

# Negligibility of small divisor effects in the normal form theory for nearly-integrable Hamiltonians with decaying non-autonomous perturbations

Alessandro Fortunati<sup>1</sup> · Stephen Wiggins<sup>1</sup>

Received: 1 October 2015 / Revised: 28 January 2016 / Accepted: 14 March 2016 /  
Published online: 1 April 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** The paper deals with the problem of the existence of a normal form for a nearly-integrable real-analytic Hamiltonian with aperiodically time-dependent perturbation decaying (slowly) in time. In particular, in the case of an isochronous integrable part, the system can be cast in an exact normal form, regardless of the properties of the frequency vector. The general case is treated by a suitable adaptation of the finite order normalization techniques usually used for Nekhoroshev arguments. The key point is that the so called “geometric part” is not necessary in this case. As a consequence, no hypotheses on the integrable part are required, apart from analyticity. The work, based on two different perturbative approaches developed by Giorgilli et al., is a generalisation of the techniques used by the same authors to treat more specific aperiodically time-dependent problems.

**Keywords** Non-autonomous Hamiltonian systems · Stability · Birkhoff normal forms · Aperiodic time dependence

**Mathematics Subject Classification** Primary: 70H08; Secondary: 37J40 · 37J25

## 1 Introduction

The problem of casting an analytic nearly-integrable Hamiltonian system into normal form is deeply related to Poincaré’s challenging *problème général de la dynamique*, Poincaré (1892). Nowadays, normal forms are still one of the main technical tools used to deal with the issue raised by Poincaré in this context.

---

✉ Alessandro Fortunati  
alessandro.fortunati@bristol.ac.uk

Stephen Wiggins  
s.wiggins@bristol.ac.uk

<sup>1</sup> School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

The particular case in which the unperturbed part is supposed to be linear in the actions (isochronous case), already investigated by Birkhoff (and for this reason also known as the *Birkhoff problem*) Birkhoff (1927), has a peculiar interest. The first rigorous statement concerning its stability can be found in Gallavotti (1986). The possibility to cast the considered Hamiltonian in normal form, up to some finite order<sup>1</sup>  $r$  and to obtain, as a consequence, a stability time estimate “à la Nekhoroshev”, is directly related to a particularly simple small-divisors analysis: the non-resonant (Diophantine) hypothesis on the frequency vector  $\omega$  of the unperturbed system is sufficient in order to ensure the resolvability of the (standard) homological equation arising in the normalization algorithm. An extensive bibliography on this problem goes beyond the purposes of this paper, we only mention the recent generalisations for the planetary problem of Pinzari (2013) and of Bambusi (2005) for infinite dimensional systems.

It is well known that the extension to the non-isochronous case requires a careful analysis (geometric part, see Nekhoroshev 1977, 1979; Benettin and Gallavotti 1986) on the regions of the phase space in which the actions  $I$  are such that  $\omega = \omega(I)$  is non-resonant (non-resonant domains).

The problem of dealing with time-dependent perturbations without any hypothesis on the time dependence (e.g. periodic or quasi-periodic) has peculiar technical difficulties. After the pioneering works of Pustyl'nikov (1974) and Giorgilli and Zehnder (1992), the interest for this class of problems has been recently renewed in Bounemoura (2013), Fortunati and Wiggins (2014a) and subsequent papers. Examples of more general (i.e. *aperiodic*) non-autonomous perturbation in the context of the Lagrangian transport theory for fluids have been pointed out in Wiggins and Mancho (2014). Despite towards a different direction it is worth mentioning the stochastic perturbations of the Kepler problem discussed in Cresson et al. (2015), naturally arising in some Celestial Mechanics models.

From a technical point of view, the presence of an aperiodic time dependence, requires a different treatment of the homological equation which takes the form of a linear PDE. A first approach consists in keeping the terms involving the time derivative of the generating function (also called extra-terms) in the normal form and then providing a bound for them. This approach, originally suggested in Giorgilli and Zehnder (1992) then used in Fortunati and Wiggins (2014a), yields a normal form result for the case a of slow time dependence. This hypothesis provides a smallness condition for the mentioned extra-terms. Alternatively, those terms can be removed by including them into the homological equation, which turns out to be, in this way, a linear ODE in time. This has been profitably used in Fortunati and Wiggins (2014b), Fortunati and Wiggins (2015a) and in Fortunati and Wiggins (2015b) but requires (except for a particular case described in Fortunati and Wiggins 2015b) an important assumption. More precisely, it is necessary to suppose that the perturbation, as a function of  $t$ , belongs to the class of summable functions over the real semi-axis.<sup>2</sup> As in (3), those functions exhibiting a (slow) exponential decay will be used as a paradigmatic case. It will be shown that the consequences of this assumption in the isochronous case are remarkable: the normalization algorithm can be iterated an infinite number of times by means of a *superconvergent method* borrowed from KAM type arguments, see e.g. Chierchia (2009). The procedure leads to the so-called *strong normal form* i.e. in which the normalized Hamiltonian has the same form of the integrable part of the initial problem. Furthermore, no restrictions are imposed on  $\omega$ , hence flows with arbitrary frequencies persist in the transformed system.

<sup>1</sup> It is easy to see that any attempt to consider the limit  $r \rightarrow \infty$  would imply the degeneration into a trivial problem, (i.e. in which the allowed perturbation size reduces to zero, see also Giorgilli and Galgani 1985, formula (46), P. 105).

<sup>2</sup> We stress that this hypothesis is usually not satisfied in the case of periodic or quasi-periodic time dependence.

As it would be likely to expect, this phenomenon has an important consequence also in the non-isochronous case. The possibility to disregard the problems related to the small divisors implies that the well known *geography of the resonances* analysis, a key step of the Nekhoroshev theorem, is not necessary in this case and the results that can be stated are purely “analytic”. In such a way, the classical assumptions on the unperturbed part of the Hamiltonian (such as steepness, convexity etc.), are no longer required. As a common feature with the isochronous case, the obtained normal form does not exhibit *resonant terms*, as these have been annihilated in the normalization by using the time-dependent homological equation. This implies that, in this case, the *plane of fast drift* (see e.g. [Giorgilli 2003](#)) degenerates to a point. The paper uses in a concise but self-contained form, the tools developed in the above mentioned papers of the same authors, especially of [Fortunati and Wiggins \(2015b\)](#) in which the concept of “family” of canonical transformations parametrised by  $t$  is introduced. The proofs are entirely constructed by using the language and the tools of the Lie series and Lie transform methods developed by Giorgilli et al., see e.g. [Giorgilli \(2003\)](#).

## 2 Setting and main results

Consider the following nearly integrable Hamiltonian

$$H(I, \varphi, \eta, t) = h(I) + \eta + \hat{\varepsilon} f(I, \varphi, t), \tag{1}$$

with  $(I, \varphi, \eta, t) \in G \times \mathbb{T}^n \times \mathbb{R} \times \mathbb{R}^+$ , where  $G \subset \mathbb{R}^n$  and  $\hat{\varepsilon} > 0$  is a small parameter, which is the “autonomous equivalent” in the extended phase space of Hamiltonian  $\mathcal{H}(I, \varphi, t) = h(I) + \hat{\varepsilon} f(I, \varphi, t)$ .

We define, for all  $t \in \mathbb{R}^+ := [0, +\infty)$ , the following complexified domain  $\mathcal{D}_{\rho, \sigma} := \mathcal{G}_\rho \times \mathbb{T}_\sigma^n \times \mathcal{S}_\rho$ , where  $\mathcal{G}_\rho := \bigcup_{I \in G} \Delta_\rho(I)$  and

$$\Delta_\rho(I) := \left\{ \hat{I} \in \mathbb{C}^n : |\hat{I} - I| \leq \rho \right\}, \quad \mathbb{T}_\sigma^n := \{ \varphi \in \mathbb{C}^n : |\Im \varphi| \leq \sigma \}, \quad \mathcal{S}_\rho := \{ \eta \in \mathbb{C} : |\Im \eta| \leq \rho \},$$

with  $\rho, \sigma \in (0, 1)$ . For all  $g : \mathcal{G}_\rho \times \mathbb{T}_\sigma^n \times \mathbb{R}^+ \rightarrow \mathbb{C}$ , write  $g = \sum_{k \in \mathbb{Z}^n} g_k(I, t) e^{ik \cdot \varphi}$ , then define the *Fourier norm* (parametrized by  $t$ )

$$\|g\|_{\rho, \sigma} := \sum_{k \in \mathbb{Z}^n} |g_k(I, t)|_\rho e^{|k|\sigma}, \tag{2}$$

with  $|\cdot|_\rho$  is the usual supremum norm over  $\mathcal{G}_\rho$  and  $|k| := \sum_{l=1}^n |k_l|$ . For all  $w : \mathcal{G}_\rho \times \mathbb{T}_\sigma^n \times \mathbb{R}^+ \rightarrow \mathbb{C}^n$  we shall set  $\|w\|_{\rho, \sigma} := \sum_{l=1}^n \|w_l\|_{\rho, \sigma}$ . The standard framework (see eg. [Benettin et al. 1984](#)) is the space  $\mathfrak{C}_{\rho, \sigma}$ , of continuous functions on  $\mathcal{G}_\rho \times \mathbb{T}_\sigma^n$ , holomorphic in its interior for some  $\rho, \sigma$  and real on  $G \times \mathbb{T}^n$  for all<sup>3</sup>  $t \in \mathbb{R}^+$ . We shall suppose  $h(I) \in \mathfrak{C}_{\rho, \cdot}$  and  $f \in \mathfrak{C}_{\rho, \sigma}$  while it is sufficient to assume that, for all  $I \in \mathcal{G}_\rho$ ,  $f_k(I, \cdot) \in \mathcal{C}^1(\mathbb{R}^+)$ .

