On systems of differential equations with extrinsic oscillation

M. Condon, A. Deaño & A. Iserles

December 1, 2009

Abstract

We present a numerical scheme for an efficient discretization of nonlinear systems of differential equations subjected to highly oscillatory perturbations. This method is superior to standard ODE numerical solvers in the presence of high frequency forcing terms, and is based on asymptotic expansions of the solution in inverse powers of the oscillatory parameter ω, featuring modulated Fourier series in the expansion coefficients. Analysis of numerical stability and numerical examples are included.

1 Introduction

In this paper we are concerned with systems of ordinary differential equations (ODEs) of the form

$$\boldsymbol{y}'(t) = \boldsymbol{h}(\boldsymbol{y}(t)) + g_\omega(t)\boldsymbol{f}(\boldsymbol{y}(t)), \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0, \tag{1.1}$$

where $\boldsymbol{y}(t) : \mathbb{C} \to \mathbb{C}^d$, $\boldsymbol{f}(\boldsymbol{y}), \boldsymbol{h}(\boldsymbol{y}) : \mathbb{C}^d \to \mathbb{C}^d$ are analytic functions, and the scalar term $g_\omega(t)$ can be expressed as a modulated Fourier expansion (MFE), that is

$$g_\omega(t) = \sum_{m=-\infty}^{\infty} a_m(t)\mathrm{e}^{\mathrm{i}m\omega t} \tag{1.2}$$

– see (Cohen, Hairer & Lubich 2005, Hairer, Lubich & Wanner 2006, Sanz-Serna 2009) for applications of MFE in the theory of numerical analysis of Hamiltonian and oscillatory ODEs. Observe that we allow the coefficients $a_m(t)$ to depend on t. Therefore it is possible to think of the function $g_\omega(t)$ as periodic in the variable ωt, but not necessarily in t, which allows us to consider a wide range of different perturbations.

Typical examples of highly oscillatory forcing terms in this context are

$$g_\omega(t) = \mathrm{e}^{\mathrm{i}\omega t}, \quad g_\omega(t) = \mathrm{e}^{\eta\cos\omega t}. \tag{1.3}$$

The first example involves just a Fourier oscillator (and similarly one can consider $g_\omega(t) = \sin\omega t$ or $g_\omega(t) = \cos\omega t$), whereas the second appears in the modelling of diode and transistor circuits, see for instance (Dautbegovic, Condon & Brennan 2005). An essential difference between these two examples is that the first one is band limited and the other is not, which has an impact on the implementation of our method, as explained in Section 3.

This general problem also has a very immediate application in the context of communications and optical systems, which are governed by nonlinear differential and algebraic-differential equations (DAEs), see for example (Dautbegovic et al. 2005, Pedro & Carvalho 2002, Roychowdhury 2001). Also, oscillators are employed for frequency translation of information signals, channel selection or for synchronization. In this context, our problem corresponds with the situation where a physical system is subject to a high frequency forcing term (which could for instance be an electromagnetic wave or a mechanical excitation), and one wishes to analyse the behaviour of the system on a time scale which is much larger than the period of the forcing function.

It is also well known that this type of periodic forcing can be used to change intrinsic features of the system, see for instance an application to injection-locking of circuits in (Bartuccelli, Deane & Gentile 2009) or (O'Neill, Bourke, Ye & Kennedy 2005). This is of importance when the frequency of the external force is close to the natural frequency of the system, but investigation of oscillators subject to external perturbations which do not cause locking is also of interest (Lai & Roychowdhury 2004), for understanding the operation and bandwidth limitations of injection locked oscillators. The effect of perturbations of differing frequencies is also important for oscillator design and operation (Lai 2008). Furthermore, the effect of power supply interference or high-frequency signal interference on oscillator performance is an important phenomenon to be accounted for in design work (Lai, Zhu & Feng 2009).

The numerical challenge of this setting is underlied by the fact that in many communications systems it is common to employ high frequencies ω, which are in the MHz range and higher, in such a way that physical devices of practical dimensions can be employed. Also, a particular concern in the operation of electronic systems and oscillators is the effect of noise and spurious high-frequency signals, see (Demir, Mehrotra & Roychowdhury 2000). Such unwanted signals can result in undesired phenomena such as interchannel interference and timing jitter, see (Demir 2000).

All these applications are a formidable computational challenge since, when $\omega \gg 1$, the highly oscillatory nature of the solution imposes a very small stepsize on standard numerical methods for ODEs. The reason for the poor performance of classical numerical methods for solving ODEs in the presence of high oscillation lies at the very heart of the standard numerical theory, which is essentially based on Taylor expansion of the solution. In any numerical method of order p with step h, the error scales roughly like $h^{p+1}\boldsymbol{y}^{(p+1)}(t)$. Since the derivatives of highly oscillatory functions grow very fast, typically $\boldsymbol{y}^{(p+1)}(t) = \mathcal{O}(\omega^{p+1})$, we require h to be extremely small in order to keep the error down to an acceptable size. This usually renders standard numerical methods exceedingly expensive, see for instance (Condon, Deaño & Iserles 2009*a*).

An alternative is given by perturbation theory, albeit not in a completely standard form. The setting (1.1) belongs to the general framework of the so-called parametric perturbation or parametric modulation of differential equations. In the mathematical literature this has been a recurring topic in the field of perturbation theory, see for instance the classical reference (Bogoliubov & Mitropolsky 1961), (Verhulst 1990) or (Jordan & Smith 2007). However, most standard methods, such as averaging, are developed for systems where the perturbation is multiplied by a small parameter ε. In this context, the general idea is that the solution of the unperturbed system plus corrections in powers of ε may yield a good approximation to the solution of the perturbed problem.

In our case the perturbation is not necessarily small, but a somewhat related idea can be ap-

plied. Intuitively, for large values of ω, the exceedingly fast oscillations of $g_\omega(t)$ will produce massive cancellation between positive and negative parts of the forcing term $g_\omega(t)\boldsymbol{f}(\boldsymbol{y}(t))$, and will therefore give a small contribution, roughly speaking. More rigorously, if we consider the unperturbed system

$$\boldsymbol{z}'(t) = \boldsymbol{h}(\boldsymbol{z}(t)), \qquad \boldsymbol{z}(0) = \boldsymbol{y}_0, \tag{1.4}$$

then nonlinear variation of constants (Hairer, Nørsett & Wanner 1993) allows us to connect the two solutions

$$\boldsymbol{y}(t) - \boldsymbol{z}(t) = \int_0^t \boldsymbol{\Phi}(t-s)\boldsymbol{f}(\boldsymbol{y}(s))g_\omega(s)\,\mathrm{d}s,$$

where $\boldsymbol{\Phi}$ solves the so-called variational equation

$$\boldsymbol{\Phi}' = \frac{\partial \boldsymbol{h}(\boldsymbol{z}(t))}{\partial \boldsymbol{z}}\boldsymbol{\Phi}, \qquad \boldsymbol{\Phi}(0) = I.$$

The matrix $\boldsymbol{\Phi}$ may not be analytically available in general, but the important fact is that if the integrand is smooth enough and $g_\omega(t)$ is a trigonometric function (see the examples cited before), then integration by parts gives

$$\int_0^t \boldsymbol{\Phi}(t-s)\boldsymbol{f}(\boldsymbol{y}(s))g_\omega(s)\,\mathrm{d}s = \mathcal{O}(1/\omega), \qquad \omega \to \infty.$$

This motivates the fundamental ansatz that we propose later on, that is, that the solution $\boldsymbol{y}(t)$ admits an expansion in inverse powers of the oscillatory parameter ω. The terms in this expansion can be computed explicitly in a quite general setting, and they adopt the form of modulated Fourier expansions themselves. The expansion can be seen therefore as a correction (in inverse powers of ω this time) of the solution of the unperturbed system.

We remark that one cannot expect this approach to be satisfactory in general when the system exhibits chaotic behaviour, for example. It is known that a parametric perturbation can be used to take such a system into or out of chaotic behaviour, or to alter the chaotic states of the original problem, see for instance (Wu, Lu, Iu & Wong 2007), but the very high sensitivity of the problem to changes in the data will typically render useless any method based on perturbation theory.

The paper is organised as follows: in Sections 2.1 and 2.2 we present the basic features of our method and how to construct the expansion explicitly. Next, in Section 3 we analyse the bandwidth of the different terms in the expansion, depending on the original bandwidth of the forcing term $g_\omega(t)$. In Section 4 we pay special attention to the stability properties of the algorithm, which essentially depends, as expected, on the behaviour of the linearisation of the system around the nonoscillatory base function. We conclude the paper with several examples to illustrate the performance of the method.

