}/8}} |z\partial_z\nu_\sharp|_{\hat{D},\mathtt{r},\mathtt{s}} \leqslant 16\kappa\mu. \tag{3.37}$$

We now need the following elementary

Lemma 3. *Let $g : (-r,r) \to \mathbb{R}$ be an even function with holomorphic extension on $[0,R]_r$. Then, one can define G holomorphic on $[0,R^2]_{r^2}$ so that $G(z^2) = g(z)$.*

Proof. Since g is even, it is actually holomorphic on $[-R,R]_r$. Denoting by $D_r(0) := \{|z| < r\}$, we have that, since g is holomorphic and even on $D_r(0)$, $g(z) = \sum_{j\geqslant 0} a_{2j}z^{2j}$, where the power series has a radius of convergence $\geqslant r$. Then $G(v) := \sum_{j\geqslant 0} a_{2j}v^j$ has radius of convergence $\geqslant r^2$. It remains to define G in the set $\Omega := [0,R^2]_{r^2} \setminus D_{r^2}(0)$. Notice that $\Omega \subset \mathbb{C}\setminus(-\infty,0]$; thus, we can define $G(v) := g(\sqrt{v})$ for $v \in \Omega$, noting that $z := \sqrt{v} \in [0,R]_r$. Indeed, if $v \in D_{r^2}(v_0^2)$, with $v_0 \in \mathbb{R}$, $v_0 > r$, then $\sqrt{v} \in D_r(v_0)$, and this is equivalent to[10] $D_{r^2}(v_0^2) \subseteq S\big(D_r(v_0)\big)$, where $S(v) := v^2$.

[9] If $\mathrm{Re}\,(1+\nu) \geqslant 1/2$ (see (3.27)), then $|(1+\nu)^{-1/2} - 1| \leqslant |\nu|$.

[10] This inclusion follows noting that, for every θ, we have $|(v_0+re^{\mathrm{i}\theta})^2 - v_0^2| \geqslant r^2$. The last inequality follows noting that it is equivalent to $|re^{\mathrm{i}2\theta} + 2v_0e^{\mathrm{i}\theta}| = |re^{\mathrm{i}\theta} + 2v_0| \geqslant r$, which follows from $v_0 > r$.

Since $\nu_\sharp$ is even in z, by Lemma 3 we can define the analytic function

$$\nu_\dagger(z^2,\hat{p},\theta) := \nu_\sharp(z,\hat{p},\theta), \tag{3.38}$$

which, by (3.36), satisfies

$$\sup_{v\in(0,\mathtt{R}^2)_{\mathtt{r}^2/16}} |\nu_\dagger(v,\hat{p},\theta)|_{\hat{D},\mathtt{r},\mathtt{s}} \leqslant \mu. \tag{3.39}$$

Moreover, since $v\partial_v\nu_\dagger(v,\hat{p},\theta) = \frac{1}{2}\sqrt{v}\partial_z\nu_\sharp(\sqrt{v},\hat{p},\theta)$, by (3.37) we get

$$\sup_{v\in(0,\mathtt{R}^2)_{\mathtt{r}^2/64}} |v\partial_v\nu_\dagger(v,\cdot,\cdot)|_{\hat{D},\mathtt{r},\mathtt{s}} \leqslant 8\kappa\mu. \tag{3.40}$$

In the following we will often omit to write the dependence upon $\hat{p}$.

In view of (3.23), (3.33), (3.30), (3.35) and (3.38), we can write $I(E) = I_1^{(2j+1)}(E,\hat{p})$ in (3.8) as

$$\begin{aligned} I(E) &= \frac{1}{2\pi}\int_{\Theta_\star(E)}^{\Theta^\star(E)} \Big[\mathcal{P}\Big(\sqrt{E-\mathtt{G}(\theta)},\theta\Big) - \mathcal{P}\Big(-\sqrt{E-\mathtt{G}(\theta)},\theta\Big)\Big]\, d\theta \\ &= \frac{1}{\pi}\int_{\Theta_\star(E)}^{\Theta^\star(E)} \sqrt{E-\mathtt{G}(\theta)}\Big(1+\nu_\sharp\Big(\sqrt{E-\mathtt{G}(\theta)},\theta\Big)\Big)\, d\theta \\ &= \frac{1}{\pi}\int_{\Theta_\star(E)}^{\Theta^\star(E)} \sqrt{E-\mathtt{G}(\theta)}\Big(1+\nu_\dagger\Big(E-\mathtt{G}(\theta),\theta\Big)\Big)\, d\theta. \end{aligned} \tag{3.41}$$

Recalling the definition of $\nu_\dagger$ in (3.38), we set

$$\tilde{\nu}(v) = \tilde{\nu}(v,\hat{p},\theta) := \nu_\dagger(v) + 2v\partial_v\nu_\dagger(v), \tag{3.42}$$

which, by (3.39) and (3.40), satisfies

$$\sup_{v\in(0,R_0^2)_{r_0^2/64}} |\tilde{\nu}(v)|_{\hat{D},r_0,s_0} \leqslant 17\kappa\mu. \tag{3.43}$$

From (3.41) and (3.42) there follows

$$\partial_E I(E) = \frac{1}{2\pi}\int_{\Theta_\star(E)}^{\Theta^\star(E)} \frac{1}{\sqrt{E-\mathtt{G}(\theta)}}\Big(1+\tilde{\nu}\Big(E-\mathtt{G}(\theta),\theta\Big)\Big)\, d\theta. \tag{3.44}$$

Now, note that by (2.8) and (2.10) for real values of θ (and $\hat{p}$)

$$\mathtt{G}(\theta_{2j+1}+\theta) - \mathtt{G}(\theta_{2j+1}) = \mathtt{G}(\theta_{2j+1}+\theta) - E_{2j+1} \lessdot \varepsilon\theta^2.$$

Thus, for $E_{2j+1} < E < E_{2j}$ we get

$$\frac{1}{\sqrt{\varepsilon}}\sqrt{E-E_{2j+1}} \lessdot \Theta^\star(E) - \theta_{2j+1},\ \theta_{2j+1} - \Theta_\star(E).$$

Finally, by (3.44) and (3.43), we see that

$$\frac{1}{\sqrt{\varepsilon}} \lessdot \frac{1}{4\pi}\int_{\Theta_\star(E)}^{\Theta^\star(E)} \frac{1}{\sqrt{E-E_{2j+1}}}\, d\theta \leqslant \partial_E I(E),$$

proving (3.17).

The proof of (3.18) is completely analogous.

Step 2: *Normal forms close to hyperbolic and elliptic equilibria*

By Definition 2, (2.10), (3.2), (3.11), and the Cauchy estimates one has

$$\sup_{\hat{p}\in\hat{D}_\mathtt{r}} |\operatorname{Im}\theta_{2j}(\hat{p})| \leqslant \tfrac{2\varepsilon\mu}{\beta\mathtt{s}} \leqslant \tfrac{\mathtt{s}}{8}, \qquad \sup_{p\in D_{3\mathtt{r}/4}} |p_1\partial_{\hat{p}}\theta_{2j}(\hat{p})| \leqslant (\mathtt{R}+\mathtt{r})\tfrac{8\varepsilon\mu}{\beta\mathtt{s}\mathtt{r}} \leqslant \tfrac{\mathtt{s}}{8}. \tag{3.45}$$

Then, by (2.8) and (3.45), the functions

$$\nu_*(p,q_1) := \nu\big(p, q_1+\theta_{2j}(\hat p)\big), \quad \mathtt{G}_*(\hat p,q_1) := \mathtt{G}\big(\hat p, q_1+\theta_{2j}(\hat p)\big), \quad \bar{\mathtt{G}}_*(q_1) := \bar{\mathtt{G}}(q_1+\bar\theta_{2j}) \tag{3.46}$$

satisfy

$$\begin{aligned}&\mathtt{G}_*(\hat p,0) = E_{2j}(\hat p), \qquad \partial_{q_1}\mathtt{G}_*(\hat p,0)=0,\\ &\sup_{\hat D_{\mathtt r}\times\mathbb{T}^1_{7\mathtt{s}/8}}|\nu_*|\leqslant\mu, \qquad \sup_{\mathbb{T}^1_{\mathtt s}}|\bar{\mathtt{G}}_*|\leqslant\varepsilon, \qquad \sup_{\hat D_{\mathtt r}\times\mathbb{T}^1_{7\mathtt{s}/8}}|\mathtt{G}_*-\bar{\mathtt{G}}_*|\leqslant 17\kappa^3\varepsilon\mu.\end{aligned} \tag{3.47}$$

In particular, the last estimate follows since for $(\hat p,q_1)\in\hat D_{\mathtt r}\times\mathbb{T}^1_{7\mathtt{s}/8}$ one has

$$\begin{aligned}|\mathtt{G}_*(\hat p,q_1)-\bar{\mathtt{G}}_*(q_1)| &\leqslant |\mathtt{G}_*(\hat p,q_1)-\bar{\mathtt{G}}(q_1+\theta_{2j}(\hat p))| + |\bar{\mathtt{G}}\big(q_1+\theta_{2j}(\hat p)\big)-\bar{\mathtt{G}}(q_1+\bar\theta_{2j})|\\ &\leqslant \varepsilon\mu + \tfrac{8\varepsilon}{\mathtt s}\tfrac{2\varepsilon\mu}{\beta\mathtt s}\leqslant 17\kappa^3\varepsilon\mu\end{aligned}$$

by (3.2), (2.10) and the Cauchy estimates. Again, by the Cauchy estimates, (2.10) and (3.11) we get

$$\sup_{\hat p\in\hat D_{\mathtt r}}|\partial^2_{q_1}\mathtt{G}_*(\hat p,0)-\partial^2_{q_1}\bar{\mathtt{G}}_*(0)|\leqslant 2^6\kappa^3\varepsilon\mu\mathtt{s}^{-2}\leqslant 2^6\kappa^6\mu\beta\leqslant 2^{-10}\beta. \tag{3.48}$$

By (2.7), $\bar{\mathtt{G}}_*$ is β-Morse (and $\langle\bar{\mathtt{G}}_*\rangle=0$); in particular, it has a maximum at $q_1=0$ and, by (2.4), $-\partial^2_{q_1}\bar{\mathtt{G}}_*(0)\geqslant\beta$. Recalling (2.8) and (3.48), for $\hat p\in\hat D_{\mathtt r}$, we see that

$$\begin{aligned}&\sqrt{\beta/2}\leqslant\bar\tau := \sqrt{-\partial^2_{q_1}\bar{\mathtt{G}}_*(0)/2}\leqslant\sqrt{\varepsilon}/\mathtt{s}, \qquad \tau(\hat p) := \sqrt{-\partial^2_{q_1}\mathtt{G}_*(\hat p,0)/2},\\ &|\tau(\hat p)-\bar\tau|\leqslant 2^6\kappa^6\mu\sqrt{\beta}\leqslant 2^{-10}\sqrt{\beta}, \qquad \frac{2}{3}\sqrt{\beta}\leqslant|\tau(\hat p)|\leqslant 2\kappa\sqrt{\varepsilon}.\end{aligned} \tag{3.49}$$

Furthermore, by (3.47), (3.11), and by (3.49), (3.47) and (2.10), we get for all $\hat p\in\hat D_{\mathtt r}$:

$$\begin{aligned}&\bar\delta := \sqrt{\bar\tau},\ \delta(\hat p) := \tfrac{\sqrt{\tau(\hat p)}}{\sqrt[4]{1+\nu_*(0,\hat p,0)}}, \qquad \tfrac12\beta^{1/4}\leqslant|\delta(\hat p)|\leqslant\kappa\varepsilon^{1/4},\ |\delta-\bar\delta|\lessdot\mu,\\ &\bar g := \bar\tau,\ g(\hat p) := \sqrt{1+\nu_*(0,\hat p,0)}\,\tau(\hat p),\ \frac{\sqrt{\beta}}{3}\leqslant|g(\hat p)|\leqslant 4\kappa\sqrt{\varepsilon}\ |g-\bar g|\lessdot\sqrt{\varepsilon}\mu.\end{aligned} \tag{3.50}$$

The normal form close to hyperbolic equilibria $(p_1,q_1)=\big(0,\theta_{2j}(\hat p)\big)$ is detailed in the following

Proposition 2. *There exist positive constants $\mathbf{c_0},\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3$, depending only on κ,n and satisfying $0<\mathbf{c}_1<\mathbf{c_0}/8n\mathbf{c}_2$, such that the following holds. There exist a (near-identity) real-analytic symplectic transformation*

$$\Phi_{\rm hp}:(y,x)\in\{|y_1|<\mathbf{c}_1\varepsilon^{1/4}\}\times\hat D_{\mathtt{r}/2}\times\{|x_1|<\mathbf{c}_1\varepsilon^{1/4}\}\times\mathbb{T}^{n-1}_{\mathtt{s}/2}\ \longrightarrow\ (p,q)\in\ D_{\mathtt{r},\mathtt{s}} \tag{3.51}$$

and a function $R_{\rm hp}(z,\hat y)$ with

$$\sup_{|z|\leqslant 2\mathbf{c}_1^2,\ \hat y\in\hat D_{\mathtt{r}/2}}|R_{\rm hp}(z,\hat y)|\leqslant\mathbf{c}_2, \qquad R_{\rm hp}(0,\hat y)=0, \quad \partial_z R_{\rm hp}(0,\hat y)=0, \tag{3.52}$$

such that

$$\mathtt{H}_{\rm hp}(y,x_1) := \mathtt{H}\circ\Phi_{\rm hp}(y,x) = E_{2j}(\hat y)+g(\hat y)(y_1^2-x_1^2)+\varepsilon R_{\rm hp}\left(\tfrac{y_1^2-x_1^2}{\sqrt{\varepsilon}},\hat y\right). \tag{3.53}$$

