Higher-order averaging, formal series and numerical integration II: the quasi-periodic case

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

September 29, 2010

1 Introduction

The Krylov-Bogoliubov-Mitropolski averaging procedure is a famous perturbation method for approximating the solution of a perturbed ordinary differential equation of the form

$$\begin{cases} y'(\tau) &= \varepsilon f(y(\tau), \tau; \varepsilon) \\ y(0) &= y_0 \end{cases} \tag{1.1}$$

on intervals of length $1/\varepsilon$ and for small $\varepsilon > 0$. The usual, though fundamental, assumption adopted in regard to this system is the existence of the average

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(y, \tau; \varepsilon) d\tau.$$

In this contribution, we shall limit our investigations to the case of functions f that are periodic or quasi-periodic in τ, a sufficient condition for the above limit to exist.

Following the pioneering work of Krylov and Bogoliubov [KB34], Bogoliubov and Mitropolski gave the first comprehensive presentation [BM55] of the method of averaging for first and second orders in ε. Various authors have later considered higher order expansions and have established a systematic procedure to derive them. However, to our knowledge, the first rigorous complete account to high-order averaging was given by L. Perko. In [Per68], the author constructs a change of variable together with an autonomous differential equation whose solution satisfies (1.1) up to terms of size ε^N.

*INRIA Rennes, ENS Cachan Bretagne, Campus Ker-Lann, av. Robert Schumann, 35170 Bruz, France. Email: Philipe.Chartier@irisa.fr

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

‡Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain. Email: sanzsern@mac.uva.es

For the case of periodic vector fields (i.e. when f in (1.1) is periodic in τ), the construction offers some freedom in the definition of the change of variables, a freedom which allows to cover several specific methods of averaging encountered in the literature, amongst which most prominently *stroboscopic* (for which the change of variables coincides with the identity map at times that are multiple of the period) and *classical* averaging (for which the angular average of the change of variables is the identity map). In [CMSS10], we situated our work in a similar context and we demonstrated how B-series could be used to derive a *fully explicit* expansion of the averaged vector field obtained by stroboscopic averaging.

Whenever f **is only assumed to be quasi-periodic**, another theorem obtained by L. Perko in [Per68] establishes the existence of a change of variable which brings system (1.1) to an autonomous differential equation. Since we are specifically concerned by this case in the present paper, it seems convenient to quote this result explicitly:

Theorem 1.1 *Consider the system*

$$\left\{ \begin{array}{lllll} y' & = & \varepsilon f(y,\theta), & y(0) & = & y_0, \\ \theta' & = & \omega, & \theta(0) & = & \theta_0, \end{array} \right. \tag{1.2}$$

where f is a real-analytic function of $\theta \in \mathbb{T}^d$ for each y in a convex region G (containing y_0) and 2π-periodic in each component of θ. Assume that the vector $\omega \in \mathbb{R}^d$ satisfies the **strong non-resonnance** *condition*

$$\forall \mathbf{k} \in \mathbb{Z}^d/\{\mathbf{0}\}, \quad |\mathbf{k}\cdot\omega| \geq \mathbf{c}|\mathbf{k}|^{-\nu} \tag{1.3}$$

for some positive constants c and ν. Assume in addition that the solution of the differential equation

$$\left\{ \begin{array}{lll} Y' & = & \varepsilon F_1(Y) \\ Y(0) & = & y_0 \end{array} \right.$$

where

$$F_1(Y) = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} f(Y,\theta)d\theta,$$

has a solution $Y_1(\tau)$ which remains in G for $\tau \leq L/\varepsilon$: Then, for any $N \geq 1$, there exists $\varepsilon_N > 0$ such that for $0 < \varepsilon < \varepsilon_N$, the system (1.2) has a unique solution in G for $0 \leq \tau \leq L/\varepsilon$ satisfying

$$\|y(\tau) - U(Y_N(\tau), \theta_0 + \omega\tau)\| \leq C\varepsilon^N$$

where $Y_N(\tau)$ is the solution of the differential equation

$$Y' = \varepsilon F_1(Y) + \varepsilon^2 F_1(Y) + \ldots + \varepsilon^N F_N(Y), \quad Y(0) = \xi$$

and where the change of variable $U(Y,\theta) = Y + \varepsilon u_1(Y,\theta) + \ldots + u_{N-1}(Y,\theta)$ and the functions F_j are *defined recursively by (for $j \geq 1$)*

$$\tilde{F}_j(Y,\theta) = \sum_{k=1}^{j-1} \left[\frac{1}{k!} \sum_{i_1+\ldots+i_k=j-1} \frac{\partial^k f}{\partial y^k}(Y,\theta)\Big(u_{i_1}(Y,\theta),\ldots,u_{i_k}(Y,\theta)\Big) - \frac{\partial u_k}{\partial Y}(Y,\theta)F_{j-k}(Y)\right], \quad (1.4)$$

$$F_j(Y) = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} \tilde{F}(Y,\theta)d\theta, \quad (1.5)$$

$$\omega \cdot \nabla_\theta u_j(Y,\theta) = \tilde{F}_j(Y,\theta) - F_j(Y), \qquad \int_{\mathbb{T}^d} u_j(Y,\theta)d\theta = 0. \quad (1.6)$$

The initial condition ξ is determined implicitly by the equation

$$\xi = y_0 - \sum_{j=1}^{N-1} \varepsilon^j u_j(\xi,\theta_0) + \mathcal{O}(\varepsilon^N).$$

A noticeable difference with the single frequency situation, is that the change of variables U is constructed in such a way that it has zero average (see equation (1.6)), a procedure which we refer to as classical-averaging. It is clear that the procedure of stroboscopic averaging, for which $U(y,\tau)$ coincides with the identity map at times τ that are multiple of the period, can not be generalized straightforwardly (there is nothing such as a period in the quasi-periodic case). However, we show in this paper that it is possible, upon using B-series, to derive a change of variable U such that $U(Y,\theta_0) = Y$, so that $\xi = y_0$. We have called this averaging procedure **quasi-stroboscopic averaging** for there exists a quasi-period $T_\nu > 0$ such that $U(Y, nT\omega) = Y + \mathcal{O}(n\nu)$ for arbitrarily small ν. Actually, the change of variables can be chosen *arbitrarily*, as this will be proved in Section 3. However and quite surprisingly, this does make any difference as far as the long-term dynamical behaviour is concerned, since all averaging procedures are actually equivalent is the sense that any two averaged vector-fields so obtained are conjugate.

Another consequence of the results we get here is the construction of formal adiabatic invariants for the Fermi-Pasta-Ulam problem and more generally, for systems that are non-linear perturbations of linear ones.

	Periodic case	Quasi-periodic case
Perko's recursion	Stroboscopic aver. and classical-aver.	classical-aver.
B-series	Stroboscopic aver. and classical-aver.	Stroboscopic aver., classical-aver. ...

Figure 1: Averaging methods available for the periodic and quasi-periodic cases

As already mentioned, the expansion we get is completely explicit. In particular, we give recursive formulas for the coefficients of the series we consider that are valid in both the periodic case (and then complement our previous work in [CMSS10]) and the quasi-periodic case considered here.

2 Expansion of the exact solution as a mode-coloured B-series

We consider the multi-frequency highly-oscillatory problem (1.1) under the assumption that $f(y,\theta)$ is real-analytic on a domain of $\mathbb{R}^n \times \mathbb{T}^d$ and thus possesses a Fourier expansion of the form

$$f(y,\theta) = \sum_{\mathbf{k}\in\mathbb{Z}^d} e^{i(\mathbf{k}\cdot\theta)} f_{\mathbf{k}}(y). \tag{2.1}$$

The Fourier coefficients $f_{\mathbf{k}}(y)$ are in general complex functions, but, in order to have a real system, we assume that, for each $\mathbf{k}$, $f_{\mathbf{k}} \equiv f^*_{-\mathbf{k}}$ where * denotes the complex conjugate.

2.1 Mode-coloured trees and elementary differentials

As it has been done for the mono-frequency case in [CMSS10], the solution y of (1.1) can be expanded as a *B-series*, that is to say a formal series whose terms are indexed by (rooted) trees. In this subsection we describe a variant of the trees considered in [CMSS10] and that will feature in the present expansion of y.

