Modulated Fourier expansions of highly oscillatory differential equations

David Cohen, Ernst Hairer and Christian Lubich

October 23, 2003

Abstract

Modulated Fourier expansions are developed as a tool for gaining insight into the long-time behaviour of Hamiltonian systems with highly oscillatory solutions. Particle systems of Fermi-Pasta-Ulam type with light and heavy masses are considered as an example. It is shown that the harmonic energy of the highly oscillatory part is nearly conserved over times that are exponentially long in the high frequency. Unlike previous approaches to such problems, the technique used here does not employ nonlinear coordinate transforms and can therefore be extended to the analysis of numerical discretizations.

1 Introduction

We study the system of differential equations

$$\ddot{x} + \Omega^2 x = g(x) \qquad \text{with} \qquad \Omega = \begin{pmatrix} 0 & 0 \\ 0 & \omega I \end{pmatrix} \tag{1.1}$$

where $\omega \gg 1$ and the nonlinearity is $g(x) = -\nabla U(x)$, so that the problem is Hamiltonian with

$$H(x, \dot{x}) = \frac{1}{2}\Big(\|\dot{x}\|^2 + \|\Omega x\|^2\Big) + U(x). \tag{1.2}$$

An important property of such systems is the near-conservation over long times of the *oscillatory energy*

$$I(x, \dot{x}) \;=\; \frac{1}{2}\Big(\|\dot{x}_2\|^2 + \omega^2 \, \|x_2\|^2\Big). \tag{1.3}$$

Here, the vectors $x = (x_1, x_2)$ and $\dot{x} = (\dot{x}_1, \dot{x}_2)$ are partitioned according to the partitioning of the matrix Ω in (1.1). A possible way of studying problems of the type (1.1) is via averaging techniques and Lindstedt series, see for example Neishtadt [10], Murdock [9], Pronin and Treschev [11]. The very problem (1.1) was thoroughly studied in Benettin, Galgani and Giorgilli [3], Fassò [5], and Bambusi and Giorgilli [1], using coordinate transformations of Hamiltonian perturbation theory. In the present paper we give a variant of their result, obtained with a completely different proof. It is based on writing the solution of (1.1) as a *modulated Fourier expansion*

$$x(t) = y(t) + \sum_{k \neq 0} \mathrm{e}^{\mathrm{i}k\omega t} z^k(t), \tag{1.4}$$

where $y(t)$ and $z^k(t)$ are smoothly varying functions (i.e., their derivatives are bounded independently of ω).

Such a representation of the solution has first been proposed by Miranker and van Veldhuizen [8] [1], who derived a scheme for constructing the "envelopes" $z^k(t)$. They suggested to compute numerically these envelopes and used them for approximating the solution $x(t)$. In [6] and [7, Chap. XIII] this technique of modulated Fourier expansions has been further developed and used in the analysis of the long-time behaviour of numerical integrators when the time step is not small compared to ω^{-1}. Standard backward error analysis (see for example [7, Chap. IX]) requires $\Delta t \cdot \omega$ to be small and therefore cannot be applied. In this situation, modulated Fourier expansions provide much insight into the long-time behaviour of numerical integrators. In the present paper, they are used to obtain rigorous long-time results for the exact solution of the differential equation.

The following result states the near-conservation of the oscillatory energy over time intervals that are exponentially long in ω. Here we assume that the initial values satisfy

$$\tfrac{1}{2}\Big(\|\dot{x}(0)\|^2 + \|\Omega x(0)\|^2 \Big) \;\leq\; E, \tag{1.5}$$

where E is independent of ω. (We do not require E to be small.)

Theorem 1.1 *Assume that $g(x) = -\nabla U(x)$ is analytic and bounded by M in the complex neighbourbood $D = \{x \in \mathbb{C}^n \,;\, \|x - \xi\| \leq R$ for some ξ with $H(\xi, 0) \leq H(x(0), \dot{x}(0))\}$ of the set of energetically admissible positions. Furthermore, let the initial values $x(0), \dot{x}(0)$ satisfy (1.5). Then there exist positive constants $\gamma, C, \widehat{C}, \omega_0$ depending on E, M, and R (but not on ω) such that for $\omega \geq \omega_0$*

$$\big\| I\big(x(t), \dot{x}(t)\big) - I\big(x(0), \dot{x}(0)\big) \big\| \;\leq\; C\,\omega^{-1} \qquad \text{for} \quad 0 \leq t \leq \widehat{C}\,\mathrm{e}^{\gamma\omega}.$$

The proof of this theorem will be given in the final section of this paper. We first discuss the modulated Fourier expansion in Sect. 2, and we show that the coefficient functions of (1.4) are given by asymptotic differential and algebraic equations. The effect of truncating the asymptotic series is studied

[1] We thank an anonymous referee for pointing out this reference.

in Sect. 3. Whereas these two sections treat the general problem (1.1), the final Sect. 4 assumes that $g(x) = -\nabla U(x)$. It is shown that the coefficient functions of the modulated Fourier expansion are then exponentially close to the solution of a Hamiltonian system in an infinite dimensional space, which has two invariants: one is close to the Hamiltonian (1.2) and the other is close to the oscillatory energy (1.3).

Let us mention that the dominating fluctuation terms in the oscillatory energy can be given explicitly. Writing down the $\mathcal{O}(\omega^{-1})$ terms in $\mathcal{I}$ of (4.4) below we find that

$$J(x,\dot{x}) = \frac{1}{2}\Big(\|\dot{x}_2\|^2 + \omega^2\,\|x_2\|^2\Big) - x_2^T g_2(x_1,0) \tag{1.6}$$

satisfies

$$\big\|J\big(x(t),\dot{x}(t)\big) - J\big(x(0),\dot{x}(0)\big)\big\| \,\le\, C\,\omega^{-2}$$

on exponentially long time intervals. Since $x_2 = \mathcal{O}(\omega^{-1})$, this implies that the fluctuations in $I(x,\dot{x})$ are of size $\mathcal{O}(\omega^{-2})$ when $g_2(x_1,0) = \mathcal{O}(\omega^{-1})$.

The techniques of this paper can also be applied to the slightly more general situation where the potential $U(x)$ contains expressions of the form $\varphi_1(x_1,x_2) + \omega\varphi_2(x_1/\omega, x_2)$, such that the differential equation becomes

$$\begin{aligned} \ddot{x}_1 &= g_1(x_1,x_2) \\ \ddot{x}_2 + \omega^2 x_2 &= \omega g_2(x_1,x_2) \end{aligned}$$

with $g(x)$ depending smoothly on ω^{-1}. In this case, the quantity

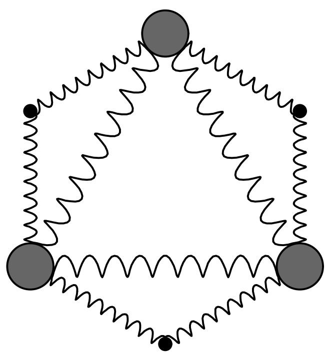
$$K(x,\dot{x}) = \frac{1}{2}\Big(\|\dot{x}_2\|^2 + \omega^2\,\|x_2\|^2\Big) \;-\; \omega x_2^T g_2(x_1,0) \;+\; \frac{1}{2}\|g_2(x_1,0)\|^2 \tag{1.7}$$

satisfies

$$\big\|K\big(x(t),\dot{x}(t)\big) - K\big(x(0),\dot{x}(0)\big)\big\| \,\le\, C\,\omega^{-1}$$

on exponentially long time intervals. Notice that the additional terms in (1.7) are in general of size $\mathcal{O}(1)$, so that the oscillatory energy exhibits fluctuations that can be large independent of the size of ω.