Similarly to [Fortunati and Wiggins \(2015b\)](#), we introduce the following

**Hypothesis 2.1** (*Time decay*) There exists  $M_f > 0$  and  $a \in (0, 1)$

$$\|f(I, \varphi, t)\|_{\rho, \sigma} \leq M_f e^{-at}. \tag{3}$$

Set  $\varepsilon := \hat{\varepsilon} M_f$ . We firstly state the following

**Theorem 2.2** (Strong aperiodic Birkhoff) *Consider Hamiltonian (1) with  $h(I) := \omega \cdot I$ , under the Hypothesis 2.1 and the described regularity assumptions. Then, for all  $a \in (0, 1)$*

<sup>3</sup> In particular, if  $g \in \mathfrak{C}_{\rho, \sigma}$  then  $|g_k|_\rho \leq \|g\|_{\rho, \sigma} \exp(-|k|\sigma)$  for all  $t \in \mathbb{R}^+$ .

there exists  $\varepsilon_a > 0$  such that the following statement holds true. For all  $\varepsilon \in (0, \varepsilon_a]$ , it is possible to find  $0 < \rho_* < \rho_0 < \rho$  and  $0 < \sigma_* < \sigma_0 < \sigma$  and an analytic, canonical,  $\varepsilon$ -close and asymptotic to the identity change of variables  $(I, \varphi, \eta) = \mathcal{B}(I^{(\infty)}, \varphi^{(\infty)}, \eta^{(\infty)})$ ,  $\mathcal{B} : \mathcal{D}_{\rho_*, \sigma_*} \rightarrow \mathcal{D}_{\rho_0, \sigma_0}$  for all  $t \in \mathbb{R}^+$ , casting Hamiltonian (1) into the strong Birkhoff normal form

$$H^{(\infty)} \left( I^{(\infty)}, \varphi^{(\infty)}, \eta^{(\infty)} \right) = \omega \cdot I^{(\infty)} + \eta^{(\infty)}. \tag{4}$$

Hence, in the new variables, the flow with frequency  $\omega$  persists for all  $\omega$ , regardless of the numerical features of this vector, i.e. more specifically, no matter if it is resonant or not. Note that the absence of a non-resonance hypothesis on  $\omega$  implies also that (4) holds also if  $\omega$  has an arbitrary number of zero components.

With a straightforward adaptation of the notational setting, the result in the general case states as follows:

**Theorem 2.3** *There exist  $\varepsilon_a^* > 0$  and  $r \in \mathbb{N} \setminus \{0\}$  such that, for all  $\varepsilon \in (0, \varepsilon_a^*]$  it is possible to find an analytic, canonical,  $\varepsilon$ -close and asymptotic to the identity change of variables  $(I, \varphi, \eta) = \mathcal{N}_r \left( I^{(r)}, \varphi^{(r)}, \eta^{(r)} \right)$ ,  $\mathcal{N}_r : \mathcal{D}_{\tilde{\rho}_*, \tilde{\sigma}_*} \rightarrow \mathcal{D}_{\tilde{\rho}_0, \tilde{\sigma}_0}$  for all  $t \in \mathbb{R}^+$ , casting Hamiltonian (1) under the Hypothesis 2.1, into the normal form of order  $r$*

$$H^{(r)} \left( I^{(r)}, \varphi^{(r)}, \eta^{(r)}, t \right) = h \left( I^{(r)} \right) + \eta^{(r)} + \mathcal{R}^{(r+1)} \left( I^{(r)}, \varphi^{(r)}, t \right), \tag{5}$$

where  $\mathcal{R}^{(r+1)}$  is “exponentially small” with respect to  $r$  and vanishes for  $t \rightarrow +\infty$ . Moreover, for all  $I(0) \in G$  one has in (I):  $|I(t) - I(0)| = \sqrt{\varepsilon} \tilde{\rho}_0 / 8$  for all  $t \in \mathbb{R}^+$ .

Similarly to Fortunati and Wiggins (2015b) (and the mentioned previous papers), no lower bounds are imposed on  $a$  so that the decay can be arbitrarily slow. The (natural) consequence is that either  $\varepsilon_a$  or  $\varepsilon_a^*$  decrease with  $a$ , see (15) and (56).

We stress that, as a difference with the classical (non-autonomous) case, the stability property following from the above stated results, is an easy consequence of (3) and it could have been possible to show it directly from the equations of motion, by means of elementary method, without the use of the normal form approach.

## Part I

### Proof of Theorem 2.2

### 3 The normalization algorithm

Given a function  $G := G(I, \varphi, t)$ , define the Lie series operator  $\exp(\mathcal{L}_G) := \text{Id} + \sum_{s \geq 1} (1/s!) \mathcal{L}_G^s$ , where  $\mathcal{L}_G F := \{F, G\} \equiv F_\varphi \cdot G_I - G_\varphi \cdot F_I - F_\eta G_t$ . The aim is to construct a generating sequence  $\{\chi^{(j)}\}_{j \in \mathbb{N}}$ , such that the formal limit

$$\mathcal{B} := \lim_{j \rightarrow \infty} \mathcal{B}^{(j)} \circ \mathcal{B}^{(j-1)} \circ \dots \circ \mathcal{B}^{(0)}, \tag{6}$$

where  $\mathcal{B}^{(j)} := \exp(\mathcal{L}_{\chi^{(j)}})$  is such that  $\mathcal{B} \circ H$  is of the form (4). The following statement shows that this is possible, at least at a formal level

---

<sup>4</sup> See bound (54).

**Proposition 3.1** *Suppose that for some  $j \in \mathbb{N}$  Hamiltonian (1) is of the form*

$$H^{(j)} = \omega \cdot I + \eta + F^{(j)}(I, \varphi, t). \tag{7}$$

*Then  $H^{(j+1)} := \mathcal{B}^{(j)} \circ H^{(j)}$  is still of the form (7) with*

$$F^{(j+1)} = \sum_{s \geq 1} \frac{s}{(s+1)!} \mathcal{L}_{\chi^{(j)}}^s F^{(j)}, \tag{8}$$

*provided that  $\chi^{(j)}$  solves the homological equation*

$$\chi_t^{(j)} + \omega \cdot \chi_\varphi^{(j)} = F^{(j)}. \tag{9}$$

Since Hamiltonian (1) is of the form (7), one can set  $H^{(0)} := H$  with  $F^{(0)} := \hat{\varepsilon} f$ . Thus, by induction, the form (7) holds for all  $j \in \mathbb{N}$ . Clearly, this does not guarantee that the objects involved in the algorithm are meaningful for all  $j$ , as it is well known their sizes can grow unboundedly as  $j$  increases, as a consequence of *small divisors* phenomena. The aim of Sect. 4 (and in particular of Lemma 4.5) is to show that this is not the case: the key ingredient is the time decay of  $f$ .

*Proof* We get  $\exp(\mathcal{L}_{\chi^{(j)}})H^{(j)} = I \cdot \omega + \eta + F^{(j)}(I, \varphi, t) + \mathcal{L}_{\chi^{(j)}}(\omega \cdot I + \eta) + \sum_{s \geq 1} (1/s!) \mathcal{L}_{\chi^{(j)}}^s F^{(j)} + \sum_{s \geq 2} (1/s!) \mathcal{L}_{\chi^{(j)}}^s (\omega \cdot I + \eta)$ . The sum between the third and fourth terms of the r.h.s. of the latter equation vanishes due to (9). As for the last two terms, by setting  $F^{(j+1)}$  as the sum of them, one gets  $F^{(j+1)} = \sum_{s \geq 1} (1/s!) \mathcal{L}_{\chi^{(j)}} [F^{(j)} + (s+1)^{-1} \mathcal{L}_{\chi^{(j)}}(\omega \cdot I + \eta)]$ , which immediately yields (8) by using (9).  $\square$

The (formal) expansions  $\chi^{(j)} = \sum_{k \in \mathbb{Z}^n} c_k^{(j)}(I, t) e^{ik \cdot \varphi}$  and  $F^{(j)} = \sum_{k \in \mathbb{Z}^n} f_k^{(j)}(I, t) e^{ik \cdot \varphi}$  yield (9) in terms of Fourier components

$$\partial_t c_k^{(j)}(I, t) + i\lambda(k) c_k^{(j)}(I, t) = f_k^{(j)}(I, t), \tag{10}$$

with  $\lambda(k) := \omega \cdot k$ . The solution of (10) is

$$c_k^{(j)}(I, t) = e^{-i\lambda(k)t} \left[ c_k^{(j)}(I, 0) + \int_0^t e^{i\lambda(k)s} f_k^{(j)}(I, s) ds \right], \tag{11}$$

where  $c_k^{(j)}(I, 0)$  will be chosen later.

