Higher-order averaging, formal series and numerical integration III: error bounds

Ph. Chartier,[*] A. Murua,[†] J. M. Sanz-Serna[‡]

January 3, 2013

Abstract

In earlier papers, it has been shown how formal series like those used nowadays to investigate the properties of numerical integrators may be used to construct high-order averaged systems or formal first integrals of Hamiltonian problems. With the new approach the averaged system (or the formal first integral) may be written down immediately in terms of (i) suitable basis functions and (ii) scalar coefficients that are computed via simple recursions. Here we show how the coefficients/basis functions approach may be used advantageously to derive exponentially small error bounds for averaged systems and approximate first integrals.

Keywords and sentences: Averaging, high-order averaging, quasi-stroboscopic averaging, highly oscillatory problems, Hamiltonian problems, formal series, first-integrals, near-integrable systems

Mathematics Subject Classification (2010) 34C29, 65L06, 34D20, 70H05

1 Introduction

This is the third and final paper of a series that explores the use within the method of averaging [13], [17], [1], [2] of the formal series expansions used nowadays to analyze numerical integrators (specially geometric numerical integrators) [12], [18], [14], [11]. While the preceding parts [5], [6] of the series were devoted to the explicit construction of high-order averaged system and formal integrals, the present contribution focuses on the derivation of error bounds.

[*]INRIA Rennes and ENS Cachan Bretagne, Campus Ker-Lann, av. Robert Schumann, 35170 Bruz, France. Email: Philippe.Chartier@inria.fr

[†]Konputazio Zientziak eta A. A. Saila, Informatika Fakultatea, UPV/EHU, E–20018 Donostia–San Sebastián, Spain. Email: Ander.Murua@ehu.es

[‡](Corresponding author) Departamento de Matemática Aplicada e IMUVA, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain. Email: sanzsern@mac.uva.es

Given a time-dependent, oscillatory dynamical system, the aim of the method of averaging is to approximate it by a simpler autonomous system (the averaged system). Standard approaches [13], [17] find the sought averaged problem after performing a sequence of changes of variables. Each of these intermediate transformations decreases the size of the time-dependence of the vector field; the composition of all intermediate changes of variable provides the overall change that relates the given problem to the resulting high-order averaged problem. Accordingly, in the standard derivations of error bounds, it is necessary to monitor carefully the effect of each intermediate transformation, a fact that complicates the analysis. The alternative technique introduced in [5], [6] yields at once the averaged system and the overall change of variables without any need to perform a sequence of intermediate transformations. More precisely, shown in [5], [6] is how to write averaged systems and changes of variables in terms of two kinds of elements:

1. Scalar *coefficients* that are *universal*, i.e. independent of the given oscillatory vector field. These coefficients may be computed by means of simple recursions.
2. *Basis functions* constructed in an explicit, systematic way in terms of the Fourier coefficients of the oscillatory vector field.

In this paper we show how the derivation of error bounds may be advantageously performed by exploiting the coefficients/basis functions structure. The article addresses two related problems: averaging oscillatory vector fields (Section 2) and constructing approximate first integrals of Hamiltonian problems (Section 3).

While it is well known that it is not possible to remove completely via changes of variables the time-dependence of a system subject to oscillatory forcing, Neishtadt [15] showed that, if the forcing is periodic, an averaged system may be constructed that approximates the given system with an error that is exponentially small (relative to the perturbation parameter). This result plays a crucial role in the theory of the numerical integration of Hamiltonian problems [18], [11], as it is the key to the backward error analysis that explains the success of symplectic methods. In [7] bounds similar to those in [15] were derived by bounding separately the coefficients and basis functions of the expansion of the averaged vector field. Simó [19] extended Neishtadt's result to the case where the time-dependence is *quasi-periodic*; this extension is far from trivial due to the presence of the *small denominator* phenomenon. In Theorem 2.7 below, we provide, by applying the coefficients/basis functions approach, bounds similar to those in [19]. The method used here makes it possible to keep track of all constants that appear in the bounds and pinpoints clearly the role played by the different hypotheses (see Section 2.3).

First integrals of Hamiltonian problems are considered in Section 3. The derivation of *formal* integrals of Hamiltonian system, particularly so within the field of Celestial Mechanics, is a task with a long history and closely related both to averaging and to the theory of normal forms [17], [1]. Different techniques are available. Since the work of Siegel it is known that, generally speaking, the series that express those formal invariants are divergent. However it is also well understood that it is possible to truncate the formal series to obtain *approximate* invariants that may provide much help in the analysis of the underlying dynamics. We refer to [8], [10] for a discussion of these issues. In [6] Section 5, we showed how to construct formal invariants via the novel approach to averaging. Here we focus on the *approximate* invariants obtained by truncating the formal series; in Theorem 3.5 we establish that, under suitable hypotheses, the time-derivative of those truncations may actually be exponentially small. As a consequence the variation undergone by a truncation may remain bounded over exponentially long time intervals (Corollary 3.6). Again the coefficients/basis functions methodology results in simple proofs where track may be kept of all intervening constants.

Before closing the introduction, it is perhaps relevant to mention that it is of course possible to implement numerically the method of averaging, see [3], [4] and their references.

2 Averaging of quasi-periodically forced problems

In this section we consider problems of the form

$$\frac{d}{dt}y = \epsilon f(y, t\omega), \tag{1}$$

$$y(0) = y_0 \in \mathbb{R}^D, \tag{2}$$

where ϵ is a small parameter, the smooth function $f = f(y, \theta)$ is 2π-periodic in each of the components θ^j, $j = 1, \dots, d$, of θ, i.e. $\theta \in \mathbb{T}^d$, and $\omega \in \mathbb{R}^d$ is a constant vector of angular frequencies. When $d = 1$ the vector field in (1) depends periodically on t, while for $d > 1$ the dependence is generically quasi-periodic.

In Section 2.1 we summarize the methodology introduced in [6] for the construction of the quasi-stroboscopic averaged system corresponding to (4); error estimates are then derived in Section 2.2 and illustrated in Section 2.3 by means of an example. In Section 2.1 we work with formal (not necessarily convergent) series in powers of ϵ and the vector ω is assumed to be non-resonant, i.e. $\mathbf{k} \cdot \omega \neq 0$ for $\mathbf{k} \in \mathbb{Z}^d \backslash \{\mathbf{0}\}$; of course, if the problem (1) is resonant it may be re-written in non-resonant form by lowering the number d of frequencies ω^j. Section 2.2 operates under stronger hypotheses on f and ω to be presented there.

2.1 The averaged system

Let the Fourier series for f be

$$f(y, \theta) = \sum_{\mathbf{k} \in \mathbb{Z}^d} e^{i\mathbf{k}\cdot\theta} f_{\mathbf{k}}(y), \tag{3}$$

($f_{\mathbf{k}}$ and $f_{-\mathbf{k}}$ are conjugate so that f takes values in $\mathbb{R}^D$). As shown in [6], the expansion in powers of ϵ of the solution of (1)–(2) involves the Fourier coefficients $f_{\mathbf{k}}$ and may be rearranged in different ways corresponding to different choices of basis functions. Specifically elementary differentials, word basis functions and iterated commutators were considered in [6]. For the purposes of this paper, word-basis functions, as in [6]

Section 3, provide the most convenient choice. Then (see [7] for a derivation different from the one presented in [6]):

$$y(t) = y_0 + \sum_{n=1}^{\infty} \epsilon^n \sum_{\mathbf{k}_1,\dots,\mathbf{k}_n \in \mathbb{Z}^d} \alpha_{\mathbf{k}_1\cdots\mathbf{k}_n}(t)\, f_{\mathbf{k}_1\cdots\mathbf{k}_n}(y_0), \tag{4}$$

where the *coefficients* $\alpha_{\mathbf{k}_1\cdots\mathbf{k}_n}(t)$ are given by

$$\alpha_{\mathbf{k}_1\cdots\mathbf{k}_n}(t) := \int_0^t e^{i\mathbf{k}_n\cdot\omega t_n} dt_n \int_0^{t_n} e^{i\mathbf{k}_{n-1}\cdot\omega t_{n-1}} dt_{n-1} \cdots \int_0^{t_2} e^{i\mathbf{k}_1\cdot\omega t_1} dt_1. \tag{5}$$

and the *word-basis functions* $f_{\mathbf{k}_1\cdots\mathbf{k}_n}(y)$ are defined recursively as

$$f_{\mathbf{k}_1\cdots\mathbf{k}_n}(y) := \partial_y f_{\mathbf{k}_2\cdots\mathbf{k}_n}(y) f_{\mathbf{k}_1}(y). \tag{6}$$

Of much importance is the fact that the coefficients depend on ω and are independent of f and the word-basis functions are independent of ω and depend on f.