Moreover, $\Phi_{\rm hp}$ has the form

$$\begin{aligned}&p_1=\delta(\hat y)\Big(y_1+\varepsilon^{1/4}a_1(\varepsilon^{-1/4}y_1,\hat y,\varepsilon^{-1/4}x_1)\Big), \quad \hat p=\hat y,\\ &q_1=\theta_{2j}(\hat y)+\frac{1}{\delta(\hat y)}\Big(x_1+\varepsilon^{1/4}a_2(\varepsilon^{-1/4}y_1,\hat y,\varepsilon^{-1/4}x_1)\Big), \quad \hat q=\hat x+\hat a(y,x_1),\end{aligned} \tag{3.54}$$

for suitable holomorphic functions $a_1, a_2, \hat{a}$, such that

$$\sup_{W_{\mathbf{c_0},\mathbf{r}}} |a_i| \leqslant \mathbf{c}_2, \qquad W_{\mathbf{c_0},\mathbf{r}} := \{|\tilde{y}_1| < \mathbf{c_0}/2\} \times \hat{D}_{\mathbf{r}/2} \times \{|\tilde{x}_1| < \mathbf{c_0}/2\}, \tag{3.55}$$

and are at least quadratic in $\tilde{y}_1, \tilde{x}_1$. Moreover, denoting $\bar{R}_{\rm hp} := R_{\rm hp}|_{\mu=0}$, we have

$$|R_{\rm hp} - \bar{R}_{\rm hp}| = O(\mu). \tag{3.56}$$

Finally, for every $\hat{y} \in \hat{D}_{\mathbf{r}/2}$, the image of the restriction of the map in (3.54)

$$(y_1, x_1) \in \{|y_1| < \mathbf{c}_1 \varepsilon^{1/4}\} \times \{|x_1| < \mathbf{c}_1 \varepsilon^{1/4}\} \to (p_1, q_1)$$

contains the (*complex*) *set*

$$\{|p_1| \leqslant 2\mathbf{c}_3 \sqrt{\varepsilon}\} \times \{|q_1 - \theta_{2j}(\hat{p})| \leqslant 2\mathbf{c}_3\}. \tag{3.57}$$

The normal form close to elliptic equilibrium $(p_1, q_1) = \big(0, \theta_{2j+1}(\hat{p})\big)$ is detailed in the following

Proposition 3. *There exist a* (*near-identity*) *real-analytic symplectic transformation $\Phi_{\rm el}$ as in* (3.51) *and* (3.54) *and a function $R_{\rm el}(z, \hat{y})$ as in* (3.52) *and* (3.56) *such that*

$$\mathtt{H}_{\rm el}(y, x_1) := \mathtt{H} \circ \Phi_{\rm el}(y, x) = E_{2j+1}(\hat{y}) + g(\hat{y})(y_1^2 + x_1^2) + \varepsilon R_{\rm el}\left(\tfrac{y_1^2 + x_1^2}{\sqrt{\varepsilon}}, \hat{y}\right). \tag{3.58}$$

Finally, for every $\hat{y} \in \hat{D}_{\mathbf{r}/2}$, the image of the restriction

$$(y_1, x_1) \in \{|y_1| < \mathbf{c}_1 \varepsilon^{1/4}\} \times \{|x_1| < \mathbf{c}_1 \varepsilon^{1/4}\} \to (p_1, q_1)$$

contains the (*complex*) *set*

$$\{|p_1| \leqslant 2\mathbf{c}_3 \sqrt{\varepsilon}\} \times \{|q_1 - \theta_{2j+1}(\hat{p})| \leqslant 2\mathbf{c}_3\}. \tag{3.59}$$

The proof of Proposition 2 is rather standard; for completeness it is included in Appendix A.

The proof of Proposition 3 is completely analogous[11)] and is omitted.

Step 3: *The action functions close to the elliptic equilibrium*

Setting $I_1 := (y_1^2 + x_1^2)/2$ and $\hat{I} := \hat{y}$, we have that, by Proposition 3, the function

$$\mathtt{E}(I) := E_{2j+1}(\hat{I}) + 2g(\hat{I}) I_1 + \varepsilon R_{\rm el}\left(\tfrac{2I_1}{\sqrt{\varepsilon}}, \hat{I}\right) \tag{3.60}$$

is well defined and holomorphic for $|I_1| \leqslant \mathbf{c}_1^2 \sqrt{\varepsilon}/2$ and $\hat{I} \in \hat{D}_{\mathbf{r}/2}$. Since $R_{\rm el}$ is (at least) quadratic in its first entry, by (3.50) and (2.10), recalling (3.6), (3.2) and taking $\mathbf{c}$ large enough in (3.11), we see that, for a suitable constant $0 < \mathbf{c}_4 < \mathbf{c}_3$ depending only on n and κ, we can invert the expression $\mathtt{E}(I) = E$ finding $I_1(E, \hat{I})$ which solves

$$\mathtt{E}\big(I_1(E, \hat{I}), \hat{I}\big) = E, \qquad \text{for} \quad |E - \bar{E}_-| < \mathbf{c}_4 \varepsilon, \quad \hat{I} \in \hat{D}_{\mathbf{r}/2}.$$

It turns out that the function I_1 above is exactly the action function introduced in (3.8). Indeed, since the map $\Phi_{\rm el}$ in (3.58) is symplectic, for every $|E - \bar{E}_-| < \mathbf{c}_4 \varepsilon$ the area enclosed by the level curve $\gamma_{2j+1}(E; \hat{I})$ in (3.9) is equal to the one included by the level curve $\mathtt{H}_{\rm el} = E$, which is simply the circle $\frac{1}{2}(x_1^2 + y_1^2) = I_1(E)$.

Hence, formula (3.16) (i. e., the analyticity of action as a function of energy close to a minimum and $I_1(E_-) = 0$) follows.

Step 4: *Away from the elliptic equilibrium*

Here, we will often omit to write the dependence on the dumb actions $\hat{p} = \hat{I} = \hat{y}$.

[11)] Indeed, in one dimension, from a complex point of view, the Birkhoff normal form is the same both in the hyperbolic and in the elliptic case.

Let us consider the action function $I(E)$ defined in (3.41) for

$$|E - \bar{E}_-| \geqslant \mathbf{c}_4 \varepsilon. \tag{3.61}$$

By Green's theorem,

$$I(E) = \frac{1}{2\pi} \int_{\Omega(E)} dq_1 dp_1, \tag{3.62}$$

where $\Omega(E)$ is the bounded portion of plane encircled by the curve $\gamma_{2j+1}(E)$ defined in (3.9), namely,[12)]

$$\Omega(E) = \Big\{ (p_1, q_1) \mid \Theta_\star(E) < q_1 < \Theta^\star(E), \tag{3.63}$$
$$\mathcal{P}\Big(-\sqrt{E - \mathtt{G}(q_1)}, q_1 \Big) \leqslant p_1 \leqslant \mathcal{P}\Big(\sqrt{E - \mathtt{G}(q_1)}, q_1 \Big) \Big\}.$$

Consider first the case in which we are away also from the hyperbolic equilibrium. By (3.41) we have

$$\int_{\Omega(E)} dq_1 dp_1 = 2 \int_{\Theta_\star(E)}^{\Theta^\star(E)} \sqrt{E - \mathtt{G}(\theta)} \Big(1 + \nu_\dagger \big(E - \mathtt{G}(\theta), \theta \big) \Big) \, d\theta,$$

which contributes to $I(E)$ with a holomorphic[13)] and bounded (by some constant depending only on κ and n) term.

Let us finally consider the case close to the hyperbolic point. Recalling (3.53), let us consider the equation

$$E_{2j}(\hat{y}) + g(\hat{y}) J + \varepsilon R_{\rm hp} \left(\tfrac{J}{\sqrt{\varepsilon}}, \hat{y} \right) = E. \tag{3.64}$$

By the inverse function theorem we construct a holomorphic function $F(z, \hat{y})$ with

$$\sup_{|z| < \mathbf{c}_5, \hat{D}_{\mathtt{r}/2}} |F(z, \hat{y})| \leqslant 1/2\mathbf{c}_5, \tag{3.65}$$

for $0 < \mathbf{c}_5 < \mathbf{c}_4 < \mathbf{c}_3$ small enough depending only on κ, n, such that the equation in (3.64) is solved by[14)]

$$J(E, \hat{y}) := -\sqrt{\varepsilon} J_{\rm hp} \left(\tfrac{E_{2j}(\hat{y}) - E}{\varepsilon}, \hat{y} \right), \qquad J_{\rm hp}(z, \hat{y}) := \tfrac{\sqrt{\varepsilon}}{g(\hat{y})} z \left(1 + z F(z, \hat{y}) \right), \tag{3.66}$$

where $J_{\rm hp}$ solves the equation

$$\frac{g(\hat{y})}{\sqrt{\varepsilon}} J_{\rm hp} - R_{\rm hp}(-J_{\rm hp}, \hat{y}) = z. \tag{3.67}$$

Recalling (3.66), we set

$$x_1(E, y) := \sqrt{-J(E, \hat{y}) + y_1^2}, \tag{3.68}$$

then, by (3.64) and (3.53),

$$\mathtt{H}_{\rm hp}\Big(y, x_1(E, y) \Big) \equiv E. \tag{3.69}$$

Fix $\bar{x}_1 := \mathbf{c}_5 \varepsilon^{1/4}$ and define

$$\bar{y}_1 = \bar{y}_1(E, \hat{y}) := \sqrt{J(E, \hat{y}) + \bar{x}_1^2}. \tag{3.70}$$

---

12) The definition of $\mathcal{P}$ is given in Lemma 2.

13) Notice that there is no problem in $\Theta_\star(E)$, $\Theta^\star(E)$ where the square root vanishes. Actually, close to these points it is convenient to write $\Omega(E)$ as a normal set with respect to q_1 and not to p_1.

14) For real values of $\hat{y}$ and E we are in the case $E < E_{2j}(\hat{y})$, namely, $z > 0$.

Consider the holomorphic functions $q_1 = q_1(y, x_1)$ and $p_1 = p_1(y, x_1)$ defined in (3.54). Set

$$\bar{p}_1^{\pm}(E, \hat{y}) := p_1\big(\pm \bar{y}_1(E, \hat{y}), \hat{y}, \bar{x}_1\big), \qquad \bar{q}_1^{\pm}(E, \hat{y}) := q_1\big(\pm \bar{y}_1(E, \hat{y}), \hat{y}, \bar{x}_1\big). \tag{3.71}$$

Note that[15)]

$$\bar{p}_1^{\pm}(E) = \mathcal{P}\Big(\pm \sqrt{E - \mathtt{G}\big(\bar{q}_1^{\pm}(E)\big)}, \bar{q}_1^{\pm}(E)\Big). \tag{3.72}$$

For every fixed $\hat{y}$ (which we will omit to write) we invert the expression $p_1 = p_1(y_1, \bar{x}_1)$, with $|y| < \mathbf{c}_1 \varepsilon^{1/4}$, finding a holomorphic function $\tilde{y}_1(p_1)$ such that

$$p_1 = p_1\big(\tilde{y}_1(p_1), \bar{x}_1\big), \qquad \bar{y}_1^{\pm}(E) = \tilde{y}_1\big(\bar{p}_1^{\pm}(E)\big). \tag{3.73}$$

Set $\tilde{q}_1(p_1) := q_1(\tilde{y}_1(p_1), \bar{x}_1)$. For real value of E and $\hat{p}$, we split the integral in (3.62) into two parts

$$\int_{\Omega(E)} dq_1 dp_1 = \int_{\Omega_1(E)} dq_1 dp_1 + \int_{\Omega_2(E)} dq_1 dp_1, \tag{3.74}$$

where

$$\Omega_1(E) := \{(p_1, q_1) \in \Omega(E) \mid q_1 \leqslant \tilde{q}_1(p_1),\ \bar{p}_1^- \leqslant p_1 \leqslant \bar{p}_1^+\}, \quad \Omega_2(E) := \Omega(E) \setminus \Omega_2(E).$$

As above, the term $\int_{\Omega_2(E)} dq_1 dp_1$ contributes to $I(E)$ with a holomorphic and bounded term.

