Available online at www.sciencedirect.com
ScienceDirect
Physica D 237 (2008) 1029–1052
PHYSICA D
www.elsevier.com/locate/physd

Analysis of a renormalization group method and normal form theory for perturbed ordinary differential equations

R.E. Lee DeVille[a], Anthony Harkin[b], Matt Holzer[c], Krešimir Josić[d], Tasso J. Kaper[c,*]

[a] *Department of Mathematics, University of Illinois, Urbana-Champaign, IL, 61801, United States*
[b] *Department of Mathematics, Rochester Institute of Technology, Rochester, NY 14623, United States*
[c] *Department of Mathematics and Statistics, Boston University, Boston, MA, 02215, United States*
[d] *Department of Mathematics, University of Houston, Houston, TX 77204-3008, United States*

Received 14 November 2006; received in revised form 18 December 2007; accepted 22 December 2007
Available online 24 January 2008

Communicated by Y. Nishiura

Abstract

For singular perturbation problems, the renormalization group (RG) method of Chen, Goldenfeld, and Oono [Phys. Rev. E. 49 (1994) 4502–4511] has been shown to be an effective general approach for deriving reduced or amplitude equations that govern the long time dynamics of the system. It has been applied to a variety of problems traditionally analyzed using disparate methods, including the method of multiple scales, boundary layer theory, the WKBJ method, the Poincaré–Lindstedt method, the method of averaging, and others. In this article, we show how the RG method may be used to generate normal forms for large classes of ordinary differential equations. First, we apply the RG method to systems with autonomous perturbations, and we show that the reduced or amplitude equations generated by the RG method are equivalent to the classical Poincaré–Birkhoff normal forms for these systems up to and including terms of $\mathcal{O}(\epsilon^2)$, where ϵ is the perturbation parameter. This analysis establishes our approach and generalizes to higher order. Second, we apply the RG method to systems with nonautonomous perturbations, and we show that the reduced or amplitude equations so generated constitute time-asymptotic normal forms, which are based on KBM averages. Moreover, for both classes of problems, we show that the main coordinate changes are equivalent, up to translations between the spaces in which they are defined. In this manner, our results show that the RG method offers a new approach for deriving normal forms for nonautonomous systems, and it offers advantages since one can typically more readily identify resonant terms from naive perturbation expansions than from the nonautonomous vector fields themselves. Finally, we establish how well the solution to the RG equations approximates the solution of the original equations on time scales of $\mathcal{O}(1/\epsilon)$.

PACS: 05.10.Cc; 05.45.-a; 02.30.Hq

Keywords: Renormalization group method; Normal form theory; Singular perturbations; Multiscale systems; Secularities; Asymptotic analysis; Near-identity coordinate changes

1. Introduction

Normal form theory offers a powerful method for putting vector fields into their simplest algebraic form in the neighborhood of invariant sets, so that every term contains essential geometric information. The nonessential terms, labeled nonresonant terms, are removed from the vector fields by sequences of near-identity coordinate changes. For example, the classical Poincaré–Birkhoff

* Corresponding author. Tel.: +1 617 353 9552.
E-mail address: tasso@math.bu.edu (T.J. Kaper).

0167-2789/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2007.12.009

normal form theory [1,2,7,14,17,33,36], which applies to C^k smooth, autonomous vector fields in the neighborhood of fixed points, begins with the linear part of the system in its simplest, Jordan canonical form and then systematically employs near-identity coordinate changes to remove the nonresonant quadratic, cubic, and higher order nonlinear terms up through order k from the vector field. Only resonant nonlinear terms remain. These are completely determined by the linear part, and they satisfy an equivariance property (see for example [19,49]).

The classical normal form theory has been extended to vector fields with periodic orbits and to vector fields with time-periodic terms by applying the classical ideas to the Poincaré maps defined by the periodic orbits. More recently, examples have been found of normal forms for a few other nonautonomous vector fields and even for a handful of partial differential equations (PDEs) (see for example [43,8,9,42]). However, as yet, there is no systematic procedure for finding normal forms for general nonautonomous vector fields or for PDEs.