In view of the structure of the right hand-side of (1.1), the vertices of the trees to be used here correspond to the functions $f_{\mathbf{k}}$ and their derivatives; hence, each vertex possesses here a label which is now a multi-index in $\mathbb{Z}^d$. The set $\mathcal{T}$ of trees may now be defined recursively by the following two rules:

1. For all $\mathbf{k} \in \mathbb{Z}^d$, $\textcircled{\scriptsize k}$ belongs to $\mathcal{T}$;
2. If $u_1, \dots, u_n$ are n trees of $\mathcal{T}$, then, the tree

$$u = [u_1, \dots, u_n]_{\mathbf{k}} \tag{2.2}$$

obtained by connecting their roots to a new root with multi-index $\mathbf{k} \in \mathbb{Z}^d$, belongs to $\mathcal{T}$.

The expansion of y will be graded according to powers of ϵ in connection with the usual notion of order: The order $|u|$ of a tree $u \in \mathcal{T}$ is simply defined as its number of nodes. We shall later use also the concept of total index $\mathcal{I}_u$, defined as the sum over all nodes of u of their multi-indices. Elementary differentials can also be defined recursively by the formulae

$$\mathcal{F}_{\textcircled{\scriptsize k}}(y_0) = f_{\mathbf{k}}(y_0) \text{ and } \mathcal{F}_{[u_1,\dots,u_n]_{\mathbf{k}}}(y_0) = \frac{\partial^n f_{\mathbf{k}}}{\partial y^n}(y_0)\Big(\mathcal{F}_{u_1}(y_0), \dots, \mathcal{F}_{u_n}(y_0)\Big).$$

2.2 Mode-coloured B-series

In accordance with previous paragraph, we now define mode-coloured B-series as power series indexed by mode-coloured trees of the set $\mathcal{T}$ of the form

$$B(\alpha, y) = \alpha_{\emptyset} y + \sum_{u\in\mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \alpha_u \mathcal{F}_u(y)$$

where σ is the *symmetry* function from $\mathcal{T}$ to $\mathbb{N}$ that acts as a normalization factor, and it is defined as its *natural* extension from standard trees (see for instance [CMSS10]) to mode-coloured trees, and where $\alpha \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$, that is,

$$\begin{array}{rcl} \alpha : \mathcal{T} \cup \{\emptyset\} & \to & \mathbb{C} \\ u & \mapsto & \alpha_u. \end{array}$$

Given two B-series $B(\alpha, y)$ and $B(\beta, y)$ (where $\alpha, \beta \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$,) consider their composition

$$B(\beta, B(\alpha, y)).$$

It is well known that that, provided that $\alpha_\emptyset = 1$, the result is still a B-series $B(\alpha \star \beta, y)$ with coefficients $\alpha \star \beta \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$. As established for standard B-series, the corresponding composition law induces a rich structure of *Hopf algebra* in which it happens to be a *convolution product*. For the sake of this exposition, it is not necessary to present the details of this construction and in particular the precise form of this law. We next fix some notation an recall some known results that are required in the present work.

We will denote $\mathcal{G} = \{\alpha \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}} \ : \ \alpha_\emptyset = 1\}$. Given $\alpha \in \mathcal{G}$ and $\beta \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$, then $\alpha \star \beta \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$. If β also belongs to $\mathcal{G}$, then $\alpha \star \beta \in \mathcal{G}$. Actually, $\mathcal{G}$ has a group structure, with neutral element $\mathbb{1}$ defined by $\mathbb{1}_\emptyset = 1$ and $\mathbb{1}_u = 0$ for all $u \in \mathcal{T}$. Obviously, the neutral element $\mathbb{1}$ in $\mathcal{G}$ corresponds to the B-series representing the identity map, that is, $B(\mathbb{1}, y) \equiv y$.

Given $\alpha \in \mathcal{G}$ and $\beta \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$, $(\alpha \star \beta)_\emptyset = \beta_\emptyset$, and for each tree $u \in \mathcal{T}$, $(\alpha \star \beta)_u$ is of the form

$$(\alpha \star \beta)_u = \alpha_u \beta_\emptyset + \beta_u + \sum_{m=2}^{|u|} \sum_{|u_1|+\ldots+|u_m|=|u|} c_{u_1,\ldots,u_m} \, \beta_{u_1} \prod_{i=2}^{m} \alpha_{u_i}$$

where the $c_{u_1,\ldots,u_m}$'s are integer coefficients.

In particular, it will be relevant here that the product $\star$ is linear in the right factor, and also that $(\alpha \star \beta)_u - \alpha_u \beta_\emptyset - \beta_u$ is a polynomial in the coefficients α_v and β_w for trees v, w with less vertices than u (which will allow to prove results by induction on the number of vertices). We will also use the fact that, when $\beta_u = 0$ for all trees with $|u| \neq 1$, then,

$$(\alpha \star \beta)_u = \beta_{\textcircled{\scriptsize k}} \alpha_{u_1} \cdots \alpha_{u_m}$$

provided that $u = [u_1 \cdots u_m]_k$ (including the case $u = [\emptyset]_k$).

2.3 Formal expansion

As stated in the introduction of this section, our aim is here to express formally the exact solution of (1.1) as

$$y(\tau) = B(\alpha(\tau), y_0), \quad \theta(\tau) = \theta_0 + \tau\omega.$$

Clearly, $\alpha(\tau)_\emptyset \equiv 1$, and thus $\alpha(\tau)$ will represent a curve in $\mathcal{G}$. Our first task is to rewrite the vector field itself as a B-series

$$\varepsilon \sum_{\mathbf{k}\in\mathbb{Z}^d} e^{i(\mathbf{k}\cdot\theta)} f_k(y) = B(\beta(\theta), y) = \sum_{u\in\mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \beta_u(\theta)\mathcal{F}_u(y) \tag{2.3}$$

with coefficients $\beta_u(\theta)$ defined for $u \in \mathcal{T} \cup \{\emptyset\}$ as follows:

$$\beta_u(\theta) = \begin{cases} e^{i(\mathbf{k}\cdot\theta)} & \text{if } u = \textcircled{\scriptsize k} \text{ for some } \mathbf{k} \in \mathbb{Z}^d, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Writing the initial value problem (1.1) in terms of B-series, we obtain

$$\begin{aligned} \frac{d}{d\tau} B(\alpha(\tau), y_0) &= B(\alpha(\tau) \star \beta(\theta(\tau)), y_0), \\ B(\alpha(0), y_0) &= y_0 = B(\mathbb{1}, y_0), \end{aligned}$$

and thus $\alpha : \mathbb{R} \to \mathcal{G}$ satisfies the initial value problem

$$\frac{d}{d\tau}\alpha(\tau) = \alpha(\tau) \star \beta(\theta(\tau)), \quad \alpha(0) = \mathbb{1}, \tag{2.5}$$

with $\theta(\tau) = \theta_0 + \tau\omega$. Since $\beta_u(\theta) = 0$ whenever $|u| \neq 1$, we obtain for each $u = [u_1 \cdots u_n]_{\mathbf{k}}$

$$\frac{d\alpha_u(\tau)}{d\tau} = \beta_{\textcircled{\scriptsize k}}(\theta)\alpha_{u_1}(\tau) \dots \alpha_{u_n}(\tau)$$

i.e., taking into account the definition of $\beta(\theta)$ and the initial condition $\alpha(0) = \mathbb{1}$ (i.e., $\alpha_u(0) = 0$ for all $u \in \mathcal{T}$), we eventually get

$$\alpha_u(\tau) = e^{i(\mathbf{k}\cdot\theta_0)} \int_0^\tau e^{is(\mathbf{k}\cdot\omega)} \alpha_{u_1}(s) \cdots \alpha_{u_n}(s) ds.$$

Clearly, the coefficients $\alpha_u(\tau)$ depend on both ω and θ_0. Whenever we want to stress the dependence of $\alpha(\tau)$ on θ_0, we will write $\alpha^{\theta_0}(\tau)$. We do not reflect the dependence on the vector of non-resonant frequencies ω, which is assumed to be fixed. A straightforward induction enables to show that for all $u \in \mathcal{T}$, one has

$$\alpha_u^{\theta_0}(\tau) = e^{i\mathcal{I}_u\cdot\theta_0} \alpha_u^0(\tau).$$

The coefficients for trees with less than four vertices are given in Figure 2.