2 Asymptotic-numerical solvers

2.1 Construction

Our basic *ansatz* in constructing asymptotic-numerical solvers is that the solution $\boldsymbol{y}(t)$ admits an asymptotic expansion in inverse powers of the oscillatory parameter ω:

$$\boldsymbol{y}(t) \sim \sum_{s=0}^{\infty} \frac{1}{\omega^s} \boldsymbol{\psi}_s(t) \qquad \omega \gg 1, \tag{2.1}$$

where the $\boldsymbol{\psi}_s(t)$s depend on ω, but $\boldsymbol{\psi}_s(t) \sim \mathcal{O}(1)$, $\omega \gg 1$, for $s \in \mathbb{Z}_+$. Given the structure of the original ODE, a reasonable assumption is that each $\boldsymbol{\psi}_s(t)$ in (2.1), except when $s = 0$, has itself the form of a modulated Fourier expansion,

$$\boldsymbol{\psi}_s(t) = \sum_{m=-\infty}^{\infty} \boldsymbol{p}_{s,m}(t) \mathrm{e}^{\mathrm{i}m\omega t}, \qquad s \geq 1. \tag{2.2}$$

Furthermore, $\boldsymbol{\psi}_0(t) = \boldsymbol{p}_{0,0}(t)$ is independent of ω, i.e. $\boldsymbol{p}_{0,m}(t) \equiv \boldsymbol{0}$ when $m \neq 0$. In order to satisfy the initial condition of the differential equation, we impose $\boldsymbol{\psi}_0(0) = \boldsymbol{y}(0) = \boldsymbol{y}_0$, which means that $\boldsymbol{\psi}_s(0) = \boldsymbol{0}$ for $s \geq 1$, or equivalently

$$\sum_{m=-\infty}^{\infty} \boldsymbol{p}_{0,m}(0) = \boldsymbol{y}_0, \qquad \sum_{m=-\infty}^{\infty} \boldsymbol{p}_{s,m}(0) = \boldsymbol{0}, \quad s \geq 1.$$

Thus our basic *ansatz* reads

$$\boldsymbol{y}(t) \sim \boldsymbol{p}_{0,0}(t) + \sum_{s=1}^{\infty} \frac{1}{\omega^s} \sum_{m=-\infty}^{\infty} \boldsymbol{p}_{s,m}(t) \mathrm{e}^{\mathrm{i}m\omega t},$$

and differentiation term by term gives, formally,

$$\boldsymbol{y}'(t) \sim \boldsymbol{p}'_{0,0}(t) + \sum_{s=1}^{\infty} \frac{1}{\omega^s} \sum_{m=-\infty}^{\infty} \left[\boldsymbol{p}'_{s,m}(t) + \mathrm{i}m\omega \boldsymbol{p}_{s,m}(t)\right] \mathrm{e}^{\mathrm{i}m\omega t}.$$

Since $\boldsymbol{h}, \boldsymbol{f} : \mathbb{C}^d \to \mathbb{C}^d$ are analytic functions, we can expand them into Taylor series around $\boldsymbol{\psi}_0$, assuming that this term will yield the main contribution,

$$\boldsymbol{h}(\boldsymbol{\psi}_0 + \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \frac{1}{n!} \boldsymbol{h}_n(\boldsymbol{\psi}_0, \overbrace{\boldsymbol{\theta}, \boldsymbol{\theta}, \ldots, \boldsymbol{\theta}}^{n \text{ times}}).$$

Here $\boldsymbol{h}_m$ is an n-tensor related to the n-th derivative of $\boldsymbol{h}$ at $\boldsymbol{\psi}_0$,

$$\begin{aligned}
\boldsymbol{h}_0(\boldsymbol{\psi}_0) &= \boldsymbol{h}(\boldsymbol{\psi}_0), \\
\boldsymbol{h}_1(\boldsymbol{\psi}_0, \boldsymbol{\theta}) &= \frac{\partial \boldsymbol{h}(\boldsymbol{\psi}_0)}{\partial \boldsymbol{y}} \boldsymbol{\theta}, \\
(\boldsymbol{h}_2(\boldsymbol{\psi}_0, \boldsymbol{\theta}, \boldsymbol{\theta}))_r &= \sum_{i=1}^{d} \sum_{j=1}^{d} \theta_i \frac{\partial^2 h_r(\boldsymbol{\psi}_0)}{\partial y_i \partial y_j} \theta_j, \qquad r = 1, 2, \ldots, d,
\end{aligned}$$

etc. In general we have

$$(\boldsymbol{h}_n(\boldsymbol{\psi}_0,\boldsymbol{\theta},\dots,\boldsymbol{\theta}))_r = \sum_{i_1=1}^{d}\cdots\sum_{i_n=1}^{d}\frac{\partial^n h_r(\boldsymbol{\psi}_0)}{\partial y_{i_1}\cdots\partial y_{i_n}}\theta_{i_1}\theta_{i_2}\cdots\theta_{i_n}, \qquad r=1,2,\dots,d.$$

Note that each $\boldsymbol{h}_n(\boldsymbol{\psi}_0,\boldsymbol{\theta},\dots,\boldsymbol{\theta})$ is linear in each of the $\boldsymbol{\theta}_k$s. It is clear that if the functions $\boldsymbol{h}$ and $\boldsymbol{f}$ are not analytic, but say C^r, we can still consider the first few terms of the expansion and adapt everything accordingly.

In the sequel, we will use the following notation,

$$\boldsymbol{h}(\boldsymbol{y}) \sim \boldsymbol{h}(\boldsymbol{p}_{0,0}) + \sum_{s=1}^{\infty}\frac{1}{\omega^s}\sum_{n=1}^{s}\frac{1}{n!}\sum_{\boldsymbol{k}\in\mathbb{I}_{n,s}}\boldsymbol{h}_n(\boldsymbol{p}_{0,0},\boldsymbol{\chi}_{k_1},\cdots,\boldsymbol{\chi}_{k_n}), \tag{2.3}$$

where

$$\boldsymbol{\chi}_k(t) = \sum_{m=-\infty}^{\infty}\boldsymbol{p}_{k,m}(t)\mathrm{e}^{\mathrm{i}m\omega t} \tag{2.4}$$

and

$$\mathbb{I}_{n,s} = \{(k_1,\dots,k_n)\in\mathbb{N}^n \,:\, |\boldsymbol{k}| = s\}, \tag{2.5}$$

with the standard notation for multi-indices $|\boldsymbol{k}| = k_1 + k_2 + \ldots + k_n$. A similar formula applies to the function $\boldsymbol{f}$. Furthermore, we can express this expansion in the following way,

$$\boldsymbol{h}(\boldsymbol{y}) \sim \boldsymbol{h}(\boldsymbol{p}_{0,0}) + \sum_{s=1}^{\infty}\frac{1}{\omega^s}\sum_{n=1}^{s}\frac{1}{n!}\sum_{m=-\infty}^{\infty}\mathrm{e}^{\mathrm{i}m\omega t}\sum_{\boldsymbol{k}\in\mathbb{I}_{n,s}}\sum_{\boldsymbol{l}\in\mathbb{K}_{n,m}}\boldsymbol{h}_n(\boldsymbol{p}_{0,0},\boldsymbol{p}_{k_1,l_1},\cdots,\boldsymbol{p}_{k_n,l_n}), \tag{2.6}$$

where

$$\mathbb{K}_{n,m} = \{(l_1,\dots,l_n)\in\mathbb{Z}^n \,:\, |\boldsymbol{l}| = m\}. \tag{2.7}$$

Observe that for simplicity of notation, we have omitted the dependence on t of the different terms $\boldsymbol{p}_{s,m}$ in the expansion.

Putting all the ingredients together, we can equate both terms and we obtain

$$\begin{aligned}
&\boldsymbol{p}'_{0,0} + \sum_{s=1}^{\infty}\frac{1}{\omega^s}\sum_{m=-\infty}^{\infty}\left[\boldsymbol{p}'_{s,m} + \mathrm{i}m\omega\boldsymbol{p}_{s,m}\right]\mathrm{e}^{\mathrm{i}m\omega t}\\
&= \boldsymbol{h}(\boldsymbol{p}_{0,0}) + \sum_{s=1}^{\infty}\frac{1}{\omega^s}\sum_{n=1}^{s}\frac{1}{n!}\sum_{m=-\infty}^{\infty}\mathrm{e}^{\mathrm{i}m\omega t}\sum_{\boldsymbol{k}\in\mathbb{I}_{n,s}}\sum_{\boldsymbol{l}\in\mathbb{K}_{n,m}}\boldsymbol{h}_n(\boldsymbol{p}_{0,0},\boldsymbol{p}_{k_1,l_1},\cdots,\boldsymbol{p}_{k_n,l_n})\\
&\quad + g_\omega(t)\left[\boldsymbol{f}(\boldsymbol{p}_{0,0}) + \sum_{s=1}^{\infty}\frac{1}{\omega^s}\sum_{n=1}^{s}\frac{1}{n!}\sum_{m=-\infty}^{\infty}\mathrm{e}^{\mathrm{i}m\omega t}\sum_{\boldsymbol{k}\in\mathbb{I}_{n,s}}\sum_{\boldsymbol{l}\in\mathbb{K}_{n,m}}\boldsymbol{f}_n(\boldsymbol{p}_{0,0},\boldsymbol{p}_{k_1,l_1},\cdots,\boldsymbol{p}_{k_n,l_n})\right].
\end{aligned}$$

Now we identify coefficients in two levels: first we consider orders of magnitude (inverse powers of ω), and then frequencies (values of m) within each order of magnitude. The first level (corresponding to $s=0$) is clear:

$$\boldsymbol{p}'_{0,0} + \mathrm{i}\sum_{m=-\infty}^{\infty}m\boldsymbol{p}_{1,m}\mathrm{e}^{\mathrm{i}m\omega t} = \boldsymbol{h}(\boldsymbol{p}_{0,0}) + \sum_{m=-\infty}^{\infty}a_m(t)\mathrm{e}^{\mathrm{i}m\omega t}\boldsymbol{f}(\boldsymbol{p}_{0,0}). \tag{2.8}$$

Separation of frequencies yields a differential equation for the coefficient $\boldsymbol{p}_{0,0}(t)$:

$$\boldsymbol{p}'_{0,0} = \boldsymbol{h}(\boldsymbol{p}_{0,0}) + a_0(t)\boldsymbol{f}(\boldsymbol{p}_{0,0}), \tag{2.9}$$

together with the initial condition $\boldsymbol{p}_{0,0}(0) = \boldsymbol{y}(0) = \boldsymbol{y}_0$, and additionally

$$\boldsymbol{p}_{1,m} = -\frac{\mathrm{i}a_m(t)}{m}\boldsymbol{f}(\boldsymbol{p}_{0,0}), \qquad m \neq 0. \tag{2.10}$$

Observe that the ODE for $\boldsymbol{p}_{0,0}$ is nonoscillatory. Hence, even if it is not solvable explicitly (due to the nonlinear terms $\boldsymbol{h}$ and $\boldsymbol{f}$), it can be efficiently solved using standard numerical methods. Note also that the components $\boldsymbol{p}_{1,m}(t)$ depend on t in general, even if the coefficients $a_m(t)$ are independent of t (that is, if the forcing term $g_\omega(t)$ is periodic and has a classical Fourier expansion).