The notation in (4)–(6) and expressions that will be needed later may be simplified by using *words* $\mathbf{k}_1\cdots\mathbf{k}_n$, made of *letters* $\mathbf{k}_r$, $r = 1,\dots,n$, $n = 1,2,\dots$, taken from the alphabet $\mathbb{Z}^d$. The set of words with n letters will be represented by $\mathcal{W}_n$. If $\mathbf{k} \in \mathbb{Z}^d$ and $n = 1,2,\dots$, then $\mathbf{k}^n$ means $\mathbf{k}\cdots\mathbf{k} \in \mathcal{W}_n$. Two words $w = \mathbf{k}_1\cdots\mathbf{k}_n$ and $w' = \mathbf{k}'_1\cdots\mathbf{k}'_m$ may be concatenated to give rise to a new word $ww' = \mathbf{k}_1\cdots\mathbf{k}_n\mathbf{k}'_1\cdots\mathbf{k}'_m \in \mathcal{W}_{n+m}$. It is also convenient to introduce an empty word $\emptyset$ such that $\emptyset w = w\emptyset = w$ for each w. The set of all words (including the empty word) will be denoted by $\mathcal{W}$. For $\mathbf{k} \in \mathbb{Z}^d$, $|\mathbf{k}|$ means the sum-norm of $\mathbf{k}$; for $w = \mathbf{k}_1\cdots\mathbf{k}_n$ we define $|w| = \sum_j |\mathbf{k}_j|$. With this notation the coefficient of ϵ^n in (4) is simply

$$\sum_{w\in\mathcal{W}_n} \alpha_w(t)\, f_w(y_0). \tag{7}$$

Following the terminology of [7], we then say that (4) is the *word series* associated with the coefficients $\alpha_w(t)$ defined in (5) for $w \neq \emptyset$ ($\alpha_\emptyset(t) = 1$ as the coefficient of y_0 in (4) is unity).

It turns out that for each $w \in \mathcal{W}$

$$\alpha_w(t) = \gamma_w(t, t\omega),$$

where $\gamma_w(t,\theta)$ is a suitable polynomial in the variables t, $e^{i\theta_1}$, …, $e^{i\theta_d}$, $e^{-i\theta_1}$, …, $e^{-i\theta_d}$. The (f-independent) coefficients $\gamma_w(t,\theta)$, $w \in \mathcal{W}$, may be computed recursively, see [6] Proposition 4.1 or [7] Proposition 2, and are used to define three new families of coefficients ($w \in \mathcal{W}$, $t \in \mathbb{R}$, $\theta \in \mathbb{T}^d$):

$$\bar{\alpha}_w(t) := \gamma_w(t, 0), \tag{8}$$

$$\bar{\beta}_w := \frac{d}{dt}\bar{\alpha}_w(t)\Big|_{t=0}, \tag{9}$$

$$\kappa_w(\theta) := \gamma_w(0, \theta). \tag{10}$$

After these preliminaries, we may write the following averaging result (see [7], Theorem 2.1) that expresses formally the oscillatory solution $y(t)$ in terms of the solution $Y(t)$ of an averaged autonomous problem:

Theorem 2.1 *The solution of (1)–(2) may be written as*

$$y(t) = U(Y(t), t\omega, \epsilon),$$

where U is the change of variables parameterized by $\theta \in \mathbb{T}^d$

$$y = Y + \epsilon\check{U}(Y, \theta, \epsilon); \quad \check{U}(Y, \theta, \epsilon) := u_1(Y, \theta) + \cdots + \epsilon^{n-1} u_n(Y, \theta) + \cdots \tag{11}$$

with

$$u_n(Y, \theta) := \sum_{w \in \mathcal{W}_n} \kappa_w(\theta)\, f_w(Y), \qquad n = 1, 2, \ldots \tag{12}$$

and

$$Y(t) = y_0 + \sum_{n=1}^{\infty} \epsilon^n \sum_{w \in \mathcal{W}_n} \bar{\alpha}_w(t)\, f_w(y_0) \tag{13}$$

is the solution with initial condition $Y(0) = y_0$ of the autonomous (averaged) system

$$\frac{d}{dt} Y = \epsilon F(Y, \epsilon), \quad F(Y) := F_1(Y) + \epsilon F_2(Y) + \cdots + \epsilon^{n-1} F_n(Y) + \cdots, \tag{14}$$

with

$$F_n(y) := \sum_{w \in \mathcal{W}_n} \bar{\beta}_w\, f_w(Y), \qquad n = 1, 2, \ldots \tag{15}$$

Thus the vector field ϵF in the averaged system (14), the averaged solution $Y(t)$ in (13) and the change of variables U in (11) are the word series that correspond to the coefficients $\bar{\beta}_w$, $\bar{\alpha}_w(t)$ and $\kappa_w(\theta)$ defined in (9), (8) and (10) respectively. These coefficients may be computed recursively as described in [6], Section 4, and are *universal* in the sense that they are completely independent of the vector field f of the oscillatory problem (1) being averaged. For instance the $\bar{\beta}_w$ may be computed from the relations

$$\begin{aligned}
\bar{\beta}_{\mathbf{k}} &= 0, \\
\bar{\beta}_{\mathbf{0}} &= 1, \\
\bar{\beta}_{\mathbf{0}}^{r+1} &= 0, \\
\bar{\beta}_{\mathbf{0}^r\mathbf{k}} &= \frac{i}{\mathbf{k}\cdot\omega}(\bar{\beta}_{\mathbf{0}^{r-1}\mathbf{k}} - \bar{\beta}_{\mathbf{0}^r}), \\
\bar{\beta}_{\mathbf{k}\mathbf{l}_1\cdots\mathbf{l}_s} &= \frac{i}{\mathbf{k}\cdot\omega}(\bar{\beta}_{\mathbf{l}_1\cdots\mathbf{l}_s} - \bar{\beta}_{(\mathbf{k}+\mathbf{l}_1)\mathbf{l}_2\cdots\mathbf{l}_s}), \\
\bar{\beta}_{\mathbf{0}^r\mathbf{k}\mathbf{l}_1\cdots\mathbf{l}_s} &= \frac{i}{\mathbf{k}\cdot\omega}(\bar{\beta}_{\mathbf{0}^{r-1}\mathbf{k}\mathbf{l}_1\cdots\mathbf{l}_s} - \bar{\beta}_{\mathbf{0}^r(\mathbf{k}+\mathbf{l}_1)\mathbf{l}_2\cdots\mathbf{l}_s}),
\end{aligned}$$

where $r, s \geq 1$, $\mathbf{k} \in \mathbb{Z}^d \backslash \{0\}$, and $\mathbf{l}_1, \ldots, \mathbf{l}_s \in \mathbb{Z}^d$. These relations clearly bear out the *small denominator* phenomenon: if $d > 1$, there are (large) values of $|\mathbf{k}|$ for which the the denominator $\mathbf{k}\cdot\omega$ is arbitrarily small.

2.2 Error bounds

Our aim is the derivation of error bounds that are exponentially small as $\epsilon \to 0$. Bounds of size $\mathcal{O}(\epsilon^N)$ with N large but fixed require weaker hypotheses and have been known since the work of Perko [16]. In this subsection we assume that $d > 1$, i.e. that (1) is quasi-periodic. In the periodic case, the derivation of exponentially small error bounds by using word series has been carried out in [7] and uses assumptions weaker than those required here.