Example. Inspired by an example of Bambusi and Giorgilli [1] we consider a closed chain of an even number of particles with alternate light and heavy masses. They interact through springs which are harmonic up to small perturbations, and neighbouring heavy particles interact also through arbitrary anharmonic springs (see the picture to the right). More precisely,



we consider the Hamiltonian system with

$$H(\xi,\dot\xi) = \sum_{i=1}^{2N} \frac{\dot\xi_i^2}{2m_i} + \frac{1}{2}\sum_{i=1}^{2N}(\xi_i-\xi_{i-1})^2 + \sum_{j=1}^{N}\varphi_j(\xi_{2j}-\xi_{2j-2}) + \sum_{i=1}^{2N}\psi_i\big(\sqrt{m}\,(\xi_i-\xi_{i-1})\big),$$

where $m_{2j-1} = m \ll 1$ and $m_{2j} = 1$ for $j = 1,\dots,N$, and $\xi_0 = \xi_{2N}$. Applying the symplectic change of coordinates $\xi_i \mapsto \sqrt{m_i}\,\xi_i$, $\dot\xi_i \mapsto \dot\xi_i/\sqrt{m_i}$, and using the notation $\omega = 1/\sqrt{m}$, the Hamiltonian becomes

$$\begin{aligned} H(\xi,\dot\xi) &= \frac{1}{2}\sum_{i=1}^{2N}\dot\xi_i^2 + \frac{1}{2}\sum_{j=1}^{N}\Big((\xi_{2j}-\omega\xi_{2j-1})^2 + (\omega\xi_{2j-1}-\xi_{2j-2})^2\Big) \\ &+ \sum_{j=1}^{N}\varphi_j(\xi_{2j}-\xi_{2j-2}) + \sum_{j=1}^{N}\Big(\psi_{2j}\Big(\frac{\xi_{2j}}{\omega}-\xi_{2j-1}\Big) + \psi_{2j-1}\Big(\xi_{2j-1}-\frac{\xi_{2j-2}}{\omega}\Big)\Big). \end{aligned}$$

We then consider an orthogonal linear transformation $\xi^* = Q\xi$ that takes the harmonic part of the Hamiltonian to diagonal form. It is given by

$$\begin{aligned} \xi^*_{2j-1} &= \xi_{2j-1} - \frac{1}{2\omega}\big(\xi_{2j}+\xi_{2j-2}\big) + \mathcal{O}(\omega^{-2}), \\ \xi^*_{2j} &= \xi_{2j} + \frac{1}{2\omega}\big(\xi_{2j+1}+\xi_{2j-1}\big) + \mathcal{O}(\omega^{-2}). \end{aligned}$$

Omitting the stars, the Hamiltonian becomes (in the new variables)

$$H(\xi,\dot\xi) = \frac{1}{2}\sum_{i=1}^{2N}\dot\xi_i^2 + \omega^2\sum_{j=1}^{N}\xi_{2j-1}^2 + \Phi_1(\xi) + \Phi_2(\xi_1,\xi_2/\omega,\xi_3,\xi_4/\omega,\dots),$$

which is of the form treated above.

Numerical Experiment. For a concrete example we put $N = 3$, $\omega = 50$, we let $\varphi_j(s) = \chi(\sqrt[6]{2} - s/\omega)$ with $\chi(s) = s^{-12} - s^{-6}$ be the Lennard-Jones potential, we take $\psi_{2j}(s) = s^2/2 + s^4/4$ for $j = 1,\dots,N-1$, and $\psi_i(s) = 0$ else.

Figure 1 shows the components ξ_2, ξ_4, ξ_6, and $10\xi_5$ on the interval $0 \le t \le 10$. The factor 10 multiplying ξ_5 is included to show more clearly the oscillations of size $\mathcal{O}(\omega^{-1})$ in the numerical solution.

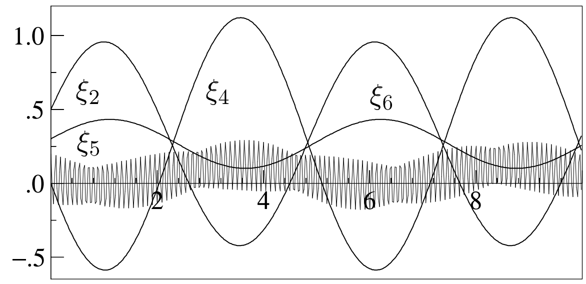


Figure 1: Solution components, where the non-zero initial positions are $\xi_2(0) = 0.5, \xi_3(0) = (2\omega)^{-1}, \xi_5(0) = \omega^{-1}, \xi_6(0) = 0.3$ and the non-zero initial velocities are $\dot{\xi}_1(0) = -\dot{\xi}_3(0) = \omega^{-1}, \dot{\xi}_2(0) = 0.8, \dot{\xi}_4(0) = -1, \dot{\xi}_6(0) = 0.2$.

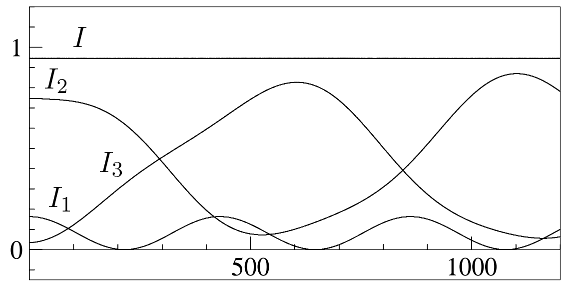


Figure 2: Oscillatory energy for the solution with initial values as in Fig. 1.

In Fig. 2 we plot the energies $I_j(\xi^*, \dot{\xi}^*) = \frac{1}{2}(\dot{\xi}^*_{2j-1})^2 + \omega^2(\xi^*_{2j-1})^2$ together with the oscillatory energy $I = I_1 + I_2 + I_3$ (cf. (1.3)) along the numerical solution on the interval $0 \le t \le 1200$. For this example, the expression $g_2(x_1, 0)$ is of size ω^{-1}, so that the oscillatory energy is conserved up to terms of size ω^{-2} (see (1.6)). Therefore, the oscillations cannot be observed in Fig. 2.

2 The Modulated Fourier Expansion

We write the system (1.1) in the equivalent form

$$\begin{aligned} \ddot{x}_1 &= g_1(x_1, x_2) \\ \ddot{x}_2 + \omega^2 x_2 &= g_2(x_1, x_2), \end{aligned} \tag{2.1}$$

where $\omega \gg 1$ represents the dominant frequency of the system. In this section we do not assume that $g(x)$ is the gradient of a potential. Our aim is to present a technique that allows us to separate the smooth and the oscillating parts of the solution of (2.1) and to write it in the form

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} + \sum_{k \neq 0} \mathrm{e}^{\mathrm{i}k\omega t} \begin{pmatrix} z_1^k(t) \\ z_2^k(t) \end{pmatrix}, \tag{2.2}$$

where $y_i(t)$ and $z_i^k(t)$ are smoothly varying functions (i.e., their derivatives are bounded independently of ω). The functions $y_i(t)$ are real-valued and $z_i^k(t)$ are complex-valued. Since the solution $x_i(t)$ is real-valued, we have to require that $z_i^{-k} = \overline{z_i^k}$. We also use the notations $z_2 := z_2^1$ and $z_2^0 := y_2$.

Inserting (2.2) into (1.1), expanding the nonlinearity into a Taylor series around $\big(y_1(t), 0\big)$, and comparing the coefficients of $\mathrm{e}^{\mathrm{i}k\omega t}$ yields differential equations for the coefficient functions $y_i(t)$ and $z_i^k(t)$. With the exception of $y_1(t)$ they are of singular perturbation type. We have to find smooth solutions of these equations. As explained in [6], the functions y_1 and z_2 are seen to be given by differential equations of the form

$$\ddot{y}_1 = \sum_{l\geq 0} \omega^{-l} F_{1l}(y_1, \dot{y}_1, z_2), \qquad \dot{z}_2 = \sum_{l\geq 1} \omega^{-l} F_{2l}(y_1, \dot{y}_1, z_2), \tag{2.3}$$

and the remaining functions by algebraic relations

$$z_i^k = \sum_{l\geq 0} \omega^{-l} G_{il}^k(y_1, \dot{y}_1, z_2). \tag{2.4}$$

Observe that $y_2 = z_2^0$, so that we also have an algebraic relation for y_2. Furthermore, for $i = 2$ and $k = 1$, we have the trivial identity $z_2^1 = z_2$ which implies

$$G_{20}^1(y_1, \dot{y}_1, z_2) = z_2, \qquad G_{2l}^1(y_1, \dot{y}_1, z_2) = 0 \quad \text{for} \quad l \geq 1. \tag{2.5}$$

Remember that z_i^{-k} is the complex conjugate of z_i^k, so that also G_{il}^{-k} is the complex conjugate of G_{il}^k.