For $s \geq 1$ and $m \in \mathbb{Z}$, we have

$$\boldsymbol{p}'_{s,m} + \mathrm{i}m\boldsymbol{p}_{s+1,m} = \boldsymbol{b}_{s,m}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{s,m-r}[\boldsymbol{f}], \tag{2.11}$$

where

$$\boldsymbol{b}_{s,m}[\boldsymbol{h}] = \sum_{n=1}^{s} \frac{1}{n!} \sum_{\boldsymbol{k}\in\mathbb{I}_{n,s}} \sum_{\boldsymbol{l}\in\mathbb{K}_{n,m}} \boldsymbol{h}_n(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{k_1,l_1}, \cdots, \boldsymbol{p}_{k_n,l_n}), \qquad m \in \mathbb{Z}, \tag{2.12}$$

and similarly with $\boldsymbol{b}_{s,m}[\boldsymbol{f}]$. Note that once again we have omitted the dependence on t for brevity. Observe that for $s \geq 1$ we have on the one hand the ODE

$$\boldsymbol{p}'_{s,0} = \boldsymbol{b}_{s,0}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{s,-r}[\boldsymbol{f}] \tag{2.13}$$

with initial condition

$$\boldsymbol{p}_{s,0}(0) = -\sum_{m\neq 0} \boldsymbol{p}_{s,m}(0), \tag{2.14}$$

since we have imposed $\boldsymbol{\psi}_s(0) = \boldsymbol{0}$, and on the other hand the recursion

$$\boldsymbol{p}_{s+1,m} = -\frac{\mathrm{i}}{m}\left[-\boldsymbol{p}'_{s,m} + \boldsymbol{b}_{s,m}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{s,m-r}[\boldsymbol{f}]\right] \tag{2.15}$$

for $m \neq 0$. This is the pattern that we will find for each value of $s \geq 1$. From a computational perspective, an alternative to this scheme would be to solve a systems of DAEs involving all the terms $\boldsymbol{p}_{s,m}(t)$ up to the desired value of s. In this context, the analysis of bandwidth presented in Section 3 is relevant to determine the number of terms that we need within each level.

2.2 Explicit form of the asymptotic expansion

In this section we derive explicitly the first few terms of our asymptotic expansion. We first note that there is an important simplification in the above formulas when the number of $a_r(t)$ terms that are different from 0 is finite (in other words when the input function $g_\omega(t)$ is band limited). As an example, we illustrate the case where the perturbation is of the form $g_\omega(t) = \mu \sin \omega t$. In that case it is clear that $a_{-1}(t) = \mathrm{i}\mu/2$, $a_1(t) = -\mathrm{i}\mu/2$ and $a_m(t) \equiv 0$ if $|m| \neq 1$. Thus, the original bandwidth is $\varrho = 1$. Similar results hold for the case $g_\omega(t) = \mu \cos \omega t$.

2.2.1 The zeroth term

As explained before, the term $\boldsymbol{p}_{0,0}(t)$ obeys the differential equation

$$\boldsymbol{p}'_{0,0} = \boldsymbol{h}(\boldsymbol{p}_{0,0}) + a_0(t)\boldsymbol{f}(\boldsymbol{p}_{0,0}),$$

together with the initial condition $\boldsymbol{p}_{0,0}(0) = \boldsymbol{y}(0) = \boldsymbol{y}_0$. Equating terms with the same frequency, we obtain additionally

$$\boldsymbol{p}_{1,m} = -\frac{\mathrm{i}a_m(t)}{m}\boldsymbol{f}(\boldsymbol{p}_{0,0}), \qquad m \neq 0.$$

With $g_\omega(t) = \mu \sin \omega t$, the previous ODE can be reduced to

$$\boldsymbol{p}'_{0,0} = \boldsymbol{h}(\boldsymbol{p}_{0,0}), \qquad \boldsymbol{p}_{0,0}(0) = \boldsymbol{0},$$

and also

$$\boldsymbol{p}_{1,1} = \boldsymbol{p}_{1,-1} = -\frac{\mu}{2}\boldsymbol{f}(\boldsymbol{p}_{0,0}), \tag{2.16}$$

together with $\boldsymbol{p}_{1,m}(t) \equiv 0$ when $|m| \geq 2$.

2.2.2 The first term

When $s = 1$ we obtain

$$\boldsymbol{p}'_{1,0} = \boldsymbol{b}_{1,0}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{1,-r}[\boldsymbol{f}]. \tag{2.17}$$

Here

$$\boldsymbol{b}_{1,m}[\boldsymbol{h}] = \boldsymbol{h}_1(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{1,m}), \qquad m \in \mathbb{Z},$$

and similarly for $\boldsymbol{f}$. Additionally we have the initial condition

$$\boldsymbol{p}_{1,0}(0) = -\sum_{m\neq 0} \boldsymbol{p}_{1,m}(0),$$

which follows from $\boldsymbol{\psi}_1(0) = \boldsymbol{0}$. Furthermore, we get

$$\boldsymbol{p}_{2,m} = -\frac{\mathrm{i}}{m}\left[-\boldsymbol{p}'_{1,m} + \boldsymbol{b}_{1,m}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{1,m-r}[\boldsymbol{f}]\right], \qquad m \neq 0. \tag{2.18}$$

When $g_\omega(t) = \mu \sin \omega t$ then (2.17), the differential equation for $\boldsymbol{p}_{1,0}$, reads

$$\boldsymbol{p}'_{1,0} = \boldsymbol{b}_{1,0}[\boldsymbol{h}] + \frac{\mathrm{i}\mu}{2}\left(\boldsymbol{b}_{1,1}[\boldsymbol{f}] - \boldsymbol{b}_{1,-1}[\boldsymbol{f}]\right) = \boldsymbol{h}_1(\boldsymbol{p}_{0,0}, p_{1,0}) = J[\boldsymbol{h}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,0}.$$

Here $J[\boldsymbol{h}](\boldsymbol{p}_{0,0})$ is the Jacobian matrix of $\boldsymbol{h}$ evaluated at $\boldsymbol{p}_{0,0}$. It is not difficult to check that $\boldsymbol{b}_{1,1}[\boldsymbol{f}] - \boldsymbol{b}_{1,-1}[\boldsymbol{f}] = \boldsymbol{0}$, because of the symmetry of the coefficients $\boldsymbol{p}_{1,-1}$ and $\boldsymbol{p}_{1,1}$. Since we impose $\boldsymbol{\psi}_1(0) = \boldsymbol{0}$, the initial condition for $\boldsymbol{p}_{1,0}(t)$ is

$$\boldsymbol{p}_{1,0}(0) = -\sum_{m\neq 0} \boldsymbol{p}_{1,m}(0) = -\boldsymbol{p}_{1,-1}(0) - \boldsymbol{p}_{1,1}(0) = \mu\boldsymbol{f}(\boldsymbol{p}_{0,0}).$$

Putting all the information together, and taking into account (2.16), we obtain

$$\boldsymbol{\psi}_1(t) = \boldsymbol{p}_{1,0} - \mu \boldsymbol{f}(\boldsymbol{p}_{0,0}) \cos \omega t. \tag{2.19}$$

Furthermore, we can compute $\boldsymbol{p}_{2,m}$ for $m \neq 0$ from (2.18). According to Section 3, the predicted bandwidth is equal to $2\varrho = 2$, and therefore we only need to compute the coefficients $\boldsymbol{p}_{2,\pm 1}$ and $\boldsymbol{p}_{2,\pm 2}$. Thus

$$\begin{aligned}\boldsymbol{p}_{2,1} &= \mathrm{i}\boldsymbol{p}'_{1,1} - \mathrm{i}\boldsymbol{b}_{1,1}[\boldsymbol{h}] - \mathrm{i}\left(a_{-1}(t)\boldsymbol{b}_{1,2}[\boldsymbol{f}] + a_1(t)\boldsymbol{b}_{1,0}[\boldsymbol{f}]\right) \\ &= -\frac{\mathrm{i}\mu}{2}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{f}(\boldsymbol{p}_{0,0}) - \mathrm{i}J[\boldsymbol{h}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,1} - \frac{\mu}{2}J[\boldsymbol{f}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,0}\end{aligned}$$

since $\boldsymbol{b}_{1,2}[\boldsymbol{f}](t) \equiv 0$.

Because of symmetry of the $\boldsymbol{p}_{1,m}$ coefficients, we deduce that

$$\boldsymbol{p}_{2,-1} = \frac{\mathrm{i}\mu}{2}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{f}(\boldsymbol{p}_{0,0}) + \mathrm{i}J[\boldsymbol{h}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,-1} - \frac{\mu}{2}J[\boldsymbol{f}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,0}.$$

Now

$$\boldsymbol{p}_{2,2} = -\frac{\mathrm{i}}{2}\left[-\boldsymbol{p}'_{1,2} + \boldsymbol{b}_{1,2}[\boldsymbol{h}] + a_{-1}(t)\boldsymbol{b}_{1,3}[\boldsymbol{f}] + a_1(t)\boldsymbol{b}_{1,1}[\boldsymbol{f}]\right] = -\frac{\mu}{4}J[\boldsymbol{f}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,1},$$

and clearly $\boldsymbol{p}_{2,-2}(t) = \boldsymbol{p}_{2,2}(t)$. Therefore, we have

$$\begin{aligned}\boldsymbol{\psi}_2(t) = \boldsymbol{p}_{2,0} &+ \left[\mu\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{f}(\boldsymbol{p}_{0,0}) + 2J[\boldsymbol{h}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,1}\right]\sin\omega t - \mu J[\boldsymbol{f}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,0}\cos\omega t \\ &- \frac{\mu}{2}J[\boldsymbol{f}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,1}\cos 2\omega t\end{aligned} \tag{2.20}$$

The equation and initial conditions for $\boldsymbol{p}_{2,0}$ are obtained when analysing the $\mathcal{O}(\omega^{-2})$ terms.