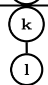
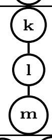
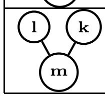
u	$\mathcal{F}_u(y)$	$\alpha_u(\tau)$	σ_u
	$f_{\mathbf{k}}(y)$	$\int_0^\tau e^{i(\mathbf{k}\cdot\omega)\tau_1}\, d\tau_1$	1
	$f'_{\mathbf{l}}(y) f_{\mathbf{k}}(y)$	$\int_0^\tau \int_0^{\tau_2} e^{i(\mathbf{k}\tau_1+\mathbf{l}\tau_2)\cdot\omega}\, d\tau_1\, d\tau_2$	1
	$f'_{\mathbf{m}}(y) f'_{\mathbf{l}}(y) f_{\mathbf{k}}(y)$	$\int_0^\tau \int_0^{\tau_3} \int_0^{\tau_2} e^{i(\mathbf{k}\tau_1+\mathbf{l}\tau_2+\mathbf{m}\tau_3)\cdot\omega}\, d\tau_1\, d\tau_2\, d\tau_3$	1
	$f''_{\mathbf{m}}(y)(f_{\mathbf{l}}(y), f_{\mathbf{k}}(y))$	$\int_0^\tau e^{i\tau_3(\mathbf{m}\cdot\omega)}\Big(\int_0^{\tau_3} e^{i\tau_1(\mathbf{k}\cdot\omega)}\, d\tau_1 \int_0^{\tau_3} e^{i\tau_2(\mathbf{l}\cdot\omega)}\, d\tau_2\Big) d\tau_3$	$1+\delta_{\mathbf{k},\mathbf{l}}$

Figure 2: Trees of orders less or equal to 3 with their associated elementary differentials and coefficients.

3 Extension of the ODE-solution to the associated transport PDE

Now, one can observe that for any $u \in \mathcal{T}$, $\alpha_u^0(\tau)$ is a combined algebraic-trigonometric Laurent polynomial of the form

$$\alpha_u^0(\tau) = P_u(\tau, e^{i\tau\omega_1}, \ldots, e^{i\tau\omega_d}, e^{-i\tau\omega_1}, \ldots, e^{-i\tau\omega_d})$$

where P_u is a Laurent polynomial of $\mathbb{C}[X, Z_1, \ldots, Z_d, Z_1^{-1}, \ldots, Z_d^{-1}]$ defined **uniquely** as soon as ω is non-resonant (i.e. such that for all $\in \mathbb{Z}^{\mathbf{d}}/\{\mathbf{0}\}$, $\mathbf{k}\cdot\omega \neq 0$). Let us define for each $u \in \mathcal{T}$,

$$\gamma_u(\tau, \theta) = P_u(\tau, e^{i\theta_1}, \ldots, e^{i\theta_d}, e^{-i\theta_1}, \ldots, e^{-i\theta_d}) \tag{3.1}$$

so that $\gamma(\tau, \tau\omega) = \alpha_u^0(\tau)$, and in particular, $\gamma(0, \mathbf{0}) = \alpha^0(\tau) = \mathbb{1}$. It should be emphasized here that $\gamma_u(\tau, \theta)$ depends polynomially in τ (i.e. with no negative powers).

3.1 The transport equation and its B-series solution

The fact that $\alpha^0(\tau) = \gamma(\tau, \tau\omega)$ is the solution of (2.5) implies that

$$\partial_\tau \gamma(\tau, \theta) + \omega \cdot \nabla_\theta \gamma(\tau, \theta) = \gamma(\tau, \theta) \star \beta(\theta) \tag{3.2}$$

provided that $\theta = \tau\omega$. Conversely, if there exists $\gamma : \mathbb{R} \times \mathbb{T}^d \to \mathcal{G}$ such that $\gamma(0, \mathbf{0}) = \mathbb{1}$ and (3.2) holds for all (τ, θ) in $\mathbb{R} \times \mathbb{T}^d$, then one has that $\alpha^0(\tau) = \gamma(\tau, \tau\omega)$. However, the solution of the transport equation (3.2) with initial condition $\gamma(0, \mathbf{0}) = \mathbb{1}$ is not unique[1]: if relation $\gamma(\tau, \tau\omega) = \alpha^0(\tau)$ holds true

[1] Generally speaking, solutions of (3.2) on $\mathbb{R} \times \mathbb{T}^d$ are recursively defined by

$$\gamma(\tau, \theta) = \chi(\theta - \tau\omega) + \int_0^\tau \gamma(s, \theta + (s-\tau)\omega) \star \beta(\theta + (s-\tau)\omega) ds$$

for an arbitrary initial value function χ of $\theta \in \mathbb{T}^d$ satisfying $\chi(0) = \mathbb{1}$.

for a given function γ, then for any real λ, it also holds for the function

$$(\tau, \theta) \mapsto \gamma\Big(\tau, \lambda\theta + \tau(1-\lambda)\omega\Big),$$

which is thus solution of (3.2) with, for $\tau = 0$, $\gamma\Big(\tau, \lambda\theta + \tau(1-\lambda)\omega\Big) = \gamma(0, \mathbf{0}) = \mathbb{1}$. Uniqueness is nonetheless recovered as soon as we impose that γ depends polynomially in τ, as this is the case in (3.1).

Definition 3.1 *We will say that a smooth function $w : \mathbb{R} \times \mathbb{T}^d \to \mathbb{C}$ is* polynomial *if it depends polynomially on τ. A function $\gamma : \mathbb{R} \times \mathbb{T}^d \to \mathcal{G}$ is* polynomial *if γ_u is polynomial for all u's in $\mathcal{T}$.*

Lemma 3.2 *There exists a unique polynomial solution $\gamma : \mathbb{R} \times \mathbb{T}^d \to \mathcal{G}$ of (3.2) such that $\gamma(0, \mathbf{0}) = \mathbb{1}$. Furthermore, for each $u \in \mathcal{T}$, $\gamma_u(\tau, \theta)$ has degree at most $|u|$ in τ, and it is a trigonometric Laurent polynomial in each component of θ.*

Proof: We proceed by induction and assume that for all u's in $\mathcal{T} \cup \{\emptyset\}$ with degrees less or equal to $p-1$, $p \geq 1$, γ_u is of the form

$$\gamma_u(\tau, \theta) = \sum_{\mathbf{l} \in S_u} \hat{\gamma}_{u,\mathbf{l}}(\tau) e^{i\mathbf{l}\cdot\theta} \text{ with } \sum_{\mathbf{l} \in S_u} \hat{\gamma}_{u,\mathbf{l}}(0) = \delta_{\emptyset,u},$$

where $\hat{\gamma}_{u,\mathbf{l}}(\tau)$ is a polynomial of degree less or equal to $|u|$ and S_u a finite subset of $\mathbb{Z}^d$. Consider now a tree $u = [u_1, \ldots, u_n]$ of degree p: γ_u is solution of (3.2), i.e. of an equation of the form

$$\sum_{\mathbf{l} \in S_u} \left(\frac{d}{d\tau}\hat{\gamma}_{u,\mathbf{l}}(\tau) + i(\mathbf{l}\cdot\omega)\hat{\gamma}_{u,\mathbf{l}}(\tau)\right) e^{i\mathbf{l}\cdot\theta} = \sum_{\mathbf{l} \in S_u^{\Pi}} \gamma_{u,\mathbf{l}}^{\Pi}(\tau) e^{i\mathbf{l}\cdot\theta}$$

where $S_u^{\Pi} := \{\mathbf{l}_1 + \ldots + \mathbf{l}_n : \mathbf{l}_1 \in S_{u_1}, \ldots, \mathbf{l}_n \in S_{u_n}\}$ and where the $\gamma_{u,\mathbf{l}}^{\Pi}(\tau)$'s are polynomials of degree less or equal to $p-1$. This implies that $S_u = S_u^{\Pi} \cup \{\mathbf{0}\}$ and the equation is then decomposed into the system

$$\forall \mathbf{l} \in S_u, \quad \frac{d}{d\tau}\hat{\gamma}_{u,\mathbf{l}}(\tau) + i(\mathbf{l}\cdot\omega)\hat{\gamma}_{u,\mathbf{l}}(\tau) = \gamma_{u,\mathbf{l}}^{\Pi}(\tau).$$

For $\mathbf{l} \neq 0$, this determines a unique polynomial solution, which is of degree less or equal to $p-1$. For $\mathbf{l} = \mathbf{0}$, the solution $\hat{\gamma}_{u,\mathbf{0}}(\tau)$ is determined up to a constant C which is then obtained in a unique way from $\sum_{\mathbf{l} \in S_u} \hat{\gamma}_{u,\mathbf{l}}(0) = 0$. It is again a polynomial, of degree less or equal to p. $\square$