The series (2.3) and (2.4) are asymptotic expansions and do not converge in general. Later, we shall truncate them suitably in order to get rigorous statements.

2.1 Recurrence Relations for the Coefficient Functions

For a computation of the functions F_{il} and G_{il}^k in (2.3) and (2.4) it is convenient to introduce the Lie operator $\mathcal{L}_l$. It can be applied to smooth functions $G(y_1, \dot{y}_1, z_2)$ and it is defined for $l \geq 0$ by

$$\mathcal{L}_l G = D_2 G \cdot F_{1l} + D_3 G \cdot F_{2l} + \begin{cases} D_1 G \cdot \dot{y}_1 & \text{if} \ \ l = 0 \\ 0 & \text{if} \ \ l \geq 1, \end{cases} \tag{2.6}$$

where D_j denotes the partial derivative with respect to the jth argument of $G(y_1, \dot{y}_1, z_2)$. This definition is motivated by the fact that, whenever $y_1(t)$ and $z_2(t)$ are a solution of the differential equation (2.3), then we have

$$\frac{d}{dt} G\big(y_1(t), \dot{y}_1(t), z_2(t)\big) = \sum_{l\geq 0} \omega^{-l} \mathcal{L}_l G\big(y_1(t), \dot{y}_1(t), z_2(t)\big). \tag{2.7}$$

Lemma 2.1 *The function* $\big(x_1(t), x_2(t)\big)$ *of (2.2) with* $y_i(t)$ *and* $z_i^k(t)$ *given by (2.3) and (2.4) represents a formal solution of (2.1) if the coefficient functions* F_{il} *and* G_{il}^k *satisfy the following recurrence relations (for* $l \geq 0$*):*

$$\begin{aligned}
F_{1l} &= S_1(0,l) \\
G_{1l}^k &= \frac{1}{k^2}\Big(\sum_{m+n+j=l-2} \mathcal{L}_m\mathcal{L}_n G_{1j}^k + 2\mathrm{i}k \sum_{m+j=l-1} \mathcal{L}_m G_{1j}^k - S_1(k,l-2)\Big) \\
F_{2l} &= \frac{1}{2\mathrm{i}}\Big(S_2(1,l-1) - \sum_{m+j=l-1} \mathcal{L}_m F_{2j}\Big) \\
G_{2l}^k &= \frac{1}{1-k^2}\Big(S_2(k,l-2) - \sum_{m+n+j=l-2} \mathcal{L}_m\mathcal{L}_n G_{2j}^k - 2\mathrm{i}k \sum_{m+j=l-1} \mathcal{L}_m G_{2j}^k\Big).
\end{aligned}$$

The sums are over $m \geq 0, n \geq 0, j \geq 0$*, and we have used the abbreviation*

$$S_i(k,l) = \sum_{m,n\geq 0} \frac{1}{m!\,n!} \sum_{\substack{\alpha,\beta \\ s(\alpha)+s(\beta)=k}} \sum_{\substack{e,f \\ s(e)+s(f)=l}} D_1^m D_2^n g_i(y_1,0)(G_{1e}^\alpha, G_{2f}^\beta).$$

Here, $\alpha = (\alpha_1,\ldots,\alpha_m), \beta = (\beta_1,\ldots,\beta_n), e = (e_1,\ldots,e_m), f = (f_1,\ldots,f_n)$ *are multi-indices with* $\alpha_i \neq 0$*,* β_i *arbitrary,* $e_i \geq 0$*,* $f_i \geq 0$*, and* $(G_{1e}^\alpha, G_{2f}^\beta) = (G_{1,e_1}^{\alpha_1},\ldots,G_{1,e_m}^{\alpha_m}, G_{2,f_1}^{\beta_1},\ldots,G_{2,f_n}^{\beta_n})$*. We use the abbreviation* $s(\alpha) = \sum_{i=1}^m \alpha_i$ *and similarly for the other multi-indices.*

Proof. Inserting the relation (2.2) into the first equation of the system (2.1), and expanding the nonlinearity into a Taylor series around $(y_1,0)$, we obtain

$$\begin{aligned}
\ddot{y}_1 \;&+\; \sum_{k\neq 0} \mathrm{e}^{\mathrm{i}k\omega t}\big(\ddot{z}_1^k + 2\mathrm{i}k\omega \dot{z}_1^k - k^2\omega^2 z_1^k\big) \\
&= \sum_{m,n\geq 0} \frac{1}{m!\,n!} \sum_{\alpha,\beta} \mathrm{e}^{\mathrm{i}\omega t(s(\alpha)+s(\beta))} D_1^m D_2^n g_1(y_1,0)(z_1^\alpha, z_2^\beta),
\end{aligned}$$

where $(z_1^\alpha, z_2^\beta) = (z_1^{\alpha_1},..,z_1^{\alpha_m}, z_2^{\beta_1},..,z_2^{\beta_n})$, and the last sum is over all multi-indices α,β with $\alpha_i \neq 0$. We now insert our ansatz (2.3) for $\ddot{y}_1$ and (2.4) for z_i^k, we use the Lie derivative for expressing the derivatives of z_1^k, and thus obtain

$$\begin{aligned}
\sum_{l\geq 0} \omega^{-l} F_{1l} \;&+\; \sum_{k\neq 0} \mathrm{e}^{\mathrm{i}k\omega t}\Big(\sum_{m,n,j\geq 0} \omega^{-m-n-j} \mathcal{L}_m\mathcal{L}_n G_{1j}^k \\
&\qquad + 2\mathrm{i}k \sum_{m,j\geq 0} \omega^{-m-j+1} \mathcal{L}_m G_{1j}^k - k^2 \sum_{j\geq 0} \omega^{-j+2} G_{1j}^k\Big) \\
&= \sum_{m,n\geq 0} \frac{1}{m!\,n!} \sum_{\alpha,\beta} \mathrm{e}^{\mathrm{i}\omega t(s(\alpha)+s(\beta))} D_1^m D_2^n g_1(y_1,0) \\
&\qquad\qquad \Big(\sum_{e\geq 0} \omega^{-s(e)} G_{1e}^\alpha, \sum_{f\geq 0} \omega^{-s(f)} G_{2f}^\beta\Big).
\end{aligned}$$

We just have to compare the coefficients of $\mathrm{e}^{\mathrm{i}k\omega t}$ and ω^{-l} (resp. ω^{-l+2}) to obtain the recurrence relations for the functions F_{1l} and G_{1l}^k. This implies

$$G_{10}^k = 0, \quad G_{11}^k = 0 \qquad \text{for all} \quad k \neq 0, \tag{2.8}$$

so that the series expansions (2.4) for all z_1^k start with the ω^{-2}-term.