Recalling (3.53) and (3.57) and setting $\tilde{\Omega}_1(E) := \Phi_{\mathrm{hp}}^{-1}\big(\Omega_1(E)\big)$, we have that

$$\int_{\Omega_1(E)} dq_1 dp_1 = \int_{\tilde{\Omega}_1(E)} dx_1 dy_1. \tag{3.75}$$

Note that by the above construction

$$\tilde{\Omega}_1(E) = \{(y_1, x_1)\ :\ -\bar{y}_1(E) \leqslant y_1 \leqslant \bar{y}_1(E),\ x_1(E, y_1) \leqslant x_1 \leqslant \bar{x}_1\},$$

then

$$\int_{\tilde{\Omega}_1(E)} dx_1 dy_1 = 2\bar{x}_1 \bar{y}_1(E) - \int_{-\bar{y}_1}^{\bar{y}_1} x_1(E, y_1)\, dy_1. \tag{3.76}$$

On the other hand, by (3.68) and (3.70)

$$\begin{aligned}
\int_{-\bar{y}_1}^{\bar{y}_1} x_1(E, y_1)\, dy_1 &= 2\int_0^{\bar{y}_1} \sqrt{-J(E) + y_1^2}\, dy_1 \\
&= \bar{y}_1 \sqrt{-J(E) + \bar{y}_1^2} - J(E)\Big(\log \tfrac{\sqrt{-J(E) + \bar{y}_1^2} + \bar{y}_1}{\varepsilon^{1/4}} - \log \tfrac{\sqrt{-J(E)}}{\varepsilon^{1/4}}\Big) \\
&= \bar{y}_1 \bar{x}_1 \ - J(E) \log \tfrac{\bar{x}_1 + \bar{y}_1}{\varepsilon^{1/4}} + \tfrac{1}{2} J(E) \log \tfrac{-J(E)}{\sqrt{\varepsilon}}.
\end{aligned}$$

Note that the first two terms above are holomorphic functions of E up to $E = 0$; instead, the term $\frac{1}{2} J(E) \log(-J(E)/\sqrt{\varepsilon})$ contains the singular term. Recalling (3.6) and setting $z = (E_{2j} - E)/\varepsilon = (E_+ - E)/\varepsilon$, the last term is transformed into (recalling (3.66))

$$\begin{aligned}
\tfrac{1}{2} J(E) \log \tfrac{-J(E)}{\sqrt{\varepsilon}} &= -\tfrac{\sqrt{\varepsilon}}{2} J_{\mathrm{hp}}(z) \log J_{\mathrm{hp}}(z) \\
&= -\tfrac{\sqrt{\varepsilon}}{2} J_{\mathrm{hp}}(z) \log\Big(\tfrac{\sqrt{\varepsilon}}{g(\hat{y})}\,(1 + zF(z, \hat{y}))\Big) - \tfrac{\sqrt{\varepsilon}}{2} J_{\mathrm{hp}}(z) \log z,
\end{aligned}$$

where the last term is the singular one, namely,

$$-\tfrac{\sqrt{\varepsilon}}{2} J_{\mathrm{hp}}(z) \log z = -\tfrac{\varepsilon}{2g(\hat{y})}\,(1 + zF(z, \hat{y}))\ z \log z.$$

Recalling (3.62), (3.74), (3.75) and (3.76), this implies that the singular term in $I(E_+ - \varepsilon z)$ in (3.12) is

$$\tfrac{\varepsilon}{4\pi g(\hat{y})}\,(1 + zF(z, \hat{y}))\ z \log z,$$

15) Omitting $\hat{y}$.

namely,

$$\psi_+^{(2j+1)}(z) = \tfrac{\varepsilon}{4\pi g(\hat{y})}\left(1 + zF(z,\hat{y})\right). \tag{3.77}$$

This proves (3.12).

By taking $\mathbf{c}$ in (3.11) large enough, by (3.65) and (3.50) the first estimate in (3.13) and the first and third estimates in (3.15) follow for $\psi_+^{(2j+1)}$.

Now consider the corresponding functions when $\mu = 0$ (namely, $\bar{I}, \bar{J}_{\rm hp}, \bar{\psi}_+^{(2j+1)}$, etc.). Observe, in particular[16), that

$$\bar{J}_{\rm hp}(z) := \tfrac{\sqrt{\varepsilon}}{\bar{g}} z\left(1 + z\bar{F}(z)\right)$$

is the solution of the equation

$$\tfrac{1}{\sqrt{\varepsilon}}\,\bar{g}\bar{J}_{\rm hp} - \bar{R}_{\rm hp}(-\bar{J}_{\rm hp}) = z \tag{3.78}$$

corresponding to (3.67). Then, recalling (3.66),

$$\begin{aligned} z\left(1 + z\bar{F}(z)\right) - \bar{R}_{\rm hp}(-\bar{J}_{\rm hp}) &= \tfrac{\bar{g}}{\sqrt{\varepsilon}}\bar{J}_{\rm hp} - \bar{R}_{\rm hp}(-\bar{J}_{\rm hp}) = \tfrac{g(\hat{y})}{\sqrt{\varepsilon}}J_{\rm hp} - R_{\rm hp}(-J_{\rm hp},\hat{y}) \\ &= z\left(1 + zF(z,\hat{y})\right) - R_{\rm hp}(-J_{\rm hp},\hat{y}). \end{aligned}$$

By (3.50) and (3.56), we get

$$|F - \bar{F}| \lessdot \mu, \qquad |J_{\rm hp} - \bar{J}_{\rm hp}| \lessdot \mu. \tag{3.79}$$

Since the unperturbed singular term is

$$\bar{\psi}_+^{(2j+1)}(z) = \tfrac{\varepsilon}{4\pi\bar{g}}\left(1 + z\bar{F}(z)\right)$$

by (3.77), (3.50) and (3.79) we get the second estimate in (3.14) in the + case when $i = 2j+1$; the other cases are analogous.

Since $\bar{\psi}_+^{(2j+1)}$ is independent of $\hat{I}$, by (3.14) and the Cauchy estimates, we get the second estimate in (3.13) for $\psi_+^{(2j+1)}$; the other estimates are analogous.

It remains to prove (3.21). In proving (3.21) we consider only the crucial zone close to maximal energies; in particular, we consider the domain

$$\mathcal{E}^i_\lambda \cap \{|E - \bar{E}_+| \leqslant \varepsilon/2\mathbf{c}\}. \tag{3.80}$$

Indeed, in the other parts the estimates are simpler and can be directly derived from the representation formula (3.25), (3.44) and the estimate (3.43); noting also that, by (3.61) and (2.8), the function $\Theta_\star(E)$, resp. $\Theta^\star(E)$, and $\bar{\Theta}_\star(E)$, resp. $\bar{\Theta}^\star(E)$, are close:

$$|\Theta_\star(E) - \bar{\Theta}_\star(E)|,\ |\Theta^\star(E) - \bar{\Theta}^\star(E)| \lessdot \varepsilon\mu.$$

Let us consider the domain in (3.80), where we can use the representation (3.12) and estimates in (3.13)–(3.14). The first and second estimate in (3.21) directly follow from (3.13). Let us now consider the third estimate in (3.21). Denote

$$\begin{cases} z := (E_{2j} - E)/\varepsilon, \quad z_* := (\bar{E}_{2j} - E)/\varepsilon, \\ f(z) := \phi'(z) + \psi'(z) z\log z + \psi(z)(1 + \log z), \\ \bar{f}(z) := \bar{\phi}'(z) + \bar{\psi}'(z) z\log z + \bar{\psi}(z)(1 + \log z) \end{cases}$$

and observe that, by (3.2), $|z - z_*| = |E_{2j} - \bar{E}_{2j}|/\beta \lessdot \mu$. Then, recalling (3.80),

$$\varepsilon\left|\partial_E I_1^{(2j)}(E) - \partial_E \bar{I}_1^{(2j)}(E)\right| = |f(z) - \bar{f}(z_*)| \lessdot \sqrt{\varepsilon}\mu/\lambda$$

by (3.13), (3.14) and the Cauchy estimates. The proof of Theorem 1 is complete. □

16) Recall (3.66).

4. THE COMPLEX ARNOL'D–LIOUVILLE TRANSFORMATION

In this section we discuss the complex properties (including analyticity radii) of the Arnol'd–Liouville transformation, which allow, in particular, an upper bound to be given on the derivatives of the energy functions in a complex domain.

For every fixed $\hat{p} \in \hat{D}$, given $I_1^{(i)}(E, \hat{p})$ as in (3.8), the action function $(p_1, q_1) \to I_1^{(i)}(\mathtt{H}(p, q_1), \hat{p})$ can be symplectically completed[17] with the angular term $(p_1, q_1) \to \varphi_1^{(i)}(p, q_1)$. We shall call $\check{\Phi}^i = \check{\Phi}^i(I, \varphi_1)$ the inverse of the map

$$(p, q_1) \to (I, \varphi_1) := \Big(I_1^{(i)}(\mathtt{H}(p, q_1), \hat{p}), \hat{p}, \varphi_1^{(i)}(p, q_1)\Big).$$

Note that the Arnol'd–Liouville "suspended" symplectic transformation $\check{\Phi}^i$ integrates $\mathtt{H}$, i. e.,

$$\mathtt{H} \circ \check{\Phi}^i(I, \varphi_1) = \mathtt{E}^{(i)}(I), \tag{4.1}$$

where $\mathtt{E}^{(i)}$ is the inverse of $I_1^{(i)}$, namely,

$$\mathtt{E}^{(i)}\Big(I_1^{(i)}(E, \hat{I}), \hat{I}\Big) = E. \tag{4.2}$$

Next, we introduce suitable decreasing subdomains $\mathcal{B}^i(\lambda)$ of $\mathcal{B}^i$ depending on a nonnegative parameter λ so that $\mathcal{B}^i(0) = \mathcal{B}^i$ and such that the map Φ^i has, for positive λ, a holomorphic extension on a suitable complex neighborhood of $\mathcal{B}^i(\lambda) \times \mathbb{T}^n$.

Define

$$\lambda_{\max} = \lambda_{\max}(\hat{I}) := \Big(E_+(\hat{I}) - E_-(\hat{I})\Big)/\varepsilon, \qquad \bar{\lambda}_{\max} := \Big(\bar{E}_+ - \bar{E}_-\Big)/\varepsilon. \tag{4.3}$$

Notice that, by (2.10), the definition of β and (2.8), one has

$$\frac{1}{\kappa} \leqslant \frac{\beta}{\varepsilon} \leqslant \bar{\lambda}_{\max} \leqslant 2; \tag{4.4}$$

notice also that, by (3.2), we have

$$|\lambda_{\max} - \bar{\lambda}_{\max}| \leqslant 6\kappa^3\mu, \tag{4.5}$$

so that, since $\mu \leqslant 1/\mathbf{c}^2$ and $\mathbf{c} \geqslant 2^8\kappa^3$ (compare Theorem 1), one has

$$\lambda_{\max} \geqslant 1/2\kappa. \tag{4.6}$$

Next, for $0 \leqslant \lambda \leqslant \lambda_{\max}$ define:

$$\begin{aligned}
&a_-^{(2j-1)}(\hat{I}; \lambda) := 0, \ \forall 1 \leqslant j \leqslant N,\\
&a_-^{(2j)}(\hat{I}; \lambda) := I_1^{(2j)}\Big(E_-^{(2j)}(\hat{I}) + \lambda\varepsilon, \hat{I}\Big), \ \forall 0 \leqslant j \leqslant N,\\
&a_+^{(i)}(\hat{I}; \lambda) := I_1^{(i)}\Big(E_+^{(i)}(\hat{I}) - \lambda\varepsilon, \hat{I}\Big), \ \forall 0 < i < 2N,\\
&a_+^{(i)}(\hat{I}; \lambda) := I_1^{(i)}\Big(\mathtt{R}^2 + \mathtt{Rr}, \hat{I}\Big), \ i = 0, 2N,
\end{aligned} \tag{4.7}$$

and, for $0 \leqslant i \leqslant 2N$,

$$\mathcal{B}^i(\lambda) := \Big\{I = (I_1, \hat{I}) \mid \hat{I} \in \hat{D}, \ \ a_-^{(i)}(\hat{I}; \lambda) < I_1 < a_+^{(i)}(\hat{I}; \lambda)\Big\} \subseteq \mathbb{R}^n, \ \ \mathcal{B}^i := \mathcal{B}^i(0). \tag{4.8}$$

Note that[18]

$$\operatorname{diam} \mathcal{B}^i(0) \leqslant 2\Big(\mathtt{R} + \operatorname{diam} \hat{D}\Big), \qquad \forall \ 0 \leqslant i \leqslant 2N. \tag{4.9}$$

---

[17] Uniquely fixing, e. g., $\varphi_1^{(i)}(p, 0) = 0$.

[18] Recall (2.8)–(2.10).