2.2.3 The second term

When $s = 2$ we obtain

$$\boldsymbol{p}'_{2,0} = \boldsymbol{b}_{2,0}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{2,-r}[\boldsymbol{f}]. \tag{2.21}$$

Here

$$\boldsymbol{b}_{2,m}[\boldsymbol{h}] = \boldsymbol{h}_1(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{2,m}) + \frac{1}{2}\sum_{l=-\infty}^{\infty}\boldsymbol{h}_2(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{1,l}, \boldsymbol{p}_{1,m-l}), \qquad m \in \mathbb{Z},$$

and similarly for $\boldsymbol{f}$, together with

$$\boldsymbol{p}_{2,0}(0) = -\sum_{m\neq 0}\boldsymbol{p}_{2,m}(0),$$

and

$$\boldsymbol{p}_{3,m} = -\frac{\mathrm{i}}{m}\left[-\boldsymbol{p}'_{2,m} + \boldsymbol{b}_{2,m}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r \boldsymbol{b}_{2,m-r}[\boldsymbol{f}]\right], \qquad m \neq 0. \tag{2.22}$$

When $g_\omega(t) = \mu \sin \omega t$ then we have the following ODE for $\boldsymbol{p}_{2,0}(t)$, in accordance with (2.21):

$$\boldsymbol{p}'_{2,0} = \boldsymbol{b}_{2,0}[\boldsymbol{h}] + \frac{\mathrm{i}\mu}{2}\left(\boldsymbol{b}_{2,1}[\boldsymbol{f}] - \boldsymbol{b}_{2,-1}[\boldsymbol{f}]\right), \tag{2.23}$$

where we have used $a_1(t) = -a_{-1}(t)$. Now

$$\begin{aligned}\boldsymbol{b}_{2,0}[\boldsymbol{h}] &= \boldsymbol{h}_1(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{2,0}) + \frac{1}{2}\sum_{l=-\infty}^{\infty} \boldsymbol{h}_2(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{1,l}, \boldsymbol{p}_{1,-l}) \\ &= \boldsymbol{h}_1(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{2,0}) + \frac{1}{2}\boldsymbol{h}_2(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{1,0}, \boldsymbol{p}_{1,0}) + \boldsymbol{h}_2(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{1,1}, \boldsymbol{p}_{1,-1})\end{aligned}$$

Next, we have

$$\boldsymbol{b}_{2,1}[\boldsymbol{f}] = \boldsymbol{f}_1(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{2,1}) + \boldsymbol{f}_2(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{1,0}, \boldsymbol{p}_{1,1}).$$

The difference $\boldsymbol{b}_{2,1}[\boldsymbol{f}] - \boldsymbol{b}_{2,-1}[\boldsymbol{f}]$ can be simplified using the symmetry of the coefficients:

$$\begin{aligned}\boldsymbol{b}_{2,1}[\boldsymbol{f}] - \boldsymbol{b}_{2,-1}[\boldsymbol{f}] &= \boldsymbol{f}_1(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{2,1}) - \boldsymbol{f}_1(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{2,-1}) \\ &= -J[\boldsymbol{f}](\boldsymbol{p}_{0,0})\left[\mathrm{i}\mu\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{f}(\boldsymbol{p}_{0,0}) + 2\mathrm{i}J[\boldsymbol{h}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,-1}\right].\end{aligned}$$

Furthermore, we have the initial condition

$$\boldsymbol{p}_{2,0}(0) = -\sum_{m\neq 0} \boldsymbol{p}_{2,m}(0).$$

Equation (2.22) gives the coefficients $\boldsymbol{p}_{3,m}(t)$. Notice that now the bandwidth is $3\varrho = 3$, so we only need these terms when $|m| \leq 3$.

It is clear that the process can be iterated, at the price of increasingly cumbersome expressions. We also note that in many relevant cases the functions $\boldsymbol{h}$ and $\boldsymbol{f}$ are quite simple, for example multivariate polynomials of low degree, and hence many terms involving high order derivatives vanish identically. We omit any further steps for brevity.

3 Bandwidth and blossoming

There is an important observation to be made regarding the bandwidth of the different functions $\boldsymbol{\psi}_s(t)$. Due to the nonlinearity of the original equation, the number of nonzero frequencies (i.e. values of m) increases as we move to higher values of s. We call this phenomenon *blossoming*, and it is relevant when discussing efficiency issues, since it quantifies the number of terms that we need to compute in each step of the algorithm.

Let us suppose that there exists $\varrho \in \mathbb{N}$ such that $\boldsymbol{a}_m \equiv \mathbf{0}$ when $|m| \geq \varrho{+}1$. In other words, we suppose that the forcing term $g_\omega(t)$ is band limited. Let θ_s be the maximum bandwidth of the term $\boldsymbol{\psi}_s(t)$, that is

$$\theta_s := \max\{m \in \mathbb{Z} : \boldsymbol{p}_{s,|m|} \neq \mathbf{0}\}.$$

It can be checked that the first few values are

$$\theta_0 = 0, \qquad \theta_1 = \varrho.$$

For the general case, we can prove the following result.

Theorem 1 *For $s \geq 0$, the maximum bandwidth θ_s of the term $\boldsymbol{\psi}_s(t)$ is*

$$\theta_s = s\varrho.$$

Proof We shall use formula (2.15),

$$\boldsymbol{p}_{s+1,m}(t) = -\frac{\mathrm{i}}{m}\left[-\boldsymbol{p}'_{s,m}(t) + \boldsymbol{b}_{s,m}[\boldsymbol{h}](t) + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{s,m-r}[\boldsymbol{f}](t)\right].$$

We note that differentiation does not alter the bandwidth, and also that the component $\boldsymbol{b}_{s,m}[\boldsymbol{h}](t)$ is a combination of terms of the form

$$\sum_{k\in\mathbb{I}_{n,s}} \sum_{l\in\mathbb{K}_{n,m}} \boldsymbol{h}_n(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{k_1,l_1}, \ldots, \boldsymbol{p}_{k_n,l_n}).$$

Since

$$s = \sum_{j=1}^{n} k_j,$$

we obtain by induction on s that

$$\sum_{j=1}^{n} \theta_{k_j} = \varrho \sum_{j=1}^{n} k_j = s\varrho.$$

This is the maximum bandwidth contributed by the term $\boldsymbol{b}_{s,m}[\boldsymbol{h}](t)$. However, when $m = (s+1)\varrho$ and $r = \varrho$ then

$$a_r(t)\boldsymbol{b}_{s,m-r}[\boldsymbol{f}](t) = a_\varrho(t)\boldsymbol{b}_{s,s\varrho}[\boldsymbol{f}](t)$$

is different from 0 in general, since for example it contains the element

$$\boldsymbol{f}_s(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{1,\varrho}, \ldots, \boldsymbol{p}_{1,\varrho}).$$

If we try $m > (s+1)\varrho$, then $m - r > s\varrho$, and according to the result given before we get $\boldsymbol{b}_{s,m-r}[\boldsymbol{f}](t) \equiv 0$. A similar argument can be used with negative frequencies, setting $m = -(s+1)\varrho$ and $r = -\varrho$. Therefore, the maximum total bandwidth of $\boldsymbol{\psi}_{s+1}(t)$ is indeed $(s+1)\varrho$.

□

It is important to observe that this result corresponds to the maximal possible bandwidth, and that the actual one could well be smaller. For instance, if the function $\boldsymbol{f}$ is constant (which corresponds to the case of a system of ODEs with an oscillatory forcing term), then $\boldsymbol{b}_{s,m}[\boldsymbol{f}]$ is identically 0 for $s \geq 1$. Therefore, we would have

$$\boldsymbol{p}_{s+1,m}(t) = -\frac{\mathrm{i}}{m}\left[-\boldsymbol{p}'_{s,m}(t) + \boldsymbol{b}_{s,m}[\boldsymbol{h}](t)\right], \qquad s \geq 1. \tag{3.24}$$

The corresponding result is the following.

Theorem 2 *Let $\boldsymbol{f}(\boldsymbol{y})$ be constant, then we have $\theta_0 = 0$, $\theta_1 = \varrho$ and*

$$\theta_s = (s-1)\varrho, \qquad s \geq 2.$$

Proof We shall use formula (3.24) and proceed along the lines of the proof of Theorem 4.1 in (Condon, Deaño & Iserles 2009*b*). We recall that we have a combination of terms of the form

$$\sum_{k \in \mathbb{I}_{n,s}} \sum_{l \in \mathbb{K}_{n,m}} \boldsymbol{h}_n(\boldsymbol{p}_{0,0}, \boldsymbol{p}_{k_1,l_1}, \ldots, \boldsymbol{p}_{k_n,l_n}),$$

and we write

$$\{k_1, \ldots, k_n\} = \{\beta_1, \ldots, \beta_{n_1}\} \cup \{\gamma_1, \ldots, \gamma_{n_2}\},$$

where $\beta_1 = \ldots = \beta_{n_1} = 1$. We thus have

$$n = n_1 + n_2, \qquad s = \sum_{j=1}^{n} k_j = n_1 + \sum_{j=1}^{n_2} \gamma_j.$$

By induction,

$$\sum_{j=1}^{n} \theta_{k_j} = n_1\varrho + \sum_{j=1}^{n_2} \theta_{\gamma_j} = n_1\varrho + \varrho \sum_{j=1}^{n_2} (\gamma_j - 1) = s\varrho - n_2\varrho \leq s\varrho,$$

with equality when $n_2 = 0$, which corresponds to the case of the multi-index $\{(1, 1, \ldots, 1)\}$. □

Further reduction of the actual bandwidth can occur in some situations, particularly if higher order derivatives of $\boldsymbol{h}$ and $\boldsymbol{f}$ vanish identically. This is the case if these functions are (multivariate) polynomials, which is typical in many applications. Compare with the analysis in (Condon et al. 2009*b*), where blossoming appears at a slower rate, though consistent with the results presented above.