Looking at the second equation of the system (2.1), we obtain

$$\begin{aligned} \ddot{y}_2 + \omega^2 y_2 \ &+ \sum_{k\neq 0} \mathrm{e}^{\mathrm{i}k\omega t}\big(\ddot{z}_2^k + 2\mathrm{i}k\omega \dot{z}_2^k + (1-k^2)\omega^2 z_2^k\big) \\ &= \sum_{m,n\geq 0} \frac{1}{m!\,n!} \sum_{\alpha,\beta} \mathrm{e}^{\mathrm{i}\omega t(s(\alpha)+s(\beta))} D_1^m D_2^n g_2(y_1,0)(z_1^\alpha, z_2^\beta). \end{aligned}$$

We insert the ansatz (2.3) for $\dot{z}_2$ and (2.4) for z_i^k, and in the same way as above we get the recurrence relations for the functions F_{2l} and G_{2l}^k. They imply

$$G_{20}^k = 0, \quad G_{21}^k = 0 \qquad \text{for} \quad k \neq \pm 1, \tag{2.9}$$

so that also the expansions (2.4) for z_2^k $(k \neq \pm 1)$ start with the ω^{-2}-term. □

2.2 Estimates for the functions F_{ij} and G_{ij}^k

Our next aim is to get upper bounds for the coefficient functions F_{ij} and G_{ij}^k of (2.3) and (2.4). Since they depend on the derivatives of $g_i(x_1,x_2)$, it is natural to require $g(x)$ to be analytic and bounded (by M) in a suitable complex domain, say in $\{(x_1,x_2)\,;\, \|x_1-y_{10}\| \leq 4R, \|x_2\| \leq 3R\}$. Cauchy's estimates then imply

$$\|D_1^m D_2^n g_i(y_1,0)\| \leq m!\,n!\,M\,(3R)^{-m-n} \qquad \text{for} \quad \|y_1-y_{10}\| \leq R \tag{2.10}$$

and for all $n,m \geq 0$. This is our main assumption of this section. To obtain the desired estimates for the coefficient functions we combine and adapt the techniques of [2] and [7, Sect. IX.5].

We fix a value $\mathcal{Y}_0 = (y_{10}, \dot{y}_{10}, 0)$, and we consider the complex ball

$$B_\rho(\mathcal{Y}_0) = \big\{(y_1,\dot{y}_1,z_2)\,;\, \|y_1-y_{10}\| \leq \rho R, \|\dot{y}_1-\dot{y}_{10}\| \leq \rho M, \|z_2\| \leq \rho R\big\}. \tag{2.11}$$

For a function $G(y_1,\dot{y}_1,z_2)$ defined on $B_\rho(\mathcal{Y}_0)$ we let

$$\|G\|_\rho = \max\big\{\|G(y_1,\dot{y}_1,z_2)\|\,;\, (y_1,\dot{y}_1,z_2) \in B_\rho(\mathcal{Y}_0)\big\}. \tag{2.12}$$

Since the coefficient functions are defined via expressions of the form $\mathcal{L}_l G$, the following lemma will be useful.

Lemma 2.2 *Let G be analytic and bounded on $B_\rho(\mathcal{Y}_0)$, and let F_{1l} and F_{2l} be bounded on $B_\sigma(\mathcal{Y}_0)$ with $0 \le \sigma < \rho$. Then we have*

$$\begin{aligned} \|\mathcal{L}_0 G\|_\sigma &\le \tfrac{1}{\rho-\sigma}\cdot \|G\|_\rho \cdot \max\bigl(\|F_{10}\|_\sigma/M, \|\dot y_1\|_\sigma/R\bigr), \\ \|\mathcal{L}_l G\|_\sigma &\le \tfrac{1}{\rho-\sigma}\cdot \|G\|_\rho \cdot \max\bigl(\|F_{1l}\|_\sigma/M, \|F_{2l}\|_\sigma/R\bigr) \qquad \textit{for} \quad l \ge 1. \end{aligned}$$

Proof. Consider $\alpha(\zeta) = G\bigl(y_1, \dot y_1 + \zeta F_{1l}(y_1,\dot y_1,z_2), z_2 + \zeta F_{2l}(y_1,\dot y_1,z_2)\bigr)$, where $(y_1,\dot y_1,z_2) \in B_\sigma(\mathcal{Y}_0)$. This function is analytic for $|\zeta| \le \varepsilon$ with $\varepsilon := (\rho - \sigma)/\max\bigl(\|F_{1l}\|_\sigma/M, \|F_{2l}\|_\sigma/R\bigr)$. Since $\alpha'(0) = (\mathcal{L}_l G)(y_1,\dot y_1,z_2)$, Cauchy's estimate yields

$$\|(\mathcal{L}_l G)(y_1,\dot y_1,z_2)\| = \|\alpha'(0)\| \le \frac{1}{\varepsilon}\sup_{|\zeta|\le\varepsilon}\|\alpha(\zeta)\| \le \frac{1}{\varepsilon}\|G\|_\rho\,,$$

which proves the statement for $l \ge 1$. For $l = 0$ we have to consider the function $\alpha(\zeta) = G\bigl(y_1 + \zeta\dot y_1, \dot y_1 + \zeta F_{10}(y_1,\dot y_1,z_2), z_2\bigr)$, because $F_{20} = 0$ by Lemma 2.1. □

The use of Lemma 2.2 implies that we cannot work with only one norm $\|\cdot\|_\rho$ for finding estimates of the coefficient functions. We therefore fix a positive integer L, we put $\delta = 1/(2L)$, and we consider the norms corresponding to balls with shrinking radius $\rho = 1 - l\delta$ $(0 \le l \le L)$.

Lemma 2.3 *Let $\mathcal{Y}_0 = (y_{10}, \dot y_{10}, 0)$ be given, and assume that (2.10) holds. The functions F_{ij} and G_{ij}^k of Lemma 2.1 satisfy*

$$\begin{aligned} &\|F_{10}\|_1 \le a_0 M, \qquad \|\dot y_1\|_1 \le a_0 R \\ &\|F_{1l}\|_{1-l\delta} \le a_l M, \qquad \|F_{2l}\|_{1-l\delta} \le a_l R, \qquad 1 \le l \le L \\ &\|G_{20}^{-1}\|_1 + \|G_{20}^{1}\|_1 \le b_0 R \\ &\max\Bigl(\sum_{k\ne 0} k^2 \|G_{1l}^k\|_{1-l\delta}\;,\; \sum_{k\in\mathbb{Z}} |1-k^2|\, \|G_{2l}^k\|_{1-l\delta}\Bigr) \le b_l R, \qquad 1 \le l \le L, \end{aligned}$$

where $a_0 = \max(9, (\|\dot y_{10}\|_1 + M)/R)$, $b_0 = 2$, and the generating functions $a(\zeta) = \sum_{l\ge1} a_l\zeta^l$ and $b(\zeta) = \sum_{l\ge1} b_l \zeta^l$ are implicitly given by

$$\begin{aligned} a(\zeta) &= -9 + 9\bigl(1 + \tfrac{M\zeta}{2R}\bigr)\bigl(1 - b(\zeta)\bigr)^{-2} + \tfrac{\zeta}{2\delta}\bigl(a_0 + a(\zeta)\bigr)a(\zeta), \\ b(\zeta) &= \tfrac{9M\zeta^2}{R}\,\bigl(1 - b(\zeta)\bigr)^{-2} + \tfrac{2\zeta}{\delta}\bigl(a_0 + a(\zeta)\bigr)\bigl(b_0 + b(\zeta)\bigr) \\ &\qquad + \tfrac{\zeta^2}{\delta^2}\bigl(a_0 + a(\zeta)\bigr)^2\bigl(b_0 + b(\zeta)\bigr). \end{aligned} \tag{2.13}$$

Proof. (a) In this proof we shall use the shorthand notation

$$\|G\|_l := \|G\|_{1-l\delta} = \max\bigl\{\|G(y_1,\dot y_1,z_2)\|\,;\,(y_1,\dot y_1,z_2)\in B_{1-l\delta}(\mathcal{Y}_0)\bigr\}. \tag{2.14}$$

Observe that $\|G\|_l$ is a decreasing function of l.

To obtain the desired statement, we begin with some estimations and then we prove the result of this Lemma by induction on l.