Setting[19)]

$$\check{\mathcal{M}}^i := \{(p,q_1)\in\mathbb{R}^n\times\mathbb{T}\ \text{ s.t. }\ (p_1,q_1)\in\mathcal{M}^i(\hat p),\ \hat p\in\hat D\},$$

we have

$$\check{\mathcal{M}}^i = \check{\mathcal{M}}^i(0) = \check\Phi^i(\mathcal{B}^i\times\mathbb{T}) = \bigcup_{0<\lambda\leqslant 1/\mathbf{c}} \check{\mathcal{M}}^i(\lambda), \quad \text{where} \quad \check{\mathcal{M}}^i(\lambda) := \check\Phi^i(\mathcal{B}^i(\lambda)\times\mathbb{T}). \tag{4.10}$$

Theorem 2. *Under the hypotheses of Theorem* 1 *there exists* $\hat{\mathbf{c}} = \hat{\mathbf{c}}(n,\kappa)\geqslant 4\mathbf{c}^2$ *depending only on* n *and* κ *such that, taking*

$$\mu \leqslant 1/\hat{\mathbf{c}}, \tag{4.11}$$

for any $0\leqslant i\leqslant 2N$*, the symplectic transformation* $\check\Phi^i$ *extends, for any* $0<\lambda\leqslant 1/\hat{\mathbf{c}}$*, to a real-analytic map*

$$\Phi^i : \big(\mathcal{B}^i(\lambda)\big)_\rho \times \mathbb{T}^n_\sigma \ \to\ D_{\mathtt{r}}\times\mathbb{T}^n_{\mathtt{s}/4} \tag{4.12}$$

with

$$\rho = \tfrac{\sqrt{\varepsilon}}{\hat{\mathbf{c}}}\,\lambda|\log\lambda|, \qquad \sigma = \tfrac{1}{\hat{\mathbf{c}}|\log\lambda|}. \tag{4.13}$$

Moreover, on $\big(\mathcal{B}^i(\lambda)\big)_\rho$ *we have*

$$\begin{aligned}
&\Big|\partial_{I_1}\mathtt{E}^i\Big| \leqslant \hat{\mathbf{c}}\sqrt{\varepsilon+|\mathtt{E}^i|}, \quad \Big|\partial^2_{I_1}\mathtt{E}^i\Big| \leqslant \tfrac{\hat{\mathbf{c}}}{\hat\lambda}, \qquad\qquad (\hat\lambda := \lambda|\log\lambda|^3)\\
&\Big|\partial^2_{I_1\hat I}\mathtt{E}^i\Big| \leqslant \hat{\mathbf{c}}\tfrac{\mu_{\rm o}}{\hat\lambda}, \quad \Big|\partial^2_{\hat I}\mathtt{E}^i\Big| \leqslant \hat{\mathbf{c}}\Big(\tfrac{\sqrt{\varepsilon}}{\mathtt{r}}I_1^i + \tfrac{\mu_{\rm o}}{\hat\lambda}\Big)\mu_{\rm o}.
\end{aligned} \tag{4.14}$$

Finally, we have

$$\operatorname{meas}\Big(\big(D^\flat\times\mathbb{T}\big)\ \setminus \bigcup_{0\leqslant i\leqslant 2N}\check{\mathcal{M}}^i(\lambda)\Big) \leqslant \hat{\mathbf{c}}\,\sqrt{\varepsilon}\operatorname{meas}(\hat D)\ \lambda|\log\lambda|, \tag{4.15}$$

where $D^\flat := (-\mathtt{R}-\mathtt{r}/3, \mathtt{R}+\mathtt{r}/3)\times\hat D$.

Remark 4. (i) The complete symplectic action-angle map $\Phi^i : (I,\varphi)\to(p,q)$ has the form

$$\Phi^i(I,\varphi) = \begin{cases} (\eta^i, \hat I, \psi^i, \hat\varphi+\chi^i), & \text{if } 0<i<2N,\\ (\eta^i,\hat I,\varphi_1+\psi^i,\hat\varphi+\chi^i), & \text{if } i=0,2N, \end{cases} \tag{4.16}$$

where η^i,χ^i,ψ^i are functions of (I,φ_1) only and are 2π-periodic in φ_1, and, in the case $i=0,2N$, $\sup|\partial_{\varphi_1}\psi^i|<1$.
Notice that, since $\mathbf{c}_1\geqslant 4\kappa\geqslant 16$, by (2.10) and (4.13) we get $\rho\leqslant 2^{-8}\mathtt{r}$ and $\sigma\leqslant\mathtt{s}/4$. By (4.12) we also get $|\operatorname{Im}\psi^i|_{\rho,\sigma}\leqslant\mathtt{s}/2$ for every $0\leqslant i\leqslant 2N$. Analogously,[20)] $|\operatorname{Im}\chi^i_j|_{\rho,\sigma}\leqslant\mathtt{s}/2$ for every $j=2,\dots,n$.

(ii) Notice the different topologies of this map: For $1\leqslant i\leqslant 2N-1$ the motion is *librational*, i. e., the q_1-coordinate oscillates around relative (stable) equilibria, while for $i=0$ and $i=2N$ the motion is *rotational*, corresponding to the q_1-coordinate rotating in the unbounded regions of phase space "outside" separatrices; such regions correspond to the labels $i=2N$ (upper unbounded region) and $i=0$ (lower unbounded region).

(iii) For related estimates on the analyticity strip in the angles, see [22].

19) Recall (3.3) and (3.5).

20) Actually, a better estimate holds: it is smaller than some constant by $\mu_{\rm o}\mathbf{s}$, where $\mu_{\rm o}$ was defined in (3.13).

Proof (*of Theorem* 2). The fact that the map $\check{\Phi}^i$ extends a complete symplectic transformation Φ^i directly follows by the Arnold–Liouville theorem. Here we have only to evaluate the analyticity radii. For brevity we will often drop the index i and the dumb actions and will often write $I(E)$ instead of $I_1^{(i)}(E)$. As above, we will consider only the case of odd i, the other one being similar.

In order to prove (4.12)–(4.13), we introduce energy-time (E,t) (symplectic) coordinates, which are a simple rescaling of action-angle variables (I,φ). Indeed, considering the integrable Hamiltonian $\mathtt{E}(I)=E$, we have that the action and the angular velocity are constant: $\dot\varphi=\partial_I\mathtt{E}(I)$, so that $\varphi(t)=\partial_I\mathtt{E}(I)t$ or, using (E,t) as independent variables

$$I=I(E),\qquad \varphi=\varphi(E,t)=\frac{t}{\partial_E I(E)}. \tag{4.17}$$

We restrict to the zone around hyperbolic points where one has worst estimates. Here we first pass to (y_1,x_1) coordinates, obtaining the Hamiltonian $\mathtt{H}_{\rm hp}(y_1,x_1)$ in (3.53). Secondly, we pass to coordinates (E,t), setting

$$E:=\mathtt{H}_{\rm hp}(y_1,x_1),\qquad t:=\frac{1}{w(E)}\mathrm{arctanh}\left(\frac{y_1}{\sqrt{-J(E)+y_1^2}}\right), \tag{4.18}$$

where $w(E):=2g\big(1+\sqrt{\varepsilon}\partial_z R_{\rm hp}(J(E)/\sqrt{\varepsilon})\big)$. Indeed, by the Hamilton equations for $\mathtt{H}_{\rm hp}$ we are led to

$$\dot y_1=-\partial_{x_1}\mathtt{H}_{\rm hp}(y_1,x_1)=2gx_1\Big(1+\sqrt{\varepsilon}\partial_z R_{\rm hp}\big((y_1^2-x_1^2)/\sqrt{\varepsilon}\big)\Big)=w(E)\sqrt{-J(E)+y_1^2},$$

which can be easily integrated (by separation of variables), giving the expression for the time in (4.18).

Lemma 4. *There exists a small constant* $0<\mathbf{c}_6\leqslant\min\{\mathbf{c}_1^2,1\}/2$ *depending only on* κ *and* n *such that, taking*

$$0<\lambda\leqslant\mathbf{c}_6\quad\textit{and}\quad \tilde E:=E_+-\varepsilon\lambda, \tag{4.19}$$

the map $(y_1,x_1)\in\{|y_1|<\mathbf{c}_1\varepsilon^{1/4}\}\times\{|x_1|<\mathbf{c}_1\varepsilon^{1/4}\}\ \to\ (E,t)$ *in* (4.18) *is invertible for*

$$|E-\tilde E|\leqslant\mathbf{c}_6\varepsilon\lambda,\qquad |t|<\mathbf{c}_6/\sqrt{\varepsilon}. \tag{4.20}$$

Proof. Inverting the second expression in (4.18), we get

$$y_1(E,t)=\sqrt{-J(E)}\sinh(w(E)t). \tag{4.21}$$

Moreover, by (3.66), (3.68) and (3.69) we have

$$x_1(E,t)=\sqrt{-J(E)+\big(y_1(E,t)\big)^2}. \tag{4.22}$$

We have to check that the above functions y_1 and x_1 are defined on the set in (4.20). Set

$$\sqrt{\varepsilon}\lambda\lessdot J_*:=\frac{\varepsilon}{\bar g}\lambda\lessdot\sqrt{\varepsilon}\lambda, \tag{4.23}$$

with $\bar g$ defined in (3.50) (recall also (2.10)). Recalling the definition of $J(E)$ in (3.66), we have

$$J(\tilde E)=-\frac{\varepsilon}{g}\lambda\big(1+\lambda F(\lambda)\big)$$

and, by (3.50) and (3.65),

$$|J(\tilde E)+J_*|=\frac{\varepsilon}{\bar g}\lambda\left|\frac{\bar g}{g}\big(1+\lambda F(\lambda)\big)-1\right|\lessdot\sqrt{\varepsilon}\lambda(\mu+\lambda).$$

Finally, since by (3.66), (3.65) and (3.50) $|\partial_E J|\lessdot 1/\sqrt{\varepsilon}$ we have, for $|E-\tilde E|\leqslant\mathbf{c}_6\varepsilon\lambda$,

$$|J(E)-J(\tilde E)|\lessdot\mathbf{c}_6\sqrt{\varepsilon}\lambda,$$

we get

$$|J(E) + J_*| \lessdot (\mathbf{c}_6 + \mu)\sqrt{\varepsilon}\lambda. \tag{4.24}$$

By (4.23) and taking $\mathbf{c}_6$ small enough and $\hat{\mathbf{c}}$ large enough, we obtain

$$-\operatorname{Re} J(E) \geqslant 4\mathbf{c}_6 \sqrt{\varepsilon}\lambda, \tag{4.25}$$

for any $|E - \tilde{E}| \leqslant \mathbf{c}\varepsilon\lambda$. Moreover,

$$|w(E)| \overset{(3.50),\,(3.52)}{\lessdot} \sqrt{\varepsilon}, \qquad |J(E)| \overset{(4.24),\,(4.23)}{\lessdot} \sqrt{\varepsilon}\lambda. \tag{4.26}$$

Then, recalling (4.21), we get

$$|y_1(E,t)| \leqslant \sqrt{\mathbf{c}_6}\varepsilon^{1/4}\sqrt{\lambda} < \mathbf{c}_1\varepsilon^{1/4},$$

for E, t satisfying (4.20). Consequently, by (4.25), the function $x_1(E,t)$ defined in (4.22) is holomorphic[21)] for E, t satisfying (4.20), taking $\mathbf{c}_6$ small enough. □

We finally pass to action-angle variables defined in (4.17). First we observe that, by (3.12) and (3.15), for E as in (4.20) we get

$$|\log\lambda|/\sqrt{\varepsilon} \lessdot |\partial_E I| \lessdot |\log\lambda|/\sqrt{\varepsilon}. \tag{4.27}$$

Note that, by (4.7) and (4.19), $a_+(\lambda) = I(\tilde{E})$. By (4.27), the image of the ball $|E - \tilde{E}| \leqslant \mathbf{c}_6\varepsilon\lambda$ through the function $I(E)$ contains the ball $|I - a_+(\lambda)| \leqslant \rho$ (defined in (4.13)), taking $\hat{\mathbf{c}}$ large enough. Analogously by (4.27), for every $|E - \tilde{E}| \leqslant \mathbf{c}_6\varepsilon\lambda$, the image of the ball $|t| < \mathbf{c}_6/\sqrt{\varepsilon}$ through the function $t \to t/\partial_E I(E)$ contains the ball $|\varphi| \leqslant \sigma$, taking $\hat{\mathbf{c}}$ large enough. Recalling (4.8), this completes the proof[22)] of (4.13).

Let us now prove (4.14). First, by the chain rule we get

$$\begin{aligned}
&\partial_{I_1}\mathtt{E} = \frac{1}{\partial_E I_1}, \qquad \partial_{\hat{I}}\mathtt{E} = -\frac{\partial_{\hat{I}} I_1}{\partial_E I_1}, \qquad \partial^2_{I_1 I_1}\mathtt{E} = -\frac{\partial^2_{EE} I_1}{(\partial_E I_1)^3},\\
&\partial^2_{I_1\hat{I}}\mathtt{E} = \frac{\partial^2_{EE} I_1 \partial_{\hat{I}} I_1}{(\partial_E I_1)^3} - \frac{\partial^2_{E\hat{I}} I_1}{(\partial_E I_1)^2},\\
&\partial^2_{\hat{I}\hat{I}}\mathtt{E} = -\frac{\partial^2_{\hat{I}\hat{I}} I_1}{\partial_E I_1} + \frac{\partial^T_{\hat{I}} I_1\ \partial_{\hat{I}}(\partial_E I_1) + \partial^T_{\hat{I}}(\partial_E I_1)\ \partial_{\hat{I}} I_1}{(\partial_E I_1)^2} - \frac{\partial^2_{EE} I_1\ \partial^T_{\hat{I}} I_1\ \partial_{\hat{I}} I_1}{(\partial_E I_1)^3},
\end{aligned} \tag{4.28}$$

where the derivatives of $\mathtt{E}$ and I_1 are evaluated in $\big(I_1(E,\hat{I}), \hat{I}\big)$ and $(E, \hat{I})$, respectively.