Remark 3.3 *Lemma 3.2 also holds if $\beta_u(\theta)$ is for each $u \in \mathcal{T}$ an arbitrary trigonometric polynomial in $\theta = (\theta_1, \ldots, \theta_d)$ (not necessarily defined as in (2.4)) and $\beta_\emptyset(\theta) \equiv 0$. Even more generally and provided that the vector of frequencies ω satisfies the Diophantine conditions (1.3), the result carries over to the case where $\beta_u : \mathbb{T}^d \to \mathbb{C}$ is an arbitrary analytic function for each $u \in \mathcal{T}$.*

From now on, we will denote as $\gamma^0(\tau, \theta)$ the unique polynomial solution of (3.2) with $\gamma(0, \mathbf{0}) = \mathbb{1}$. Consider now equation (3.2) with $\beta(\theta)$ replaced by $\beta(\theta + \theta_0)$, and denote as $\gamma^{\theta_0}(\tau, \theta)$ its solution with initial value $\gamma^{\theta_0}(0, \mathbf{0}) = \mathbb{1}$. The linearity of $\star$ with respect to the right factor and Lemma 3.2 imply the following Corollary.

Corollary 3.4 *For any prescribed $\chi \in \mathcal{G}$, equation (3.2) complemented with the initial condition $\gamma(\tau_0, \theta_0) = \chi$ has a unique polynomial solution given by*

$$\gamma(\tau, \theta) = \chi \star \gamma^{\theta_0}(\tau - \tau_0, \theta - \theta_0).$$

In the particular case with $\chi = \mathbb{1}$, we get that $\gamma(\tau, \theta) = \gamma^{\theta_0}(\tau - \tau_0, \theta - \theta_0)$ is the unique polynomial solution of (3.2) with $\gamma(\tau_0, \theta_0) = \mathbb{1}$.

As $\beta(\theta)$ is given by (2.4), replacing $\beta(\theta)$ by $\beta(\theta + \theta_0)$ in (2.3) is equivalent to considering (2.3) with $f_k(y)$ replaced by $e^{i(\mathbf{k}\cdot\theta_0)} f_k(y)$, and hence $\mathcal{F}_u(y)$ replaced by $e^{i(\mathcal{I}_u\cdot\theta_0)}\mathcal{F}_u(y)$. This shows that

$$\forall u \in \mathcal{T}, \quad \gamma_u^{\theta_0}(\tau, \theta) = e^{i\mathcal{I}_u\cdot\theta_0}\gamma_u^0(\tau, \theta),$$

or more briefly, $\gamma^{\theta_0}(\tau, \theta) = e^{i\mathcal{I}\cdot\theta_0}\gamma^0(\tau, \theta)$.

3.2 Extended flow-maps and their composition

System (1.1) being **autonomous** in the phase-space $(y, \theta) \in \mathbb{R}^n \times \mathbb{T}^d$, its flow map obeys group laws. Hence, denoting

$$\begin{array}{rcl} \Phi_\tau : \mathbb{R}^n \times \mathbb{T}^d & \to & \mathbb{R}^n \times \mathbb{T}^d \\ (y, \theta) & \mapsto & \big(B(\alpha^\theta(\tau), y), \theta + \tau\omega\big) \end{array}$$

one has for all $(\tau_1, \tau_2) \in \mathbb{R}^2$, $\Phi_{\tau_1} \circ \Phi_{\tau_2} = \Phi_{\tau_1+\tau_2}$, where

$$\begin{array}{rcl} \Phi_{\tau_1} \circ \Phi_{\tau_2} : \mathbb{R}^n \times \mathbb{T}^d & \to & \mathbb{R}^n \times \mathbb{T}^d \\ (y, \theta) & \mapsto & \big(B(\alpha^\theta(\tau_1) \star \alpha^{\theta+\tau_1\omega}(\tau_2), y), \theta + (\tau_1 + \tau_2)\omega\big) . \end{array}$$

This means that the following relation

$$\alpha^\theta(\tau_1 + \tau_2) = \alpha^\theta(\tau_1) \star \alpha^{\theta+\tau_1\omega}(\tau_2).$$

holds true for all $(\tau_1, \tau_2) \in \mathbb{R}^2$, or equivalently that

$$\gamma^\theta(\tau_1 + \tau_2, (\tau_1 + \tau_2)\omega) = \gamma^\theta(\tau_1, \tau_1\omega) \star \gamma^{\theta+\tau_1\omega}(\tau_2, \tau_2\omega).$$

Actually, such a relation is true in a wider sense.

Theorem 3.5 *For all $\tau_1, \tau_2 \in \mathbb{R}$ and all $\theta_0, \theta_1, \theta_2 \in \mathbb{T}^d$,*

$$\gamma^{\theta_0}(\tau_1 + \tau_2, \theta_1 + \theta_2) = \gamma^{\theta_0}(\tau_1, \theta_1) \star \gamma^{\theta_0+\theta_1}(\tau_2, \theta_2).$$

Proof: Consider $\gamma : \mathbb{R} \times \mathbb{T}^d \to \mathcal{G}$ defined as

$$\gamma(\tau, \theta) = \gamma^{\theta_0}(\tau, \theta - \theta_0). \tag{3.3}$$

Clearly, $\gamma(\tau, \theta)$ is a polynomial solution of (3.2), which trivially satisfies the initial condition $\gamma(\tau_1, \theta_0 + \theta_1) = \gamma^{\theta_0}(\tau_1, \theta_1)$. By virtue of Corollary 3.4 (with $\chi := \gamma^{\theta_0}(\tau_1, \theta_1)$ and $(\tau_0, \theta_0) := (\tau_1, \theta_0 + \theta_1)$),

$$\gamma(\tau, \theta) = \gamma^{\theta_0}(\tau_1, \theta_1) \star \gamma^{\theta_0+\theta_1}(\tau - \tau_1, \theta - \theta_0 - \theta_1). \tag{3.4}$$

so that the statement follows from (3.3) and (3.4) with $(\tau, \theta) = (\tau_1 + \tau_2, \theta_0 + \theta_1 + \theta_2)$. □

Example 3.6 *For instance, consider the tree* $u = \begin{smallmatrix} (\mathbf{l}) \\ | \\ (\mathbf{k}) \end{smallmatrix}$ *with* $\mathbf{k} = -\mathbf{l}$, *for which we have*

$$\alpha_u^0(\tau) = \int_0^\tau e^{i(\mathbf{l}\cdot\omega)s} \frac{e^{-i(\mathbf{l}\cdot\omega)s} - 1}{-i(\mathbf{l}\cdot\omega)} ds = \int_0^\tau \frac{1 - e^{i(\mathbf{l}\cdot\omega)s}}{-i(\mathbf{l}\cdot\omega)} ds = \frac{i\tau}{\mathbf{l}\cdot\omega} + \frac{e^{i(\mathbf{l}\cdot\omega)\tau} - 1}{(i(\mathbf{l}\cdot\omega))^2} = \frac{i\tau}{\mathbf{l}\cdot\omega} + \frac{1 - e^{i(\mathbf{l}\cdot\omega)\tau}}{(\mathbf{l}\cdot\omega)^2}$$

so that $\gamma_u^0(\tau,\theta) = \frac{i\tau}{\mathbf{l}\cdot\omega} + \frac{1-e^{i\mathbf{l}\cdot\theta}}{(\mathbf{l}\cdot\omega)^2}$. *Theorem 3.5 gives*

$$\begin{aligned} \gamma_u^0(\tau_1+\tau_2, \theta_1+\theta_2) &= \gamma_u^0(\tau_1,\theta_1) + \gamma_u^{\theta_1}(\tau_2,\theta_2) + \gamma^0{}_{(\mathbf{k})}(\tau_1,\theta_1)\gamma^{\theta_1}{}_{(\mathbf{l})}(\tau_2,\theta_2) \\ &= \gamma_u^0(\tau_1,\theta_1) + \gamma_u^0(\tau_2,\theta_2) + \gamma^0{}_{(\mathbf{k})}(\tau_1,\theta_1)e^{il\cdot\theta_1}\gamma^0{}_{(\mathbf{l})}(\tau_2,\theta_2). \end{aligned}$$