### 4 Convergence

The classical argument requires the construction of a sequence of nested domains  $\mathcal{D}_{\rho_{j+1}, \sigma_{j+1}} \subset \mathcal{D}_{\rho_j, \sigma_j} \ni (I^{(j)}, \varphi^{(j)}, \eta^{(j)})$ , such that  $\mathcal{B}_j : \mathcal{D}_{j+1} \rightarrow \mathcal{D}_j$ . The resulting progressive restriction is essential in order to use standard Cauchy tools, see Proposition 4.1. The estimates found in Lemma 4.2, concerning the solution of Eq. (9), will be used to prove Lemma 4.5, providing in this way the bound on  $F^{(j)}$  defined in Proposition 3.1. This is achieved for a suitable sequence of domains prepared in Lemma 4.4 via  $\{\rho_j\}$  and  $\{\sigma_j\}$ . This allows us to conclude that the perturbation term is actually removed in the limit (6).

The final step consists of showing that  $\mathcal{B}$  defines an analytic map  $\mathcal{B} : \mathcal{D}_{\rho_*, \sigma_*} \ni (I^{(\infty)}, \varphi^{(\infty)}, \eta^{(\infty)}) \rightarrow \mathcal{D}_{\rho_0, \sigma_0} \ni (I^{(0)}, \varphi^{(0)}, \eta^{(0)}) \equiv (I, \varphi, \eta)$ , where  $\rho_* \leq \rho_j$  and  $\sigma_* \leq \sigma_j$  for all  $j \in \mathbb{N}$ . This property is shown in Lemma 4.6. As  $\mathcal{D}_{\rho_*, \sigma_*}$  will be the domain of analyticity of the transformed Hamiltonian via  $\mathcal{B}$ , it will be essential to require that  $\rho_*, \sigma_* > 0$ .

### 4.1 Some preliminary results

**Proposition 4.1** *Let  $F, G : \mathcal{G}_\rho \times \mathbb{T}_\sigma^n \times \mathbb{R}^+ \rightarrow \mathbb{C}$  such that  $\|F\|_{(1-d')(\rho,\sigma)}$  and  $\|G\|_{(1-d'')(\rho,\sigma)}$  are bounded for some  $d', d'' \in [0, 1)$ . Then, defining  $\delta := |d' - d''|$  and  $\hat{d} := \max\{d', d''\}$ , for all  $\tilde{d} \in (0, 1 - \hat{d})$  one has for all  $s \in \mathbb{N} \setminus \{0\}$*

$$\|\mathcal{L}_G^s F\|_{(1-\tilde{d}-\hat{d})(\rho,\sigma)} \leq \frac{s!}{e^2} \left( \frac{2e}{\tilde{d}(\tilde{d} + \tilde{\delta}_s)} \|G\|_{(1-d'')(\rho,\sigma)} \right)^s \|F\|_{(1-d')(\rho,\sigma)}, \tag{12}$$

where  $\tilde{\delta}_s = \delta$  if  $s = 1$  and is zero otherwise.

*Proof* Straightforward from (Giorgilli 2003, Lemmas 4.1, 4.2). □

**Lemma 4.2** *Suppose that  $F^{(j)}$  satisfies  $\|F^{(j)}\|_{[\hat{\sigma}, \hat{\rho}]} \leq M^{(j)} \exp(-at)$  for some  $M^{(j)} > 0$ ,  $\hat{\rho} \leq \rho$  and  $\hat{\sigma} \leq \sigma$ . Define  $C_\omega := 1 + |\omega|$ , then for all  $\delta \in (0, 1)$  the solution of (9) satisfies*

$$\|\chi^{(j)}\|_{(1-\delta)(\hat{\rho}, \hat{\sigma})} \leq \frac{M^{(j)}}{a} \left(\frac{e}{\delta \hat{\sigma}}\right)^{2n} e^{-at}, \quad \|\chi_t^{(j)}\|_{(1-\delta)(\hat{\rho}, \hat{\sigma})} \leq C_\omega \frac{M^{(j)}}{a} \left(\frac{e}{\delta \hat{\sigma}}\right)^{2n} e^{-at}. \tag{13}$$

*Proof* First of all, by hypothesis  $|f_k^{(j)}(I, t)| \leq M^{(j)} \exp(-|k|\hat{\sigma} - at)$ , in particular, by choosing  $c_k^{(j)}(I, 0) := -\int_{\mathbb{R}^+} \exp(i\lambda(k)s) f_k^{(j)}(I, s) ds$  we have that  $|c_k^{(j)}(I, 0)| < +\infty$  for all  $I \in \mathcal{G}_\rho$ . Substituting  $c_k^{(j)}(I, 0)$  in (11) one gets  $|c_k^{(j)}(I, t)| \leq \int_t^\infty |f_k^{(j)}(I, s)| ds \leq (M^{(j)}/a) \exp(-|k|\hat{\sigma} - at)$  which yields<sup>5</sup> the first of (13). As for the second of (13), it is sufficient to use (10), which implies,  $|\partial_t c_k^{(j)}(I, t)| \leq (M^{(j)}/a)(1 + |\omega||k|) \exp(-|k|\hat{\sigma} - at)$  then proceed similarly. □

*Remark 4.3* It is immediate to notice that a hypothesis of non-resonance on  $\omega$  does not substantially improve the bounds (13). A more careful computation yields

$$|c_k^{(j)}(I, t)| \leq M^{(j)} (a^2 + (\omega \cdot k)^2)^{-\frac{1}{2}} e^{-|k|\sigma_j - at},$$

Hence the estimate cannot be refined due to the presence of  $|c_0^{(j)}(I, t)|$ , no matter what the minimum value of  $(\omega \cdot k)$  is.

### 4.2 A suitable sequence of domains

**Lemma 4.4** *Let  $\{d_j\}_{j \in \mathbb{N}}$  be a (real valued) sequence such that  $0 \leq d_j \leq 1/6$ . Consider, for all  $j \in \mathbb{N}$ , the following sequences*

$$\epsilon_{j+1} := K a^{-1} d_j^{-\tau} \epsilon_j^2, \quad (\rho_{j+1}, \sigma_{j+1}) := (1 - 3d_j) (\rho_j, \sigma_j), \tag{14}$$

with  $K > 0$  and  $\tau := 2n + 3$ . Then, for all  $0 < \rho_0 \leq \rho, 0 < \sigma_0 \leq \sigma$  and  $\epsilon_0 \leq \epsilon_a$  where

$$\epsilon_a \leq a K^{-1} (2\pi)^{-2\tau}, \tag{15}$$

it is possible to construct  $\{d_j\}_{j \in \mathbb{N}}$  such that  $(\rho_*, \sigma_*) = (1/2)(\rho_0, \sigma_0)$ , in particular they are strictly positive. Furthermore  $\lim_{j \rightarrow \infty} \epsilon_j = 0$ .

<sup>5</sup> Recall (2), then use the inequality  $\sum_{k \in \mathbb{Z}^n} \exp(-\delta|k|\hat{\sigma}) \leq (e\delta^{-1}\hat{\sigma}^{-1})^{2n}$ . Its variant  $\sum_{k \in \mathbb{Z}^n} (1 + |\omega||k|) \exp(-\delta|k|\hat{\sigma}) \leq C_\omega (e\delta^{-1}\hat{\sigma}^{-1})^{2n}$  is used to obtain the second of (13).

*Proof* Choose  $\epsilon_j := \epsilon_0(j + 1)^{-2\tau}$  (so that  $\lim_{j \rightarrow \infty} \epsilon_j = 0$  by construction). By the first of (14) one gets

$$d_j = (\epsilon_0 K a^{-1})^{\frac{1}{\tau}} (j + 2)^2 / (j + 1)^4, \tag{16}$$

hence, by (15),  $d_j \leq \pi^{-2}(j + 1)^{-2}$ . This implies  $\sum_{j \geq 0} d_j \leq 1/6$  and then, trivially,  $d_j \leq 1/6$  for all  $j \in \mathbb{N}$ . Now we have<sup>6</sup>  $\ln \prod_{j \geq 0} (1 - 3d_j) = \sum_{j \geq 0} \ln(1 - 3d_j) \geq -6 \ln 2 \sum_{j \geq 0} d_j = -\ln 2$ , hence  $\lim_{j \rightarrow \infty} \rho_j = \rho_0 \prod_{j \geq 0} (1 - 3d_j) \geq \rho_0 / 2 =: \rho_*$ . Analogously  $\sigma_* := \sigma_0 / 2$ .  $\square$

### 4.3 Bounds on the formal algorithm

**Lemma 4.5** *There exists  $K = K(\rho_0, \sigma_0) > 0$  such that, if  $\epsilon \leq \epsilon_a$  where  $\epsilon_a$  satisfies (15), then*

$$\|F^{(j)}\|_{(\rho_j, \sigma_j)} \leq \epsilon_j e^{-at}, \tag{17}$$

for all  $j \in \mathbb{N}$ . Hence, the transformed Hamiltonian  $\mathcal{B} \circ H$  is in the form (4).