(b) Because of (2.8), (2.9) and (2.5), the above estimates for G_{il}^k also imply

$$\sum_{k\neq 0} \|G_{1l}^k\|_l \leq b_l R, \qquad \sum_{k\in\mathbb{Z}} \|G_{2l}^k\|_l \leq b_l R \qquad \text{for} \quad l \geq 0. \tag{2.15}$$

Using these relations and the analyticity assumption (2.10), we are able to majorize the $S_i(k,l)$ as follows:

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} \|S_i(k,l)\|_l &\leq \sum_{m,n\geq 0} \frac{m!\,n!}{m!\,n!} \sum_{\substack{\alpha,\beta\\ \alpha_i\neq 0}} \sum_{\substack{s(e)+s(f)\\ =l}} M(3R)^{-m-n} \|G_{1e_1}^{\alpha_1}\|_l \ldots \|G_{2f_1}^{\beta_1}\|_l \ldots \\
&\leq M \sum_{m,n\geq 0} \sum_{s(e)+s(f)=l} 3^{-m-n} b_{e_1} \ldots b_{e_m} b_{f_1} \ldots b_{f_n} \\
&\leq M \sum_{j\geq 0} (j+1) \sum_{d_1+\ldots+d_j=l} 3^{-j} b_{d_1} \ldots b_{d_j} \;=\; M c_l,
\end{aligned}$$

where c_l ($l \geq 0$) are the coefficients of the generating function

$$\sum_{l\geq 0} c_l \zeta^l = c(\zeta) = \frac{1}{\left(1 - \frac{b_0+b(\zeta)}{3}\right)^2} = \frac{9}{\left(1 - b(\zeta)\right)^2}.$$

We have used $\|G_{1e_1}^{\alpha_1}\|_l \leq \|G_{1e_1}^{\alpha_1}\|_{e_1}$ and $\|G_{2f_1}^{\beta_1}\|_l \leq \|G_{2f_1}^{\beta_1}\|_{f_1}$, which are a consequence of $e_1 \leq l$ and $f_1 \leq l$.

(c) For $m+n+j = l-2$ a twofold application of Lemma 2.2 yields

$$\|\mathcal{L}_m \mathcal{L}_n G_{ij}^k\|_l \leq \frac{1}{\delta^2} \|G_{ij}^k\|_j \, a_m \, a_n \quad \text{and} \quad \sum_{k\neq 0} \|\mathcal{L}_m \mathcal{L}_n G_{1j}^k\|_l \leq \frac{R}{\delta^2} \, b_j \, a_m \, a_n.$$

This implies

$$\sum_{k\neq 0} \sum_{m+n+j=l-2} \|\mathcal{L}_m \mathcal{L}_n G_{1j}^k\|_l \leq \frac{R}{\delta^2} \, d_{l-2},$$

where the generating function of the d_l is

$$d(\zeta) = \sum_{l\geq 0} d_l \zeta^l = \big(b_0 + b(\zeta)\big)\big(a_0 + a(\zeta)\big)^2.$$

The same estimate is obtained for $\sum_{k\in\mathbb{Z}} \sum_{m+n+j=l-2} \|\mathcal{L}_m \mathcal{L}_n G_{2j}^k\|_l$.

(d) In order to estimate $|k| \, \|\mathcal{L}_m G_{ij}^k\|_l$ for $m+j = l-1$, we observe that similar to (2.15) also

$$\sum_{k\in\mathbb{Z}} |k| \, \|G_{1l}^k\|_l \leq b_l R, \qquad \sum_{k\in\mathbb{Z}} |k| \, \|G_{2l}^k\|_l \leq b_l R \qquad \text{for} \quad l \geq 0 \tag{2.16}$$

holds. As in part (b) we thus obtain

$$\sum_{k\in\mathbb{Z}} |k| \sum_{m+j=l-1} \|\mathcal{L}_m G_{ij}^k\|_l \;\le\; \frac{R}{\delta}\, q_{l-1},$$

where the generating function for the q_l is

$$q(\zeta) = \sum_{l\ge 0} q_l \zeta^l = \big(b_0 + b(\zeta)\big)\big(a_0 + a(\zeta)\big).$$

(e) After these preparations the statement can be proved by induction on l. The bounds a_0 and b_0 are defined just to satisfy the estimates for $l = 0$. The form of the generating functions for a_l and b_l are a consequence of the recurrence relations of Lemma 2.1 and of parts (b), (c) and (d) of this proof. □

To get bounds on the expressions of Lemma 2.3, we have to majorize a_l and b_l. This can be done with the help of Cauchy's inequalities, because the generating functions $a(\zeta)$ and $b(\zeta)$ are analytic in a neighbourhood of the origin. Since the equations (2.13) depend on δ, R and M, we have to be careful in determining the radius of the disc of analyticity. In the following we assume $M \ge R$. This can be done without loss of generality, because we can always increase M without violating (2.10) or, even better, we can rescale time in the differential equation and thus multiply $g(x)$ by a scalar factor.

Theorem 2.4 *We fix $\mathcal{Y}_0 = (y_{10}, \dot y_{10}, 0)$, and we assume that the nonlinearity $g(x)$ satisfies (2.10) with $M \ge R$, and that $\|\dot y_{10}\| \le M$. The coefficient functions of Lemma 2.1 then satisfy for $l \ge 1$*

$$\|F_{1l}\|_{1/2} \le \mu M\Big(\frac{\nu l M}{R}\Big)^l, \qquad \|F_{2l}\|_{1/2} \le \mu R\Big(\frac{\nu l M}{R}\Big)^l,$$
$$\max\Big(\sum_{k\ne 0} k^2 \|G_{1l}^k\|_{1/2}, \sum_{k\in\mathbb{Z}} |1-k^2|\,\|G_{2l}^k\|_{1/2}\Big) \le \mu R\Big(\frac{\nu l M}{R}\Big)^l,$$

where μ and ν only depend on an upper bound of M/R but not on the other data of the differential equation. The norm is that of (2.12).

Proof. We multiply the ζ in (2.13) either by $\frac{1}{\delta} \ge 1$ or by $\frac{M}{R} \ge 1$ so that the relations only depend on $\frac{\zeta M}{\delta R}$, $a(\zeta)$, and $b(\zeta)$. This makes the coefficients a_l and b_l at worst larger, so that the estimates of Lemma 2.3 still hold. We then introduce the new variables $\hat\zeta = \zeta M/\delta R$, $\widehat a(\hat\zeta) = a(\zeta)$, and $\widehat b(\hat\zeta) = b(\zeta)$, so that (2.13) becomes

$$\begin{aligned} \widehat a(\hat\zeta) &= -9 + 9\big(1+\tfrac{\hat\zeta}{2}\big)\big(1-\widehat b(\hat\zeta)\big)^{-2} + \tfrac{\hat\zeta}{2}\big(a_0+\widehat a(\hat\zeta)\big)\widehat a(\hat\zeta), \\ \widehat b(\hat\zeta) &= 9\,\hat\zeta^2\big(1-\widehat b(\hat\zeta)\big)^{-2} + 2\hat\zeta\big(a_0+\widehat a(\hat\zeta)\big)\big(2+\widehat b(\hat\zeta)\big) \\ &\qquad + \hat\zeta^2\big(a_0+\widehat a(\hat\zeta)\big)^2\big(2+\widehat b(\hat\zeta)\big). \end{aligned} \tag{2.17}$$

Observe that $a_0 \le \max(9, 2M/R)$, which is a consequence of $\|\dot y_{10}\| \le M$.