4 Stability

As a direct consequence of the construction of the method in the previous subsections, one obtains $\boldsymbol{y}(t) - \boldsymbol{\psi}_0(t) = \mathcal{O}(1/\omega)$, where $\boldsymbol{y}(t)$ is the solution of the perturbed system and $\boldsymbol{\psi}_0(t)$ is the solution of the unperturbed one.

More precisely, suppose that $\boldsymbol{y}(t)$ is a solution of the perturbed system

$$\boldsymbol{y}' = \boldsymbol{h}(\boldsymbol{y}) + g_\omega(t)\boldsymbol{f}(\boldsymbol{y}), \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0,$$

and consider the unperturbed system

$$\boldsymbol{\psi}_0'(t) = \boldsymbol{h}(\boldsymbol{\psi}_0), \qquad \boldsymbol{\psi}_0(0) = \boldsymbol{y}_0,$$

Then, if we write $\boldsymbol{y}(t) = \boldsymbol{\psi}_0(t) + \boldsymbol{w}(t)$, assuming that $\boldsymbol{w}(t) = \mathcal{O}(1/\omega)$ for large ω, we obtain to leading order

$$\boldsymbol{w}' = [A(t) + B(t)g_\omega(t)]\,\boldsymbol{w} + g_\omega(t)\boldsymbol{f}(\boldsymbol{\psi}_0), \qquad \boldsymbol{w}(0) = \boldsymbol{0}, \tag{4.25}$$

where $A(t)$ and $B(t)$ are respectively the Jacobian matrices of $\boldsymbol{h}$ and $\boldsymbol{f}$ evaluated at $\boldsymbol{\psi}_0(t)$. This is nothing but a local linearisation of the solution around $\boldsymbol{\psi}_0$. Now we split the matrices $A(t) + B(t)g_\omega(t) = U + V(t)$, where U is a constant matrix, and compare (4.25) with the system

$$\boldsymbol{z}' = U\boldsymbol{z}, \qquad \boldsymbol{z}(0) = \boldsymbol{0},$$

with trivial solution $\boldsymbol{z} \equiv \boldsymbol{0}$. Using standard variation of constants it follows that

$$\boldsymbol{w}(t) = \boldsymbol{0} + \int_0^t \Phi(t-s)\boldsymbol{F}(s)\,\mathrm{d}s = \int_0^t \mathrm{e}^{(t-s)U}\boldsymbol{F}(s)\,\mathrm{d}s, \tag{4.26}$$

where $\Phi(t) = \mathrm{e}^{tU}$ is the fundamental matrix of the system and

$$\boldsymbol{F}(s) = V(s)\boldsymbol{w}(s) + g_\omega(s)\boldsymbol{f}(\boldsymbol{\psi}_0(s)).$$

It is clear from (4.26) that $\boldsymbol{w}(t)$ represents a deviation from the zero solution $\boldsymbol{v}(t)$, and therefore the behaviour of $\boldsymbol{w}(t)$ as $t > 0$ is related to the stability of this zero solution. This in turn is governed by the eigenvalues of the fundamental matrix U, see for instance (Verhulst 1990, Ch. 6). More explicitly, we can state the following result.

Theorem 3 *If*

- *all the eigenvalues of the matrix C, say λ_k, $k = 1, \ldots, d$, satisfy that $\operatorname{Re}\lambda_k \leq 0$, and those eigenvalues with zero real part are simple, and*
- *it is true that for any $t > 0$ there exist constants $c_1, c_2 > 0$ such that*

$$\int_0^t \|V(s)\|\,\mathrm{d}s < c_1, \qquad \int_0^t \|g_\omega(s)\boldsymbol{f}(\boldsymbol{\psi}_0(s))\|\,\mathrm{d}s < c_2,$$

then the zero solution solution $\boldsymbol{v}(t)$ is stable in the sense of Lyapunov, and $\boldsymbol{w}(t)$ is bounded.

Proof The proof follows along the lines of the one given in (Verhulst 1990, Th. 6.2). Since

$$\boldsymbol{w}(t) = \int_0^t \Phi(t-s)V(s)\boldsymbol{w}(s)\,\mathrm{d}s + \int_0^t \Phi(t-s)g_\omega(s)\boldsymbol{f}(\boldsymbol{\psi}_0(s))\,\mathrm{d}s,$$

we get the immediate bound

$$\|\boldsymbol{w}\| \leq \int_0^t \|\Phi(t-s)\| \cdot \|V(s)\| \cdot \|\boldsymbol{w}(s)\|\,\mathrm{d}s + \int_0^t \|\Phi(t-s)\| \cdot \|g_\omega(s)\boldsymbol{f}(\boldsymbol{\psi}_0(s))\|\,\mathrm{d}s,$$

Because of the first condition, we have $\|\Phi\| < c_3$, for a certain constant c_3, and therefore

$$\|\boldsymbol{w}\| \leq c_3 \int_0^t \|V(s)\| \cdot \|\boldsymbol{w}(s)\| \, \mathrm{d}s + c_3 c_2,$$

Applying Gronwall's inequality, we get

$$\|\boldsymbol{w}\| \leq c_3 c_2 \mathrm{e}^{c_3 \int_0^t \|V(s)\| \, \mathrm{d}s} \leq c_3 c_2 \mathrm{e}^{c_3 c_1}.$$

□

It is possible to obtain a stronger result if we impose strict negativity of the real part of all the eigenvalues λ_k of the matrix U. In that case, one has the bound

$$\|\Phi(t)\| \leq c_4 \mathrm{e}^{-\nu t},$$

for suitable positive constants c_4 and ν, see (Verhulst 1990). Thus the zero solution $\boldsymbol{z}(t)$ is asymptotically stable and $\boldsymbol{w}(t)$ will tend exponentially fast to $\mathbf{0}$ with t. This fact, together with large values of ω will make our method very effective.

In the particular case where the system has a linear part, that is

$$\boldsymbol{y}' = M\boldsymbol{y} + g_\omega(t)\boldsymbol{f}(\boldsymbol{y}), \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0,$$

for some constant matrix $M \in \mathbb{C}^{n \times n}$, then clearly $M = C$ and stability is determined by the eigenvalues of the linear part of the system.

If the conditions of the theorem are not met, we expect the solution $\boldsymbol{w}(t)$ to grow unboundedly in t, and then the difference between $\boldsymbol{y}(t)$ and $\boldsymbol{\psi}_0(t)$ can be very large. A typical example of this situation occurs when the system exhibits chaotic behaviour (like in the case of the Lorenz system), and as such is very sensitive to small perturbations in the data. However, we remark that from a computational perspective, if the size of the eigenvalues is moderate, it may happen that the difference between the perturbed solution $\boldsymbol{y}(t)$ and the unperturbed one $\boldsymbol{\psi}_0(t)$ is small enough to be acceptable when ω is large. Some examples further on will illustrate this last point.

5 Examples

In this section we present several examples that illustrate the construction and properties of the expansion that we have presented in previous sections. In all cases we will compare the approximation given by the first few terms of the asymptotic-numerical solver with the exact solution (which is either analytically available or computed numerically with standard MATLAB routines up to prescribed accuracy). We will normally use the standard ODE solver `ode45` in MATLAB, with an absolute and relative tolerance equal to 10^{-12}.

We stress that the values of ω that we use are much smaller than the ones normally present in applications. This restriction is essentially imposed by the fact that the comparison with the exact solution should be reliable and affordable. Increasing ω will benefit the asymptotic-numerical solver, since the approximation with a fixed number of terms is more accurate, while the computational cost is roughly similar.

We use the notation

$$e_s = \left| \boldsymbol{y}(t) - \sum_{n=0}^{s} \frac{\boldsymbol{\psi}_n(t)}{\omega^n} \right|, \qquad s \geq 0,$$

for the errors, taken componentwise.

5.1 A linear system

As a first example, we can consider a simple forced oscillator with damping. This system is modelled by a simple second order ODE:

$$x''(t) + bx'(t) + kx(t) = \mu \cos \omega t, \qquad x(0) = x_0, \qquad x'(0) = x'_0,$$

where b is the damping coefficient, k the spring constant and we have set the mass to 1 for simplicity. We introduce a forcing term with amplitude μ and frequency ω, and we assume that $\omega \gg \omega_0$, where ω_0 is the natural frequency of the unperturbed oscillator in the underdamped case. In a matrix form

$$\boldsymbol{x}'(t) = \begin{bmatrix} 0 & 1 \\ -k & -b \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mu \cos \omega t, \tag{5.1}$$

thereby, using our notation,

$$\boldsymbol{h}(\boldsymbol{x}) = \begin{bmatrix} 0 & 1 \\ -k & -b \end{bmatrix} \boldsymbol{x}, \qquad \boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

In this case, since the system is already linear, the matrix U is directly

$$U = \begin{bmatrix} 0 & 1 \\ -k & -b \end{bmatrix},$$

with eigenvalues

$$\lambda_{\pm} = \frac{-b \pm \sqrt{b^2 - 4k}}{2}.$$

Since both $b, k > 0$, the real parts of both eigenvalues are always negative and we have asymptotic stability according to Section 4. The construction of the asymptotic expansion is particularly simple in this case, since we have after brief computation

$$\boldsymbol{p}'_{0,0} = U\boldsymbol{p}_{0,0}, \qquad \boldsymbol{p}_{0,0}(0) = \boldsymbol{x}(0),$$

together with

$$\boldsymbol{p}_{1,-1} = \frac{\mathrm{i}\mu}{2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\boldsymbol{p}_{1,1}.$$

Actually, because of the function $\boldsymbol{f}$ being constant, we have from (2.13) and (2.14) that for $s \geq 1$

$$\boldsymbol{p}'_{s,0} = U\boldsymbol{p}_{s,0}, \qquad \boldsymbol{p}_{s,0}(0) = -\sum_{m \neq 0} \boldsymbol{p}_{s,m}(0).$$