Then, we split $\big(\mathcal{B}(\lambda)\big)_\rho$ into two subsets: the region where I_1 is close to $a_+(\lambda)$ (near the hyperbolic equilibrium) and the region far away from it. More precisely, recalling (3.20), we set

$$\begin{aligned}
&\mathcal{E}_{\rm cl} := \Big\{|E - E_+| < \varepsilon/\mathbf{c}\ \big|\ \operatorname{Im}(E - E_+) = 0 \implies \operatorname{Re}(E - E_+) > 0\Big\} \times \hat{D}_{\mathtt{r}/4},\\
&\mathcal{E}_{\rm aw} := \mathcal{E}_{1/4\mathbf{c}} \cap \{|\operatorname{Im} E| < \varepsilon/(3\mathbf{c})^5\} \times \hat{D}_{\mathtt{r}/4}.
\end{aligned} \tag{4.29}$$

Then,

$$\big(\mathcal{B}(\lambda)\big)_\rho \subset \mathcal{E}_{\rm cl} \cap \mathcal{E}_{\rm aw},$$

---

21) We are considering the square root as a holomorphic function in the complex plane excluding the negative real axis.

22) Close to a hyperbolic point the estimates for the action analyticity radius in (4.13) is $\rho = \lambda|\log\lambda|\sqrt{\varepsilon}/C$ since $\partial_E I \sim |\log\lambda|/\sqrt{\varepsilon}$ (see (4.27)), $\lambda\varepsilon$ being the distance in energy from the critical energy of the hyperbolic point (see (4.30) below). Far away from the hyperbolic point the derivative is smaller (namely, $\partial_E I \sim 1/\sqrt{\varepsilon}$), but the distance in energy is bigger (being $\sim \varepsilon$).

taking $\hat{\mathbf{c}}$ large enough. Regarding the first region we start noting that by the Cauchy estimates, (3.12) and (3.13), for $I_1 = I_1\big(E_\pm(\hat{I}) \mp \varepsilon z,\, \hat{I}\big)$, with $|z| < 1/\mathbf{c}$ not belonging to the negative real semiaxis and $\hat{I} \in \hat{D}_{\mathtt{r}/4}$, we have

$$\begin{aligned}&\frac{|\log z|}{\sqrt{\varepsilon}} \lessdot |\partial_E I_1| \lessdot \frac{|\log z|}{\sqrt{\varepsilon}}, \qquad \frac{1}{\varepsilon^{3/2}|z|} \lessdot |\partial_{EE} I_1| \lessdot \frac{1}{\varepsilon^{3/2}|z|},\\ &|\partial_{\hat{I}} I_1| \lessdot \mu_{\rm o}, \qquad |\partial^2_{E,\hat{I}} I_1| \lessdot \mu_{\rm o}|\log z|, \qquad |\partial^2_{\hat{I},\hat{I}} I_1| \lessdot \frac{\mu_{\rm o}}{\mathtt{r}},\end{aligned} \tag{4.30}$$

where the first line follows by (3.15). Then (4.14) directly follows from (4.28).

Consider now the second region, namely, $\mathcal{E}_{\rm aw}$. By (3.21) with $\lambda = 1/\mathbf{c}$ we have $|\partial_E I_1| \leqslant 5\mathbf{c}^3/\sqrt{\varepsilon}$ on $\mathcal{E}_{1/2\mathbf{c}} \times \hat{D}_{\mathtt{r}/2}$. Then, by the Cauchy estimates, we get $|\partial^2_{EE} I_1| \leqslant 40\mathbf{c}^4/\varepsilon^{3/2}$ on $\mathcal{E}_{\rm aw}$. Then by (3.17) we get $|\partial_E I_1| \geqslant 1/2\mathbf{c}\sqrt{\varepsilon}$ on $\mathcal{E}_{\rm aw}$. Using this lower bound, (4.28), (3.21) and the Cauchy estimates, (4.14) follows also in $\mathcal{E}_{\rm aw}$.

We finally prove (4.15). Since the maps $\check{\Phi}^i$ preserve the $(n+1)$-dimensional measure $dId\varphi_1$, recalling (4.7)–(4.8) and using (3.12)–(3.13), (2.8), (2.10), one obtains, by Fubini's theorem,

$$\begin{aligned}&\operatorname{meas}\Big(\big(D^\flat \times \mathbb{T}\big) \ \setminus \bigcup_{0\leqslant i\leqslant 2N} \check{\mathcal{M}}^i(\lambda)\Big) \leqslant 2\pi \sum_{0\leqslant i\leqslant 2N} \operatorname{meas}(\hat{D}) \operatorname{meas}\big(\mathcal{B}^i(0)\setminus \mathcal{B}^i(\lambda)\big)\\ &\leqslant 2\pi \sum_{0\leqslant i\leqslant 2N} \operatorname{meas}(\hat{D}) \sup_{\hat{I}\in\hat{D}} \Big(a_+(\hat{I};0) - a_+(\hat{I};\lambda) + a_-(\hat{I};\lambda) - a_-(\hat{I};0)\Big)\\ &\overset{(3.21)}{\lessdot} \ \operatorname{meas}(\hat{D})\sqrt{\varepsilon}\int_0^\lambda |\log z|\, dz \leqslant \ \operatorname{meas}(\hat{D})\sqrt{\varepsilon}\lambda|\log\lambda|.\end{aligned}$$

□

5. CONVEXITY ENERGY ESTIMATES

In this section we investigate the convexity of the energy functions $I_1 \to \mathtt{E}^i(I_1, \hat{p})$ defined as the inverse functions of the action functions[23] $E \to I_1^i(E, \hat{p})$.

We also let $\bar{I}_1^i := I_1^i|_{\mu=0}$ denote the "unperturbed action function", and its inverse $\bar{\mathtt{E}}^i := \mathtt{E}^i|_{\mu=0}$, the "unperturbed energy function".

Remark 5. Observe that $\bar{I}_1^{(0)}(E) = \bar{I}_1^{(2N)}(E)$ and $\bar{\mathtt{E}}^{(0)}(I_1) = \bar{\mathtt{E}}^{(2N)}(I_1)$.

In general, the energy functions have inflection points[24], however, there are some cases in which the convexity is definite, namely, in the outer regions ($i = 0, 2N$) and in the case where the reference potential $\bar{\mathtt{G}}$ is "close" to a cosine (in which case $N = 2$) in the sense of the following

Definition 3 (Cosine-like functions). Let $0 < \mathtt{g} < 1/4$. We say that a real-analytic function $G : \mathbb{T}_1 \to \mathbb{C}$ is $\mathtt{g}$-cosine-like if, for some $\eta > 0$ and $\theta_0 \in \mathbb{R}$, one has

$$\sup_{\theta\in\mathbb{T}_1} |G(\theta) - \eta\cos(\theta+\theta_0)| \leqslant \eta\mathtt{g}. \tag{5.1}$$

Proposition 4. (i) *If $i = 0, 2N$, then, for every $E > \bar{E}_i$, one has: $\partial^2_{I_1}\bar{\mathtt{E}}^i\big(\bar{I}_1^i(E)\big) \geqslant 2$.*

(ii) *If $\bar{\mathtt{G}}$ is cosine-like with $\mathtt{g} \leqslant 2^{-40}$, then*

$$\partial^2_{I_1}\bar{\mathtt{E}}^1\big(\bar{I}_1^1(E)\big) \leqslant -\frac{1}{27}, \qquad \forall E \in (\bar{E}_1, \bar{E}_2). \tag{5.2}$$

---

[23] Observe that the action function $E \to I_1^i(E, \hat{p})$ is strictly increasing and hence invertible.
[24] Compare [9].

Proof. (i) Let us consider now the zone above separatrices. First observe that the cases $i=0$ and $i=2N$ are identical by Remark 5. Let us then consider the case $i=2N$. By definition,

$$\bar I_1^{2N}(E)=\frac{1}{2\pi}\int_0^{2\pi}\sqrt{E-\bar{\mathtt G}(x)}dx,$$

thus, by Jensen's inequality

$$\big(2\partial_E \bar I_1^i(E)\big)^3=\Big(\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{\sqrt{E-\bar{\mathtt G}(x)}}dx\Big)^3\leqslant\frac{1}{2\pi}\int_0^{2\pi}\frac{1}{\big(E-\bar{\mathtt G}(x)\big)^{3/2}}dx=-4\partial_E^2\bar I_1^i(E),$$

and the claim follows by

$$\partial^2_{I_1I_1}\bar{\mathtt E}^{(i)}\big(\bar I_1^{(i)}(E)\big)=-\frac{\partial^2_{EE}\bar I_1^{(i)}(E)}{\big(\partial_E\bar I_1^{(i)}(E)\big)^3}. \tag{5.3}$$

(ii) First we note that, up to a phase translation, we can take $\theta_0=0$ in (5.1). Then set

$$M:=\max_{\mathbb R}\bar{\mathtt G},\quad m:=\min_{\mathbb R}\bar{\mathtt G},\quad L(y):=\frac{2y-M-m}{M-m},\quad V:=L\circ\bar{\mathtt G}. \tag{5.4}$$

Note that $\max_{\mathbb R}V=1$, $\min_{\mathbb R}V=-1$.

The idea is to study the action variable of the Hamiltonian $p_1^2+V(q_1)$ which is strictly related to the one of $p_1^2+\bar{\mathtt G}(q_1)$, see (5.6) below. Denoting $|\cdot|_r:=\sup_{\mathbb T_r}|\cdot|$, we have the following

Lemma 5. *If $\bar{\mathtt G}$ satisfies* (5.1)*, then V in* (5.4) *satisfies $|V-\cos z|_1\leqslant 4\mathtt g$.*

Proof. By (5.1) and (5.4), we have

$$-\mathtt g\hat\eta\ \leqslant\ M-\hat\eta,\ m+\hat\eta\ \leqslant\ \mathtt g\hat\eta,\qquad \left|\frac{2\hat\eta}{M-m}-1\right|\leqslant\frac{\mathtt g}{1-\mathtt g}. \tag{5.5}$$

By (5.1) and (5.4), we get

$$\left|V(z)-\frac{2\hat\eta}{M-m}\cos z\right|_1\leqslant\frac{2}{M-m}\Big(\mathtt g\hat\eta+\frac{M+m}{2}\Big)\leqslant\frac{4\mathtt g\hat\eta}{M-m}\leqslant\frac{2\mathtt g}{1-\mathtt g}.$$

Then

$$\sup_{\mathbb T_1}|V(z)-\cos z|\leqslant\frac{2\mathtt g}{1-\mathtt g}+\left|\frac{2\hat\eta-M+m}{M-m}\cos z\right|_1\leqslant\frac{2+\cosh 1}{1-\mathtt g}\leqslant 4\mathtt g.$$

□

Next, we need a representation lemma whose proof is given in Appendix B.2:

Lemma 6. *Let $0<\mathtt g_{\rm o}\leqslant 2^{-10}$ and let w be a real-analytic 2π-periodic function satisfying*

$$\max_{\mathbb R}w=1,\quad \min_{\mathbb R}w=-1,\quad |w(z)-\cos z|_1\leqslant\mathtt g_{\rm o}.$$

Then, there exists a unique real-analytic 2π-periodic function b such that

$$w(z)=\cos\big(z+b(z)\big),\qquad |b|_{1/4}\leqslant 9\sqrt{\mathtt g_{\rm o}}.$$

Lemma 6 can be applied to the potential V in (5.4), so that, in particular, V has only two critical points (a maximum and a minimum) on a period.

For $i=0,1,2$ let $\bar I_1^{(i)}(E)$ and $\tilde I_1^{(i)}(E)$ denote the action variable of the Hamiltonian $p_1^2+\bar{\mathtt G}(q_1)$ and $p_1^2+V(q_1)$, respectively, in the three zones below $(i=0)$, inside $(i=1)$ and above $(i=2)$ separatrices.

Now, the relation between the action $\bar{I}_1^{(i)}$ of $p_1^2+\bar{\mathtt{G}}(q_1)$ and the action $\tilde{I}_1^{(i)}$ of $p_1^2+V(q_1)$ is given by the following formula:

$$\bar{I}_1^{(i)}(E)=\sqrt{\frac{M-m}{2}}\tilde{I}_1^{(i)}\big(L(E)\big),\qquad i=0,1,2. \tag{5.6}$$

Indeed, considering the case $i=2$ (the other ones being analogous), and recalling (5.4), one finds

$$\begin{aligned}\bar{I}_1^{(i)}(E)&=\frac{1}{2\pi}\int_0^{2\pi}\sqrt{E-\bar{\mathtt{G}}(x)}dx=\frac{1}{2\pi}\int_0^{2\pi}\sqrt{L^{-1}\big(L(E)\big)-L^{-1}(V)(x)}dx\\&=\sqrt{\frac{M-m}{2}}\frac{1}{2\pi}\int_0^{2\pi}\sqrt{L(E)-V(x)}dx=\sqrt{\frac{M-m}{2}}\tilde{I}_1^{(i)}\big(L(E)\big),\end{aligned}$$

which proves (5.6).