In the equations (2.17) we obtain $\widehat a = 0$, $\widehat b = 0$ for $\hat\zeta = 0$, and the implicit function theorem can be applied. This proves the existence of constants μ and ν, such that $\widehat a(\hat\zeta)$ and $\widehat b(\hat\zeta)$ are analytic in the disc $|\hat\zeta| \le 2/\nu$ and bounded by μ. Cauchy's inequalities thus prove that the lth coefficient of these generating functions is bounded by $\mu(\nu/2)^l$. This yields

$$a_l\Big(\frac{\delta R}{M}\Big)^l \le \mu\Big(\frac{\nu}{2}\Big)^l, \qquad b_l\Big(\frac{\delta R}{M}\Big)^l \le \mu\Big(\frac{\nu}{2}\Big)^l.$$

Putting $l = L$ in the estimates of Lemma 2.3 and inserting the just obtained upper bounds for a_L and b_L, proves the theorem. We use the fact that $1 - L\delta = 1/2$. □

3 Exponentially Small Error Estimates

In general, the series expansions in (2.3) and (2.4) diverge, even for arbitrarily large ω. For obtaining rigorous statements we have to truncate these series. We thus consider

$$\ddot y_1 = \sum_{0\le l\le N} \omega^{-l} F_{1l}(y_1, \dot y_1, z_2), \quad \dot z_2 = \sum_{1\le l\le N} \omega^{-l} F_{2l}(y_1, \dot y_1, z_2), \tag{3.1}$$

$$z_i^k = \sum_{2\le l\le N} \omega^{-l} G_{il}^k(y_1, \dot y_1, z_2). \tag{3.2}$$

The choice of the truncation index will be made on the basis of the estimates of Theorem 2.4. The lth term in the expansions (2.3) and (2.4) is majorized by $\mathit{Const}\,(\nu l M/\omega R)^l$, which is minimal for $\nu l M/\omega R = 1/\mathrm{e}$. We therefore choose the integer truncation index N such that

$$N \le \frac{\omega R}{\mathrm{e}\nu M} < N + 1. \tag{3.3}$$

Using the inequality

$$\sum_{2\le l\le N} l^2 \Big(\frac{\nu l M}{\omega R}\Big)^{l-2} \le \sum_{2\le l\le N} l^2 \Big(\frac{l}{\mathrm{e}N}\Big)^{l-2} \le 8.65,$$

which can be checked numerically for small N, and the left-hand expression of which is a decreasing function of N for large N, it immediately follows from Theorem 2.4 that

$$\sum_{k\ne 0} k^2 \sum_{2\le l\le N} \omega^{-l} \|G_{1l}^k\|_{1/2} \le 8.65\,\mu\, R \Big(\frac{\nu M}{\omega R}\Big)^2 \le \mathit{Const}\cdot R\Big(\frac{M}{\omega R}\Big)^2. \tag{3.4}$$

The remaining bounds of Theorem 2.4 yield similar estimates also for G_{2l}^k, F_{1l}, and F_{2l}.

3.1 Initial Values for the Modulated Fourier Expansion

In this section we consider the function

$$\begin{pmatrix} \widetilde{x}_1(t) \\ \widetilde{x}_2(t) \end{pmatrix} = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} + \sum_{k\neq 0} \mathrm{e}^{\mathrm{i}k\omega t} \begin{pmatrix} z_1^k(t) \\ z_2^k(t) \end{pmatrix}, \tag{3.5}$$

where $y_i(t)$ and $z_i^k(t)$ are solutions of the truncated system (3.1)–(3.2). The sum over k is still infinite.

In the following we consider the differential equation (2.1) with initial values $x_1(0) = x_{10}$, $\dot{x}_1(0) = \dot{x}_{10}$, $x_2(0) = x_{20}$, $\dot{x}_2(0) = \dot{x}_{20}$, and we assume that the harmonic energy of these initial values is bounded by E independent of ω, see (1.5). We first show that to these initial values there correspond (locally) unique initial values for the system (3.1), such that $\widetilde{x}(0) = x(0)$ and $\dot{\widetilde{x}}(0) = \dot{x}(0)$. We then show that the function (3.5), obtained with these initial values for y_1, $\dot{y}_1$ and z_2, has an exponentially small defect when it is inserted into (2.1).

Lemma 3.1 *Consider the differential equation (2.1) with initial values $x(0) = (x_{10}, x_{20})$, $\dot{x}(0) = (\dot{x}_{10}, \dot{x}_{20})$ satisfying (1.5). Assume that the nonlinearity $g(x)$ is analytic in a ball $\{(x_1, x_2) \mid \|x_1 - x_{10}\| \le 4R, \|x_2\| \le 3R\}$ and bounded by M, with $M \ge R$. For sufficiently large ω ($M/\omega R \le \gamma$, where γ does not depend on ω) there exist (locally) unique initial values $y_1(0) = y_{10}$, $\dot{y}_1(0) = \dot{y}_{10}$, $z_2(0) = z_{20}$ for the system (3.1), such that*

$$x(0) = \widetilde{x}(0), \qquad \dot{x}(0) = \dot{\widetilde{x}}(0) \tag{3.6}$$

with $\widetilde{x}(t)$ from (3.5). These initial values satisfy

$$\begin{aligned} x_{10} &= y_{10} + \mathcal{O}(R\omega^{-2}), & x_{20} &= z_{20} + \overline{z}_{20} + \mathcal{O}(R\omega^{-2}), \\ \dot{x}_{10} &= \dot{y}_{10} + \mathcal{O}(R\omega^{-1}), & \dot{x}_{20} &= \mathrm{i}\omega z_{20} - \mathrm{i}\omega \overline{z}_{20} + \mathcal{O}(R\omega^{-1}), \end{aligned}$$

where the constant symbolizing the $\mathcal{O}(\cdot)$ can depend on M/R and on the harmonic energy E, but not on ω.

Proof. Using the truncated relations (3.2) and the Lie operator $\mathcal{L}_k$, the condition (3.6) becomes

$$
\begin{aligned}
x_{10} &= y_{10} + \sum_{k\neq 0} \sum_{2\le l\le N} \omega^{-l}\, G_{1l}^k(y_{10}, \dot y_{10}, z_{20}, \overline{z}_{20}) \\
x_{20} &= z_{20} + \overline{z}_{20} + \sum_{|k|\neq 1} \sum_{2\le l\le N} \omega^{-l}\, G_{2l}^k(y_{10}, \dot y_{10}, z_{20}, \overline{z}_{20}) \\
\dot x_{10} &= \dot y_{10} + \sum_{k\neq 0} \sum_{2\le l\le N} \omega^{-l}\Big((\mathrm{i}k\omega) G_{1l}^k(y_{10}, \dot y_{10}, z_{20}, \overline{z}_{20}) \\
&\qquad + \sum_{0\le s\le N} \omega^{-s}\, (\mathcal{L}_s G_{1l}^k)(y_{10}, \dot y_{10}, z_{20}, \overline{z}_{20})\Big) \\
(\mathrm{i}\omega)^{-1}\dot x_{20} &= z_{20} - \overline{z}_{20} + (\mathrm{i}\omega)^{-1} \sum_{|k|\neq 1} \sum_{2\le l\le N} \omega^{-l}\Big((\mathrm{i}k\omega) G_{2l}^k(y_{10}, \dot y_{10}, z_{20}, \overline{z}_{20}) \\
&\qquad + \sum_{0\le s\le N} \omega^{-s}\, (\mathcal{L}_s G_{2l}^k)(y_{10}, \dot y_{10}, z_{20}, \overline{z}_{20})\Big),
\end{aligned}
$$

Collecting the unknown variables into a vector $\mathcal{Y}_0 = (y_{10}, \dot y_{10}, z_{20}, \overline{z}_{20})$, this system can be readily brought to the form $\mathcal{Y}_0 = \mathcal{F}(\mathcal{Y}_0)$. Using Cauchy's inequalities and (3.4), we have $\|\mathcal{F}'(\mathcal{Y})\| \le Const \cdot (\frac{M}{\omega R}) < 1$ if $M/\omega R$ is sufficiently small. This implies, by the Mean Value Theorem, that $\mathcal{F}$ is a contraction on the closed ball

$$
B = \{(y_1, \dot y_1, z_2) \mid \|y_1 - x_{10}\| \le R/4, \|\dot y_1 - \dot x_{10}\| \le M/4, \|z_2\| \le R/4\}.
$$

Furthermore, by (1.5), (3.4) and using the fact that $M/\omega R$ is sufficiently small, we have $\mathcal{F}(B) \subset B$. To conclude the proof, we apply the Banach Fixed Point Theorem to solve the nonlinear system $\mathcal{Y} = \mathcal{F}(\mathcal{Y})$. □

3.2 Estimation of the Defect

After having found suitable initial values for the differential equation (3.1), wich exist for $\omega \ge \omega_0$ with a sufficiently large ω_0, we investigate the length of the time-interval such that the solution exists and remains in the ball

$$
B = \{(y_1, \dot y_1, z_2) \mid \|y_1 - y_{10}\| \le R/2, \|\dot y_1 - \dot y_{10}\| \le M/2, \|z_2\| \le R/2\}.
$$