Here, we have used that

$$\gamma_u^{\theta_1}(\tau_2,\theta_2) = e^{i(\mathbf{l}+\mathbf{k})\cdot\theta_1}\gamma_u^0(\tau_2,\theta_2) = \gamma_u^0(\tau_2,\theta_2).$$

4 Perko's Theorem revisited

In this section, we will show how Theorem 3.5 may be used to exhibit very easily a modified vector field and a change of variables in the spirit of Perko's Theorem 1.1. In particular, we will see that the (formal) vector field $F(Y) := \varepsilon F_1(Y) + \varepsilon^2 F_2(Y) + \ldots$, the (formal) change of variable $U(Y,\theta) = Y + \varepsilon u_1(Y,\theta) + \varepsilon^2 u_2(Y,\theta) + \cdots$ of Theorem 1.1, and the (formal) solution $Y(\tau)$ of $Y' = F(Y)$ can all be expanded in B-series

$$F(Y) = B(\bar\beta, Y), \quad U(Y,\theta) = B(\kappa(\theta), Y), \quad Y(\tau) = B(\bar\alpha(\tau), Y(0)),$$

where $\bar\beta \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$ with $\bar\beta_\emptyset = 0$, $\kappa : \mathbb{T}^d \to \mathcal{G}$, $\bar\alpha : \mathbb{R} \to \mathcal{G}$. Recall that the solution $y(\tau)$ of (1.1) can be written as $y(\tau) = B(\alpha(\tau), y_0)$, where $\alpha : \mathbb{R} \to \mathcal{G}$ is the solution of (2.5). The formal part of the statement of Theorem 1.1 can then be re-interpreted as follows: $y(\tau) = B(\kappa(\theta(\tau)), Y(\tau))$, where $\theta(\tau) = \theta_0 + \tau\omega$, and $Y(\tau) = B(\bar\alpha(\tau), Y_0)$ is the solution of

$$\frac{d}{d\tau}B(\bar\alpha(\tau), Y_0) = B(\bar\beta, B(\bar\alpha(\tau), Y_0)), \quad B(\bar\alpha(0), Y_0) = Y_0, \tag{4.1}$$

where $Y_0 = B(\kappa(\theta_0)^{-1}, y_0)$, and

$$\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} B(\kappa(\theta), Y) d\theta = Y = B(\mathbb{1}, Y).$$

In terms of $\alpha, \bar\alpha, \bar\beta, \kappa$, this means that

$$\alpha(\tau) = \bar\alpha(\tau) \star \kappa(\theta_0 + \tau\omega), \quad \frac{d}{d\tau}\bar\alpha(\tau) = \bar\alpha(\tau) \star \bar\beta, \tag{4.2}$$

where $\bar\alpha(0) = \kappa(\theta_0)^{-1}$, and $\kappa(\theta)$ is such that

$$\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} \kappa(\theta) d\theta = \mathbb{1}. \tag{4.3}$$

In the sequel, we give an algebraic proof of the existence of $\bar\beta, \kappa(\theta), \bar\alpha(\tau)$ satisfying (4.2) such that (4.3), and extend it by allowing different choices for determining κ other than (4.3).

$$
\begin{array}{ccc}
\left\{\begin{array}{ll} \frac{dy}{d\tau} = \varepsilon f(y,\theta), & y(0) = y_0 \\ \frac{d\theta}{d\tau} = \omega, & \theta(0) = \theta_0 \end{array}\right.
& \xleftrightarrow[\varepsilon f(y,\theta) = B(\beta(\theta), y)]{y(\tau) = B(\alpha(\tau), y_0)} &
\left\{\begin{array}{ll} \frac{d\alpha}{d\tau} = \alpha \star \beta(\theta), & \alpha(0) = \mathbb{1} \\ \frac{d\theta}{d\tau} = \omega, & \theta(0) = \theta_0 \end{array}\right. \\
y = U(Y,\theta) \quad \updownarrow & & \updownarrow \quad \alpha = \bar{\alpha} \star \kappa(\theta) \\
\left\{\begin{array}{ll} \frac{dY}{d\tau} = \varepsilon F(Y), & Y(0) = Y_0 \\ \frac{d\theta}{d\tau} = \omega, & \theta(0) = \theta_0 \end{array}\right.
& \xleftrightarrow[\varepsilon F(Y) = B(\bar{\beta}, Y)]{Y(\tau) = B(\bar{\alpha}(\tau), Y_0)} &
\left\{\begin{array}{ll} \frac{d\bar{\alpha}}{d\tau} = \bar{\alpha} \star \bar{\beta}, & \bar{\alpha}(0) = \kappa^{-1}(\theta_0) \\ \frac{d\theta}{d\tau} = \omega, & \theta(0) = \theta_0 \end{array}\right.
\end{array}
$$

Figure 3: Perko's Theorem in terms of B-series

4.1 Quasi-stroboscopic averaging

Although the initial value $Y(0)$ in Theorem 1.1 does not coincide with $y(0)$, we now show that it is in fact possible to choose κ in such a way that $\kappa(\mathbf{0}) = \mathbb{1}$, and thus $Y(0) = y(0)$. In this sense, we extend stroboscopic averaging to the multi-frequency case, and call this generalization *quasi-stroboscopic averaging*. Note however, that recovering the exact solution of the original system (1.1) from the solution of the averaged differential equation is not as straightforward as for stroboscopic averaging, since nothing such as a period exists here.

Consider the solution $\alpha(\tau)$ of (2.5) with $\theta(\tau) = \theta_0 + \tau\omega$, so that $\alpha(\tau) = \gamma^{\theta_0}(\tau, \tau\omega)$. Application of Theorem 3.5 with $\tau_1 = \tau$, $\tau_2 = 0$, $\theta_1 = 0$, and $\theta_2 = \Delta\theta$ gives

$$\gamma^{\theta_0}(\tau, \Delta\theta) = \gamma^{\theta_0}(\tau, 0) \star \gamma^{\theta_0}(0, \Delta\theta). \tag{4.4}$$

Let us denote $\bar{\alpha}(\tau) = \gamma^{\theta_0}(\tau, 0)$, and $\kappa(\theta) = \gamma^{\theta_0}(0, \theta - \theta_0)$, so that (4.4) with $\Delta\theta = \tau\omega$ gives the factorization

$$\gamma^{\theta_0}(\tau, \Delta\theta) = \bar{\alpha}(\tau) \star \kappa(\theta_0 + \Delta\theta), \tag{4.5}$$

hence, $\alpha(\tau) = \gamma^{\theta_0}(\tau, \tau\omega) = \bar{\alpha}(\tau) \star \kappa(\theta_0 + t\omega)$.

By considering $\theta_1 = \theta_2 = 0$ in Theorem 3.5, we get

$$\gamma^{\theta_0}(\tau_1 + \tau_2, 0) = \gamma^{\theta_0}(\tau_1, 0) \star \gamma^{\theta_0}(\tau_2, 0),$$

which leads, after differentiating with respect to τ_2 in both sides and substitution $\tau_1 = \tau$, $\tau_2 = 0$, to

$$\frac{\partial}{\partial \tau}\gamma^{\theta_0}(\tau, 0) = \gamma^{\theta_0}(\tau, 0) \star \frac{\partial}{\partial \tau}\gamma^{\theta_0}(0, \mathbf{0}), \tag{4.6}$$

and thus second equality in (4.2) holds true with

$$\bar{\beta} = \frac{\partial}{\partial \tau}\gamma^{\theta_0}(0, \mathbf{0}).$$

In addition, we have that $\bar{\alpha}(0) = \gamma^{\theta_0}(0, 0) = \mathbb{1} = \alpha(0)$, and we obtain precisely the condition required for what we have called quasi-stroboscopic averaging.