*Proof* By induction. Note that (17) is true for  $j = 0$  setting  $\epsilon_0 := \epsilon$ . The condition on  $\epsilon$  ensures the validity of Lemma 4.4. Hence, supposing (17), by Lemmas 4.2 and 4.4, we get

$$\|\chi^{(j)}\|_{(1-d_j)(\rho_j, \sigma_j)} \leq \epsilon_j (e/\sigma_*)^{2n} a^{-1} d_j^{-2n} e^{-at}. \tag{18}$$

By (8) and Proposition 4.1 with  $d' = d_j$ ,  $d'' = 0$  and  $\tilde{d} = d_j$  (the condition  $d_j \leq 1 - d_j$  holds as  $d_j \leq 1/6$ )

$$\|F^{(j+1)}\|_{(1-2d_j)(\rho_j, \sigma_j)} \leq \sum_{s \geq 1} \frac{1}{s!} \|\mathcal{L}_{\chi^{(j)}}^s F^{(j)}\|_{(1-2d_j)(\rho_j, \sigma_j)} \leq 2^{-1} \Theta \|F^{(j)}\|_{(\rho_j, \sigma_j)}, \tag{19}$$

where<sup>7</sup>

$$\Theta := 2\epsilon_j n C_\omega (e/\sigma_*)^\tau \rho_*^{-1} a^{-1} d_j^{-2n-2} e^{-at} \leq 1/2 \tag{20}$$

is a sufficient condition for the convergence of the operator  $\exp(\mathcal{L}_{\chi^{(j)}})$ , from which  $\sum_{s \geq 1} \Theta^s \leq 2\Theta$ . Hence, by (19), (20), then by (18) one gets (use also  $\sigma_*, \rho_*, d_j < 1$ )

$$\|F^{(j+1)}\|_{(1-2d_j)(\rho_j, \sigma_j)} \leq \epsilon_j^2 n C_\omega (e/\sigma_*)^\tau \rho_*^{-1} a^{-1} d_j^{-\tau} e^{-at}. \tag{21}$$

The latter is valid *a fortiori* in  $\mathcal{D}_{(1-3d_j)(\rho_j, \sigma_j)}$ .

In conclusion, by choosing  $K := n C_\omega (e/\sigma_*)^\tau \rho_*^{-1} = 2^{\tau+1} n C_\omega (e/\sigma_0)^\tau \rho_0^{-1}$ , from the first of (14), we have that (17) is satisfied for  $j \rightarrow j + 1$ . Furthermore, by the first of (14), condition (20) yields  $1 \geq 4\epsilon_j K d_j a^{-1} d_j^{-\tau} e^{-at} = 4d_j (\epsilon_{j+1}/\epsilon_j) e^{-at}$ . The latter is trivially true for all  $t \in \mathbb{R}^+$  by the monotonicity of  $\epsilon_j$  and as  $d_j \leq 1/6$ . Furthermore this implies

$$\Theta \leq 2d_j e^{-at}. \tag{22}$$

Hence  $\exp(\mathcal{L}_{\chi^{(j)}})$  is well defined for all  $j \in \mathbb{N}$ .  $\square$

In this way the value of  $\epsilon_a$  mentioned in the statement of Theorem 2.2 is determined once and for all.

<sup>6</sup> Use the inequality  $\ln(1 - x) \geq -2x \ln 2$ , valid for all  $x \in [0, 1/2]$ .

<sup>7</sup> The reason for using  $n C_\omega$  in the definition of  $\Theta$  will be clear in the proof of Lemma 4.6.

### 4.4 Estimates on the transformation of coordinates

**Lemma 4.6** *The limit (6) exists, it is  $\varepsilon$ -close to the identity and satisfies*

$$\left| I^{(\infty)} - I \right|, \left| \eta^{(\infty)} - \eta \right| \leq (\rho_0/6) e^{-at}, \quad \left| \varphi^{(\infty)} - \varphi \right| \leq (\sigma_0/6) e^{-at}, \tag{23}$$

in particular it defines an analytic map  $\mathcal{B} : \mathcal{D}_{\rho_*, \sigma_*} \rightarrow \mathcal{D}_{\rho_0, \sigma_0}$  and  $H^{(\infty)}$  is an analytic function on  $\mathcal{D}_{\rho_*, \sigma_*}$  for all  $t \in \mathbb{R}^+$ .

*Proof* Let us start with  $I$ . Note that  $\left\| \mathcal{L}_{\chi^{(j)}} I^{(j+1)} \right\|_{(1-2d_j)(\rho_j, \sigma_j)} \leq n(ed_j \rho_j)^{-1} \left\| \chi^{(j)} \right\|_{(1-d_j)(\rho_j, \sigma_j)}$  by a Cauchy estimate [see (Giorgilli 2003, Lemma 4.1)], so that the presence of  $n$  in (20) is justified. Hence use Proposition 4.1 with  $F \leftarrow \mathcal{L}_{\chi^{(j)}} I^{(j+1)}$ ,  $s \leftarrow s - 1$ , obtaining  $\left\| \mathcal{L}_{\chi^{(j)}}^s \varphi^{(j+1)} \right\|_{(1-3d_j)(\rho_j, \sigma_j)} \leq e^{-2s} s! \Theta^s \rho_0$ . This implies

$$\left| I^{(j+1)} - I^{(j)} \right| \leq e^{-2} \sum_{s \geq 1} (1/s!) \left\| \mathcal{L}_{\chi^{(j)}}^s I^{(j+1)} \right\|_{(1-3d_j)(\rho_j, \sigma_j)} \leq 2^{-1} \Theta \rho_0 \leq d_j \rho_0 e^{-at},$$

by (22). In particular  $|I^{(j+1)} - I^{(j)}|$  is  $\varepsilon$ -close to the identity by (16) for all  $j \in \mathbb{N}$ , hence  $|I^{(\infty)} - I| \leq \sum_{j \geq 0} |I^{(j+1)} - I^{(j)}|$  is. It is now sufficient to recall  $\sum_{j \geq 0} d_j \leq 1/6$  in order to conclude.

The argument for  $\varphi$  is analogous while the variable  $\eta$  requires a slight modification. In particular, as one needs to set  $F \leftarrow \mathcal{L}_{\chi^{(j)}} \eta = -\chi_t^{(j)}$ , the use of the second of (13) requires the contribution of  $C_\omega$  in (20).

In conclusion, the obtained composition of analytic maps is uniformly convergent in any compact subset of  $\mathcal{D}_{\rho_*, \sigma_*}$ . This implies that  $\mathcal{B}$  is analytic on  $\mathcal{D}_{\rho_*, \sigma_*}$  by the Weierstraß Theorem and hence the image of  $H$  via  $\mathcal{B}$  is an analytic function in the same domain.  $\square$

### 5 Further perturbation examples

In this section we consider two alternative examples of perturbation. The main purpose is to show that the hypothesis of summability in time over the semi-axis is the only key requirement for the argument beyond the proof of Theorem 2.2.

In particular, we shall firstly consider a decay which is assumed to be quadratic in time, while in the second example a perturbation exhibiting a finite number of (differentiable) bumps is examined. The procedure is fully similar, with the exception of some bounds that will be explicitly given below.

#### 5.1 Quadratic decay

Let us suppose that (3) is modified as

$$\|f(I, \varphi, t)\|_{\rho, \sigma} \leq M_f(t + 1)^{-2}.$$

In the same framework, it is immediate to show that the analogous of Lemma 4.2 yields the following estimates

$$\begin{aligned} \left\| \chi^{(j)} \right\|_{(1-\delta)(\hat{\rho}, \hat{\sigma})} &\leq M^{(j)} (e\delta^{-1}\hat{\sigma}^{-1})^{2n} (t + 1)^{-1}, \quad \left\| \chi_t^{(j)} \right\|_{(1-\delta)(\hat{\rho}, \hat{\sigma})} \\ &\leq M^{(j)} C_\omega (e\delta^{-1}\hat{\sigma}^{-1})^{2n} (t + 1)^{-1}. \end{aligned}$$



Clearly, in this case, the integration has led to a “loss of a power” in the decay. This is harmless as, by (19),  $\|F^{(j+1)}\|_{(1-2d_j)(\rho_j, \sigma_j)} = O(F^j)O(\chi^{(j)}) + h.o.t.$  and then  $F^{(j+1)} \sim (t+1)^{-3} \leq (t+1)^{-2}$  so that the scheme can be iterated.<sup>8</sup>

The rest of the proof is analogous provided that the term  $e^{-at}$  is replaced with 1 in the remaining estimates.