Coming back to the proof of (5.2), let $\bar{\mathtt{E}}^{(i)}(I_1)$ and $\tilde{\mathtt{E}}^{(i)}(I_1)$ denote the inverse function of $\bar{I}_1^{(i)}(E)$ and $\tilde{I}_1^{(i)}(E)$, respectively. By[25] (5.3), (5.6) and (5.4), we get

$$\partial^2_{I_1I_1}\bar{\mathtt{E}}^{(i)}\big(\bar{I}_1^{(i)}(E)\big)=\partial^2_{I_1I_1}\tilde{\mathtt{E}}^{(i)}\Big(\tilde{I}_1^{(i)}\big(L(E)\big)\Big). \tag{5.7}$$

Let us consider first the zone inside separatrices and, to simplify notation, denote $\tilde{I}_1^{(1)}(E)$ by $A(E)$. Then,

$$A(E)=A_-(E)+A_+(E):=\frac{1}{\pi}\int_{x_m-2\pi}^{V_-^{-1}(E)}\sqrt{E-V(x)}dx+\frac{1}{\pi}\int_{V_+^{-1}(E)}^{x_m}\sqrt{E-V(x)}dx,$$

where $V_-^{-1}(E)$ and $V_+^{-1}(E)$ are the inverse of $V(x)=\cos\big(\psi(x)\big)=E$, with $\psi(x):=x+b(x)$, in the intervals $[x_m-2\pi,x_M]$ and $[x_M,x_m]$, respectively; namely, $V_-^{-1}(E)=\psi^{-1}\big(-\arccos(E)\big)$ and $V_+^{-1}(E)=\psi^{-1}\big(\arccos(E)\big)$. Recall that $\psi(x_M)=0$ and $\psi(x_m)=\pi$. Since $|b|_{1/4}\leqslant 18\sqrt{\mathtt{g}}$ by the Cauchy estimates we have that ψ is invertible with inverse $\psi^{-1}(y)=y+u(y)$ for a suitable[26] 2π-periodic real-analytic u satisfying

$$|u|_{1/5}\leqslant 18\sqrt{\mathtt{g}}. \tag{5.8}$$

We get

$$A_+'(E)=\frac{1}{2\pi}\int_{V_+^{-1}(E)}^{x_m}\frac{dx}{\sqrt{E-V(x)}}=\frac{1}{2\pi}\int_0^1\frac{1+u'\Big(\arccos\big(g(E,t)\big)\Big)}{\sqrt{t-t^2}\sqrt{1+t-E(1-t)}}dt,$$

making the substitution $x=x(t):=V_+^{-1}\big(g(E,t)\big)$ with $g(E,t):=E-(1+E)t$. Analogously,

$$A_-'(E)=\frac{1}{2\pi}\int_0^1\frac{1+u'\Big(-\arccos\big(g(E,t)\big)\Big)}{\sqrt{t-t^2}\sqrt{1+t-E(1-t)}}dt.$$

Then, taking the even part v of u', namely, $v(y):=\frac{1}{2}\big(u'(y)+u'(-y)\big)$, we have

$$A'(E)=\frac{1}{\pi}\int_0^1\frac{1+v\Big(\arccos\big(g(E,t)\big)\Big)}{\sqrt{t-t^2}\sqrt{1+t-E(1-t)}}dt.$$

Note that by the Cauchy estimates $|v|_{1/6}\leqslant 540\sqrt{\mathtt{g}}$. Deriving, we get

$$A''(E)=\frac{1}{2\pi}\int_0^1\sqrt{\frac{1-t}{t}}\frac{1+v_0(E,t)}{\big(1+t-E(1-t)\big)^{3/2}}dt,$$

[25] And the analogous formula for $\partial^2_{I_1I_1}\tilde{\mathtt{E}}^{(i)}\big(\tilde{I}_1^{(i)}(E)\big)$.

[26] u is the solution of the fixed point equation $u(y)=-b\big(y+u(y)\big)$ in the space of 2π-periodic real-analytic function with holomorphic extension on the strip $\{|\operatorname{Im} y|<1/5\}$ and $|u|_{1/5}\leqslant 18\sqrt{\mathtt{g}}$.

with

$$v_0(E,t) := v\Big(\arccos\big(g(E,t)\big) - 2\tilde v(E,t),$$
$$\tilde v(E,t) := \frac{v'\Big(\arccos\big(g(E,t)\big)\Big)\sqrt{1+t-E(1-t)}}{\sqrt{1-t}\sqrt{1+E}}.$$

Since v is 2π-periodic and even, we have $v'(\pi)=0$. Then, by the Cauchy estimates, we get

$$|v'(\xi)| \leqslant 39880\cdot\sqrt{\mathtt{g}}\,|\xi-\pi|, \qquad \forall\, \xi\in\mathbb{R}.$$

Note that

$$0 \leqslant \pi - \arccos(-1+\xi) \leqslant \frac{\pi}{\sqrt{2}}\sqrt{\xi}, \qquad \forall\, 0\leqslant\xi\leqslant 2.$$

Therefore, since $g(E,t)+1=(1-t)(1+E)$, for $0<t<1$ and $-1<E<1$, one has

$$|v'\Big(\arccos\big(g(E,t)\big)\Big)| \leqslant 39880\sqrt{\mathtt{g}}|\pi-\arccos\big(g(E,t)\big)| \leqslant 19440\pi\sqrt{2\mathtt{g}}\sqrt{1-t}\sqrt{1+E},$$

which implies

$$|\tilde v(E,t)| \leqslant 244292\sqrt{\mathtt{g}} \qquad \text{and} \qquad |v_0(E,t)| \leqslant 244292\sqrt{\mathtt{g}} \leqslant 2^{18}\sqrt{\mathtt{g}}.$$

Taking $\mathtt{g}\leqslant 2^{-38}$, we have $|v_0(E,t)|\leqslant 1/2$ and therefore for every $-1<E<1$

$$\frac{1}{2}A_0'(E) \leqslant A'(E) \leqslant \frac{3}{2}A_0'(E), \qquad \frac{1}{2}A_0''(E) \leqslant A''(E) \leqslant \frac{3}{2}A_0''(E), \tag{5.9}$$

where $A_0(E)$ is the action variable with exactly cosine potential (namely, when $\mathtt{g}=0$), namely,

$$A_0'(E) = \frac{1}{\pi}\int_0^1 \frac{1}{\sqrt{t-t^2}\sqrt{1+t-E(1-t)}}dt,$$
$$A_0''(E) = \frac{1}{2\pi}\int_0^1 \sqrt{\frac{1-t}{t}}\,\frac{1}{\big(1+t-E(1-t)\big)^{3/2}}\,dt.$$

Then, since $\tilde{\mathtt{E}}^{(1)}(I_1)$ is the inverse of $\tilde I_1^{(1)}(E)=A(E)$, for every $-1<E<1$

$$-\partial^2_{I_1}\tilde{\mathtt{E}}^{(1)}\big(\tilde I_1^{(1)}(E)\big) = \frac{A''(E)}{\big(A'(E)\big)^3} \geqslant \frac{4}{27}\frac{A_0''(E)}{\big(A_0'(E)\big)^3} \geqslant \frac{1}{27}, \tag{5.10}$$

since, as is easy to check, the function $\frac{A_0''(E)}{\big(A_0'(E)\big)^3}$ is increasing and has limit $1/4$ for $E\to -1^+$. $\square$

APPENDIX A. PROOFS OF PROPOSITION 2

First, recalling (3.45) and (3.46), we define the symplectic transformation

$$\Phi_* : (\mathtt{p},\mathtt{q})\in D_{7\mathtt{r}/8,7\mathtt{s}/8} \longrightarrow \Big(\mathtt{p},\mathtt{q}_1+\theta_{2j}(\hat{\mathtt{p}}),\hat{\mathtt{q}}+\mathtt{p}_1\partial_{\hat{\mathtt{p}}}\theta_{2j}(\hat{\mathtt{p}})\Big)\in D_{\mathtt{r},\mathtt{s}}, \tag{A.1}$$

transforming the Hamiltonian $\mathtt{H}$ in (2.6) into

$$\mathtt{H}_* := \mathtt{H}\circ\Phi_*(\mathtt{p},\mathtt{q}) =: \big(1+\nu_*(\mathtt{p},\mathtt{q}_1)\big)\mathtt{p}_1^2 + \mathtt{G}_*(\hat{\mathtt{p}},\mathtt{q}_1). \tag{A.2}$$

By Taylor expansion at $(\mathtt{p},\mathtt{q}_1)=(0,\hat{\mathtt{p}},0)$, recalling (A.2), (3.47), (3.49) and (2.10), we get

$$\mathtt{H}_* = E_{2j}(\hat{\mathtt{p}}) + \big(1+\nu_*(0,\hat{\mathtt{p}},0)\big)\mathtt{p}_1^2 - \lambda^2(\hat{\mathtt{p}})\mathtt{q}_1^2 + R_*(\mathtt{p},\mathtt{q}_1), \qquad \text{with}$$
$$R_*(\mathtt{p},\mathtt{q}_1) := \big(\nu_*(\mathtt{p},\mathtt{q}_1)-\nu_*(0,\hat{\mathtt{p}},0)\big)\mathtt{p}_1^2 + \mathtt{G}_*(\hat{\mathtt{p}},\mathtt{q}_1) - \frac{1}{2}\partial^2_{\mathtt{q}_1}\mathtt{G}_*(\hat{\mathtt{p}},0)\mathtt{q}_1^2. \tag{A.3}$$

Then, the following trivial lemma holds.

Lemma 7. *There exists a constant $0 < \mathbf{c_0} < 1/8$, depending only on κ, n, such that, defining the symplectic transformation*

$$\Phi_0 : \{|Y_1| < \mathbf{c_0}\varepsilon^{1/4}\} \times \hat{D}_{3\mathtt{r}/4} \times \{|X_1| < \mathbf{c_0}\varepsilon^{1/4}\} \times \mathbb{T}^{n-1}_{3\mathtt{s}/4} \longrightarrow D_{7\mathtt{r}/8, 7\mathtt{s}/8}, \tag{A.4}$$

$$\mathtt{p}_1 = \delta(\hat{Y})Y_1, \qquad \hat{\mathtt{p}} = \hat{Y}, \qquad \mathtt{q}_1 = \frac{1}{\delta(\hat{Y})}X_1, \qquad \hat{\mathtt{q}} = \hat{X} - \frac{\partial_{\hat{Y}}\delta(\hat{Y})}{\delta(\hat{Y})}Y_1X_1,$$

we have that $\mathtt{H}_0 := \mathtt{H}_* \circ \Phi_0$ has the form

$$\mathtt{H}_0 = E_{2j}(\hat{Y}) + g(\hat{Y})(Y_1^2 - X_1^2) + \varepsilon R_0(\varepsilon^{-1/4}Y_1, \hat{Y}, \varepsilon^{-1/4}X_1), \tag{A.5}$$

where $R_0(\tilde{Y}_1, \hat{Y}, \tilde{X}_1)$ is holomorphic on

$$\{|\tilde{Y}_1| < \mathbf{c_0}\} \times \hat{D}_{3\mathtt{r}/4} \times \{|\tilde{X}_1| < \mathbf{c_0}\},$$

with $|R_0| \lessdot 1$ and, finally, it is at least cubic in $\tilde{Y}_1, \tilde{X}_1$.

Proof. The fact that Φ_0 is well defined on its domain follows by the explicit expression in (A.4), by (3.50), (2.10) and (2.8) (in particular, $\varepsilon \leqslant \mathtt{r}^2/2^{16}$). Eq. (A.5) follows by (A.3), setting

$$R_0(\tilde{Y}_1, \hat{Y}, \tilde{X}_1) := \varepsilon^{-1}R_*\left(\delta(\hat{Y})\varepsilon^{1/4}\tilde{Y}_1,\ \hat{Y},\ \frac{\varepsilon^{1/4}}{\delta(\hat{Y})}\tilde{X}_1\right). \tag{A.6}$$

Finally, the estimate $|R_0| \lessdot 1$ follows from (3.47) and (3.50).

Next, we shall use the following well-known result, whose proof can be found, e.g., in[27] [12] or in [14].

Lemma 8. *Given a Hamiltonian $\mathtt{H}_0$ as in* (A.5)*. For suitable constants $0 < \mathbf{c}_1 < \mathbf{c_0}/8n\mathbf{c}_2$, depending only on κ, n, there exist a (near-identity) symplectic transformation*

$$\begin{aligned} \Phi_1 :\ & \{|y_1| < \mathbf{c}_1\varepsilon^{1/4}\} \times \hat{D}_{\mathtt{r}/2} \times \{|x_1| < \mathbf{c}_1\varepsilon^{1/4}\} \times \mathbb{T}^{n-1}_{\mathtt{s}/2} \longrightarrow \\ & \{|Y_1| < \mathbf{c_0}\varepsilon^{1/4}\} \times \hat{D}_{3\mathtt{r}/4} \times \{|X_1| < \mathbf{c_0}\varepsilon^{1/4}\} \times \mathbb{T}^{n-1}_{3\mathtt{s}/4}, \end{aligned} \tag{A.7}$$

and a function $R_{\rm hp}(z, \hat{y})$ satisfying (3.52) *and* (3.56)*, such that $\mathtt{H}_{\rm hp}(y, x_1) := \mathtt{H}_0 \circ \Phi_1(y, x)$ satisfies* (3.53)*. Moreover, Φ_1 has the form*

$$\begin{aligned} &Y_1 = y_1 + \varepsilon^{1/4}a_1(\varepsilon^{-1/4}y_1, \hat{y}, \varepsilon^{-1/4}x_1), \quad \hat{Y} = \hat{y}, \\ &X_1 = x_1 + \varepsilon^{1/4}a_2(\varepsilon^{-1/4}y_1, \hat{y}, \varepsilon^{-1/4}x_1), \quad \hat{X} = \hat{x} + \sqrt{\varepsilon}\mathtt{r}^{-1}a_3(\varepsilon^{-1/4}y_1, \hat{y}, \varepsilon^{-1/4}x_1), \end{aligned} \tag{A.8}$$

for suitable functions $a_i(\tilde{y}_1, \hat{y}, \tilde{x}_1)$, $i = 1, 2, 3$, which are holomorphic and bounded by $\mathbf{c}_2$ on

$$\{|\tilde{y}_1| < \mathbf{c_0}/2\} \times \hat{D}_{\mathtt{r}/2} \times \{|\tilde{x}_1| < \mathbf{c_0}/2\},$$

moreover, a_1, a_2 and a_3 are, respectively, at least quadratic and cubic in $\tilde{y}_1, \tilde{x}_1$.