4.2 General averaging

We next show that classical averaging can be effectively described using B-series. That is, that there exist $\kappa : \mathbb{T}^d \to \mathcal{G}$, $\bar{\alpha} : \mathbb{R} \to \mathcal{G}$, and $\bar{\beta} \in \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$ with $\bar{\beta}_\emptyset = 0$, such that (4.2) holds with (4.3). Indeed, starting from the fundamental decomposition (4.4), we introduce an arbitrary $\nu \in \mathcal{G}$, so that

$$\gamma^{\theta_0}(\tau, \Delta\theta) = \gamma^{\theta_0}(\tau, \mathbf{0}) \star \nu^{-1} \star \nu \star \gamma^{\theta_0}(0, \Delta\theta),$$

and then define $\kappa(\theta) = \nu \star \gamma(0, \theta - \theta_0)$ and $\bar{\alpha}(\tau) = \gamma^{\theta_0}(\tau, \mathbf{0}) \star \nu^{-1}$, so that we effectively have the factorization (4.5). We now determine $\nu = \kappa(\theta_0) = \bar{\alpha}(0)^{-1}$ by imposing (4.3). By applying angular averages in both sides of (4.5), and taking into account the linearity of $\star$ with respect to the right factor, (4.3) implies that

$$\bar{\alpha}(\tau) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \gamma^{\theta_0}(\tau, \theta) d\theta,$$

and hence

$$\nu^{-1} = \bar{\alpha}(0) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \gamma(0, \theta) d\theta.$$

From equation (4.6) we obtain

$$\frac{\partial \gamma}{\partial \tau}(\tau, \mathbf{0}) \star \nu^{-1} = \gamma(\tau, \mathbf{0}) \star \nu^{-1} \star \nu \star \frac{\partial \gamma}{\partial \tau}(0, \mathbf{0}) \star \nu^{-1}$$

which gives the second equality in (4.2) for $\bar{\beta} = \nu \star \frac{\partial \gamma}{\partial \tau}(0, \mathbf{0}) \star \nu^{-1}$. Observe that, we also have that

$$\bar{\beta} = \frac{d}{d\tau} \bar{\alpha}(\tau) \Big|_{\tau=0} = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \frac{\partial}{\partial \tau} \gamma^{\theta_0}(0, \theta) d\theta.$$

4.3 Equivalence of averaging procedures

From previous subsections, it is clear that there exists a whole class of possible averaging procedures depending on the specific choice of $\nu = \kappa(\theta_0)$, i.e., on the choice of the change of variable at $\theta(0) = \theta_0$

$$U_0(Y) = U(Y, \theta_0) = B(\nu, Y).$$

All of such averaging procedures are conjugate to each other, that is, they give rise to a formal change of variable $y = U(Y, \theta) = B(\kappa(\theta), Y)$, an averaged vector field $F(Y) = B(\bar{\beta}, Y)$, and an expansion of the averaged solution $Y(\tau) = B(\bar{\alpha}(\tau), y(0))$, whose B-series coefficients are related to the B-series coefficients of stroboscopic averaging as follows:

$$\kappa(\theta) = \nu \star \gamma(0, \theta - \theta_0), \quad \bar{\alpha}(\tau) = \gamma^{\theta_0}(\tau, \mathbf{0}) \star \nu^{-1}, \quad \bar{\beta} = \nu \star \frac{\partial \gamma}{\partial \tau}(0, \mathbf{0}) \star \nu^{-1}.$$

This is equivalent to the relations

$$U(Y, \theta) = B(\gamma^{\theta_0}(0, \theta - \theta_0), U_0(Y)), \tag{4.7}$$

$$Y(\tau) = U_0^{-1}(B(\gamma^{\theta_0}(\tau, \mathbf{0}), y(0))), \tag{4.8}$$

$$F(Y) = \frac{\partial}{\partial Y} U_0^{-1}(U_0(Y)) B(\tfrac{\partial}{\partial \tau} \gamma^{\theta_0}(0, \mathbf{0}), U_0(Y)). \tag{4.9}$$

Theorem 4.1 *Consider the differential system (1.1), and an averaged vector field $F(Y)$ and a change of variable $y = U(Y, \theta)$, 2π-periodic in each angle θ_i, not necessarily expanded as B-series, such that*

$$y(\tau) = U(Y(\tau), \theta(\tau)), \quad \theta(\tau) = \theta_0 + \tau\omega,$$

where

$$\frac{dY}{d\tau} = F(Y) \text{ with } U(Y(0), \theta_0) = y_0.$$

Then, the relations (4.7)–(4.9) hold true, where $U_0(Y) := U(Y, \theta_0)$.

5 An Arnold-Liouville-like theorem for quasi-periodic systems and its consequences

The emphasis in this section is put, in accordance with the remaining of this paper, on formal transformations (expanded in B-series) and formal results and not on the equally important aspects of obtaining error estimates.

Theorem 5.1 *Assume that ω is a non-resonant vector of $\mathbb{R}^d$ and consider the associated coefficients γ. Consider now $\Delta_1\theta$, ..., $\Delta_m\theta$, m vectors of $\mathbb{T}^d$ and define, for each $\Delta_i\theta$, $i = 1, \ldots, m$, the following flow:*

$$\begin{array}{rcl} \Psi_\tau^{\Delta_i\theta} : \mathbb{R}^n \times \mathbb{T}^d & \to & \mathbb{R}^n \times \mathbb{T}^d \\ (y, \theta) & \mapsto & \big(B(\gamma^\theta(\tau, \Delta_i\theta), y), \theta + \Delta_i\theta\big) \end{array}$$

Then, the following relation holds true for all $(\tau_1, \ldots, \tau_m) \in \mathbb{R}^m$:

$$\Psi_{\tau_1}^{\Delta_1\theta} \circ \cdots \circ \Psi_{\tau_m}^{\Delta_m\theta} = \Psi_{\tau_1+\ldots+\tau_m}^{\Delta_1\theta+\ldots+\Delta_m\theta}.$$

In particular, any two such flows commute.

Proof: This is a straightforward consequence of Theorem 3.5. □

Note that for $\Delta\theta = \tau\tilde{\omega}$, $\Psi_\tau^{\Delta\theta}$ is the flow map of the autonomous differential equation

$$\left\{ \begin{array}{rcl} y' & = & \varepsilon f^{\tilde{\omega}}(y, \theta), \\ \theta' & = & \tilde{\omega}, \end{array} \right. \tag{5.1}$$

where

$$f^{\tilde{\omega}}(y, \theta) = \sum_{u \in \mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \beta^{\tilde{\omega}} e^{\mathcal{I}_u \cdot \theta} \mathcal{F}_u(y),$$

where, as before, $\beta^{\tilde{\omega}} = \left.\frac{d\gamma_u(\tau, \tau\tilde{\omega})}{d\tau}\right|_{\tau=0}$. The first coefficients of the expansion are given in Table 1.

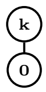
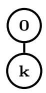
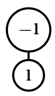
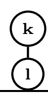
Tree	μ	Index range	Order	σ_u	$\beta_u^{\tilde{\omega}}$
	I	$\emptyset$	1	1	1
	II	$\mathbf{k} \neq \mathbf{0}$	1	1	$\frac{\mathbf{k}\cdot\tilde{\omega}}{\mathbf{k}\cdot\omega}$
	III	$\mathbf{k} \neq \mathbf{0}$	2	1	$i\frac{\mathbf{k}\cdot(\omega-\tilde{\omega})}{(\mathbf{k}\cdot\omega)^2}$
	IV	$\mathbf{k} \neq \mathbf{0}$	2	1	$i\frac{\mathbf{k}\cdot(\tilde{\omega}-\omega)}{(\mathbf{k}\cdot\omega)^2}$
	V	$\mathbf{l} \neq \mathbf{0}$	2	1	$i\frac{\mathbf{l}\cdot(\omega-\tilde{\omega})}{(\mathbf{l}\cdot\omega)^2}$
	VI	$\mathbf{k} \neq \mathbf{0}, \mathbf{l} \neq \mathbf{0}, \mathbf{k} \neq -\mathbf{l}$	2	1	$i\frac{(\mathbf{l}\cdot\tilde{\omega})(\mathbf{k}\cdot\omega)-(\mathbf{l}\cdot\omega)(\mathbf{k}\cdot\tilde{\omega})}{(\mathbf{l}\cdot\omega)(\mathbf{k}\cdot\omega)((\mathbf{l}+\mathbf{k})\cdot\omega)}$

Table 1: Coefficients of the $\tilde{\omega}$-averaged solution for first- and second-order trees

Example 5.2 *Consider the following vector field:*

$$f(y,\theta) = \cos^2(\omega_1\tau + \theta_1) + \cos(\omega_2\tau + \theta_2)\, y.$$