In this article, we demonstrate that renormalization group theory (RG) of the type presented by Chen, Goldenfeld, and Oono in [5,6] offers a systematic method for finding normal forms for large classes of finite-dimensional vector fields. We now state our results separately for autonomous and nonautonomous systems.

1.1. Statement of the results for autonomous systems

In the first part of this article, we apply the RG method to autonomous differential equations

$$x' = Ax + \epsilon f(x), \tag{1.1}$$

where $x \in \mathbb{C}^n$, $0 < \epsilon \ll 1$, A is a constant, diagonal $n \times n$ matrix, with purely imaginary eigenvalues, and f is smooth. Many of the perturbation problems treated in [5,6] are of this type or can be recast in this form; it will be clear below how to extend the results to the case in which A is diagonalizable.

We prove that the reduced equation generated by the RG method is equivalent to the classical Poincaré–Birkhoff normal form of (1.1). Moreover, we prove that the near-identity coordinate changes employed in both methods – the coordinate change that renormalizes the naive expansion in RG theory, and the change of dependent variables in the vector field in NF theory – are equivalent, up to translation between the spaces in which they are defined. We carry out the calculations explicitly up to and including $\mathcal{O}(\epsilon^2)$ and note that the procedure may be generalized to higher order.

For autonomous systems (1.1), the RG method has advantages over NF theory. Chief among these is the fact that the secular terms can be identified more readily by inspection of the naive asymptotic expansion than by inspection of the vector field. We also illustrate all of these analytical results on the example of the Rayleigh oscillator.

It is important to note that, to first order in ϵ, the connection between the RG method and Poincaré–Birkhoff NF theory for autonomous equations (1.1) was established in [49]. Our article extends the results of [49] in that we show that the RG and NF methods are equivalent to second and higher order. Furthermore, we extend the analysis of the RG method to systems with nonautonomous perturbations, as we now describe.

1.2. Statement of the results for nonautonomous systems

In the second part of this article, we apply the RG method to nonautonomous differential equations of the form

$$x' = Ax + \epsilon f(x, t), \tag{1.2}$$

where again $x \in \mathbb{C}^n$, $0 < \epsilon \ll 1$, A is a constant, diagonal $n \times n$ matrix, with purely imaginary eigenvalues, and f is smooth. We show that the reduced or amplitude equations generated by the RG method constitute time-asymptotic normal forms for these systems, which are based on a Krylov–Bogolyubov–Mitropolsky (KBM) average.

Along the way to proving this, we also show that the coordinate change used to renormalize the naive expansion in the RG method naturally generates the near-identity change of dependent variables needed to put the nonautonomous vector field into the asymptotic normal form. Then, as a result of this equivalence, one can also show using standard methods that the solutions of the reduced equations are $\mathcal{O}(\phi(\epsilon))$ close to those of the original equation on time scales of $\mathcal{O}(1/\epsilon)$ for (1.1) and (1.2), where $0 < \phi(\epsilon) \ll 1$ and ϕ depends on the order to which the method is applied, as well as on the rate at which the KBM average is attained.

Finally, we observe that the RG method offers an advantageous technique for deriving normal forms for these nonautonomous systems, because it is typically simpler to spot resonant terms in the naive expansions than directly from the time-dependent vector fields. This advantage is more pronounced here than it is for autonomous systems (1.1).