### 5.2 Differentiable bumps

Let  $L \in \mathbb{N} \setminus \{0\}$  and  $h > 0$ . Consider an increasing sequence  $\{t_l\}_{l=1, \dots, L} \in \mathbb{R}^+$  such that  $t_{l+1} - t_l > 2h$ , then the following function

$$\xi_l(t) := \begin{cases} (a_l/h^4) [(t - t_l + h)(t - t_l - h)]^2 & t \in [t_l - h, t_l + h] \\ 0 & \text{otherwise} \end{cases}$$

where  $a_l \in \mathbb{R}$ . Considering a function  $\tilde{f}(I, \varphi) \in \mathfrak{C}_{\rho, \sigma}$ , we set as

$$f(I, \varphi, t) := \tilde{f}(I, \varphi) \sum_{l=1}^L \xi_l(t).$$

In such case we find

$$\|\chi^{(j)}\|_{(1-\delta)(\hat{\rho}, \hat{\sigma})} \leq 2AM^{(j)}h(e\delta^{-1}\hat{\sigma}^{-1})^{2n}, \quad \|\chi_t^{(j)}\|_{(1-\delta)(\hat{\rho}, \hat{\sigma})} \leq M^{(j)}C_\omega(e\delta^{-1}\hat{\sigma}^{-1})^{2n},$$

with  $A := \sum_{l=1}^L |a_l|$ . The remaining part of the proof is straightforward with the obvious modifications. In particular, as for the proof of Lemma 4.5, one finds  $K = 2nC_\omega(e/\sigma_*)^\tau hA\rho_*^{-1}$ .

## Part II

### Proof of Theorem 2.3

In order to simplify the notation, we shall use  $(\rho_H, \sigma_H)$  in place of  $(\rho, \sigma)$  and  $(\rho, \sigma)$  in place of  $(\hat{\rho}_0, \hat{\sigma}_0)$  from now on.

## 6 Formal algorithm

As in Giorgilli (2003), we write Hamiltonian (1) in the form

$$H(I, \varphi, \eta, t) = H_0(I, \eta) + H_1(I, \varphi, t) + H_2(I, \varphi, t) + \dots$$

<sup>8</sup> A similar (and even stronger) phenomenon could have been noticed in the original setting. Namely, suppose by induction that  $\|F^{(j)}\|_{(\rho_j, \sigma_j)} \leq \epsilon_j \exp(-a_j t)$ . By Lemma 4.2 and (19), one finds that  $\|F^{(j+1)}\|_{(\rho_{j+1}, \sigma_{j+1})} \leq \epsilon_{j+1} \exp(-2a_j t)$  and so on. This leads to a remarkable rate of decay ( $a_j = 2^j a$ ) but not to a substantial improvement of the estimates and of the threshold (15) of  $\epsilon_a$ , as these are uniform in  $j$ .

where

$$H_0(I, \eta) := h(I) + \eta, \quad H_s(I, \varphi, t) := \sum_{k \in \Lambda_s} f_k(I, t) e^{ik \cdot \varphi},$$

where  $\Lambda_s := \{k \in \mathbb{Z}^n : (s - 1)N \leq |k| < sN\}$  and  $N \in \mathbb{N} \setminus \{0\}$  is meant to be determined.

Given a sequence of functions  $\{\chi^{(s)}\}_{s \geq 1} : \mathfrak{C}_{\rho, \sigma} \rightarrow \mathbb{C}$ , the Lie transform operator is defined as

$$T_\chi := \sum_{s \geq 0} E_s, \quad E_s := \begin{cases} \text{Id} & s = 0 \\ \frac{1}{s} \sum_{j=1}^s j \mathcal{L}_{\chi^{(j)}} E_{s-j} & s \geq 1 \end{cases} \quad (24)$$

Let  $r \in \mathbb{N} \setminus \{0\}$  to be determined. A *finite* generating sequence of order  $r$ , denoted with  $\chi^{[r]}$ , is such that  $\chi^{(s)} \equiv 0$  for all  $s > r$ . Our aim is to determine it in such a way the effect of  $H_1, \dots, H_r$  is removed, i.e.

$$H^{(r)} := T_{\chi^{[r]}} H = H_0 + \mathcal{R}^{(r+1)}(I, \varphi, t), \quad (25)$$

where the *remainder*  $\mathcal{R}^{(r+1)}$  contains  $H_{>r}$  and a multitude of terms produced during the normalization, which Fourier harmonics lie on  $\Lambda_{>r}$ . The smallness of the remainder is an immediate consequence of the decay property of the coefficients of an analytic function. The procedure is standard: condition (25), with the use of (24), yields a well known diagram which  $s$ -th *level*<sup>9</sup> is of the form

$$\mathcal{E}_s := E_s H_0 + \sum_{l=1}^{s-1} E_{s-l} H_l + H_s = 0, \quad (26)$$

if  $s = 2, \dots, r$  and  $E_1 H_0 + H_1 = 0$  if  $s = 1$ . As sum of all the “non-normalised” levels, the remainder easily reads as

$$\mathcal{R}^{(r+1)} = \sum_{s>r} \mathcal{E}_s. \quad (27)$$

By writing the first term of (26) in the form  $E_s = \mathcal{L}_{\chi^{(s)}} + \sum_{j=1}^{s-1} (j/s) \mathcal{L}_{\chi^{(j)}} E_{s-j}$  and using the manipulation described in (Giorgilli 2003, Chapter 5), one obtains a remarkable cancellation of the contribution of  $H_0$ . In this way, the generating sequence is determined as a solution of

$$\mathcal{L}_{H_0} \chi^{(s)} = \Psi_s, \quad \Psi_s := \begin{cases} H_1 & s = 1 \\ H_s + \sum_{j=1}^{s-1} \frac{j}{s} E_{s-j} H_j & s \geq 2 \end{cases} \quad (28)$$

A formal expansion of  $\chi^{(j)}$  and of  $\Psi_s := \sum_{k \in \mathbb{Z}^n} \psi_k^{(s)}(I, t) e^{ik \cdot \varphi}$  yields for all  $s = 1, \dots, r$

$$\partial_t c_k^{(s)}(I, t) + i(\omega(I) \cdot k) c_k^{(s)}(I, t) = \psi_k^{(s)}(I, t), \quad k \in \Lambda_s, \quad (29)$$

where, as usual,  $\omega(I) := \partial_I h(I)$ .

*Remark 6.1* As a substantial difference with the isochronous case, the function  $\omega(I)$  is a complex valued vector as  $I \in \mathcal{G}_\rho$ . In this way the exponent  $\lambda(k)t$  appearing in formula (11)

<sup>9</sup> Namely, those terms of the diagram which Fourier harmonics belong to  $\Lambda_s$ .

is no longer purely complex. More precisely, one finds a term of the form  $\exp((\omega_C(I) \cdot k)t)$ , having denoted  $\omega(I) = \omega_R(I) + i\omega_C(I)$ ,  $\omega_{R,C}(I) \in \mathbb{R}^n$ . The size of this term cannot be controlled without a cut-off on  $k$ . By restricting the analysis on the levels  $\Lambda_s$  and using the fact that  $|\omega_C(I)| \rightarrow 0$  as  $\rho \rightarrow 0$ , a loss “of part of time decay” at each step (see Lemma 7.1) will be the key ingredient to overcome this difficulty. The mentioned elements are clear obstructions to the limit  $r \rightarrow \infty$ .

## 7 Convergence

### 7.1 Set-up and some preliminary results

The use of the analytic tools requires the usual construction of a sequence of nested domains. We shall choose, for all  $s = 1, \dots, r$ , the rule

$$d_s := d(s - 1)/r, \tag{30}$$

with  $d \in (0, 1/4]$ . Clearly  $d_s < d$  for all  $s = 1, \dots, r$ . Consider also the monotonically decreasing sequence of non-negative real numbers  $\{a_s\}$  defined as follows

$$a_{s+1} := a_s(2r - s)/(2r), \quad a_1 := a. \tag{31}$$

Given the analyticity domain of  $H$  expressed by  $(\rho_H, \sigma_H)$ , set  $\sigma := \sigma_H/2$ . Now consider the function  $\Omega(\rho) := \sup_{I \in \mathcal{G}_\rho} |\omega_C(I)|$ , clearly  $\Omega(0) = 0$ . From now on we shall suppose that  $\rho$  satisfies the following condition

$$4rN\Omega(\rho) \leq a. \tag{32}$$

The analyticity<sup>10</sup> of  $h(I)$  implies the existence of  $C_h \in [1, +\infty)$  such that the value of  $\rho$  can be determined as

$$\rho := \min \{ \rho_H, a(4rNC_h)^{-1} \}, \tag{33}$$

once  $r$  and  $N$  will be chosen.