Assume that (1) is to be studied in a domain $\mathcal{K} \subset \mathbb{R}^D$ with closure $\bar{\mathcal{K}}$. If $\|\cdot\|$ denotes both a norm in $\mathbb{C}^D$ and the associated norm for $D \times D$ complex matrices, for $\rho \geq 0$ we consider the set $\mathcal{K}_\rho \subset \mathbb{C}^D$ given by

$$\mathcal{K}_\rho = \{y + z \in \mathbb{C}^D : y \in \bar{\mathcal{K}},\ ||z|| \leq \rho\},$$

and for vector or matrix-valued bounded functions ϕ defined in $\mathcal{K}_\rho$ we write

$$\|\phi\|_\rho = \sup_{y \in \mathcal{K}_\rho} \|\phi(y)\|.$$

We shall work under the following assumption on f:
Assumption A. There exist $R > 0$, $\mu > 0$ and an open set $\mathcal{U} \supset \mathcal{K}_R$, such that, for each $\theta \in \mathbb{T}^d$, $f(\cdot, \theta)$ may be extended to a map $\mathcal{U} \to \mathbb{C}^D$ that is analytic at each point $y \in \mathcal{K}_R$. Furthermore the Fourier coefficients $f_{\mathbf{k}}$ of f have bounds

$$\forall \mathbf{k} \in \mathbb{Z}^d, \qquad \|f_{\mathbf{k}}\|_R \leq a_{\mathbf{k}} e^{-\mu|\mathbf{k}|}, \qquad a_{\mathbf{k}} \geq 0, \tag{16}$$

where the $a_{\mathbf{k}}$ are such that

$$M := \sum_{\mathbf{k} \in \mathbb{Z}^d} a_{\mathbf{k}} < \infty.$$

Remark 2.2 Under Assumption A, the Fourier series (3) converges absolutely and uniformly for $y \in \mathcal{K}_R$, $\theta \in \mathbb{T}^d_\mu$, where

$$\mathbb{T}^d_\mu = \{\zeta = \theta + i\eta \in \mathbb{C}^d : \theta \in \mathbb{T}^d, \eta \in \mathbb{R}^d, |\eta_1| \leq \mu, \dots, |\eta_d| \leq \mu\},$$

and, accordingly, f may be extended to a function analytic at each $(y, \theta) \in \mathcal{K}_R \times \mathbb{T}^d_\mu$. Note that also

$$\forall \theta \in \mathbb{T}^d, \qquad \|f(\cdot, \theta)\|_R \leq M.$$

Conversely, suppose that, for a suitable $\mu' > 0$, f may be extended to a function analytic at each $(y, \theta) \in \mathcal{K}_R \times \mathbb{T}^d_{\mu'}$. In that case, $\|f_{\mathbf{k}}\|_R \leq M' e^{-\mu'|\mathbf{k}|}$, where M' is the maximum of $\|f\|$ in $\mathcal{K}_R \times \mathbb{T}^d_{\mu'}$, and Assumption A is satisfied for any $\mu < \mu'$ after taking $a_{\mathbf{k}} = M' \exp\big(-(\mu' - \mu)|\mathbf{k}|\big)$.

The exponential decay in (16) is used to overcome the effect of the small divisors; this will be clearly illustrated in the example presented in Section 2.3. In the periodic case, treated in [7], it is possible to allow $\mu = 0$ in Assumption A; in that scenario f need not be analytic with respect to θ.

An additional hypothesis required to deal with small denominators is the assumption that the vector $\omega \in \mathbb{R}^d$, $d > 1$, satisfies a *strong non-resonance* condition

$$\forall \mathbf{k} \in \mathbb{Z}^d \backslash \{\mathbf{0}\}, \qquad |\mathbf{k} \cdot \omega| \geq c|\mathbf{k}|^{-\nu} \tag{17}$$

for some constants $c > 0$ and $\nu \geq 1$. It is well known that, for fixed $\nu > d - 1$, the measure of the set of vectors $\omega \in \mathbb{R}^d$ that do not satisfy (17) for any $c > 0$ is zero.

It may be proved (similarly to Theorem 3.1 of [7]) that, if f satisfies Assumption A and ω is strongly non-resonant, then for each $\rho \in [0, R)$, $n = 1, 2, \dots$, $y_0 \in \mathcal{K}_\rho$, the series (7) converges absolutely and uniformly. Furthermore the expansion in powers of ϵ in (4) converges for $y_0 \in \mathcal{K}_\rho$ and ϵt sufficiently small.

In the following two propositions we derive bounds for the coefficients and word-basis functions respectively. While the strong non-resonance condition makes it possible to estimate the coefficients from above in spite of the small divisors, the bounds here, that grow with $|w|$, are considerably worse than those obtained in [7] Proposition 5 for the periodic case, that are independent of $|w|$.

Proposition 2.3 *Assume that $\omega \in \mathbb{R}^d$, $d > 1$, satisfies the strong non-resonance condition (17). Then*

1. *The coefficients κ_w in (10) satisfy $|\kappa_w(\theta)| \leq 2^n c^{-n} |w|^{\nu n}$ for $w \in \mathcal{W}_n$, $n = 1, 2, \dots$, and $\theta \in \mathbb{T}$.*

2. *For $w \in \mathcal{W}_n$, $n = 1, 2, \dots$, and $\theta \in \mathbb{T}$, $|(\omega \cdot \partial_\theta)\kappa(\theta)| \leq 2^{n-1} c^{-(n-1)} |w|^{\nu(n-1)}$.*

3. *The coefficients $\bar{\beta}_w$ in (9) satisfy $|\bar{\beta}_w| \leq 2^{n-1} c^{-(n-1)} |w|^{\nu(n-1)}$, for $w \in \mathcal{W}_n$, $n = 1, 2, \dots$*

Proof: By induction in the recursions for $\gamma_w(t, \theta)$ and $\bar{\beta}_w$ given in [6] Proposition 4.1 and Theorem 4.2. □

Proposition 2.4 *If f satisfies Assumption A, then, for $0 \leq \rho < R$, $n = 1, 2, \dots$ and $w \in \mathcal{W}_n$, the following bounds hold:*[1]

$$\|f_w\|_\rho \leq \frac{(n-1)^{n-1}}{(R-\rho)^{n-1}} a_w e^{-\mu|w|} \tag{18}$$

and

$$\|\partial_y f_w\|_\rho \leq \frac{n^n}{(R-\rho)^n} a_w e^{-\mu|w|}. \tag{19}$$

Here $a_w = a_{\mathbf{k}_1} \cdots a_{\mathbf{k}_n}$ if $w = \mathbf{k}_1 \cdots \mathbf{k}_n$.

Proof: Use Cauchy estimates as in the proof of Proposition 6 of [7]. □

We now investigate the convergence of the series (12) and (15). It is at this point that we exploit the exponential factor $\exp(-\mu|\mathbf{k}|)$ in Assumption A to overcome the growth with $|w|$ of the bounds in Proposition 2.3. Note that the leading term in (20) is

[1] Throughout the paper it is understood that for $n = 1$ the expression $(n-1)^{a(n-1)}$, $a > 0$ takes the value 1.

$(n-1)^{(\nu+1)(n-1)}$; a lower value of ν in (17) (corresponding to vectors of frequencies ω 'further away' from resonance) leads to a better error estimate. In the periodic case, see [7], the term $(n-1)^{(\nu+1)(n-1)}$ is replaced by $(n-1)^{n-1}$. Similar comments are valid for (21)–(23).

Proposition 2.5 *Suppose that f satisfies the requirements in Assumption A and $\omega \in \mathbb{R}^d$, $d > 1$, satisfies the strong non-resonance condition (17). For $n = 1, 2, \dots$, $0 \le \rho < R$, $y \in \mathcal{K}_\rho$, $\theta \in \mathbb{T}$, the series in (15) and (12) are absolutely and uniformly convergent. Furthermore, the functions $F_n(y)$, $u_n(y, \theta)$ defined by those series satisfy*

$$||F_n||_\rho \le \frac{ML^{n-1}(n-1)^{(\nu+1)(n-1)}}{(R-\rho)^{n-1}}, \tag{20}$$

$$||u_n(\cdot,\theta)||_\rho \le \frac{L^n n^{\nu n}(n-1)^{n-1}}{(R-\rho)^{n-1}}, \tag{21}$$

$$||(\omega\cdot\partial_\theta)u_n(\cdot,\theta)||_\rho \le \frac{ML^{n-1}(n-1)^{(\nu+1)(n-1)}}{(R-\rho)^{n-1}}, \tag{22}$$

$$||\partial_y u_n(\cdot,\theta)||_\rho \le \frac{L^n n^{(\nu+1)n}}{(R-\rho)^n}, \tag{23}$$

where

$$L = \frac{2\,M\,\nu^\nu}{c\,\mu^\nu\,e^\nu}. \tag{24}$$

Proof: We only prove (20), since the cases of (21)–(23) are similar. From (15), Proposition 2.3 and (18)

$$\|F_n\|_\rho \le \sum_{w\in\mathcal{W}_n} |\bar\beta_w| \|f_w\|_\rho \le \frac{2^{n-1}c^{-(n-1)}(n-1)^{n-1}}{(R-\rho)^{n-1}} \sum_{w\in\mathcal{W}_n} a_w e^{-\mu|w|}|w|^{\nu(n-1)}.$$

The maximum of $e^{-\mu x}x^{\nu(n-1)}$ for $x > 0$ is achieved at $x = \nu(n-1)/\mu$, and therefore,

$$e^{-\mu|w|}|w|^{\nu(n-1)} \le \big(\nu(n-1)/\mu\big)^{\nu(n-1)} \exp\big(-\nu(n-1)\big),$$

so that

$$\|F_n\|_\rho \le \left(\frac{2\,\nu^\nu}{c\,\mu^\nu\,e^\nu}\right)^{n-1} \frac{(n-1)^{(\nu+1)(n-1)}}{(R-\rho)^{n-1}} \sum_{w\in\mathcal{W}_n} a_w.$$

Since

$$\sum_{w\in\mathcal{W}_n} a_w = \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d} a_{\mathbf{k}}\Big)^n = M^n,$$

the result follows. □

The bounds (20) and (21) are too weak to imply the convergence of the series (11) and (14) and in fact it is well known that those series are typically divergent (see Section 2.3). The next proposition studies polynomial truncations of the change of variables (11).