In particular, that means that

$$\boldsymbol{p}'_{1,0} = U\boldsymbol{p}_{1,0}, \qquad \boldsymbol{p}_{s,0}(0) = \mathbf{0},$$

which leads to $\boldsymbol{p}_{1,0} \equiv \mathbf{0}$. Hence we conclude that in this case the first term is

$$\boldsymbol{\psi}_1(t) = \begin{bmatrix} 0 \\ \mu \end{bmatrix} \sin \omega t.$$

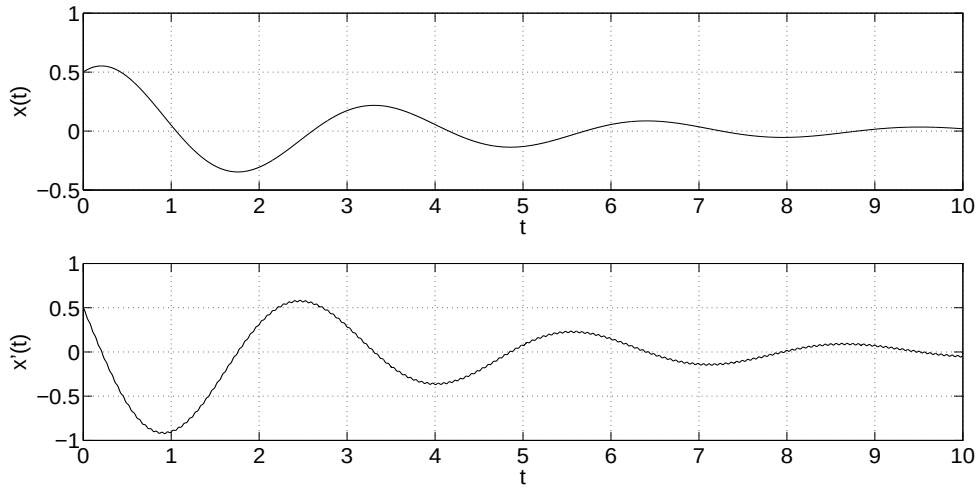

Figure 5.1: Solution of the perturbed system (5.1) for $\omega = 100$.

Note that the order $\mathcal{O}(1/\omega)$ term is 0 for the first component of the solution. In other words, $x(t)$ is superimposed with tiny oscillations of amplitude $\mathcal{O}(1/\omega^2)$, whereas in the case of the derivative $x'(t)$ these are of order $\mathcal{O}(1/\omega)$. This is intuitively consistent with what can be observed in Figure 5.1. Analogously,

$$\boldsymbol{p}_{2,-1} = -\frac{\mu}{2} \begin{bmatrix} 1 \\ -b \end{bmatrix} = \boldsymbol{p}_{2,1},$$

and since the bandwidth in this example is $\theta_2 = 1$, see Section 3, we know that $\boldsymbol{p}_{2,m} \equiv \mathbf{0}$ if $|m| > 1$. Furthermore

$$\boldsymbol{p}'_{2,0} = U\boldsymbol{p}_{2,0}, \qquad \boldsymbol{p}_{2,0}(0) = -2\boldsymbol{p}_{2,1}(0) = \mu \begin{bmatrix} 1 \\ -b \end{bmatrix},$$

hence

$$\boldsymbol{\psi}_2(t) = \boldsymbol{p}_{2,0} + \mu \begin{bmatrix} -1 \\ b \end{bmatrix} \cos \omega t.$$

In Figures 5.1 and 5.2 we plot the solution of the perturbed system with parameters $k = 4.2$, $b = 0.6$, $\mu = 0.8$, $\omega = 100$ and initial values $x(0) = x'(0) = 1/2$, and the errors when we compare with the first terms of the approximation, respectively.

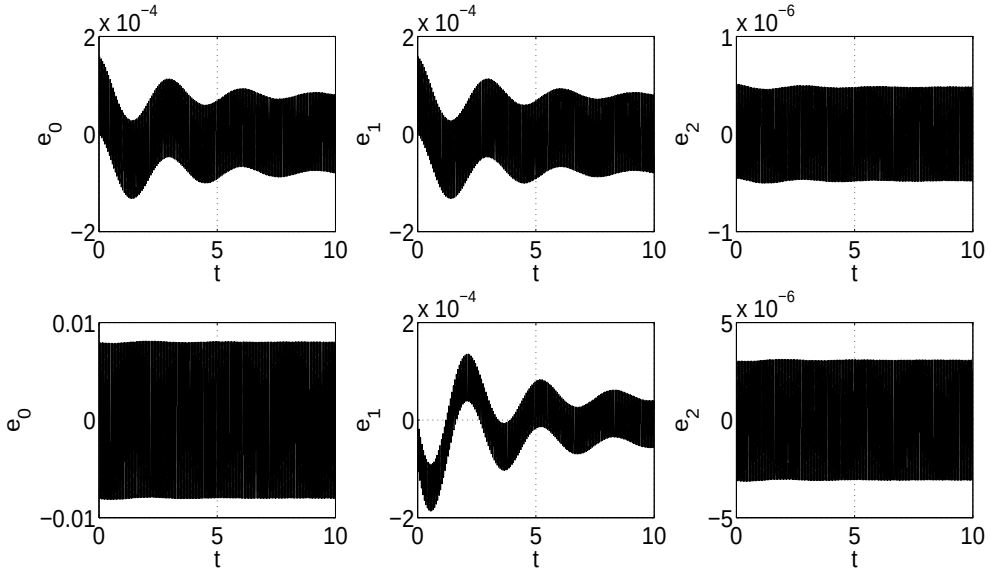

Figure 5.2: Absolute errors in the approximation of the solution of the perturbed system (5.1) for $k = 4.2$, $b = 0.6$, $\mu = 0.8$ and $\omega = 100$. Top row, errors in $x(t)$ using the zeroth term (left), using up to the first term of the approximation (centre) and using up to the second term of the approximation (right). Bottom row, same for the derivative $x'(t)$.

5.2 A model for an injection-locked frequency divider

Next we consider a nonlinear example, given by the following system, which is used in (O'Neill et al. 2005) and (Bartuccelli et al. 2009),

$$C\frac{\mathrm{d}V_C}{\mathrm{d}t} = I_L + f(V_C), \qquad L\frac{\mathrm{d}I_L}{\mathrm{d}t} = -RI_L - V_C,$$

where

$$f(V_C) = AV_C\left(1 - \frac{V_C^2}{V_{dd}^2}\right),$$

and A, V_{dd}, C, L and R are parameters of the system. A periodic perturbation can be introduced as follows,

$$f(V_C, t) = (A + B\sin\Omega t)V_C\left(1 - \frac{V_C}{V_{dd}^2}\right), \qquad A > 0, \quad B \in \mathbb{R}.$$

After normalisation and scaling, this system can be written as

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \alpha v + \Phi(t)u(1-u^2), \qquad \frac{\mathrm{d}v}{\mathrm{d}t} = -u - v, \tag{5.2}$$

where

$$\alpha = \frac{L}{R^2C}, \quad \beta = \frac{LA}{RC}, \quad \mu = \frac{LB}{RC}, \quad \Phi(t) = \beta + \mu\sin\omega t, \quad \omega = \frac{\Omega L}{R}$$

Hence, using our notation, we have

$$\boldsymbol{h}(u,v) = \begin{bmatrix} \alpha v + \beta u(1-u^2) \\ -u - v \end{bmatrix},$$

and

$$\boldsymbol{f}(u,v) = \begin{bmatrix} u(1-u^2) \\ 0 \end{bmatrix},$$

together with $g_\omega(t) = \mu \sin \omega t$, and thus $a_{-1}(t) = \mathrm{i}\mu/2$ and $a_1(t) = -\mathrm{i}\mu/2$.

The first term of the expansion, following (2.9), solves the system

$$\boldsymbol{p}'_{0,0} = \boldsymbol{h}(\boldsymbol{p}_{0,0}),$$

since $a_0 \equiv 0$, and we also get from (2.10)

$$\boldsymbol{p}_{1,-1} = -\frac{\mu}{2}\boldsymbol{f}(\boldsymbol{p}_{0,0}) = \boldsymbol{p}_{1,1},$$

together with $\boldsymbol{p}_{1,m} \equiv \mathbf{0}$ when $|m| > 1$. From (2.17), the coefficient $\boldsymbol{p}_{1,0}$ satisfies the ODE

$$\boldsymbol{p}'_{1,0} = \boldsymbol{b}_{1,0}[\boldsymbol{h}] + a_{-1}\boldsymbol{b}_{1,1}[\boldsymbol{f}] + a_1\boldsymbol{b}_{1,-1}[\boldsymbol{f}] = \boldsymbol{b}_{1,0}[\boldsymbol{h}] + a_{-1}\left(\boldsymbol{b}_{1,1}[\boldsymbol{f}] - \boldsymbol{b}_{1,-1}[\boldsymbol{f}]\right) = \boldsymbol{b}_{1,0}[\boldsymbol{h}],$$

because of the parity of the coefficients $\boldsymbol{p}_{1,\pm 1}$. Here

$$\boldsymbol{b}_{1,0}[\boldsymbol{h}] = J[\boldsymbol{h}](\boldsymbol{p}_{0,0})\boldsymbol{p}_{1,0}.$$

Furthermore, the initial condition is

$$\boldsymbol{p}_{1,0}(0) = -\boldsymbol{p}_{1,-1}(0) - \boldsymbol{p}_{1,1}(0) = \mu\boldsymbol{f}(\boldsymbol{p}_{0,0}(0)).$$

Putting everything together, we have

$$\boldsymbol{\psi}_1(t) = \boldsymbol{p}_{1,0} - \mu\boldsymbol{f}(\boldsymbol{p}_{0,0})\cos\omega t.$$

The term $\boldsymbol{\psi}_2(t)$ can be obtained from (2.20) substituting all the data corresponding to the functions $\boldsymbol{h}$ and $\boldsymbol{f}$ in this example.