Remark 6. $\mathtt{H}_{\rm hp}$ is simply the hyperbolic Birkhoff normal form of $\mathtt{H}_0$. Any canonical transformation of the form $y_1 = \alpha\tilde{y}_1 + \beta\tilde{x}_1$, $x_1 = \beta\tilde{y}_1 + \alpha\tilde{x}_1$, with $\alpha^2 - \beta^2 = 1$ and $\hat{y} = \hat{\tilde{y}}$ leaves $\mathtt{H}_{\rm hp}$ invariant since $y_1^2 - x_1^2 = \tilde{y}_1^2 - \tilde{x}_1^2$. Namely, the integrating transformation Φ_1 is not unique. However, as is well known, the form of the integrated Hamiltonian $\mathtt{H}_{\rm hp}$ in (3.53) is unique, in the sense that E_{2j}, g and R are unique.

Note also that the map Φ_1 is close to the identity, for small $\mathbf{c}_1$, since its Jacobian is the identity plus a matrix whose entries are (by the Cauchy estimates) uniformly bounded on its domain in (3.51) by $2\mathbf{c}_2\mathbf{c}_1/\mathbf{c_0} \leqslant 1/4n$.

[27] See, in particular, Lemma 0 and Appendix A.3 in [12].

Let us come back to the proof of Proposition 2 and let us prove (3.56).

Evaluating (3.53) for $\mu = 0$, we get

$$\begin{aligned}\bar{\mathtt{H}}_1(y, x_1) &:= \bar{\mathtt{H}}_1(y, x_1)|_{\mu=0} = \bar{\mathtt{H}}_0 \circ \bar{\Phi}_1(y, x)\\ &= \bar{E}_{2j} + \bar{g}(y_1^2 - x_1^2) + \varepsilon \bar{R}_{\rm hp}\left(\frac{y_1^2 - x_1^2}{\sqrt{\varepsilon}}\right) = O(\varepsilon)\end{aligned} \tag{A.9}$$

on the domain defined in (3.51). Let us denote $\bar{\mathtt{H}}_0 := \mathtt{H}_0|_{\mu=0}$. Since by (A.3), (A.6), (3.47) one has $\mathtt{H}_0 - \bar{\mathtt{H}}_0 = O(\varepsilon\mu)$,

$$\mathtt{H}_0 \circ \bar{\Phi}_1 = \bar{\mathtt{H}}_0 \circ \bar{\Phi}_1 + (\mathtt{H}_0 - \bar{\mathtt{H}}_0) \circ \bar{\Phi}_1 = \bar{\mathtt{H}}_1 + R_1, \quad \text{with} \quad R_1 = O(\varepsilon\mu),$$

namely, the system is integrated up to a small term of order $\varepsilon\mu$. Note also that, since $\bar{\Phi}_1$ has the form in (A.8), it leaves invariant the terms of order $\leqslant 2$ in (y_1, x_1), namely,

$$\mathtt{H}_0 \circ \bar{\Phi}_1 = E_{2j} + g(y_1^2 - x_1^2) + \varepsilon \bar{R} + Q, \quad \text{with} \quad Q = O(\varepsilon\mu). \tag{A.10}$$

Now we want to construct a symplectic transformation Φ_μ integrating $\mathtt{H}_0 \circ \bar{\Phi}_1$. Since $\bar{\mathtt{H}}_1$ is already in normal form, we claim that the integrating transformation Φ_μ is $O(\varepsilon^{1/4}\mu)$-close to the identity and

$$\mathtt{H}_0 \circ \bar{\Phi}_1 \circ \Phi_\mu = (\bar{\mathtt{H}}_1 + R_1) \circ \Phi_\mu =: \mathtt{H}'_{\rm hp} = \bar{\mathtt{H}}_1 + O(\varepsilon\mu), \tag{A.11}$$

where $\mathtt{H}'_{\rm hp}$ is in normal form, namely, as in (3.53). By the unicity of the Birkhoff normal form, we deduce that $\mathtt{H}_{\rm hp} = \mathtt{H}'_{\rm hp} = \bar{\mathtt{H}}_1 + O(\varepsilon\mu)$. By (3.53), (A.9), (3.50) and (3.2), we get (3.56).

It remains to prove (A.11). The crucial point here is that the generating function[28] χ of the integrating transformation Φ_μ is $O(\sqrt{\varepsilon}\mu)$ and its gradient is, by the Cauchy estimates, $O(\varepsilon^{1/4}\mu)$ in a domain $\{|y_1|, |x_1| \lessdot \varepsilon^{1/4}\}$. The fact that $\chi = O(\sqrt{\varepsilon}\mu)$ can be easily seen by passing, as is usual in Birkhoff's normal form, to the coordinate $\xi = (y_1 - x_1)/\sqrt{2}$, $\eta = (y_1 + x_1)/\sqrt{2}$. In these coordinates, recalling (A.10), we get

$$\mathtt{H}_0 \circ \bar{\Phi}_1 = E_{2j} + 2g\xi\eta + \varepsilon \bar{R}'(\xi, \eta) + Q'(\xi, \eta)$$

with $\bar{R}' = \bar{R}_{\rm hp}(2\xi\eta/\sqrt{\varepsilon}) = O(1)$ and $Q' = O(\varepsilon\mu)$. Note that the Taylor expansion of $\bar{R}'$ contains only a monomial of the form $\bar{R}'_{hh}\xi^h\eta^h$. At the first step, we have to cancel all the monomials of Q' of the form $Q'_{hk}\xi^h\eta^k$ with $h + k = 3$. The generating function$\chi^{(3)}$ of the first step is exactly

$$\chi^{(3)} = \sum_{h+k=3} \frac{Q'_{hk}}{2g(h-k)} \xi^h \eta^k \overset{(3.50)}{=} O(\sqrt{\varepsilon}\mu).$$

After this first step the Hamiltonian becomes $E_{2j} + 2g\xi\eta + \varepsilon\bar{R}'(\xi, \eta) + Q''(\xi, \eta)$ with $Q'' = O(\varepsilon\mu)$. At the second step, we have to cancel all the monomials of Q'' of the form $Q''_{hk}\xi^h\eta^k$ with $h + k = 4$, $h \neq k$. We proceed as in the first step with analogous estimates. Analogously for the other infinite steps, obtaining

$$E_{2j} + 2g\xi\eta + \varepsilon\bar{R}'(\xi, \eta) + \bar{Q}(\xi, \eta)$$

with $\bar{Q} = O(\varepsilon\mu)$ and $\bar{Q}_{hk} = 0$ for $h \neq k$, proving (A.11) (recall (3.50) and (3.2)).

We can conclude the proof of Proposition 2:

The composition of the symplectic transformations defined in (A.1), (A.4), (3.51) integrates $\mathtt{H}$, namely, (3.53) holds[29] with $\Phi_{\rm hp} := \Phi_* \circ \Phi_0 \circ \Phi_1$ satisfying (3.51), (3.54) and (3.55).
The inclusion (3.57) follows by (3.54) and (3.50).

[28] According to Lie's series method.
[29] As well as (3.52) and (3.56) by Lemma 8.

APPENDIX B. PROOFS OF TWO SIMPLE LEMMATA

B.1. Proof of Lemma 1

We know that $\partial_\theta \bar{\mathtt{G}}(\bar\theta_i) = 0$ and we want to solve the equation $\partial_\theta \mathtt{G}\big(\hat p, \theta_i(\hat p)\big) = 0$. Equivalently, for $\mu \leqslant 2^{-8}\kappa^{-6}$, we want to find a real-analytic $y = y(\hat p)$, $\hat p \in \hat D_{\mathtt{r}}$, with

$$\sup_{\hat D_{\mathtt{r}}} |y| \leqslant \rho := \tfrac{2\varepsilon\mu}{\beta \mathtt{s}} \overset{(2.10)}{\leqslant} \tfrac{\mathtt{s}}{2}, \tag{B.1}$$

by solving the equation

$$\partial_\theta \mathtt{G}\big(\hat p, \bar\theta_i + y(\hat p)\big) = 0, \tag{B.2}$$

so that $\theta_i(\hat p) = \bar\theta_i + y(\hat p)$. We have

$$\partial_\theta \mathtt{G}(\hat p, \bar\theta_i + y) = \partial_\theta \mathtt{G}(\hat p, \bar\theta_i) + g(\hat p, y) y, \quad \text{where} \quad g(\hat p, y) := \int_0^1 \partial_\theta^2 \mathtt{G}(\hat p, \bar\theta_i + ty)dt.$$

Then (B.2) can be written as the fixed point equation

$$y = \Psi(y), \quad \text{where} \quad \Psi(y) := -\frac{\partial_\theta \mathtt{G}(\hat p, \bar\theta_i)}{g(\hat p, y)}$$

to be solved in the closed set of the real-analytic functions $y = y(\hat p)$ on $\hat D_{\mathtt{r}}$ satisfying the bound (B.1). Note that, since $\partial_\theta \bar{\mathtt{G}}(\bar\theta_i) = 0$, by (2.4) we have $|\partial_\theta^2 \bar{\mathtt{G}}(\bar\theta_i)| \geqslant \beta$. Moreover, by (2.8) and the Cauchy estimates, we get for $|y| \leqslant \rho$ and $\hat p \in \hat D_{\mathtt{r}}$

$$|g - \partial_\theta^2 \bar{\mathtt{G}}(\bar\theta_i)| \leqslant \frac{4\varepsilon\mu}{\mathtt{s}^2}, \quad \text{which implies} \quad |g| \geqslant \beta - \frac{4\varepsilon\mu}{\mathtt{s}^2} \overset{(2.10)}{\geqslant} \frac{\beta}{2}. \tag{B.3}$$

Again by $\partial_\theta \bar{\mathtt{G}}(\bar\theta_i) = 0$, (2.8) and the Cauchy estimates, we find uniformly on $\hat D_{\mathtt{r}}$ that

$$|\partial_\theta \mathtt{G}(\hat p, \bar\theta_i)| \leqslant \varepsilon\mu/\mathtt{s}. \tag{B.4}$$

Then by (B.3) we obtain for $|y| \leqslant \rho$ and $\hat p \in \hat D_{\mathtt{r}}$

$$|\Psi| \leqslant \tfrac{2\varepsilon\mu}{\beta\mathtt{s}} = \rho, \tag{B.5}$$

by (2.10) and (B.1). Moreover,

$$\partial_y \Psi(y) := \frac{\partial_\theta \mathtt{G}(\hat p, \bar\theta_i)}{\big(g(\hat p, y)\big)^2} \partial_y g(\hat p, y).$$

Then for $|y| \leqslant \rho$ and $\hat p \in \hat D_{\mathtt{r}}$ we get

$$|\partial_y \Psi| < 2^6 \tfrac{\varepsilon^2\mu}{\beta^2 \mathtt{s}^4} \leqslant 2^6 \kappa^6 \mu \leqslant 1, \tag{B.6}$$

by (B.4), (B.3), (2.10) and since $|\partial_y g(\hat p, y)| < 16\varepsilon/\mathtt{s}^3$ by (2.8) and the Cauchy estimates. In conclusion, by (B.5) and (B.6) we have that Ψ is a contraction and the fixed point theorem applies proving the first estimate in (3.2).

Let us now show the second estimate in (3.2).

By (2.8), the first estimate in (3.2), (2.10) and the Cauchy estimates, we get

$$\begin{aligned} |E_i(\hat p) - \bar E_i| &\leqslant |\mathtt{G}\big(\hat p, \theta_j(\hat p)\big) - \bar{\mathtt{G}}\big(\theta_j(\hat p)\big)| + |\bar{\mathtt{G}}\big(\theta_j(\hat p)\big) - \bar{\mathtt{G}}(\bar\theta_j)| \\ &\leqslant \varepsilon\mu + \tfrac{2\varepsilon^2\mu}{\beta\mathtt{s}^2} \leqslant 3\kappa^3\varepsilon\mu, \end{aligned}$$

proving the second estimate in (3.2).

Let us prove the final claim. By (2.11) (applied to $\bar{\mathtt{G}}$) and by the Cauchy estimates, it follows that the minimal distance between two critical points of $\bar{\mathtt{G}}$ can be estimated from below by $2\beta\mathtt{s}^2/\varepsilon$.

Thus, by the first estimates in (3.2), it follows that the relative order of the critical points of $\bar{\mathtt{G}}$ is preserved, provided $8\varepsilon^3\mu^2 < \beta^3\mathtt{s}^4$, which, using (2.10), is implied by $2^3\kappa^7\mu^2 < 1$, which, in turn, is implied by the hypothesis $\mu \leqslant 1/(2\kappa)^6$.

As for critical energies, since $\bar{\mathtt{G}}$ is β-Morse, they are at least β apart; hence, from the second estimate in (3.2) the claim follows provided $3\kappa^3\varepsilon\mu < \beta$, which, by (2.10), is implied by $\mu < 1/(3\kappa^4)$, which, again, is implied by the hypothesis.