Proposition 2.6 *Suppose that f satisfies the requirements in Assumption A and $\omega \in \mathbb{R}^d$, $d > 1$, satisfies the condition (17). For $N = 1, 2, \dots$ consider the change of variables*

$$y = Y + \epsilon \check{U}^{(N)}(Y, t, \epsilon)$$

with

$$\check{U}^{(N)}(y, \theta, \epsilon) := u_1(y, \theta) + \epsilon u_2(y, \theta) + \cdots + \epsilon^{N-1} u_N(y, \theta)$$

(the functions u_n are as in the preceding proposition). Assume that $\epsilon \in \mathbb{C}$ satisfies:

$$|\epsilon| \leq \epsilon_0, \qquad \epsilon_0 = \epsilon_0(N) := \frac{R}{4LN^{\nu+1}}, \tag{25}$$

then:

1. *For each $\theta \in \mathbb{T}$, the mapping $Y \in \mathcal{K}_{R/2} \mapsto Y + \epsilon \check{U}^{(N)}(Y, \theta, \epsilon)$ is analytic and takes values in $\mathcal{K}_R$.*
2. *For each $\theta \in \mathbb{T}$, $\|(\omega \cdot \partial_\theta) \check{U}^{(N)}(\cdot, \theta, \epsilon)\|_{R/2} \leq 3M/2$.*
3. *For each $\theta \in \mathbb{T}$ and $Y \in \mathcal{K}_{R/2}$, the Jacobian matrix $I + \epsilon \partial_Y \check{U}^{(N)}(Y, \theta, \epsilon)$ is invertible and*
$$\|(I + \epsilon \partial_Y \check{U}^{(N)})^{-1}\| \leq 2.$$

Proof: The sums

$$b_N := \sum_{n=1}^{N} \frac{n^n}{2^n N^n}, \qquad b_N^* := \sum_{n=1}^{N} \frac{(n-1)^{n-1}}{2^{n-1} N^{n-1}} < 1 + b_N, \qquad N = 1, 2, \dots,$$

will appear in the proof. It is readily shown that $b_N \leq 1/2$ and therefore $b_N^* < 3/2$ (see [7] Lemma 4.3).

By (21) with $\rho = R/2$ and (25):

$$\begin{aligned} \|\epsilon \check{U}^{(N)}(\cdot, t, \epsilon)\|_{R/2} &\leq \frac{R}{4} \sum_{n=1}^{N} \frac{n^{\nu n} (n-1)^{n-1}}{2^{n-1} N^{(\nu+1)n}} \\ &\leq \frac{R}{4N} b_N^* \leq \frac{3R}{8N} \leq \frac{R}{2}, \end{aligned}$$

so that $y \in \mathcal{K}_R$ for $Y \in \mathcal{K}_{R/2}$.

The bound in the second item is proved in an analogous way by using (22) and (25).

For the third item, (23) and (25) yield $\|\epsilon \partial_Y \check{U}^{(N)}(Y, t, \epsilon)\| \leq b_N \leq 1/2$; the result then follows from the well-known estimate $\|(I + A)^{-1}\| \leq (1 - \|A\|)^{-1}$. $\Box$

The next theorem is the main result of this section. Note that the exponentially small term $\exp(-K|\epsilon|^{-1/(\nu+1)})$ in (26) increases with ν; again the estimates deteriorate if ω is closer to resonance. In the periodic case, the bound for $R^{(N)}$ in (26) is of the form $C \exp(-K/|\epsilon|)$.

Theorem 2.7 *Suppose that f satisfies the requirements in Assumption A and $\omega \in \mathbb{R}^d$, $d > 1$, satisfies the strong non-resonance condition (17). The application of the change of variables in Proposition 2.6 subject to (25) to the initial value problem (1)–(2) results in a problem*

$$\frac{d}{dt}Y = \epsilon\big(F^{(N)}(Y,\epsilon) + R^{(N)}(Y,t,\epsilon)\big), \qquad Y(0) = y_0,$$

where

$$F^{(N)}(y,\epsilon) = F_1(y) + \epsilon F_2(y) + \cdots + \epsilon^{N-1}F_N(y)$$

(the functions F_j are as defined in Proposition (2.5)). The remainder $R^{(N)}$ possesses the bound

$$\|R^{(N)}(\cdot,t,\epsilon)\|_{R/2} \leq \frac{5|\epsilon/\epsilon_0|^N}{1-|\epsilon/\epsilon_0|}\, M.$$

In particular, assume that for given ϵ, with $|\epsilon| \leq R/(4eL)$ (L is given in (24)), N is chosen as the integer part of the real number $\big(R/(4eL|\epsilon|)\big)^{1/(\nu+1)} \geq 1$. Then the following exponentially small estimate holds true:

$$\|R^{(N)}(\cdot,\theta,\epsilon)\|_{R/2} \leq \frac{5e^2}{e-1}\, M\, \exp\left(-\frac{K}{|\epsilon|^{1/(\nu+1)}}\right), \quad K = \left(\frac{R}{4eL}\right)^{1/(\nu+1)}. \quad (26)$$

Proof: It is very similar to that of Theorem 3.4 in [7]. □

Remark 2.8 It is possible to relax Assumption A and still obtain exponentially small error estimates. For instance, one may assume, instead of (16),

$$\forall \mathbf{k} \in \mathbb{Z}^d, \qquad \|f_{\mathbf{k}}\|_R \leq a_{\mathbf{k}} e^{-\mu|\mathbf{k}|^\alpha}, \qquad a_{\mathbf{k}} \geq 0,$$

with $0 < \alpha \leq 1$. All the results above may be easily adapted to this weaker hypothesis; in (26), the exponential term would be of the form $\exp(-K|\epsilon|^{-1/(\nu'+1)})$ with $\nu' = \nu/\alpha \geq \nu$.

2.3 An example

In this subsection we illustrate, by means of an example, some aspects of the analysis above. In particular, we shall show that the construction of the formal averaged system (29)–(30) cannot be carried out in general if, as a function of θ, f in (1) is only of class C^ℓ, $\ell = 0, 1, \ldots$ and that, under assumption (A), the bound in Proposition 2.5 reflects in general the true growth of $\|F_n\|$ as a function of n.

Consider the system with $D = 2$, $y = (y^1, y^2)$:

$$\frac{dy^1}{dt} = \epsilon, \quad \frac{dy^2}{dt} = \epsilon A(t\omega)g(y^1), \qquad (27)$$

where A is a real-valued function defined on $\mathbb{T}^d$, $\omega \in \mathbb{R}^d$ is non-resonant and g is a smooth real-valued function of a real variable. The solutions of (27) are easily written down in terms of a quadrature, but we do not need them here.