The scheme is constructed in such a way one can set  $(\tilde{\rho}_*, \tilde{\sigma}_*) := (1 - d)(\rho, \sigma)$ .

As a consequence of Hypothesis 2.1 and of the standard properties of analytic functions, one has

$$\|H_m\|_{\rho, \sigma} \leq \mathcal{F}h^{m-1}e^{-at}, \quad m \geq 1, \tag{34}$$

with  $\mathcal{F} := \varepsilon\tilde{\mathcal{F}}$ , where [see (Giorgilli 2003, Lemma 5.2)]  $\tilde{\mathcal{F}} := [(1 + \exp(-\sigma/2))/(1 - \exp(-\sigma/2))]^n$  and

$$h := \exp(-N\sigma/2). \tag{35}$$

**Lemma 7.1** *Suppose that  $\|\Psi_s\|_{(1-d_s)(\rho, \sigma)} \leq M^{(s)} \exp(-a_s t)$ , for some  $M^{(s)} > 0$ . Then the solution of (28) satisfies*

$$4a \left\| \chi^{(s)} \right\|_{(1-d_{s+1/2})(\rho, \sigma)}, 4 \left\| \partial_t \chi^{(s)} \right\|_{(1-d_{s+1/2})(\rho, \sigma)} \leq C_r M^{(s)} e^{-a_{s+1} t}, \tag{36}$$

where  $C_r := 2^{2n+4}(r/d)^n$ .

<sup>10</sup> Obviously,  $\Omega(\rho) \equiv 0$  for all  $\rho$  in the case of an isochronous system, so that (32) would impose no restrictions on  $\rho$ .

*Proof* Use (29). Similarly to Lemma 4.2, we choose  $c_k^{(s)}(I, 0) := -\int_{\mathbb{R}^+} \exp(-(\omega(I) \cdot k)\tau)\psi_k^{(s)}(I, \tau)d\tau$ . Note that  $|c_k^{(s)}(I, 0)| \leq M^{(s)} \exp(-(1 - d_s)|k|\sigma) \int_{\mathbb{R}^+} \exp(|\omega_C(I)||k| - a)\tau d\tau < +\infty$  on  $\Lambda_s$  by (32). By using again (32) one gets

$$|c_k^{(s)}(I, t)| \leq M^{(s)} e^{-(1-d_s)|k|\sigma} e^{\frac{as}{4r}t} \int_t^\infty e^{as\left(\frac{s-4r}{4r}\right)\tau} d\tau \leq \frac{4}{a} M^{(s)} e^{-(1-d_s)|k|\sigma} e^{-a_s(1-\frac{s}{2r})t}. \tag{37}$$

The first of (36) is easily recognised<sup>11</sup> by (31). The second of (36) follow from (37) and from (29). □

**Lemma 7.2** *Let  $A, \Gamma, \tau > 0$  and consider the real-valued sequences  $\{\kappa_s\}_{s \geq 1}$  and  $\{\gamma_l\}_{l \geq 0}$  defined as*

$$\kappa_s := A\tau^{s-1} + \Gamma \sum_{j=1}^{s-1} \tau^{j-1} \kappa_{s-j}, \quad \gamma_l := \Gamma \sum_{j=1}^l \tau^{j-1} \gamma_{l-j}, \tag{38}$$

where  $\kappa_1$  and  $\gamma_0$  are given. Define  $\Delta := \tau + \Gamma$ , then for all  $s \geq 2$  and  $l \geq 1$

$$\kappa_s = (\Gamma\kappa_1 + \tau A) \Delta^{s-2}, \quad \gamma_l = \gamma_0 \Gamma \Delta^{l-1}. \tag{39}$$

*Proof* We shall denote with (38a) and (38b) the first and the second of (38), respectively. The same for (39). Let us suppose for a moment that (39a) is proven, then choose  $A = \Gamma\gamma_0$  and  $\kappa_1 = \Gamma\gamma_0 = \gamma_1$ . By substituting in (39a) one immediately gets (39b). Hence we need only to prove (39a).

For this purpose we use the well-known *generating function* method (see e.g. Wilf 2006). Namely, define  $g(z) := \sum_{n=1}^\infty w_n z^n$ , multiply each equation obtained from (38a) by  $z^s$  as  $s$  varies, then “sum” all the equations. This leads to  $g(z) = [1 - \Delta z]^{-1} (\kappa_1(z - \tau z^2) + A\tau z^2) = (1 + \Delta z + \Delta^2 z^2 + \dots)(\kappa_1(z - \tau z^2) + A\tau z^2) = \kappa_1 z + (\Gamma\kappa_1 + \tau A) \sum_{n \geq 2} \Delta^{n-2} z^n$ , which is the (39a). □

### 7.2 Bounds on the generating function

**Proposition 7.3** *For all  $s \leq r$ , the following estimate holds*

$$\|\chi_s\|_{(1-d_{s+1/2})(\rho, \sigma)} \leq (4a)^{-1} C_r \beta_s \mathcal{F} e^{-a_s+1t}, \tag{40}$$

where the sequence  $\{\beta_s\}_{s=1, \dots, r} \in \mathbb{R}^+$  is determined by the following system

$$\begin{cases} \beta_s = h^{s-1} + \frac{\Gamma}{s} \sum_{j=1}^{s-1} j \theta_{s-j} \\ \theta_l = \frac{\Gamma}{l} \sum_{j=1}^l j \beta_j \theta_{l-j} \end{cases} \tag{41}$$

with  $\{\theta_l\}_{l=0, \dots, r-1} \in \mathbb{R}^+$  and

$$\Gamma := 16nr^2 C_r \mathcal{F} (ad^2 \rho \sigma)^{-1}, \tag{42}$$

under the conditions<sup>12</sup>  $\beta_1 = \theta_0 = 1$ .

<sup>11</sup> Use the inequality  $\sum_{|k| \geq (s-1)N} \exp(-\delta|k|\sigma) \leq \exp(-N\delta n(s-1)\sigma) (\sum_{m=0}^{+\infty} \exp(-\delta m\sigma))^n \leq (2/\delta)^n$ , where in this case  $\delta := d_{s+1/2} - d_s = d/(2r)$ .

<sup>12</sup> From a “computational” point of view, first compute  $\theta_1$  then proceed with  $\beta_s, \theta_s$  for all  $s = 2, \dots, r$ .

First of all note that by (24) and (34), one has  $\|\Psi_1\|_{(1-d_1)(\rho,\sigma)} \leq \mathcal{F} \exp(-a_1 t)$  and  $\|E_0 H_m\|_{(1-d)(\rho,\sigma)} \leq \mathcal{F} h^{m-1} \exp(-a_1 t)$  (recall (31)). Hence, given by  $s \leq r$ , we can suppose by induction to know  $\beta_1, \dots, \beta_{s-1}$  and  $\tilde{\theta}_{0,m}, \dots, \tilde{\theta}_{s-2,m}$ , for all  $m \geq 1$ , with  $\beta_1 = 1$  and  $\tilde{\theta}_{0,m} = h^{m-1}$ , such that the the following bounds hold for all  $j = 1, \dots, s - 1$  and  $l = 0, \dots, s - 2$

$$\|\Psi_j\|_{(1-d_j)(\rho,\sigma)} \leq \beta_j \mathcal{F} e^{-a_j t}, \tag{43a}$$

$$\|E_l H_m\|_{(1-d_{l+1})(\rho,\sigma)} \leq \tilde{\theta}_{l,m} \mathcal{F} e^{-a_{l+1} t}, \tag{43b}$$

By (43a) and Lemma 7.1, the bound (40) holds with  $j$  in place of  $s$ . Hence by Proposition 4.1 with  $G = \chi^{(j)}$ ,  $F = E_{s-j-1} H_m$  then  $\hat{d} = \max_{j=1, \dots, s-1} \{d_{j+1/2}, d_{s-j}\} = d_{s-1/2}$  and finally  $\tilde{d} := d_s - d_{s-1/2} = d/(2r)$ , one has (by setting  $\delta = 0$ )

$$\begin{aligned} \|\mathcal{L}_{\chi^{(j)}} E_{s-j-1} H_m\|_{(1-d_s)(\rho,\sigma)} &\leq 8r^2 (ed^2 \rho \sigma)^{-1} \|\chi^{(j)}\|_{(1-d_{j+1/2})(\rho,\sigma)} \\ &\quad \|E_{l-j} H_0\|_{(1-d_{l-j+1/2})(\rho,\sigma)} \\ &\leq \Gamma \mathcal{F} \beta_j \gamma_{l-j} e^{-a_{l+1} t} \end{aligned} \tag{44}$$

where the property  $a_{j+1} + a_{l-j+1} \geq a_{l+1}$  has been used. Recalling (24), we have that (43b) holds also for  $l = s - 1$ , where

$$\tilde{\theta}_{l,m} = \frac{\Gamma}{l} \sum_{j=1}^l j \beta_j \tilde{\theta}_{l-j,m}. \tag{45}$$

Furthermore, it is easy to show from the latter that  $\tilde{\theta}_{l,m} = h^{m-1} \tilde{\theta}_{l,1}$  in such a way, defined  $\theta_l := \tilde{\theta}_{l,1}$  one gets  $\tilde{\theta}_{l,m} = h^{m-1} \theta_l$ , and then the second of (41), provided  $\theta_0 = 1$ . In conclusion, by using (34), and the second of (41) in the definition of  $\Psi_s$  as in (28), we get that (43a) is satisfied if  $\beta_s$  is defined as in the first of (41). Bound (40) follows from Lemma 7.1.