B.2. Proof of Lemma 6

First denote $R(z) := w(z) - \cos z$, so that $|R|_1 \leqslant \mathtt{g}_{\rm o}$. We note that, on the real line, w has exactly two critical points: a maximum x_M (with $w(x_M) = 1$) and a minimum x_m (with $w(x_m) = -1$) in the interval $[-\pi/2, 3\pi/2)$. Indeed, since by the Cauchy estimates $\sup_R |w'| \leqslant \mathtt{g}_{\rm o}$, the equation $w'(x) = -\sin x + R'(x) = 0$ in the interval $[-\pi/2, 3\pi/2)$ has only two solutions x_M, x_m with $|x_M|, |x_m - \pi| \leqslant 1.0001\mathtt{g}_{\rm o} \leqslant 0.001$. Obviously, $x_M + b(x_M) = 0$ and $x_m + b(x_m) = \pi$.

On the real line the function b is given by the 2π-periodic continuous[30] function defined in the interval $[x_m - 2\pi, x_m]$ by the expression

$$b(x) := \operatorname{sign}(x - x_M) \arccos\big(w(x)\big) - x.$$

Let us consider first the complex domain $\Omega_0 := \{0.4 < \operatorname{Re} z < \pi - 0.4,\ |\operatorname{Im} z| < 1/4\}$ where $b(z)$ is clearly extendible to a holomorphic function. Here we have $\sup_{\Omega_0} |\cos z| \leqslant 0.913$ and, therefore, $\sup_{\Omega_0} |\cos z| + |R(z)| \leqslant 0.914$. Then for $z \in \Omega_0$ we get

$$|b(z)| = |\arccos\big(\cos z + R(z)\big) - z| \leqslant \int_0^1 \left| \frac{R(z)}{\sqrt{1 - \big(\cos z + tR(z)\big)^2}} \right| dt \leqslant 6.1\mathtt{g}_{\rm o}.$$

We now prove that $b(z)$ is extendible to a holomorphic function for $|z| < 1/2$. First we prove that there exists a real-analytic positive function d with holomorphic extension on $|z| < 1/2$ such that $w(z) = 1 - \frac{1}{2}\big((z - x_M)d(z)\big)^2$. By Taylor's expansion at $z = x_M$ we have that $d^2(z) = -2\int_0^1 (1-t) w''\big(x_M + t(z - x_M)\big)dt$ and, therefore, for $|z| < 1/2$

$$|d^2(z) - 1| \leqslant 1 - \cos x_M + \sup_{|z|<1/2} |\sin z||z - x_M| + 2\mathtt{g}_{\rm o} \leqslant 0.55.$$

Then we can take the principle square root[31] of $d^2(z)$, obtaining the function $d(z)$. Now consider the holomorphic function $a(z)$ defined for $|z| < 2$ such that $a'(z) = 1/\sqrt{1 - (z/2)^2}$ and $a(0) = 0$. Then for real x we get $a(x) = \operatorname{sign}(x) \arccos(1 - x^2/2)$ and also (with $d(x) > 0$)

$$b(x) := \operatorname{sign}(x - x_M) \arccos\Big(1 - \frac{1}{2}\big((x - x_M)d(x)\big)^2\Big) - x = a\big((x - x_M)d(x)\big) - x.$$

Then $a\big((z - x_M)d(z)\big) - z$ is a holomorphic extension of b for $|z| < 2$. An analogous argument holds for $|z - \pi| < 2$.

In the following we will estimate $b(z)$ for a strip $|z| < 1/2$, analogous arguments holds for $|z - \pi| < 1/2$. We will often use that[32]

$$|z| \leqslant 1 \quad \Longrightarrow \quad 0.45|z|^2 \leqslant |1 - \cos z| \leqslant 0.55|z|^2. \tag{B.7}$$

Now we prove that there exists a unique function $b(z)$ defined for

$$\Omega_1 := \{3\sqrt{\mathtt{g}_{\rm o}} < |z| < 1/2\}$$

---

[30] Since $b(x_m - 2\pi) = b(x_m) = \pi - x_m$.

[31] Namely, taking a cut in the negative real line.

[32] Using that $\frac{1}{2}|z|^2 - (\cosh|z| - 1 - \frac{1}{2}|z|^2) \leqslant |1 - \cos z| \leqslant \cosh|z| - 1$.

satisfying $\sup_{\Omega_1}|b| \leqslant \frac{3}{2}\sqrt{\mathrm{g_o}}$, such that $w(z) = \cos\big(z + b(z)\big)$, as a fixed point of the equation

$$b(z) = \Psi(b)(z) := 2\arcsin\left(\frac{-R(z)}{2\sin(z + b(z)/2)}\right).$$

Indeed,

$$\cos\big(z + b(z)\big) - \cos z = -2\sin\left(z + b(z)/2\right)\sin\left(b(z)/2\right) = R(z).$$

For $z \in$ we have $|z + b(z)/2| \geqslant \frac{3}{2}\sqrt{\mathrm{g_o}}$, which implies[33] $|\sin(z + b(z)/2)| \geqslant \frac{6}{5}\sqrt{\mathrm{g_o}}$ and[34] $\sup_{\Omega_1}|\Psi(b)(z)| < \sqrt{\mathrm{g_o}}$. Finally, Ψ is a contraction since[35]

$$\begin{aligned}&\sup_{\Omega_1}|\Psi(b) - \Psi(b')| \leqslant \tfrac{2}{\sqrt{3}}\mathrm{g_o}\sup_{\Omega_1}\left|\tfrac{1}{\sin(z+b(z)/2)} - \tfrac{1}{\sin(z+b'(z)/2)}\right| \\ &\leqslant \tfrac{2}{\sqrt{3}}\tfrac{5^2}{6^2}2\left|\sin\left(\tfrac{b'(z)-b(z)}{4}\right)\cos\left(z + \tfrac{b'(z)+b(z)}{4}\right)\right| \leqslant \tfrac{5}{6}\sup_{\Omega_1}|b - b'|.\end{aligned}$$

In conclusion, we get $\sup_{\Omega_1}|b| \leqslant \frac{3}{2}\sqrt{\mathrm{g_o}}$.

Next, we claim that in the domain $\Omega_2 := \{|z| \leqslant 3\sqrt{\mathrm{g_o}}\}$ we have that $|b(z)| < 9\sqrt{\mathrm{g_o}}$. Indeed, by contradiction, assume that there exists $z_0 \in \Omega_2$ such that for every $|z| < |z_0|$ we have $|b(z)| < 9\sqrt{\mathrm{g_o}}$, but $|b(z_0)| = 9\sqrt{\mathrm{g_o}}$. Then $|z_0 + b(z_0)| \leqslant 12\sqrt{\mathrm{g_o}}$ and by (B.7) and since $\cos\big(z_0 + b(z_0)\big) - 1 = \cos z_0 - 1 + R(z_0)$ we get

$$\begin{aligned}&16\mathrm{g_o} \leqslant 0.45(|b(z_0)| - |z_0|)^2 \leqslant 0.45|z_0 + b(z_0)|^2 \leqslant |\cos\big(z_0 + b(z_0)\big) - 1| \\ &\leqslant |\cos z_0 - 1| + |R(z_0)| \leqslant 0.55|z_0|^2 + \mathrm{g_o} \leqslant 6\mathrm{g_o},\end{aligned}$$

which is a contradiction. Thus, $\sup_{\Omega_2}|b(z)| \leqslant 9\sqrt{\mathrm{g_o}}$. □

ACKNOWLEDGMENTS

The authors are grateful to A. Neishtadt for providing parts of his thesis ([20], in Russian) related to the present paper.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Arnol'd, V. I., On the Nonstability of Dynamical Systems with Many Degrees of Freedom, *Soviet Math. Dokl.,* 1964, vol. 5, no. 3, pp. 581–585; see also: *Dokl. Akad. Nauk SSSR,* 1964, vol. 156, no. 1, pp. 9–12.
2. Arnol'd, V. I., *Mathematical Methods of Classical Mechanics,* 2nd ed., Grad. Texts in Math., vol. 60, New York: Springer, 1997.
3. Arnol'd, V. I., Kozlov, V. V., and Neishtadt, A. I., *Mathematical Aspects of Classical and Celestial Mechanics,* 3rd ed., Encyclopaedia Math. Sci., vol. 3, Berlin: Springer, 2006.
4. Bambusi, D., Fusè, A., and Sansottera, M., Exponential Stability in the Perturbed Central Force Problem, *Regul. Chaotic Dyn.,* 2018, vol. 23, nos. 7–8, pp. 821–841.
5. Bernard, P., Kaloshin, V., and Zhang, K., Arnold Diffusion in Arbitrary Degrees of Freedom and Normally Hyperbolic Invariant Cylinders, *Acta Math.,* 2016, vol. 217, no. 1, pp. 1–79.
6. Biasco, L. and Chierchia, L., On the Topology of Nearly-Integrable Hamiltonians at Simple Resonances, *Nonlinearity,* 2020, vol. 33, no. 7, pp. 3526–3567.
7. Biasco, L. and Chierchia, L., Quasi-Periodic Motions in Generic Nearly-Integrable Mechanical Systems, *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur.,* 2022, vol. 33, no. 3, pp. 575–580.

[33] Using that for $|z| < 1$ we have $\frac{4}{5}|z| \leqslant |\sin z| \leqslant \frac{6}{5}|z|$.
[34] Using that for $|z| \leqslant 1/2$ we have $|\arcsin z| \leqslant \frac{2}{\sqrt{3}}|z|$.
[35] Using that for $|z| \leqslant 1$ we have $|\cos z| \leqslant \sqrt{3}$.

8. Biasco, L. and Chierchia, L., *Global Properties of Generic Real-Analytic Nearly-Integrable Hamiltonian Systems,* in preparation (2023).
9. Biasco, L. and Chierchia, L., *Singular KAM Theory,* in preparation (2023).
10. Chen, Q. and de la Llave, R., Analytic Genericity of Diffusing Orbits in *a priori* Unstable Hamiltonian Systems, *Nonlinearity,* 2022, vol. 35, no. 4, pp. 1986–2019.
11. Chierchia, L., Kolmogorov – Arnold – Moser (KAM) Theory, in *Mathematics of Complexity and Dynamical Systems: Vol. 2,* R. A. Meyers (Ed.), New York: Springer, 2012, pp. 810–836.
12. Chierchia, L. and Gallavotti, G., Drift and Diffusion in Phase Space, *Ann. Inst. H. Poincaré Phys. Théor.,* 1994, vol. 60, no. 1, 144 pp. (Erratum: Drift and Diffusion in Phase Space, *Ann. Inst. H. Poincaré Phys. Théor.,* 1998, vol. 68, no. 1, p. 135.)
13. Delshams, A., de la Llave, R., and Seara, T. M., Instability of High Dimensional Hamiltonian Systems: Multiple Resonances Do Not Impede Diffusion, *Adv. Math.,* 2016, vol. 294, pp. 689–755.
14. Giorgilli, A., Unstable Equilibria of Hamiltonian Systems, *Discrete Contin. Dynam. Systems,* 2001, vol. 7, no. 4, pp. 855–871.
15. Hofer, H. and Zehnder, E., *Symplectic Invariants and Hamiltonian Dynamics,* Birkhäuser Advanced Texts: Basler Lehrbücher, Basel: Birkhäuser, 1994.
16. Kaloshin, V. and Zhang, K., *Arnold Diffusion for Smooth Systems of Two and a Half Degrees of Freedom,* Ann. Math. Stud., vol. 208, Princeton, N.J.: Princeton Univ. Press, 2020.
17. de la Llave, R. and Wayne, C. E., Whiskered and Low Dimensional Tori in Nearly Integrable Hamiltonian Systems, *Math. Phys. Electron. J.,* 2004, vol. 10, Paper 5, 45 pp.
18. Mather, J. N., Arnold Diffusion by Variational Methods, in *Essays in Mathematics and Its Applications,* Heidelberg: Springer, 2012, pp. 271–285.
19. Medvedev, A. G., Neishtadt, A. I., and Treschev, D. V., Lagrangian Tori near Resonances of Near-Integrable Hamiltonian Systems, *Nonlinearity,* 2015, vol. 28, no. 7, pp. 2105–2130.
20. Neishtadt, A. I., Problems of Perturbation Theory for Nonlinear Resonant Systems, *Doctoral Dissertation,* Moscow State University, Moscow, 1988, 342 pp. (Russian).
21. Neishtadt, A. I., On the Change in the Adiabatic Invariant on Crossing a Separatrix in Systems with Two Degrees of Freedom, *J. Appl. Math. Mech.,* 1987, vol. 51, no. 5, pp. 586–592; see also: *Prikl. Mat. Mekh.,* 1987, vol. 51, no. 5, pp. 750–757.
22. Okunev, A., *On the Fourier Coefficients of the Perturbation Written Using the Angle Variable Near a Separatrix Loop,* Loughborough University, https://doi.org/10.17028/rd.lboro.12278741.v1 (2020).
23. Treschev, D., Arnold Diffusion Far from Strong Resonances in Multidimensional *a priori* Unstable Hamiltonian Systems, *Nonlinearity,* 2012, vol. 25, no. 9, pp. 2717–2757.
24. Zhang, K., Speed of Arnold Diffusion for Analytic Hamiltonian Systems, *Invent. Math.,* 2011, vol. 186, no. 2, pp. 255–290.