If

$$A(\theta) = \sum_{\mathbf{k}\in\mathbb{Z}^d} e^{i\mathbf{k}\cdot\theta} A_{\mathbf{k}}, \tag{28}$$

is the Fourier expansion of A, then the Fourier coefficients $f_{\mathbf{k}}$ in (3) are

$$f_{\mathbf{0}}(y) = \left[\begin{array}{c} 1 \\ A_{\mathbf{0}}\, g(y^1) \end{array}\right], \qquad f_{\mathbf{k}}(y) = \left[\begin{array}{c} 0 \\ A_{\mathbf{k}} g(y^1) \end{array}\right], \qquad \mathbf{k} \neq \mathbf{0}.$$

In view of (6), for words $w \in \mathcal{W}_{n+1}$, $n = 1, 2, \dots$, the basis function f_w vanishes identically except in cases where $w = \mathbf{0}^n\mathbf{k}$, $\mathbf{k} \in \mathbb{Z}^d$. For these words, $f_{\mathbf{0}^n\mathbf{k}} = [0, A_{\mathbf{k}} g^{(n)}(y^1)]^T$. Furthermore, from the recurrence relations for the coefficients, $\bar\beta_{\mathbf{0}} = 1$, $\kappa_{\mathbf{0}}(\theta) = 0$, and for $\mathbf{k} \neq \mathbf{0}$, $\bar\beta_{\mathbf{k}} = 0$, $\kappa_{\mathbf{k}}(\theta) = i/(\mathbf{k}\cdot\omega)(1 - e^{i\mathbf{k}\cdot\theta})$, $\bar\beta_{\mathbf{0}^n\mathbf{k}} = -(i/(\mathbf{k}\cdot\omega))^n$ and $\kappa_{\mathbf{0}^n\mathbf{k}} = (i/(\mathbf{k}\cdot\omega))^{n+1}(1 - e^{i\mathbf{k}\cdot\theta})$. Thus the change of variables and averaged system are, respectively,

$$\begin{aligned} y^1 &= Y^1, \\ y^2 &= Y^2 + \sum_{n=1}^{\infty} \epsilon^n i^n \left(\sum_{\mathbf{k}\neq 0} \left(\frac{1}{\mathbf{k}\cdot\omega}\right)^n (1 - e^{i\,t\,\mathbf{k}\cdot\omega}) A_{\mathbf{k}} \right) g^{(n-1)}(Y^1), \end{aligned}$$

and

$$\frac{dY^1}{dt} = \epsilon, \tag{29}$$

$$\frac{dY^2}{dt} = \epsilon A_{\mathbf{0}}\, g(Y^1) - \sum_{n=2}^{\infty} \epsilon^n i^{n-1} \left(\sum_{\mathbf{k}\neq\mathbf{0}} \left(\frac{1}{\mathbf{k}\cdot\omega}\right)^{n-1} A_{\mathbf{k}} \right) g^{(n-1)}(Y^1). \tag{30}$$

For reasons of brevity in what follows we shall confine the attention to (29)–(30), the discussion of the change of variables is similar and leads to the same conclusions.

2.3.1 Convergence of the inner sums

For (30) to make sense as a formal series in powers of ϵ it is necessary that all the inner sums

$$\sum_{\mathbf{k}\neq\mathbf{0}} \left(\frac{1}{\mathbf{k}\cdot\omega}\right)^{n-1} A_{\mathbf{k}}, \qquad n = 2, 3, \dots \tag{31}$$

be convergent. In view of the presence of the (possibly small) denominators $\mathbf{k}\cdot\omega$, such requirement is much stronger than the convergence of the Fourier series (28). This will be illustrated in the case where the vector of frequencies has

$$d = 2, \quad \omega^1 = \phi, \quad \omega^2 = 1, \tag{32}$$

where $\phi = (1+\sqrt{5})/2$ is the golden ratio. It is well-known that ω satisfies the strong non-resonance condition (17) with $c = 1$, $\nu = 1$. We denote by $\mathcal{F}_j$ the j-th Fibonnacci number ($\mathcal{F}_0 = 0$, $\mathcal{F}_1 = 1$, $\mathcal{F}_{j+2} = \mathcal{F}_{j+1} + \mathcal{F}_j$, $j = 0, 1, \dots$), and for given $\ell =$

$0, 1, \ldots$, we define a function A as follows. If $\mathbf{k} = \pm(\mathcal{F}_j, -\mathcal{F}_{j+1})$, $j = 0, 1, \ldots$, we set in (28) $A_\mathbf{k} = |\mathbf{k}|^{-(\ell+1)}$; all the remaining $A_\mathbf{k}$, $\mathbf{k} \in \mathbb{Z}^2$ are taken to be zero. Then

$$A(\theta) = 2\sum_{j=0}^{\infty} \cos(\mathcal{F}_j\theta^1 - \mathcal{F}_{j+1}\theta^2)\, \mathcal{F}_{j+2}^{-(\ell+1)}$$

is a function of class C^ℓ of the variable $\theta \in \mathbb{T}^2$, as one sees by considering the series obtained by differentiating term by term. However, for n odd, the inner sum (31) is given by

$$2\sum_{j=0}^{\infty} \left|\frac{1}{\omega^1\mathcal{F}_j - \omega^2\mathcal{F}_{j+1}}\right|^{n-1} \mathcal{F}_{j+2}^{-(\ell+1)}.$$

It is elementary to show, by using the closed-form expression for $\mathcal{F}_j$ as a function of j, that

$$\left|\frac{1}{\omega^1\mathcal{F}_j - \omega^2\mathcal{F}_{j+1}}\right| \geq \frac{\phi}{2}(\mathcal{F}_j + \mathcal{F}_{j+1}) = \frac{\phi}{2}\mathcal{F}_{j+2}, \qquad j = 0, 1, \ldots \tag{33}$$

and therefore (31) is divergent for $n - 1 \geq \ell + 1$. This shows that the construction of the formal averaged system (29)–(30) cannot be carried out in general if, as a function of θ, f in (1) is only of class C^ℓ, $\ell = 0, 1, \ldots$

2.3.2 Growth of the functions F_n

We now keep ω as in (32) but change our choice of A. We take $A_\mathbf{k} = \exp(-\mu'|\mathbf{k}|)$, $\mu' > 0$, for $\mathbf{k} = \pm(\mathcal{F}_j, -\mathcal{F}_{j+1})$, $j = 0, 1, \ldots$, and $A_\mathbf{k} = 0$ for all other $\mathbf{k} \in \mathbb{Z}^2$; with this exponential decay of the $A_\mathbf{k}$ the inner series (31) are convergent. We consider in $\mathcal{K} = \{|y^1| < 1/2\} \subset \mathbb{R}^2$, the system (27) with $g(y^1) = 1/(1 - y^1)$; in $\mathbb{R}^2$ or $\mathbb{C}^2$ we employ the maximum norm. Since $|g(y^1)| \leq 2/(1 - 2R)$ for $|y^1| \leq 1/2 + R$, $0 < R < 1/2$, Assumption A holds for any μ, $0 < \mu < \mu'$, by taking $a_\mathbf{0} = 1$, $a_\mathbf{k} = (2/(1 - 2R))\exp(-(\mu' - \mu)|\mathbf{k}|)$ if $\mathbf{k} = \pm(\mathcal{F}_j, -\mathcal{F}_{j+1})$ and $a_\mathbf{k} = 0$ for the remaining $\mathbf{k}$. Then

$$M = M(\mu', \mu, R) = 1 + \frac{4}{1 - 2R}\sum_{j=0}^{\infty} \exp\big(-(\mu' - \mu)\mathcal{F}_{j+2}\big)$$

and the bound (20) yields, for $\rho = 0$,

$$\|F_n\|_0 \leq M\mathcal{L}^{n-1}(n-1)^{2(n-1)}, \qquad n = 1, 2, \ldots, \tag{34}$$

with $\mathcal{L} = 2M/(R\mu e)$.