We assume that the nonlinearity $g(x)$ satisfies (2.10) with $M \ge R$ and that $\|\dot y_{10}\| \le M$ (this assumption is essentially a definition of M and R). Similar to (3.4), the estimates of Theorem 2.4 then yield

$$
\begin{aligned}
\sum_{0\le l\le N} \omega^{-l}\, \|F_{1l}(y_1, \dot y_1, z_2)\|_{1/2} &\le Const \cdot M \\
\sum_{1\le l\le N} \omega^{-l}\, \|F_{2l}(y_1, \dot y_1, z_2)\|_{1/2} &\le Const \cdot R \Big(\frac{M}{\omega R}\Big) \le Const \cdot M \cdot \omega^{-1}
\end{aligned}
\tag{3.7}
$$

for $(y_1, \dot y_1, z_2) \in B$. As long as the solution of (3.1) remains in B, we thus have the estimates

$$\begin{aligned} \|y_1(t) - y_{10}\| &\le t\,\|\dot y_{10}\| + t^2 M\,Const \\ \|\dot y_1(t) - \dot y_{10}\| &\le t\,M\,Const \\ \|z_2(t) - z_{20}\| &\le t\,M\omega^{-1}\,Const. \end{aligned} \tag{3.8}$$

This proves the existence of a $T > 0$ such that $\big(y_1(t), \dot y_1(y), z_2(t)\big) \in B$ for $0 \le t \le T$. As the generic constant *Const*, also T only depends on an upper bound of M/R.

In the following we denote

$$y^0(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}, \qquad y^k(t) = \mathrm{e}^{\mathrm{i}k\omega t} \begin{pmatrix} z_1^k(t) \\ z_2^k(t) \end{pmatrix}, \tag{3.9}$$

where $y_i(t)$ and $z_i^k(t)$ are the solution of the system (3.1)–(3.2). The approximate solution $\widetilde{x}(t)$ of (3.5) is thus equal to $\sum_k y^k(t)$. Without any truncation of the series in (3.1)–(3.2), the functions $y^k(t)$ are formally a solution of

$$\ddot y^k + \Omega^2 y^k = \sum_{m\ge 0} \frac{1}{m!} \sum_{s(\alpha)=k,\alpha_i\ne 0} g^{(m)}(y^0)\big(y^{\alpha_1}, \ldots, y^{\alpha_m}\big), \tag{3.10}$$

because they are obtained by comparing the coefficients of $\mathrm{e}^{\mathrm{i}k\omega t}$ (see the proof of Lemma 2.1). Let us study here the effect of the truncation.

Theorem 3.2 *Consider the differential equation (2.1) with initial values $x(0)$ and $\dot x(0)$ satisfying (1.5). Assume that the nonlinearity $g(x)$ is analytic in the complex ball $\{(x_1, x_2) \mid \|x_1 - x_1(0)\| \le 4R, \|x_2\| \le 4R\}$ and bounded by M with $M \ge R$ and let $\|\dot y_{10}\| \le M$. Let the truncation index N in (3.1) and (3.2) be determined by (3.3). Then, there exist $\gamma > 0, T > 0$ and $\omega_0 > 0$ such that the defect*

$$\delta_k(t) = \ddot y^k(t) + \Omega^2 y^k(t) - \sum_{m\ge 0} \frac{1}{m!} \sum_{s(\alpha)=k,\alpha_i\ne 0} g^{(m)}(y^0(t))\big(y^{\alpha_1}(t), \ldots, y^{\alpha_m}(t)\big)$$

satisfies for $0 \le t \le T$ and for $\omega \ge \omega_0$

$$\sum_{k\in\mathbb{Z}} \|\delta_k(t)\| \le C\,M\,\mathrm{e}^{-\gamma\omega}.$$

The constants C, γ, T, ω_0 only depend on an upper bound of M/R but not on ω.

Proof. First we let N and ω be independent variables (for the time being not related by (3.3)), and we consider the defect as a function of t, N, and ω^{-1}, i.e., $\delta_k(t) = \delta_k(t, N, \omega^{-1})$. By the construction of the coefficient functions y^k, the defect δ_k is an analytic function of $\zeta = \omega^{-1}$ in a neighbourhood of the origin, and moreover, $\delta_k = \mathcal{O}(\omega^{-N-1})$. Therefore, the following function is analytic in a neighbourhood of the origin:

$$F(\zeta) = \sum_{|k|\le m} u_k^* \, \delta_k(t, N, \zeta)\, \zeta^{-(N+1)},$$

where m is an arbitrary integer, and the u_k are arbitrary vectors of unit norm. For $t \le T$, with T sufficiently small (see (3.8)), the function $F(\omega^{-1})$ is well defined for $|\omega^{-1}| \le \varepsilon_N$, where

$$\varepsilon_N := \frac{R}{2\nu M N} \ ,$$

so that the Maximum Principle can be applied on this disk. For $|\omega^{-1}| = \varepsilon_N$, i.e., for $|\omega|$ and N related like in (3.3) but with 2 instead of e in the denominator, the bounds (3.4) and (3.7) are still valid (except that the constant 8.65 increases to 12.4).

For $t \le T$, we have $\|y^0(t) - x(0)\| \le R$ and Cauchy's estimates yield

$$\begin{aligned}\sum_{k\in\mathbb{Z}} \ &\Big\| \sum_{m\ge 0} \frac{1}{m!} \sum_{s(\alpha)=k, \alpha_i\ne 0} g^{(m)}(y^0(t))\big(y^{\alpha_1}(t), \dots, y^{\alpha_m}(t)\big)\Big\| \\ &\le M \sum_{m\ge 0} \frac{1}{m!} \sum_{\alpha_i\ne 0} \dots \sum_{\alpha_m\ne 0} m!(3R)^{-m} \|y^{\alpha_1}\| \dots \|y^{\alpha_m}\| \le Const\cdot M.\end{aligned}$$

The last inequality is a consequence of (3.4) and (3.7), which yield

$$\sum_{\alpha\ne 0} \|y^\alpha\| \le Const\cdot M \cdot \omega^{-1}$$

which is smaller than $2R$ for $\omega \ge \omega_0$ (take ω_0 greater if necessary). Again by (3.4) and (3.7), we obtain

$$\sum_{k\in\mathbb{Z}} \|\ddot y^k + \Omega^2 y^k\| = \sum_{k\in\mathbb{Z}} \|\ddot z^k + 2ik\omega \dot z^k - k^2\omega^2 z^k + \Omega^2 z^k\| \le Const\cdot M.$$

Putting this together, we obtain the bound

$$\sum_{k\in\mathbb{Z}} \|\delta_k(t, N, \zeta)\| \le Const\cdot M \qquad \text{for} \quad |\zeta| = \varepsilon_N.$$

With the Maximum Principle, this gives for $|\omega^{-1}| \le \varepsilon_N$

$$\begin{aligned}|F(\omega^{-1})| \ &\le \ \max_{|\zeta|=\varepsilon_N} |F(\zeta)| \\ &\le \ \max_{|\zeta|=\varepsilon_N} \sum_{k\in\mathbb{Z}} \|\delta_k(t, N, \zeta)\| \cdot \varepsilon_N^{-(N+1)} \le Const\cdot M \cdot \varepsilon_N^{-(N+1)}.\end{aligned}$$

Choosing now $u_k = \delta_k(t,N,\omega^{-1})/\|\delta_k(t,N,\omega^{-1})\|$ in the definition of $F(\zeta)$ and letting $m \to \infty$ gives

$$\sum_{k\in\mathbb{Z}} \|\delta_k(t,N,\omega^{-1})\| \le Const \cdot M \cdot (\omega\varepsilon_N)^{-(N+1)}.$$

For ω and N related by (3.3) we have $(\omega\varepsilon_N)^{-1} \le 2/\mathrm{e} = \mathrm{e}^{-\alpha}$ with $\alpha = 1 - \ln 2 > 0$, so that in this case

$$\sum_{k\in\mathbb{Z}} \|\delta_k(t)\| \le Const \cdot M \cdot \mathrm{e}^{-\alpha(N+1)} \le Const \cdot M \cdot \mathrm{e}^{-\gamma\omega}$$

holds with the exponent $\gamma = \frac{\alpha R}{\nu M \mathrm{e}}$. □

4 The Hamiltonian Case

Sections 2 and 3 treated general second order differential equations with rapid oscillations. Our main interest is in Hamiltonian systems, where $g(x) = -\nabla U(x)$ and $U(x)$ is an analytic potential. The Hamiltonian $H(x,\dot{x})$ of the system (2.1) is then given by (1.2).