**Proposition 7.4** *The sequence  $\beta_s$  defined by (41) satisfies*

$$\beta_s \leq \tau^{s-1}/s, \tag{46}$$

for  $s = 1, \dots, r$ , if

$$\tau := eh, \quad \Gamma \leq h/(2r^2). \tag{47}$$

*Proof* The property (46) is trivially true for  $s = 1$ , hence let us suppose it for  $j = 1, \dots, s - 1$  and proceed by induction with  $\tau$  to be determined. Define  $\tilde{\theta}_l := \theta_l (\beta_j)_{|\beta_j = \tau^{j-1}/j}$ , then  $\hat{\theta}_l := \tilde{\theta}_l/l$ , obtaining  $\hat{\theta}_l = \Gamma \sum_{j=1}^l \tau^{j-1} \hat{\theta}_{l-j}$ . Clearly  $\theta_l \leq \tilde{\theta}_l \leq \hat{\theta}_l/l$ , furthermore  $\theta_0 = \tilde{\theta}_0 = \hat{\theta}_0 = 1$ . Hence, by Lemma 7.2 we have

$$\theta_l \leq \Gamma \Delta^{l-1}/l. \tag{48}$$

Now choose  $\tau, \Gamma$  as in (47). By using (34) and (48) in the first of (41) one gets that (46) is satisfied simply by checking that the inequality

$$y(s) := s + \frac{(s-1)}{2r^2} \left( e + \frac{1}{2r^2} \right)^{s-1} \leq e^{s-1} \tag{49}$$

holds true for all<sup>13</sup>  $s = 1, \dots, r$ . □

<sup>13</sup> Clearly (49) holds for  $s \leq r$  if  $y(r) \leq \exp(r - 1)$  for all  $r \geq 3$  (let it be directly checked for  $r = 1, 2$ ). Hence set  $r = n + 1$  and prove that  $y(r)_{r=n+1} \leq \exp(n)$  for all  $n \geq 2$ , conclusion that is immediate as one can find that  $y(n) \leq n + 1 + 3e^n/(4n)$ .

### 7.3 Estimates on the coordinates transformation

From now on we shall suppose that  $h$  and  $\varepsilon$  are chosen in such a way

$$8eh \leq 1 \tag{50a}$$

$$2r^2\Gamma \leq \sqrt{\varepsilon}h \tag{50b}$$

In particular, by definition and by (47), this immediately implies that

$$4\Delta \leq 1 \tag{51}$$

As in [Giorgilli \(2003\)](#) it is used that, despite the generating sequence is finite, one can use the bound obtained from 7.3

$$\left\| \chi^{(s)} \right\|_{(1-d)(\rho,\sigma)} \leq (4a)^{-1} C_r \mathcal{F} \beta_s e^{-a_{r+1}t}, \tag{52}$$

with  $\beta_s$  satisfying (46) for all  $s$ , as it would be, trivially,  $\beta_{>r} = 0$ .

**Proposition 7.5** Define  $(I^{(r)}, \varphi^{(r)}, \eta^{(r)}) := T_{\chi^{[r]}}(I, \varphi, \eta)$ . Then the following estimates hold

$$\begin{aligned} \left\| I - I^{(r)} \right\|_{(1-d)(\rho,\sigma)}, \left\| \eta - \eta^{(r)} \right\|_{(1-d)(\rho,\sigma)} &\leq \frac{d\rho}{8} e^{-a_{r+1}t}, \\ \left\| \varphi - \varphi^{(r)} \right\|_{(1-d)(\rho,\sigma)} &\leq \frac{d\sigma}{8} e^{-a_{r+1}t}. \end{aligned} \tag{53}$$

*Proof* Let us start from the variable  $I$ . Firstly, note that  $\left\| I - T_{\chi^{[r]}} I \right\|_{(1-d)(\rho,\sigma)} \leq \sum_{s \geq 1} \|E_s I\|_{(1-d)(\rho,\sigma)}$ . In addition

$$\begin{aligned} \|E_1 I\|_{(1-d_2)(\rho,\sigma)} &= \left\| \partial_\varphi \chi^{(1)} \right\|_{(1-d_2)(\rho,\sigma)} \leq 2nr(ed\sigma)^{-1} \left\| \chi^{(1)} \right\|_{(1-d_{3/2})(\rho,\sigma)} \\ &\leq D_\sigma \mathcal{F} \exp(-a_{r+1}t), \end{aligned}$$

with  $D_\sigma := nrC_r/(2d\sigma a)$  by Prop. 7.3. Hence suppose  $\|E_l I\|_{(1-d_{l+1})(\rho,\sigma)} \leq \mathcal{F}u_l \exp(-a_{r+1}t)$  for all  $l = 1, \dots, s - 1$  with  $u_1 = D_\sigma$  and proceed by induction.

The bound of  $E_l I$  can be treated in the same way of (43b) with the difference that in this case the term  $\mathcal{L}_{\chi^{(l)}} I$  appearing in  $E_l I$  needs to be bounded separately by using (40) and a Cauchy estimate. This leads to  $u_l = \beta_l D_\sigma + \Gamma/l \sum_{j=1}^{l-1} j\beta_j u_{l-j}$ . By using the same procedure used in the proof of Proposition 7.4 for  $\theta_l$  one gets  $u_l \leq (D_\sigma/l)\Delta^{l-1}$ . The required bound easily follows as  $\mathcal{F} \sum_{s \geq 1} u_s \leq 2\mathcal{F}D_\sigma \leq \Gamma d\rho \leq \sqrt{\varepsilon}d\rho/8$ , where the second inequality follows from (51) and the last one from (50b) then from (50a). The procedure for the variables  $\varphi$  and  $\eta$  is similar. The analyticity of the transformation  $\mathcal{N}_r := T_{\chi^{[r]}}^{-1}$  easily follows from the bounds (53) and the invertibility of the Lie transform operator, see [Giorgilli \(2003\)](#).  $\square$

### 7.4 Bound on the remainder

**Proposition 7.6** Define  $A := 10\tilde{\mathcal{F}}$  then for all  $r \geq 1$

$$\left\| \mathcal{R}^{(r+1)} \right\|_{(1-2d)(\rho,\sigma)} \leq \varepsilon A e^{-(r+a_{r+1}t)}. \tag{54}$$

*Proof* Define  $(\rho', \sigma') := (1 - d)(\rho, \sigma)$ . Now recall (27) and suppose by induction, for all  $l = 1, \dots, s - 1, m = 0, \dots, s - 2$  with  $s \in \mathbb{N}$

$$\begin{aligned} \|E_l H_0\|_{(1-(l/s)d)(\rho',\sigma')} &\leq \mathcal{F}\epsilon_l \exp(-a_{r+1}t), \\ \|E_m H_n\|_{(1-(m/s)d)(\rho',\sigma')} &\leq \mathcal{F}\zeta_{m,n} \exp(-a_{r+1}t). \end{aligned} \tag{55}$$

Indeed one can set  $\zeta_{0,n} = h^{n-1}$  and  $\epsilon_1 = \beta_1 = 1$  as  $\mathcal{L}_{\chi^{(1)}}H_0 = -\Psi_1$  by (28). We stress that, despite based on the same computations, the argument is conceptually different from the previous estimates as  $s \in (r, +\infty)$  and the use of  $\delta$  in (12) plays here a key role. More precisely, use Proposition 4.1 with  $G = \chi^{(j)}$  and  $F = E_{s-j}H_0$  hence  $d'' = 0$  then  $\hat{d} = d' = \delta = d(s-j)/s$  from which  $\tilde{d} = (j/s)d$ . This leads to  $\|\mathcal{L}_{\chi^{(j)}}E_{s-j}H_0\|_{(1-d)(\rho',\sigma')} \leq \Gamma(s/j)\beta_j\epsilon_{s-j}\exp(-a_{r+1}t)$ , implying<sup>14</sup> that the first of (55) holds for  $l = s$  provided  $\epsilon_s = \beta_s + \Gamma\sum_{j=1}^{s-1}\beta_j\epsilon_{s-j} = \Delta^{s-1}$ , the latter by Lemma 7.2. This implies  $\|\sum_{l=1}^s E_{s-l}H_l\|_{(1-d)(\rho',\sigma')} \leq \mathcal{F}(s+1)\Delta^{s-1}\exp(-a_{r+1}t)$  by using (34) and the trivial bound  $h \leq \Delta$ . Similarly one finds  $\zeta_{s,n} = h^{n-1}\Delta^{s-1}$ , hence

$$(\mathcal{F}e^{-a_{r+1}t})^{-1}\mathcal{R}^{(r+1)} \leq \sum_{s>r}(2+s)\Delta^{s-1} = \Delta^r \left( \frac{r+3}{1-\Delta} + \frac{1}{1-\Delta^2} \right) \leq 2(r+4)\Delta^r,$$

by (51). Noticing that  $\mathcal{D}_{(1-2d)(\rho,\sigma)} \subset \mathcal{D}_{(1-d)^2(\rho,\sigma)}$ , the bound (54) easily follows from (51) and from the simple inequality  $(r+4)e^r \leq 5(4^r)$ . □