Denoting $\Delta\theta = (\Delta_1, \Delta_2)$ and again $\theta = (\theta_1, \theta_2)$, we have

$$\Psi_\tau^{(\Delta_1,\Delta_2)}\Big(y,\theta\Big) = y + \frac{\varepsilon\tau}{2} - \frac{\varepsilon}{4}\,\frac{(\sin(2\theta_1)\,\omega_2 + 4\,y\,\sin(\theta_2)\,\omega_1 - \omega_2\sin(2\Delta_1 + 2\theta_1) - 4\,y\,\sin(\Delta_2 + \theta_2)\,\omega_1)}{\omega_1\omega_2}$$

$$\Psi_{x\tau}^{(\Delta_1,0)}(y,(0,0)) = \left(y + \frac{\varepsilon}{2}\,\frac{(\cos(\Delta_1)\sin(\Delta_1) + x\tau\,\omega_1)}{\omega_1},\ 0+\Delta_1,\ 0+0\right)$$

$$\Psi_{(1-x)\tau}^{(0,\Delta_2)}(y,(\Delta_1,0)) = \left(y + \frac{\varepsilon}{2}\frac{(\omega_2\tau(1-x) + 2\,y\,\sin(\Delta_2))}{\omega_2},\ \Delta_1 + 0,\ 0+\Delta_2\right)$$

$$\Psi_\tau^{(\Delta_1,\Delta_2)}\Big(y,(0,0)\Big) = \left(y + \frac{\varepsilon}{2}\frac{(\omega_2\cos(\Delta_1)\sin(\Delta_1) + \tau\,\omega_1\omega_2 + 2\,y\,\sin(\Delta_2)\,\omega_1)}{\omega_1\omega_2},\ \Delta_1,\ \Delta_2\right)$$

It is interesting to notice that the case where the angle is also perturbed as follows

$$\left\{ \begin{array}{rcl} x' & = & \varepsilon k(x,\xi) \in \mathbb{R}^n \\ \xi' & = & \omega + \varepsilon g(x,\xi) \in \mathbb{T}^d \end{array} \right.$$

can simply be rewritten in the form (1.1) with $y = (y_1, y_2)$ and $f(y,\theta) = \Big(k(y_1, y_2+\theta), g(y_1, y_2+\theta)\Big)$ so that perturbed integrable systems are part of our analysis. We explore this connection a bit further in next section.

5.1 Averaging for a class of near-integrable systems

Consider a system of the form

$$z' = k(z) + \varepsilon g(z) \in \mathbb{R}^n$$

and assume that the flow Φ_τ of $z' = k(z)$ is of the form Ψ_θ with $\theta = \tau\omega$ and $\omega \in \mathbb{T}^d$, where Ψ_θ satisfies

$$\forall (\theta_1, \theta_2) \in \mathbb{T}^d \times \mathbb{T}^d, \Psi_{\theta_1+\theta_2} = \Psi_{\theta_1} \circ \Psi_{\theta_2}. \tag{5.2}$$

Writing $z(\tau) = \Psi_{\tau\omega}(y(\tau))$ we get

$$\begin{array}{rcl} z' & = & \dfrac{d\Psi_{\tau\omega}}{d\tau}(y) + \dfrac{\partial \Psi_{\tau\omega}}{\partial y}(y)\, y' \\ & = & k(\Psi_{\tau\omega}(y)) + \dfrac{\partial \Psi_{\tau\omega}}{\partial y}(y)\, y' \end{array}$$

so that

$$\left\{ \begin{array}{rcl} y' & = & \varepsilon \Big(\frac{\partial \Psi_{-\theta}}{\partial y}(\Psi_\theta(y))\Big) g(\Psi_\theta(y)) \\ \theta' & = & \omega \end{array} \right.$$

where we have used relation (5.2) with $\theta_1 = -\theta_2 = \theta$. The system for (y,θ) is thus exactly of the form considered in this paper with $f(y,\theta) = \Big(\frac{\partial \Psi_{-\theta}}{\partial y}(\Psi_\theta(y))\Big) g(\Psi_\theta(y))$ and we can compute its B-series expansion in terms of the Fourier coefficients of f, namely

$$f_{\mathbf{k}}(y) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} e^{-i(\mathbf{k}\cdot\theta)} \Big(\frac{\partial \Psi_{-\theta}}{\partial y}(\Psi_\theta(y))\Big) g(\Psi_\theta(y)) d\theta.$$

The expansion is of the form derived in previous sections, that is to say

$$\left\{ \begin{array}{rcl} y(\tau) & = & y + \sum_{u\in\mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \gamma^\theta(\tau, \tau\omega) \mathcal{F}_u(y) \\ \theta(\tau) & = & \theta + \tau\omega \end{array} \right.$$

and the corresponding quasi-stroboscopic averaged differential equations are of the form

$$\left\{ \begin{array}{rcl} \frac{dy}{d\tau} & = & \sum_{u\in\mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \left. \frac{d\gamma_u(\tau,\tau\tilde\omega)}{d\tau}\right|_{\tau=0} e^{\mathcal{I}_u\cdot\theta}\mathcal{F}_u(y) \\ \theta' & = & \tilde\omega \end{array} \right.$$

for arbitrary $\tilde\omega \in \mathbb{R}^d$. In order to get an equation for $z = \Psi_{\tau\tilde\omega}(y)$ we first have to compute

$$\begin{array}{rcl} f_{\mathbf{k}}\left(\Psi_{-\theta_2}(z)\right) & = & \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} e^{-i(\mathbf{k}\cdot\theta_1)}\Big(\frac{\partial \Psi_{-\theta_1}}{\partial y}(\Psi_{\theta_1}(\Psi_{-\theta_2}(z)))\Big) g(\Psi_{\theta_1}(\Psi_{-\theta_2}(z)))d\theta_1 \\ & = & \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} e^{-i(\mathbf{k}\cdot\theta_1)}\Big(\frac{\partial \Psi_{-\theta_1}}{\partial y}(\Psi_{\theta_1-\theta_2}(z))\Big) g(\Psi_{\theta_1-\theta_2}(z))d\theta_1 \\ & = & \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} e^{-i(\mathbf{k}\cdot(\chi+\theta_2))}\Big(\frac{\partial \Psi_{-(\chi+\theta_2)}}{\partial y}(\Psi_{\chi}(z))\Big) g(\Psi_{\chi}(z))d\chi \end{array}$$

where we have used relation (5.2). Now, using once again (5.2), we have

$$\Psi_{-(\chi+\theta_2)} \circ \Psi_\chi = \Psi_{-\theta_2}$$

so that, by differentation

$$\Big(\frac{\partial \Psi_{-(\chi+\theta_2)}}{\partial y}(\Psi_\chi(z))\Big)\cdot\Big(\frac{\partial \Psi_\chi}{\partial y}(z)\Big) = \frac{\partial \Psi_{-\theta_2}}{\partial y}(z).$$

Eventually, we get

$$f_{\mathbf{k}}\left(\Psi_{-\theta_2}(z)\right) = e^{-i\mathbf{k}\cdot\theta_2}\frac{\partial \Psi_{-\theta_2}}{\partial y}(z) f_{\mathbf{k}}(z). \tag{5.3}$$

From now on, we assume in addition that Ψ_θ is linear, so that its second derivative w.r.t. y vanishes. Then, relation (5.3) leads to

$$f'_{\mathbf{k}}\left(\Psi_{-\theta_2}(z)\right) = e^{-i\mathbf{k}\cdot\theta_2}\frac{\partial \Psi_{-\theta_2}}{\partial y}(z) f'_{\mathbf{k}}(z)\Big(\frac{\partial \Psi_{-\theta_2}}{\partial y}(z)\Big)^{-1}$$

and more generally to

$$\mathcal{F}_u\left(\Psi_{-\theta_2}(z)\right) = e^{-i\mathcal{I}_u\cdot\theta_2}\frac{\partial \Psi_{-\theta_2}}{\partial y}(z)\mathcal{F}_u(z)$$

Getting back to the differential equation for z, we obtain

$$\left\{ \begin{array}{rcl} z' & = & k^{\tilde\omega}(z) + \sum_{u\in\mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u}\beta^{\tilde\omega}\mathcal{F}_u(z) \\ \theta' & = & \tilde\omega \end{array} \right.$$

where $k^{\tilde\omega}(z) = \frac{d\Psi_{\tau\tilde\omega}}{d\tau}\left(\Psi_{-\tau\tilde\omega}(z)\right)$ is independent of τ.