In the following example, we have taken the following values of the parameters,

$$A = 2.5 \times 10^{-3}, \quad B = 1 \times 10^{-3}, \quad V_{dd} = 9, \quad R = 3.06, \quad L = 2.2 \times 10^{-4}.$$

We vary the parameters C and Ω in the examples, but C is always positive and small and Ω large enough so that the scaled frequency $\omega = \Omega L/R$ is large. We observe that the matrix U in this case is

$$U = \begin{bmatrix} \beta & \alpha \\ -1 & -1 \end{bmatrix},$$

with eigenvalues

$$\lambda_\pm = \frac{\beta - 1 \pm \sqrt{(\beta+1)^2 - 4\alpha}}{2}.$$

With the parameters given (in particular because $L \ll 1$), the term $(\beta+1)^2 - 4\alpha$ is negative, and we have complex eigenvalues. This is consistent with the observed oscillatory

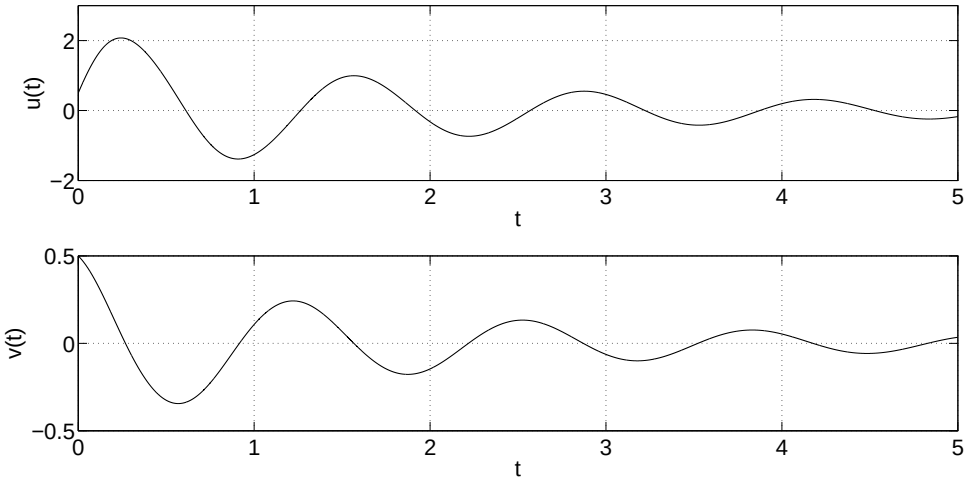

Figure 5.3: Solution of the perturbed system (5.2) for $C = 10^{-6}$ and $\Omega = 2\pi \times 10^6$

behaviour of the solutions, see Figure 5.3. The real part of the eigenvalues is given by $\frac{1}{2}(\beta-1)$, and for stability we need $\beta - 1 < 0$, see Section 4. Thus

$$\beta - 1 < 0 \qquad \Leftrightarrow \qquad LA < RC,$$

which holds for roughly $C \geq 10^{-7}$.

In Figure 5.4 we plot the solution of the unperturbed system and the absolute errors when using the first terms of the expansions when $C = 10^{-6}$, $\Omega = 2\pi \times 10^6$ (which gives $\omega = 451.73$ after scaling) and initial values $u(0) = v(0) = 1/2$.

Similar results, with smaller errors, are obtained if we consider larger values of Ω. If we use $C = 10^{-8}$ then the eigenvalues of the matrix U are $\lambda_\pm = 8.49 \pm 47.53\mathrm{i}$, therefore we do not expect numerical stability in this case. Indeed, Figure 5.5 shows that no significant improvement is obtained when adding more terms in the expansion. However, it is worth noting that the errors are quite small on the whole interval of integration, so the numerical solution might be acceptable, depending on the required accuracy.

5.3 An expcos oscillator

A more complicated example features a linear part plus a perturbation of the form $g_\omega(t) = \mathrm{e}^{\mu \cos \omega t}$. For example, in the modelling of diode circuits with inductive loads, we would find an equation of the form

$$x'(t) = -\frac{L}{RC}x(t) + \frac{I_s L}{C}\left[\exp\left(\frac{g_\omega(t) - x(t)}{V_T}\right) - 1\right] - \frac{L}{C}y(t), \qquad y'(t) = x(t),$$

where L, R, C, I_s and V_T are parameters. We will take the values $L = 10^{-4}$, $R = 100$, $C = 10^{-6}$, $I_s = 10^{-12}$ and $V_T = 0.0259$. The forcing term is $g_\omega(t) = \mu \cos \omega t$, with large

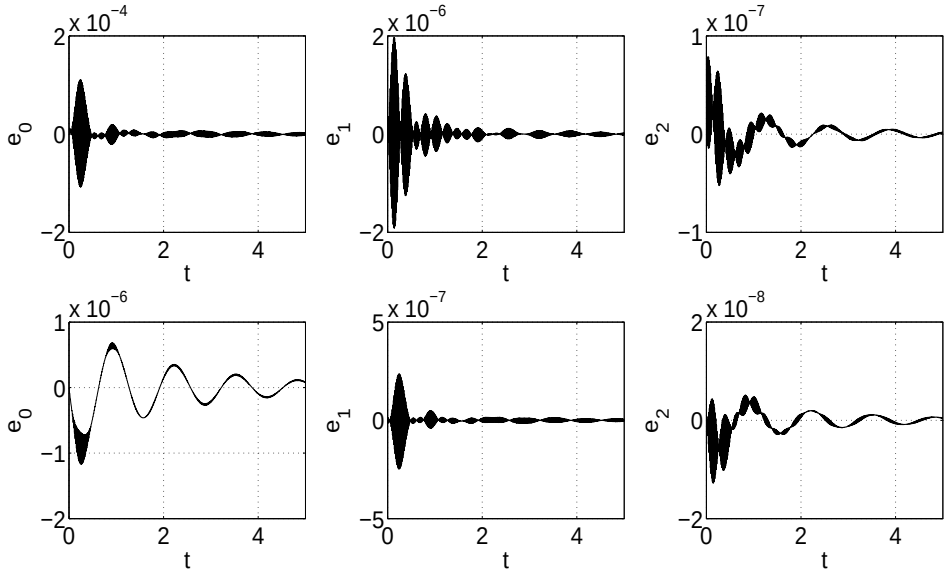

Figure 5.4: Absolute errors in the approximation of the solution of the perturbed system (5.2) for $C = 10^{-6}$ and $\Omega = 2\pi \times 10^6$. Top row, errors in $u(t)$ using the zeroth term (left), using up to the first term of the approximation (centre) and using up to the second term of the approximation (right). Bottom row, same for $v(t)$.

ω. The constant term $-I_s L/C$ can be added in the $\mathcal{O}(1)$ level in a straightforward way. Thus the resulting system is

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix}' = \begin{bmatrix} -L/RC & -L/C \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} \beta \mathrm{e}^{-x(t)/V_T} \\ 0 \end{bmatrix} \exp\left[\frac{g_\omega(t)}{V_T}\right] - \begin{bmatrix} \beta \\ 0 \end{bmatrix}, \quad (5.3)$$

where $\beta = I_s L/C$. The properties of similar types of oscillator have been analysed in (Condon et al. 2009*a*) and (Condon, Deaño, Iserles, Maczyński & Xu 2009*c*). The relevant fact is that this function can be expanded in Fourier series using modified Bessel functions, see (Abramowitz & Stegun 1964, Eq. 9.6.34)

$$\mathrm{e}^{\mu \cos \omega t} = I_0(\mu) + 2 \sum_{m=1}^{\infty} I_m(\mu) \cos m\omega t, \quad (5.4)$$

and the asymptotic behaviour of the modified Bessel functions for large orders guarantees convergence for fixed values of μ and t.

It is clear that the coefficients are $a_m = I_m(\mu)$ for $m \in \mathbb{Z}$, using the fact that for integer orders $I_m(\mu) = I_{-m}(\mu)$, see (Abramowitz & Stegun 1964, Eq. 9.6.6). The base equation follows from (2.9),

$$\boldsymbol{p}_{0,0} = \boldsymbol{h}(\boldsymbol{p}_{0,0}) + I_0(\mu)\boldsymbol{f}(\boldsymbol{p}_{0,0}), \qquad \boldsymbol{p}_{0,0}(0) = \boldsymbol{x}(0),$$

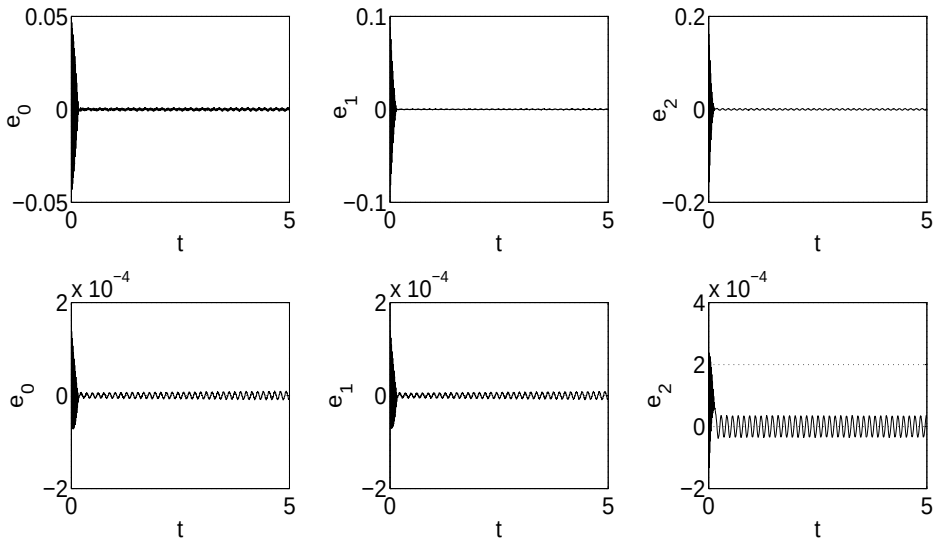

Figure 5.5: Same as Figure 5.4 but with parameters $C = 10^{-8}$ and $\Omega = 2\pi \times 10^6$, in the unstable regime.

and also

$$\boldsymbol{p}_{1,m} = -\frac{\mathrm{i}I_m(\mu)}{m}\boldsymbol{f}(\boldsymbol{p}_{0,0}), \qquad m \neq 0.$$

The differential equation for $\boldsymbol{p}_{0,0}$ cannot be analytically solved because of the nonlinearity originating in the function $\boldsymbol{f}$, but it is nonoscillatory and therefore amenable to numerical solution using standard methods.