4.1 Hamiltonian of the Modulated Fourier Expansion

It is interesting to note that the Hamiltonian structure passes over to the differential equation for the coefficients of the modulated Fourier expansion. To see this, we let

$$\mathbf{y} = (..., y^{-2}, y^{-1}, y^0, y^1, y^2, ...)$$

be a two-sided infinite sequence, and we define

$$\mathcal{U}(\mathbf{y}) = U(y^0) + \sum_{m\ge 0} \frac{1}{m!} \sum_{s(\alpha)=0,\alpha_i\neq 0} U^{(m)}(y^0)\big(y^{\alpha_1},\dots,y^{\alpha_m}\big). \tag{4.1}$$

This function is well-defined as long as $\sum_{k\neq 0}\|y^k\| \le R$. The system (3.10) then becomes

$$\ddot{y}^k + \Omega^2 y^k \;=\; -\nabla_{y^{-k}}\mathcal{U}(\mathbf{y}) \tag{4.2}$$

and is Hamiltonian with

$$\mathcal{H}(\mathbf{y},\dot{\mathbf{y}}) \;=\; \frac{1}{2}\sum_{k\in\mathbb{Z}} \Big((\dot{y}^{-k})^T\dot{y}^k + (y^{-k})^T\Omega^2 y^k\Big) + \mathcal{U}(\mathbf{y}). \tag{4.3}$$

4.2 An Almost-Invariant Close to the Oscillatory Energy

It turns out that, besides the Hamiltonian $\mathcal{H}(\mathbf{y},\dot{\mathbf{y}})$ (see [6]), the system (4.2) also has

$$\mathcal{I}(\mathbf{y},\dot{\mathbf{y}}) \;=\; -\mathrm{i}\,\omega \sum_{k\neq 0} k (y^{-k})^T \dot{y}^k \tag{4.4}$$

as a conserved quantity. This series converges if $\sum_{k\neq 0} |k|\,\|y^k\| < \infty$ and $\max_{k\neq 0}\|\dot{y}^k\| < \infty$. For the functions $y^k(t)$ of (3.9), where $y_i(t)$ and $z_i^k(t)$ are the solution of the truncated system (3.1)-(3.2), this is a consequence of (3.4).

We shall prove here that the expression $\mathcal{I}\big(\mathbf{y}(t),\dot{\mathbf{y}}(t)\big)$ is conserved up to exponentially small terms. Moreover it turns out that this expression is close to the oscillatory energy

$$I(x,\dot{x}) = \frac{1}{2}\,\|\dot{x}_2\|^2 + \frac{\omega^2}{2}\,\|x_2\|^2 \tag{4.5}$$

of the system (2.1) with $g(x) = -\nabla U(x)$.

Theorem 4.1 *Let $\mathbf{y}(t)$ be the infinite vector with components $y^k(t)$ given by (3.9) and corresponding to initial values given by Lemma 3.1. Under the assumption of Theorem 3.2 we then have*

$$\begin{aligned}
\mathcal{I}\big(\mathbf{y}(t),\dot{\mathbf{y}}(t)\big) \;&=\; \mathcal{I}\big(\mathbf{y}(0),\dot{\mathbf{y}}(0)\big) + \mathcal{O}(\mathrm{e}^{-\gamma\omega})\\
\mathcal{I}\big(\mathbf{y}(t),\dot{\mathbf{y}}(t)\big) \;&=\; I\big(x(t),\dot{x}(t)\big) + \mathcal{O}(\omega^{-1})
\end{aligned}$$

for $0\le t\le T$ and $\omega\ge\omega_0$, where the constants symbolizing the $\mathcal{O}(\cdot)$ depend on E, M, and R, but not on ω.

Proof. We use the algebraic identity

$$\sum_{k\neq 0} \mathrm{i}\,k\,(y^k)^T \nabla \mathcal{U}_{y^k}(\mathbf{y}) \;=\; 0, \tag{4.6}$$

which holds for $\sum_{k\neq 0}|k|\,\|y^k\| < \infty$. For a proof we refer to [6] and [7, Sect. XIII.6.2].

We then compute the time-derivative of $\mathcal{I}\big(\mathbf{y}(t),\dot{\mathbf{y}}(t)\big)$ with $y(t)$ of (3.9):

$$\begin{aligned}
\frac{d}{dt}\mathcal{I}\big(\mathbf{y}(t),\dot{\mathbf{y}}(t)\big) \;&=\; -\mathrm{i}\,\omega\sum_{k\neq 0} k\,\dot{y}^{-k}(t)^T\dot{y}^k(t) - \mathrm{i}\,\omega\sum_{k\neq 0} k\,y^{-k}(t)^T\ddot{y}^k(t)\\
&=\; -\mathrm{i}\,\omega\sum_{k\neq 0} k\,y^{-k}(t)^T\Big(\ddot{y}^k(t) + \Omega^2 y^k(t) + \nabla\mathcal{U}_{y^{-k}}\big(\mathbf{y}(t)\big)\Big)\\
&=\; -\mathrm{i}\,\omega\sum_{k\neq 0} k\,y^{-k}(t)^T\,\delta_k(t).
\end{aligned}$$

We have used that the terms $k\,(\dot{y}^{-k})^T\dot{y}^k$ as well as $k\,(y^{-k})^T\Omega^2 y^k$ cancel with the corresponding terms for $-k$. Furthermore, we have added the expression (4.6) to make appear the defect in the right-hand expression. The first statement now follows from Theorem 3.2, and by an integration on the interval $[0, t]$.

The second statement is obtained as in the proof of Theorem 4.3 in [6]. □

4.3 Proof of Theorem 1.1

To prove the main theorem of this article, which states that (4.5) is nearly conserved over exponentially long time, we only have to use Theorem 4.1 and change the $\mathcal{O}(\omega^{-N})$ remainders by $\mathcal{O}(\mathrm{e}^{-\gamma\omega})$ in the proof of Corollary 4.4 in [6].

Acknowledgement

We are grateful to Francesco Fassò for interesting discussions on the long-time conservation of the oscillatory energy.

References

[1] D. Bambusi and A. Giorgilli, *Exponential stability of states close to resonance in infinite dimensional Hamiltonian systems*, J. Statist. Phys. 71 (1993) 569–606.

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

[3] G. Benettin, L. Galgani and A. Giorgilli, *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part II*, Commun. Math. Phys. 121 (1989) 557–601.

[4] D. Cohen, *Développement asymptotique de la solution d'une équation différentielle à grandes oscillations*, Diploma Thesis, University of Geneva, Mai 2000.

[5] F. Fassò, *Lie series method for vector fields and Hamiltonian perturbation theory*, Z. Angew. Math. Phys. 41 (1990) 843–864.

[6] E. Hairer and Ch. Lubich, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal. 38 (2001) 414–441.

[7] E. Hairer, Ch. Lubich and G. Wanner, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, 2002.

[8] W.L. Miranker and M. van Veldhuizen, *The method of envelopes*, Math. Comp. 32 (1978), no. 142, 453–496.

[9] J.A. Murdock, *Perturbations. Theory and Methods*, a Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1991.

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

[11] A.V. Pronin and D.V. Treschev, *Continuous averaging in multi-frequency slow-fast systems*, Regul. Chaotic Dyn. 5 (2000) 157–170.