To investigate the sharpness of this estimation, we note that from (29)–(30)

$$\|F_n\|_0 \geq \sum_{\mathbf{k}\neq\mathbf{0}} \left|\frac{1}{\mathbf{k}\cdot\omega}\right|^{n-1} |A_\mathbf{k}|\, |g^{(n-1)}(1/2)|, \qquad n = 2, 3, \ldots$$

Now

$$|g^{(n-1)}(1/2)| = 2^n(n-1)!,$$

while for n odd, from (33),

$$\begin{aligned}\sum_{\mathbf{k}\neq\mathbf{0}}\left|\frac{1}{\mathbf{k}\cdot\omega}\right|^{n-1}|A_{\mathbf{k}}| &= 2\sum_{j=0}^{\infty}\left|\frac{1}{\omega^1\mathcal{F}_j-\omega^2\mathcal{F}_{j+1}}\right|^{n-1}\exp(-\mu'\mathcal{F}_{j+2})\\ &\geq 2\sum_{j=0}^{\infty}\left(\frac{\phi}{2}\right)^{n-1}\mathcal{F}_{j+2}^{n-1}\exp(-\mu'\mathcal{F}_{j+2}).\end{aligned}$$

If n is so large that $(n-1)/\mu'\geq 1=\mathcal{F}_2$, there exist a unique integer j_n such that

$$\mathcal{F}_{j_n+2}\leq\frac{n-1}{\mu'}<\mathcal{F}_{j_n+3}\leq 2\mathcal{F}_{j_n+2}$$

and we may estimate the last series as follows:

$$\begin{aligned}\sum_{j=0}^{\infty}\left(\frac{\phi}{2}\right)^{n-1}\mathcal{F}_{j+2}^{n-1}\,\exp(-\mu'\mathcal{F}_{j+2}) &\geq \left(\frac{\phi}{2}\right)^{n-1}\mathcal{F}_{j_n+2}^{n-1}\,\exp(-\mu'\mathcal{F}_{j_n+2})\\ &\geq \left(\frac{\phi}{2}\right)^{n-1}\left(\frac{n-1}{2\mu'}\right)^{n-1}e^{-(n-1)}.\end{aligned}$$

Putting everything together, we see that, for n odd and sufficiently large,

$$\|F_n\|_0\geq 4\left(\frac{\phi}{2\mu' e}\right)^{n-1}(n-1)^{n-1}(n-1)!; \qquad (35)$$

this shows that the series (30) is divergent for $\epsilon\neq 0$. Note that here each of the factors $(n-1)^{n-1}$, $(n-1)!$ on its own prevents the convergence of (30). The factor $(n-1)^{n-1}$ originates from the function A, is related to the small denominators and would not be present in a periodic problem. The factor $(n-1)!$ is introduced by our choice of g; it would not be present if we had set, e.g., $g(y^1)=\cos(y^1)$.

By Stirling's formula, the right-hand side of (35) may be bounded from below by an expression of the form $\hat{M}\hat{\mathcal{L}}^{n-1}(n-1)^{2(n-1)}$, which only differs from the bound (34) in the values of the constants $\hat{M}$ and $\hat{\mathcal{L}}$. We conclude that, for the problem at hand, the estimate (20) captures the essence of the behaviour of $\|F_n\|_0$ as a function of n.

3 Nearly conserved quantities in autonomous Hamiltonian problems

In this section we consider canonical Hamiltonian systems

$$\frac{d}{dt}x=J^{-1}\nabla\mathcal{H}(x,\epsilon),\qquad x\in\mathbb{R}^D, \qquad (36)$$

where D is even, J is the skew-symmetric $D\times D$ matrix that defines the canonical symplectic form and the smooth Hamiltonian function $\mathcal{H}$ may be written in the form

$$\mathcal{H}(x,\epsilon):=\sum_{j=1}^{d}\omega^j I_j(x)+\epsilon K(x), \qquad (37)$$

with $\omega \in \mathbb{R}^d$, $d = 1, 2, \dots$, a non-resonant vector of angular frequencies. It is assumed that:

1. The functions I_j, $j = 1, \dots, d$, are in involution: $\{I_j, I_k\} = 0$ for all j, k. Here $\{\cdot, \cdot\}$ denotes the Poisson bracket

$$\{F, G\} := -\nabla F^T J^{-1} \nabla G.$$

2. The flows $\Psi_t^{[j]}$, $j = 1, \dots, d$, of the Hamiltonian systems

$$\frac{d}{dt} x = J^{-1} \nabla I_j(x),$$

are 2π-periodic.

Note that, in the unperturbed, $\epsilon = 0$, case the solution flow of system (36) is explicitly given by $\Psi_{t\omega}$ with

$$\Psi_\theta = \Psi_{\theta_1}^{[1]} \circ \cdots \circ \Psi_{\theta_d}^{[d]}, \qquad \theta \in \mathbb{T}^d$$

(the maps in the right-hand side commute with each other). As discussed in [6], the format (36)–(37) includes many perturbations of integrable Hamiltonian problems. However here, as distinct from the situation in the theory of integrable systems, the number d of 'actions' I_j does not necessarily coincide with the number of $D/2$ of degrees of freedom. While the hypothesis that the flow $\Psi_t^{[j]}$ has a *constant* (amplitude independent) period is usually associated with linear problems, there are examples of nonlinear isochronous Hamiltonians, see [9] and its references.

3.1 Formal conserved quantities

The problem at hand may be recast in the form (1) by performing the canonical, time-dependent change of variables $x = \Psi_{t\omega}(y)$. The corresponding (pulled-back) vector field is Hamiltonian with $f(y, \theta) = J^{-1} \nabla_y K(\Psi_\theta(y))$. Furthermore each Fourier coefficient $f_{\mathbf{k}}$ in (3) is also a Hamiltonian vector field and its Hamiltonian $H_{\mathbf{k}}(y)$ is the corresponding Fourier coefficient of $K(\Psi_\theta(y))$:

$$K(\Psi_\theta(y)) = \sum_{\mathbf{k} \in \mathbb{Z}^d} e^{i\mathbf{k}\cdot\theta} H_{\mathbf{k}}(y). \tag{38}$$

As discussed in [6], Section 3, the stroboscopically averaged system (14) is once more Hamiltonian and given by

$$\frac{d}{dt} Y = \epsilon J^{-1} \nabla \tilde{K}(Y, \epsilon) \tag{39}$$

with

$$\epsilon \tilde{K} := \sum_{n=1}^{\infty} \epsilon^n \tilde{K}_n, \qquad \tilde{K}_n := \frac{1}{n} \sum_{w \in \mathcal{W}_n} \bar{\beta}_w \, \mathcal{B}_w. \tag{40}$$

Here, $\bar{\beta}_w$ is the universal coefficient in (9) and $\mathcal{B}_w$ denotes the iterated Poisson bracket

$$\mathcal{B}_w := \{\{\cdots\{\{H_{\mathbf{k}_1}, H_{\mathbf{k}_2}\}, H_{\mathbf{k}_3}\}\cdots\}, H_{\mathbf{k}_n}\}, \qquad w = \mathbf{k}_1 \cdots \mathbf{k}_n \in \mathcal{W}_n.$$

(For words with one letter, $\mathcal{B}_{\mathbf{k}} = H_{\mathbf{k}}$.)

In the unpertubed, $\epsilon = 0$, case the functions I_j have been assumed to be conserved quantities in involution of (36). It is proved in [6], Section 5.3, that, in the perturbed case there exists a formal decomposition

$$\mathcal{H}(x, \epsilon) = \sum_{j=1}^{d} \omega^j \tilde{I}_j(x, \epsilon) + \epsilon \tilde{K}(x, \epsilon), \tag{41}$$

where the $\tilde{I}_j$, $j = 1, \ldots, d$ are formal first integrals in involution of both (36) and (39). The power-series $\tilde{I}_j$ is given by

$$\tilde{I}_j(x, \epsilon) := I_j(x) + \sum_{n=1}^{\infty} \epsilon^n \tilde{I}_{j,n}(x), \qquad \tilde{I}_{j,n}(x) := \frac{1}{n} \sum_{w \in \mathcal{W}_n} \beta_w^{[j]} \mathcal{B}_w(x), \tag{42}$$

where the coefficients $\beta_w^{[j]}$ are universal (i.e. independent of the functions I_j, K in the given Hamiltonian (37)) and may be computed recursively by means of the following formulae ([6], Theorem 5.7):

$$\begin{aligned}
\beta_{\mathbf{k}}^{[j]} &= \frac{\mathbf{k} \cdot \mathbf{e}_j}{\mathbf{k} \cdot \omega}, \\
\beta_{\mathbf{0^r}}^{[j]} &= 0, \\
\beta_{\mathbf{0}^r \mathbf{k}}^{[j]} &= \frac{i}{\mathbf{k} \cdot \omega} \beta_{\mathbf{0}^{r-1} \mathbf{k}}^{[j]}, \\
\beta_{\mathbf{k}\mathbf{l}_1 \cdots \mathbf{l}_s}^{[j]} &= \frac{i}{\mathbf{k} \cdot \omega} (\beta_{\mathbf{l}_1 \cdots \mathbf{l}_s}^{[j]} - \beta_{(\mathbf{k}+\mathbf{l}_1)\mathbf{l}_2 \cdots \mathbf{l}_s}^{[j]}), \\
\beta_{\mathbf{0}^r \mathbf{k}\mathbf{l}_1 \cdots \mathbf{l}_s}^{[j]} &= \frac{i}{\mathbf{k} \cdot \omega} (\beta_{\mathbf{0}^{r-1} \mathbf{k}\mathbf{l}_1 \cdots \mathbf{l}_s}^{[j]} - \beta_{\mathbf{0}^r (\mathbf{k}+\mathbf{l}_1)\mathbf{l}_2 \cdots \mathbf{l}_s}^{[j]}).
\end{aligned}$$

Here $r, s \geq 1$, $\mathbf{k} \in \mathbb{Z}^d \backslash \{0\}$, $\mathbf{l}_1, \ldots, \mathbf{l}_s \in \mathbb{Z}^d$ and $\mathbf{e}_j$ denotes the j-th coordinate vector. Note once more the presence of small denominators for $d > 1$.