5.2 Application to the case of the FPU problem

We consider the Hamiltonian system with Hamiltonian

$$\frac{1}{2}p_1^T p_1 + \frac{1}{2}p_2^T p_2 + \frac{1}{2\varepsilon^2}q_2^T\Omega^2 q_2 + U(q_1,q_2;\varepsilon),$$

where U is a real-valued potential that may be expanded in non-negative powers of ε and Ω is a $d \times d$ symmetric positive definite matrix of the form $\Omega = Q^T\mathrm{Diag}(\omega_1,\ldots,\omega_d)Q$ with $Q^TQ = I$. When $\tau = \varepsilon^{-1}t$ is used as independent variable, the Hamiltonian function becomes

$$H(p_1,p_2,q_1,q_2) \;=\; H_1(p_1,p_2,q_1,q_2) + H_2(p_1,p_2,q_1,q_2),$$

with

$$H_1(p_1,p_2,q_1,q_2) := \frac{1}{2}\left(\varepsilon p_2^T p_2 + \frac{1}{\varepsilon}q_2^T\Omega^2 q_2\right) \text{ and } H_2(p_1,p_2,q_1,q_2) := \frac{\varepsilon}{2}p_1^T p_1 + \varepsilon U(q_1,q_2;\varepsilon), \quad (5.4)$$

and the equations of motion are then

$$\begin{cases} \frac{d}{d\tau}p_1 &= -\varepsilon\,\nabla_1 U(q_1,q_2;\varepsilon), \\ \frac{d}{d\tau}p_2 &= -\frac{1}{\varepsilon}\Omega^2 q_2 - \epsilon\,\nabla_2 U(q_1,q_2;\varepsilon), \\ \frac{d}{d\tau}q_1 &= \varepsilon\, p_1, \\ \frac{d}{d\tau}q_2 &= \varepsilon\, p_2. \end{cases}$$

Now, consider the flow Φ_τ of the system with Hamiltonian $H_1(p_1,p_2,q_1,q_2) = \frac{1}{2}\left(\varepsilon p_2^T p_2 + \frac{1}{\varepsilon}q_2^T\Omega^2 q_2\right)$:

$$\Phi_\tau \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ \cos(\tau\Omega)p_2 - \varepsilon^{-1}\Omega\sin(\tau\Omega)q_2 \\ q_1 \\ \varepsilon\Omega^{-1}\sin(\tau\Omega)p_2 + \cos(\tau\Omega)q_2 \end{pmatrix}$$

and rewrite it as $\Psi_{\tau\omega}$ with

$$\Psi_\theta \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ \cos(\Theta)p_2 - \varepsilon^{-1}\Omega\sin(\Theta)q_2 \\ q_1 \\ \varepsilon\Omega^{-1}\sin(\Theta)p_2 + \cos(\Theta)q_2 \end{pmatrix}$$

where $\Theta = Q^T\mathrm{Diag}(\theta_1,\ldots,\theta_d)Q$. Note that relation (5.2) can be easily checked. The change of variables $\Psi_{\tau\omega}(\hat{p}_1,\hat{p}_2,\hat{q}_1,\hat{q}_2) = (p_1,p_2,q_1,q_2)$ then transforms the system into

$$\begin{pmatrix} \frac{d}{d\tau}\hat{p}_1 \\ \frac{d}{d\tau}\hat{p}_2 \\ \frac{d}{d\tau}\hat{q}_1 \\ \frac{d}{d\tau}\hat{q}_2 \\ \frac{d}{d\tau}\theta \end{pmatrix} = \begin{pmatrix} -\varepsilon\,\nabla_1 U(\hat{q}_1, \cos(\Theta)\hat{q}_2 + \varepsilon\Omega^{-1}\sin(\Theta)\hat{p}_2;\varepsilon) \\ -\varepsilon\,\cos(\Theta)\,\nabla_2 U(\hat{q}_1, \cos(\Theta)\hat{q}_2 + \varepsilon\Omega^{-1}\sin(\Theta)\hat{p}_2;\varepsilon) \\ \varepsilon\,\hat{p}_1 \\ -\varepsilon^2\,\Omega^{-1}\sin(\Theta)\,\nabla_2 U(\hat{q}_1, \cos(\Theta)\hat{q}_2 + \varepsilon\Omega^{-1}\sin(\Theta)\hat{p}_2;\varepsilon) \\ \omega \end{pmatrix} = \begin{pmatrix} \varepsilon f(y,\theta) \\ \omega \end{pmatrix}$$

with $y = (\hat{p}_1, \hat{p}_2, \hat{q}_1, \hat{q}_2)$. The function f can also be written as $f(y,\theta) = J^{-1}\nabla_y \hat{H}(y,\theta)$ with

$$\hat{H}(\hat{p}_1, \hat{p}_2, \hat{q}_1, \hat{q}_2, \theta; \varepsilon) = \frac{\epsilon}{2}\hat{p}_1^T\hat{p}_1 + \epsilon\, U(\hat{q}_1, \cos(\Theta)\hat{q}_2 + \varepsilon\Omega^{-1}\sin(\Theta)\hat{p}_2; \varepsilon). \tag{5.5}$$

It is easy to verify that, given any vector $\tilde{\omega} \in \mathbb{R}^d$, we have

$$\frac{d\Psi_{\tau\tilde{\omega}}}{d\tau}\Big(\Psi_{-\tau\tilde{\omega}}(p_1, p_2, q_1, q_2)\Big) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\varepsilon^{-1}\Omega\tilde{\Omega} \\ 0 & 0 & 0 & 0 \\ 0 & \varepsilon\Omega^{-1}\tilde{\Omega} & 0 & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = k^{\tilde{\omega}}(z). \tag{5.6}$$

where $\tilde{\Omega} = Q^T \mathrm{diag}(\tilde{\omega}_1, \ldots, \tilde{\omega}_d)Q$. This is the flow of a Hamiltonian system with Hamiltonian $H_1^{\tilde{\omega}}(p_1, p_2, q_1, q_2) = \frac{1}{2}(\varepsilon p_2^T\Omega^{-1}\tilde{\Omega}p_2 + \varepsilon^{-1}q_2^T\Omega\tilde{\Omega}q_2)$. Enventually, the averaged Hamiltonian $H^{\tilde{\omega}}$ is of the form

$$\begin{aligned} H^{\tilde{\omega}}(p_1, p_2, q_1, q_2) \;&=\; \varepsilon\Big(\frac{1}{2}(p_2^T\Omega^{-1}\tilde{\Omega}p_2 + \varepsilon^{-2}q_2^T\Omega\tilde{\Omega}q_2) + \frac{1}{2}p_1^Tp_1 + U_{\mathbf{0}}(q_1, q_2, p_2) + \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\mathbf{k}\cdot\tilde{\omega}}{\mathbf{k}\cdot\omega}U_{\mathbf{k}}(q_1, q_2, p_2)\Big) \\ &\quad + \mathcal{O}(\varepsilon^2). \end{aligned}$$

Now, according to Theorem 5.1, for any pair of vectors $\tilde{\omega}$, the corresponding flows commute, and we have in particular $\varphi_\tau^\omega \circ \varphi_\tau^{\tilde{\omega}} = \varphi_\tau^\omega \circ \varphi_\tau^{\tilde{\omega}}$ so that

$$\{H^{\tilde{\omega}}, H^\omega\} = 0$$

implying that $H - H^{\tilde{\omega}}$ is constant along the exact solution of the Hamiltonian system with Hamiltonian H. This implies that all highly-oscillatory energies are preserved up to $\mathcal{O}(\varepsilon)$-terms.

References

[BM55] N.N. Bogolyubov and Yu.A. Mitropol'skii. Asymptotic methods in the theory of nonlinear oscillations. Technical report, FOREIGN TECHNOLOGY DIV WRIGHT-PATTERSON AFB OHIO, 1955.

[CMSS10] Philippe Chartier, Ander Murua, and Jesus-Maria Sanz-Serna. High-order averaging, formal series and numerical integration: B-series. *Foundations of Comput. Math.*, 2010. to appear.

[KB34] N. M. Krylov and N. N. Bogoliubov. Application des méthodes de la mécanique non linéaire à la théorie des oscillations stationnaires. *Edition de l'Académie des Sciences de la R. S. S. d'Ukraine*, 1934.

[Per68] L. Perko. Higher order averaging and related methods for perturbed periodic and quasi-periodic systems. *SIAM J. Appl. Math.*, 17:698–724, 1968.