The differential equation for $\boldsymbol{p}_{1,0}$ apparently involves an infinite number of terms,

$$\boldsymbol{p}'_{1,0} = \boldsymbol{b}_{1,0}[\boldsymbol{h}] + \sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{1,-r}[\boldsymbol{f}],$$

however

$$\sum_{r=-\infty}^{\infty} a_r(t)\boldsymbol{b}_{1,-r}[\boldsymbol{f}] = I_0(\mu)\boldsymbol{b}_{1,0}[\boldsymbol{f}] + \mathrm{i}\sum_{r\neq 0}\frac{I_r^2(\mu)}{r}\boldsymbol{f}(\boldsymbol{p}_{0,0}) = I_0(\mu)\boldsymbol{b}_{1,0}[\boldsymbol{f}],$$

the last sum being zero because of the symmetry of the modified Bessel functions with respect to the order. Hence

$$\boldsymbol{p}'_{1,0} = \boldsymbol{b}_{1,0}[\boldsymbol{h}] + I_0(\mu)\boldsymbol{b}_{1,0}[\boldsymbol{f}] = \left(J[\boldsymbol{h}] + I_0(\mu)J[\boldsymbol{f}]\right)\boldsymbol{p}_{1,0}, \qquad \boldsymbol{p}_{1,0}(0) = \boldsymbol{0},$$

which implies that $\boldsymbol{p}_{1,0} \equiv \boldsymbol{0}$. Therefore,

$$\boldsymbol{\psi}_1(t) = 2\boldsymbol{f}(\boldsymbol{p}_{0,0})\sum_{m=1}^{\infty}\frac{I_m(\mu)}{m}\sin m\omega t.$$

Observe that this last sum converges as well due to the rapid decay of the modified Bessel functions, so its numerical implementation is not problematic.

The second term $\boldsymbol{\psi}_2(t)$ can be computed from the general setting, although the different sums involving modified Bessel functions that appear can be quite expensive to evaluate. In particular, observe that $\mu = 1/V_T$, so if $V_T = 0.0259$ then μ is moderately large and the convergence of the series (5.4) can be slow. As a compensation, we note that our expansion has two important advantages, namely that its cost is essentially independent of ω (whereas any standard numerical method will need to reduce the stepsize considerably when ω grows), and that large values of ω will yield a more accurate expansion with the same number of terms.

5.4 A Lorenz-type system

A final example is provided by the Lorenz-type system, see (Chang 2009):

$$\left\{ \begin{array}{rcl} \dot{x} & = & \sigma(y-x) \\ \dot{y} & = & x(\rho - z) - y \\ \dot{z} & = & xy - \beta z \end{array} \right.$$

where σ, ρ and β are given parameters. Usual values of the parameters are $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$, for which the system exhibits chaotic behaviour and develops a strange attractor.

Thus $\boldsymbol{h} : \mathbb{R}^3 \to \mathbb{R}^3$, namely

$$\boldsymbol{h}(x,y,z) = \left[\begin{array}{c} \sigma(y-x) \\ x(\rho - z) - y \\ xy - \beta z \end{array} \right]$$

The perturbation considered in (Chang 2009) is given by

$$\boldsymbol{f}(x,y,z) = \left[\begin{array}{c} 25(y-x) \\ -35x + 29y \\ -z/3 \end{array} \right],$$

together with (using our notation) $g_\omega(t) = \frac{1}{2} \pm \frac{1}{2}\cos\omega t$. In this way $a_0 = \frac{1}{2}$, $a_{-1}(t) = a_1(t) = \frac{1}{4}$ and $a_m(t) \equiv 0$ otherwise.

Alternatively, in (Wu et al. 2007) one has

$$\boldsymbol{f}(x,y,z) = \left[\begin{array}{c} 0 \\ cy \\ 0 \end{array} \right],$$

and $g_\omega(t) = \varepsilon \sin\omega t$. It follows that $a_{-1}(t) = \mathrm{i}\varepsilon/2$, $a_1(t) = -\mathrm{i}\varepsilon/2$ and $a_m(t) \equiv 0$ when $|m| \neq 1$.

Yet another perturbation of this system is given in (Choe, Hohne, Benner & Kivshar 2005), where

$$\boldsymbol{f}(x,y,z) = \left[\begin{array}{c} 0 \\ \rho x \\ 0 \end{array} \right],$$

together with $g_\omega(t) = k\cos\omega t$, whereby $a_1(t) = a_{-1}(t) = k/2$, $a_m(t) \equiv 0$ if $|m| \neq 1$. Thus the original bandwidth is $\varrho = 1$.

The computations can be carried out in a similar way to the previous examples. Away from the chaotic regime the approximation gives good results, but when the perturbation is used for chaos suppression then one should not expect the procedure to be accurate, since the unperturbed solution will be chaotic and the perturbed one will not be so.

Acknowledgements

A. Deaño acknowledges financial support from the Spanish Ministry of Education under project MTM2006-09050. The material is based upon works supported by Science Foundation Ireland under Principal Investigator Grant No. 05/IN.1/I18.

References

Abramowitz, M. & Stegun, I. A. (1964), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Vol. 55 of *National Bureau of Standards Applied Mathematics Series*, U.S. Government Printing Office, Washington.

Bartuccelli, M. V., Deane, J. H. B. & Gentile, G. (2009), 'Frequency locking in the injection-locked frequency divider equation', *Proc. Royal Soc. A* **465**, 283–306.

Bogoliubov, N. N. & Mitropolsky, Y. A. (1961), *Asymptotic methods in the theory of non-linear oscillations*, Hindustani Publishing Corp.

Chang, W.-D. (2009), 'Digital secure communications via chaotic system', *Digital Signal Processing* **19**, 693–699.

Choe, C.-U., Hohne, K., Benner, H. & Kivshar, Y. S. (2005), 'Chaos suppression in the parametrically driven Lorenz system', *Phys. Review E* **72**, 036206.

Cohen, D., Hairer, E. & Lubich, C. (2005), 'Modulated Fourier expansions of highly oscillatory differential equations', *Found. Comput. Maths* **3**, 327–450.

Condon, M., Deaño, A. & Iserles, A. (2009*a*), 'On highly oscillatory problems arising in electronic engineering', *ESAIM: Math. Modelling & Num. Anal.* **43**, 785–804.

Condon, M., Deaño, A. & Iserles, A. (2009*b*), On second order differential equations with highly oscillatory forcing terms, Technical Report DAMTP NA2009/07, University of Cambridge.

Condon, M., Deaño, A., Iserles, A., Maczyński, K. & Xu, T. (2009*c*), 'On numerical methods for highly oscillatory problems in circuit simulation', *To appear in COMPEL*.

Dautbegovic, E., Condon, M. & Brennan, C. (2005), 'An efficient nonlinear circuit simulation technique', *IEEE Trans. Microwave Theory & Applics* **53**, 548–555.

Demir, A. (2000), 'Phase noise and timing jitter in oscillators with colored-noise sources', *IEEE Transactions Circuits & systems I: Fundamental Theory & Applics* **49**, 1782–1791.

Demir, A., Mehrotra, A. & Roychowdhury, J. (2000), 'Phase noise in oscillators: A unifying theory and numerical methods for characterization', *IEEE Trans. Circuits & Systems I: Fundamental Theory & Applics* **47**, 655–674.

Hairer, E., Lubich, C. & Wanner, G. (2006), *Geometric Numerical Integration*, 2nd edn, Springer Verlag.

Hairer, E., Nørsett, S. & Wanner, G. (1993), *Solving Ordinary Differential Equations I. Non-stiff Problems*, Springer Verlag.

Jordan, D. W. & Smith, P. (2007), *Nonlinear differential equations*, 4th edn, Oxford University Press.

Lai, X. (2008), Frequency aware PPV: A robust phase macromodel for accurate oscillator noise analysis, *in* 'Proceedings of the 2008 IEEE/ACM International Conference on Computer-Aided Design', pp. 803–806.

Lai, X. & Roychowdhury, J. (2004), 'Capturing oscillator injection locking via nonlinear phase domain macromodels', *IEEE Trans. MTT* **52**, 2251–2261.

Lai, X., Zhu, Y. & Feng, D. (2009), Steady state simulation and noise analysis for driven oscillators, *in* 'IEEE Microwave Symposium 2009', pp. 1489–1492.

O'Neill, D., Bourke, D., Ye, Z. & Kennedy, M. P. (2005), Accurate modeling and experimental validation of an injection-locked frequency divider, *in* 'Proceedings 2005 European Conference on Circuit Theory and Design', pp. 409–412.

Pedro, J. C. & Carvalho, N. (2002), 'Simulation of RF circuits driven by modulated signals without bandwidth constraints', *Proc. IEEE Internl Symp. Microwave Theory & Tech.* **3**, 2173–2176.

Roychowdhury, J. (2001), 'Analyzing circuits with widely separated time scales using numerical PDE methods', *IEEE Trans. on Circuits & Systems - I: Fundamental Theory & Applics* **48**, 578–594.

Sanz-Serna, J. (2009), 'Modulated Fourier expansions and heterogeneous multiscale methods', *IMA J. Num. Anal.* **29**(3), 595–605.

Verhulst, F. (1990), *Nonlinear Differential Equations and Dynamical Systems*, Springer Verlag.

Wu, X., Lu, J.-A., Iu, H. H. C. & Wong, S.-C. (2007), 'Suppression and generation of chaos for a three-dimensional autonomous system using parametric perturbations', *Chaos, Solitons and Fractals* **31**, 811–819.