There are two sources of analytical concern in (42). The first is the convergence of the series that define the coefficients $\tilde{I}_{j,n}$, the second the convergence of the series in powers of ϵ for $\tilde{I}_j$. These difficulties are addressed next.

3.2 Approximate conservation

We focus on the quasi-periodic, $d > 1$, situation and leave the case $d = 1$ for Remark 3.7.

We study (36) in a domain $\mathcal{K} \subset \mathbb{R}^D$ and consider the sets $\mathcal{K}_\rho \subset \mathbb{C}^D$ introduced in the preceding section. Some results below become somewhat simpler if the norm $\|\cdot\|$ in $\mathbb{C}^D$ is assumed to be the standard Euclidean norm for which J is an isometry. We hereafter work under this assumption. For bounded, scalar-valued functions defined in $\mathcal{K}_\rho$, the notation $\|\cdot\|_\rho$ refers to the familiar sup-norm.

We work under the following hypothesis:

Assumption B. There exist $R > 0$, $\mu > 0$ and an open set $\mathcal{U} \supset \mathcal{K}_R$, such that, for each $\theta \in \mathbb{T}^d$, $K(\Psi_\theta(\cdot))$ may be extended to a map $\mathcal{U} \to \mathbb{C}$ that is analytic at each point in $\mathcal{K}_R$. Furthermore the Fourier coefficients $H_\mathbf{k}$ in (38) have bounds

$$\forall \mathbf{k} \in \mathbb{Z}^d, \qquad \|H_\mathbf{k}\|_R \le a_\mathbf{k} e^{-\mu|\mathbf{k}|}, \qquad a_\mathbf{k} \ge 0,$$

where the $a_\mathbf{k}$ are such that

$$M := \sum_{\mathbf{k}\in\mathbb{Z}^d} a_\mathbf{k} < \infty.$$

Note that, in particular, $K(y) = K(\Psi_0(y))$ is assumed to be analytic.

The convergence of the series for $\widetilde{I}_{j,n}$ in (42) will be proved by estimating separately the coefficients $\beta_w^{[j]}$ and the iterated Poisson brackets $\mathcal{B}_w$. For the coefficients we have:

Proposition 3.1 *Assume that $\omega \in \mathbb{R}^d$, $d > 1$, satisfies the strong non-resonance condition (17). Then, for $w \in \mathcal{W}_n$:*

$$|\beta_w^{[j]}| \le 2^{n-1} c^{-n} |w|^{\nu n+1} \le 2^{n-1} c^{-n} |w|^{\nu(n+1)}.$$

Proof: For the first inequality, use induction in the recursions for $\beta_w^{[j]}$ given above. The second is trivial since $\nu \ge 1$. $\Box$

The estimation of the iterated brackets relies on the Cauchy estimates provided in the following Lemma:

Lemma 3.2 *Assume that $\rho \ge 0$, $\delta > 0$, $n = 1, 2, \dots$ and that F, G, $F_1, \dots F_n$ are complex analytic functions in $\mathcal{K}_{\rho+\delta}$. Then*

$$\begin{aligned} \|\nabla F\|_\rho &\le \delta^{-1} \|F\|_{\rho+\delta}, \\ \|\{F, G\}\|_\rho &\le \delta^{-1} \|\nabla F\|_\rho \, \|G\|_{\rho+\delta}, \end{aligned}$$

and

$$\|\{\{\cdots\{\{F_1, F_2\}, F_3\}\cdots\}, F_n\}\|_\rho \le \frac{(n-1)^{2(n-1)}}{\delta^{2(n-1)}(n-1)!} \|F_1\|_{\rho+\delta} \cdots \|F_n\|_{\rho+\delta}.$$

Proof: For the first inequality use the standard Cauchy estimate for the function $\tau \in \mathbb{C} \mapsto F(x + \tau v) \in \mathbb{C}$ where $v \in C^D$, $\|v\| = 1$, $x \in \mathcal{K}_\rho$ (cf. the proof of Lemma 4.2 in [7]). For the second consider analogously the function $\tau \in \mathbb{C} \mapsto G(x - \tau J^{-1}\nabla F(x)) \in \mathbb{C}$. The third inequality is trivially true for $n = 1$. If $n > 1$, consider the radii $\rho_k = \rho + k\delta/(n-1)$, $k = 0, \dots, n-1$. Then

$$\begin{aligned} &\|\{\{\cdots\{\{F_1, F_2\}, F_3\}\cdots\}, F_n\}\|_{\rho_0} \\ &\qquad \le \frac{n-1}{\delta} \|\{\{\cdots\{\{F_1, F_2\}, F_3\}\cdots\}, F_{n-1}\}\|_{\rho_1} \, \|\nabla F_n\|_{\rho_0}, \end{aligned}$$

and iterating

$$\|\{\{\cdots\{\{F_1,F_2\},F_3\}\cdots\},F_n\}\|_{\rho_0} \\ \le \frac{(n-1)^{n-1}}{\delta^{n-1}}\|F_1\|_{\rho_{n-1}}\|\nabla F_2\|_{\rho_{n-2}}\cdots\|\nabla F_n\|_{\rho_0}.$$

The proof concludes by noting that, for $k=2,\ldots,n$,

$$\|\nabla F_k\|_{\rho_{n-k}} \le \frac{n-1}{(k-1)\delta}\|F_k\|_{\rho_{n-1}}.$$

□

Proposition 3.3 *Suppose that Assumption B holds and $\omega \in \mathbb{R}^d$, $d>1$, satisfies the strong non-resonance condition (17).*

1. *For $n=1,2,\ldots$, the series for $\tilde{K}_n$ in (40) is absolutely and uniformly convergent in $\mathcal{K}_\rho$, $0\le\rho<R$.*
2. *Similarly the series for $\tilde{I}_{j,n}(x)$, $j=1,\ldots,d$, in (42) is absolutely and uniformly convergent in $\mathcal{K}_\rho$ and*

$$\|\tilde{I}_{j,n}\|_\rho \le \frac{2^{n-1}M^n(n-1)^{2(n-1)}\big(\nu(n+1)\big)^{\nu(n+1)}}{c^n(R-\rho)^{2(n-1)}\mu^{\nu(n+1)}e^{\nu(n+1)}n!}.$$

Proof: It follows the pattern of that of Proposition 2.5 and will not be given. □

Since the bounds we have just obtained for the coefficients $\tilde{I}_{j,n}(x)$ are too weak to establish the convergence of the series in powers of ϵ in (42), we have to resort to using the polynomial truncations

$$\tilde{I}_j^{(N)}(x,\epsilon) := I_j(x) + \sum_{n=1}^{N}\epsilon^n\,\tilde{I}_{j,n}(x), \qquad x\in\bar{\mathcal{K}}, \quad N=1,2,\ldots \tag{43}$$

For these we have the following result:

Lemma 3.4 *Suppose that Assumption B holds and $\omega\in\mathbb{R}^d$, $d>1$, satisfies the strong non-resonance condition (17). For $j=1,\ldots,d$, $N=1,2,\ldots$, $x\in\mathcal{K}_\rho$, $0\le\rho<R$,*

$$\{\mathcal{H},\tilde{I}_j^{(N)}\} = \epsilon^{N+1}\{K,\tilde{I}_{j,N}\}.$$

Proof: According to the first item in Proposition 3.3, we may define, for $x\in\mathcal{K}_\rho$,

$$\epsilon\tilde{K}^{(N)}(x,\epsilon) := \sum_{n=1}^{N}\epsilon^n\tilde{K}_n(x).$$