### 7.5 Choice of the parameters

Let us discuss a possible choice of the parameters in such a way the convergence conditions are satisfied. More precisely by (35), condition (50a) holds if  $N = \lceil 2\sigma^{-1}(1+3\log 2) \rceil$ , where  $\lceil \cdot \rceil$  denotes the rounding to the greater integer. This implies that  $h \geq 1/(16e)$ , hence (50b) holds if  $2^5 e r^2 \Gamma \leq \sqrt{\epsilon}$ . Hence, recalling (32) and (42), this condition is achieved by choosing (see also Giorgilli and Galgani 1985)

$$r := \left\lfloor \left( \frac{\epsilon_a^*}{\epsilon} \right)^{\frac{1}{2\gamma}} \right\rfloor, \quad \sqrt{\epsilon_a^*} := \frac{a^2 d^{n+2} \rho_H \sigma^2}{2^{2n+19} e n C_h \tilde{\mathcal{F}}}, \tag{56}$$

where<sup>15</sup>  $\gamma = 5+n$  and  $\lfloor \cdot \rfloor$  denotes the rounding to the lower integer. The condition  $\epsilon \leq \epsilon_a^*$ , as in the statement of Theorem 2.3, clearly ensures that  $r \geq 1$ . The final value of  $\rho$  is determined with (33).

Let us write the usual bound  $|I(t) - I(0)| \leq |I(t) - I^{(r)}(t)| + |I^{(r)}(t) - I^{(r)}(0)| + |I^{(r)}(0) - I(0)|$ . The first and third term of the r.h.s. are bounded by  $\sqrt{\epsilon}d\rho/8$  by (53). As for the second one, from the equations of motion  $\dot{I}^{(r)} = -\partial_\varphi H^{(r)} = -\partial_\varphi \mathcal{R}^{(r+1)}$ , furthermore  $\|\partial_\varphi \mathcal{R}^{(r+1)}\|_{(1-2d)(\rho,\sigma)} \leq \epsilon A(ed\sigma)^{-1}\exp(-(r+a_{r+1}t))$  by a Cauchy estimate and by (54). Hence

$$\left| I^{(r)}(t) - I^{(r)}(0) \right| \leq \epsilon A(ed\sigma)^{-1}e^{-r} \int_0^t e^{-a_{r+1}s} ds \leq \epsilon A(ade\sigma)^{-1}(2/e)^r, \tag{57}$$

as  $a_{r+1} = a(2r-1)(2r-2)\dots(r)/(2r)^r > a2^{-r}$ .

*Remark 7.7* The bound (57) is the key difference with the standard Nekhoroshev theorem, despite a normal form of finite order. The remainder, which is bounded by a constant in the classical Nekhoroshev estimate and then produces a linearly growing bound for the quantity

<sup>14</sup> The use of (12) with  $\delta = 0$  would have given  $(s/j)^2$  instead of  $(s/j)$ , producing in this way a troublesome factorial in the estimates.

<sup>15</sup> Note that the threshold  $\epsilon_a^*$  takes into account of the condition (33) as we have used the obvious lower bound  $\rho \geq a\rho_H(4rNC_h)^{-1}$ , immediate from (33).

$|I^{(r)}(t) - I^{(r)}(0)|$ , is now summable over  $\mathbb{R}^+$ . Hence, a restriction to exponentially large times is no longer necessary.

It is immediate from (57) that for all  $\varepsilon \leq \varepsilon_a^*$  one has  $|I^{(r)}(t) - I^{(r)}(0)| \leq 2\varepsilon_a^* A(ade^2\sigma)^{-1}$  which is clearly smaller than  $\sqrt{\varepsilon}d\rho/4$  by (56). Hence  $|I(t) - I(0)| \leq \sqrt{\varepsilon}d\rho/2$ .

**Acknowledgments** The first author is grateful to Proff. D. Bambusi, L. Biasco, A. Giorgilli and T. Penati for very useful discussions on a preliminary version of this paper. This research was supported by ONR Grant No. N00014-01-1-0769 and MINECO: ICMAT Severo Ochoa project SEV-2011-0087.

## References

- Bambusi, D.: Birkhoff normal form for some quasilinear Hamiltonian PDEs. XIVth International Congress on Mathematical Physics, pp. 273–280. World Sci. Publ., Hackensack (2005)
- Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.-M.: A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. II *Nuovo Cimento B* **79**(2), 201–223 (1984)
- Benettin, G., Gallavotti, G.: Stability of motions near resonances in quasi-integrable Hamiltonian systems. *J. Stat. Phys.* **44**(3–4), 293–338 (1986)
- Birkhoff, G.: *Dynamical Systems*. American Mathematical Society colloquium publications. American Mathematical Society, Providence (1927)
- Bounemoura, A.: Effective stability for slow time-dependent near-integrable Hamiltonians and application. *C. R. Math. Acad. Sci. Paris* **351**(17–18), 673–676 (2013)
- Chierchia, L.: Kolmogorov–Arnold–Moser (KAM) theory. In: Meyers, R.A. (ed.) *Encyclopedia of Complexity and Systems Science*, pp. 5064–5091. Springer, New York (2009)
- Cresson, J., Pierret, F., Puig, B.: The Sharma-Parthasarathy stochastic two-body problem. *J. Math. Phys.* **56**(3), 032701 (2015). 21
- Fortunati, A., Wiggins, S.: Normal form and Nekhoroshev stability for nearly integrable Hamiltonian systems with unconditionally slow aperiodic time dependence. *Regul. Chaotic Dyn.* **19**(3), 363–373 (2014a)
- Fortunati, A., Wiggins, S.: Persistence of Diophantine flows for quadratic nearly integrable Hamiltonians under slowly decaying aperiodic time dependence. *Regul. Chaotic Dyn.* **19**(5), 586–600 (2014b)
- Fortunati, A., Wiggins, S.: A Kolmogorov theorem for nearly-integrable Poisson systems with asymptotically decaying time-dependent perturbation. *Regul. Chaotic Dyn.* **20**(4), 476–485 (2015a)
- Fortunati, A., Wiggins, S.: Normal forms à la Moser for aperiodically time-dependent Hamiltonians in the vicinity of a hyperbolic equilibrium. Accepted for the publication on *Discr. Dyn. Sys.-S, Cont* (2015b)
- Gallavotti, G.: Quasi-integrable mechanical systems. In: *Phénomènes critiques. systèmes aléatoires, théories de jauge, Part I, II* (Les Houches, 1984), pp. 539–624. North-Holland, Amsterdam (1986)
- Giorgilli, A.: Exponential stability of Hamiltonian systems. *Dynamical systems. Part I*, pp. 87–198. *Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa* (2003)
- Giorgilli, A., Galgani, L.: Rigorous estimates for the series expansions of Hamiltonian perturbation theory. *Celest. Mech.* **37**(2), 95–112 (1985)
- Giorgilli, A., Zehnder, E.: Exponential stability for time dependent potentials. *Z. Angew. Math. Phys.* **43**(5), 827–855 (1992)
- Nekhoroshev, N.N.: An exponential estimate on the time of stability of nearly-integrable Hamiltonian systems. *Russ. Math. Surveys* **32**, 1–65 (1977)
- Nekhoroshev, N.N.: An exponential estimate on the time of stability of nearly-integrable Hamiltonian systems II. *Trudy Sem. Petrovs.* **5**, 5–50 (1979)
- Pinzari, G.: Aspects of the planetary Birkhoff normal form. *Regul. Chaotic Dyn.* **18**(6), 860–906 (2013)
- Poincaré, H.: *Les méthodes nouvelles de la mécanique céleste*. Gauthier-Villars, Paris (1892)
- Pustyl'nikov, L.D.: Stable and oscillating motions in nonautonomous dynamical systems. A generalization of C. L. Siegel's theorem to the nonautonomous case. *Math. Sb. (N.S.)* **94**(136), 407–429 (1974). 495
- Wiggins, S., Mancho, A.M.: Barriers to transport in aperiodically time-dependent two-dimensional velocity fields: Nekhoroshev's theorem and nearly invariant tori. *Nonlinear Process. Geophys.* **21**(1), 165–185 (2014)
- Wilf, H.: *Generating functionology: Third Edition*. Ak Peters Series. Taylor & Francis, Milton Park (2006)