Then, the decomposition (41) implies that, as a function of ϵ,

$$\sum_{k=1}^{d}\omega^k\tilde{I}_k^{(N)}(x,\epsilon) + \epsilon\tilde{K}^{(N)}(x,\epsilon), \tag{44}$$

is the N-degree leading polynomial in the power series for $\mathcal{H}$. Since (37) shows that $\mathcal{H}$ is actually a polynomial of the first degree in ϵ, then $\mathcal{H}$ and (44) must coincide and, accordingly,

$$\{\mathcal{H}, \tilde{I}_j^{(N)}\} = \left\{ \sum_{k=1}^{d} \omega^k \tilde{I}_k^{(N)} + \epsilon \tilde{K}^{(N)}, \tilde{I}_j^{(N)} \right\}. \tag{45}$$

From the formal result in [6] mentioned above, $\{\tilde{I}_k, \tilde{I}_j\} = 0$, $\{\tilde{K}, \tilde{I}_j\} = 0$ and therefore the right-hand side of (45) is $\mathcal{O}(\epsilon^{N+1})$. The left-hand side of (45) is a polynomial of degree $\leq N+1$ (bracket of a linear polynomial and a polynomial of degree $\leq N$). We conclude that the polynomial $\{\mathcal{H}, \tilde{I}_j^{(N)}\}$ coincides with its leading ϵ^{N+1} term, i.e. with the bracket of the leading terms ϵK of $\mathcal{H}$ and $\epsilon^N \tilde{I}_{j,N}$ of $\tilde{I}_j^{(N)}$. □

We are now in a position to give the main result in this section:

Theorem 3.5 *Suppose that Assumption B holds, $\omega \in \mathbb{R}^d$, $d > 1$, satisfies the strong non-resonance condition (17). Then:*

1. *For $j = 1, \dots, d$, $N = 1, 2, \dots$, $x \in \mathcal{K}$,*

$$|\{\mathcal{H}, \tilde{I}_j^{(N)}\}| \leq C|L|^{N+1}|\epsilon|^{N+1}(N+1)^{(1+\nu)(N+1)} \tag{46}$$

   with

$$C = \frac{R^2 c}{64 e^2 M} \|K\|_{R/2}, \qquad L = \frac{8M\nu^\nu}{cR^2 \mu^\nu e^{\nu-1}}.$$

2. *Assume that $|\epsilon|$ is such that $L^{-1/(\nu+1)}|\epsilon|^{-1/(\nu+1)}e \geq 2$ and choose N as the integer part of $L^{-1/(\nu+1)}|\epsilon|^{-1/(\nu+1)}e - 1$. Then for $j = 1, \dots, d$, $x \in \mathcal{K}$:*

$$|\{\mathcal{H}, \tilde{I}_j^{(N)}\}| \leq C' \exp\left(-\frac{K}{|\epsilon|^{1/(\nu+1)}} \right), \tag{47}$$

   where

$$C' = Ce^{\nu+1}, \qquad K = \frac{\nu+1}{eL^{1/(\nu+1)}}.$$

Proof: From Lemmas 3.4 and 3.2

$$|\{\mathcal{H}, \tilde{I}_j^{(N)}\}| = |\epsilon^{N+1}\{K, \tilde{I}_{j,N}\}| \leq \frac{4}{R^2} |\epsilon|^{N+1} \|K\|_{R/2} \|\tilde{I}_{j,N}\|_{R/2},$$

a result that combined with Proposition 3.3 leads to

$$|\{\mathcal{H}, \tilde{I}_j^{(N)}\}| \leq C|L\epsilon|^{N+1} \frac{4(N-1)^{2(N-1)}}{e^{N-1}N!}(N+1)^{\nu(N+1)}.$$

The bound (46) is then a consequence of the inequality

$$\frac{4(N-1)^{2(N-1)}}{e^{N-1}N!} \leq (N+1)^{N+1},$$

valid for $N \geq 1$.

The proof of the second item is an elementary computation. □

Since $\{\mathcal{H}, F\}$ is the rate of change of F along solutions of (36), the theorem implies that $\tilde{I}_j^{(N)}$ is approximately conserved by the flow of the given system over exponentially long time intervals:

Corollary 3.6 *Assume that ϵ and N are as in the second item of Theorem 3.5. Choose $\delta > 0$ arbitrarily small and set*

$$T := \frac{\delta}{C'} \exp\left(\frac{K}{|\epsilon|^{1/(\nu+1)}}\right).$$

Then for any solution $x(t)$ of (36) that remains in $\mathcal{K}$ for $0 \leq t \leq T$:

$$|\ \tilde{I}_j^{(N)}(x(t), \epsilon) - \tilde{I}_j^{(N)}(x(0), \epsilon)\ | \leq \delta, \qquad 0 \leq t \leq T, \quad j = 1, \ldots, d.$$

Remark 3.7 When $d = 1$ small denominators are not present and one obtains bounds of the form (46)–(47) with $\nu = 0$, even if Assumption B is weakened to have $\mu = 0$.

Acknowledgement. A. Murua and J.M. Sanz-Serna have been supported by projects MTM2010-18246-C03-03 and MTM2010-18246-C03-01 respectively from Ministerio de Ciencia e Innovación.

References

[1] V. I. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations, 2nd ed., Springer, New York, 1988.

[2] V. I. Arnold, Mathematical Methods of Classical Mechanics, 2nd ed., Springer, New York, 1989.

[3] M.P. Calvo, Ph. Chartier, A. Murua, J.M. Sanz-Serna, A stroboscopic method for highly oscillatory problems. In: B. Engquist, O. Runborg, R. Tsai, (eds.) Numerical Analysis of Multiscale Computations, Lect. Notes Comput. Sci. Eng. 82, pp. 73–87. Springer, Berlin, 2011.

[4] M.P. Calvo, Ph. Chartier, A. Murua, J.M. Sanz-Serna, Numerical stroboscopic averaging for ODEs and DAEs, Applied Numerical Mathematics, *Appl. Numer. Math.* **61**, 1077–1095 (2011).

[5] P. Chartier, A. Murua, J.M. Sanz-Serna, Higher-Order averaging, formal series and numerical integration I: B-series, *Found. Comput. Math.* **10**, 695–727 (2010).

[6] P. Chartier, A. Murua, J.M. Sanz-Serna, Higher-Order averaging, formal series and numerical integration II: the quasi-periodic case, *Found. Comput. Math.* **12**, 471-508 (2012).

[7] P. Chartier, A. Murua, J.M. Sanz-Serna, A formal series approach to averaging: exponentially small error estimates, *DCDS A* **32**, 3009-3027 (2012).

[8] G. Contopoulos, C. Efthymiopoulos and A. Giorgilli, Non-convergence of formal integrals of motion, *J. Phys. A* **36**, 8369–8660 (2003).

[9] A. Durga Devi, R. Gladwin Pradeep, V. K. Chandrasekar, M. Laksmanan, Method of generating N-dimensional nonsigular Hamiltonian systems, *J. Nonlinear Math. Phys.* to appear.

[10] C. Efthymiopoulos, A. Giorgilli and G. Contopoulos, Non-convergence of formal integrals: II. Improved estimates for the optimal order of truncation, *J. Phys. A* **37**, 10831–10858 (2004).

[11] E. Hairer, Ch. Lubich and G. Wanner, Geometric Numerical Integration, 2nd ed., Springer, Berlin, 2006.

[12] E. Hairer, S. P. Nørsett and G. Wanner, Solving Ordinary Differential Equations I, Nonstiff Problems, 2nd ed., Springer, Berlin, 1993.

[13] P. Lochak, C. Meunier, Multiphase Averaging for Classical Systems, with Applications to Adiabatic Theorems, Springer, New York, 1988.

[14] A. Murua, The Hopf algebra of rooted trees, free Lie algebras and Lie series, *Found. Comput. Math.* **6**, 387–426 (2006).

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

[16] L.M. Perko, Higher order averaging and related methods for perturbed periodic and quasi-periodic systems, *SIAM J. Appl. Math.* **17**, 698–724 (1968).

[17] J. A. Sanders, F. Verhulst and J. Murdock, Averaging Methods in Nonlinear Dynamical Systems (2nd. ed.), Springer, New York, 2007.

[18] J. M. Sanz-Serna and M. P. Calvo, Numerical Hamiltonian Problems, Chapman and Hall, London, 1994.

[19] C. Simó, Averaging under fast quasi-periodic forcing. In: I. Seimenis (ed.) Proceedings of the NATO-ARW Integrable and Chaotic Behaviour in Hamiltonian Systems, Torun, Poland, 1993, pp. 13–34. Plenum, New York, 1994.