1.3. Implications of the main results for normal form theory

Mathematically, the RG method may be described as follows. Let V be any suitable space of vector fields that admit a naive perturbation expansion in powers of ϵ, for example those in Eqs. (1.1) and (1.2), and let S be the space of asymptotic expansions

that formally satisfy such equations. One may assume that these expansions are truncated at a finite order, but in general they will not be asymptotically valid for all time, due to secular terms. The RG procedure consists of three steps: The first step, RG_1, is the map between the space of vector fields V and the space of truncated asymptotic expansions S, defined by taking a naive perturbation series, plugging it into the differential equation, and solving order by order. Step RG_2 maps S to S and consists of the coordinate transformation defined on the resulting asymptotic series in which all bounded, time-independent terms in the original naive expansion are absorbed into the initial condition. As we show below, this coordinate change is near-identity on time scales of $\mathcal{O}(1/\epsilon)$. The last step, RG_3, maps S back to V and is frequently referred to as the *RG condition*. In practice, the RG condition involves setting the derivative of the asymptotic series obtained in RG_2 with respect to the initial time t_0 equal to zero. These three steps produce reduced forms of the initial equations which are typically easier to solve. This formulation of the RG method is the same as that originally proposed in [5,6] for broad classes of vector fields, including (1.1) and (1.2).

Our results show that the relationship between the RG method and NF theory for systems (1.1) and (1.2) may be summarized by the following diagram:

$$\begin{array}{ccc} S & \xrightarrow{\mathsf{RG}_2} & S \\ {\scriptstyle \mathsf{RG}_1}\uparrow & & \downarrow{\scriptstyle \mathsf{RG}_3} \\ V & \xrightarrow{\mathsf{NF}} & V, \end{array} \tag{1.3}$$

where NF denotes the change of coordinates central to the NF method.

This view of the RG method reveals that the essential reductive step is the change of coordinates, RG_2. This change of coordinates is near-identity and is applied to the initial conditions in the asymptotic expansions. It removes nonresonant terms from the asymptotic expansion, and it is the analog of the coordinate change involving the dependent variables that is done in NF theory to remove nonresonant terms from the vector field. In this sense, the RG method carries out reduction on naive expansions, rather than on the vector field directly, as NF theory does.

The RG method is a *priori* much broader in scope than the current normal form theory. In particular, the RG method generates good asymptotics for many other classes of systems, beyond (1.1) and (1.2). Furthermore, it is in principle simpler to identify the divergent (resonant) terms from the naive expansions with which one works in the RG method than it is to do so from the vector fields with which one works in NF theory, as we stated above. Therefore, we think that the RG method will be useful as a technique for generating normal forms of systems for which there is as yet no normal form theory.

Remark 1. Our results about the equivalence of the RG and NF approaches for (1.1) also suggest that the normal form to all orders is a fixed point of the RG operator. In particular, the normal form of system (1.1) – when obtained to all orders in the perturbation parameter ϵ – is invariant under the sequence of near-identity coordinate changes employed in the normal form procedure, and hence it is also invariant under the RG operator. Questions about the dimensions of the stable and unstable manifolds of this fixed point of the RG operator are the subject of current work.

1.4. Relation of this analysis to other results about the RG method in the literature

The RG method [5,6] offers a unified, formal approach to deriving asymptotic expansions for the solutions of a wide variety of singularly perturbed ordinary differential equations (ODEs).

It is motivated by renormalization group methods used in solid state physics, quantum field theory, and other areas of physics, see for example [11], and it has been applied to derive reduced approximating equations of ordinary and partial differential equations in problems with boundary layers, with fast and slow time scales, with and without turning points, and numerous others. It offers a versatile alternative to classical perturbation methods, such as the Poincaré–Lindstedt method [26,34], the method of matched asymptotic expansions [15,18,25,39], the method of multiple time scales [15,18,29,39], the method of averaging [3,41], and the WKBJ method [35] which were each developed for specific types of problems. In numerous examples the results obtained using the RG method are shown to agree [5,6,30,37,49] with those obtained from classical methods. Moreover, it apparently automatically introduces – where needed – the appropriate gauge functions, such as fractional powers of ϵ and logarithmic terms in ϵ, in the expansions, avoiding the need for the user to 'see' that they should be used.

There are several other articles in which the RG condition and the RG method have been analyzed and applied, and we conclude this introduction with brief descriptions.

Independently and concurrently to [5,6], Woodruff in [45,46] developed a method based on an infinitesimal *invariance condition* that is analogous to the RG condition. This method has many similar features to the RG method, and it yields similar results to the RG method. Starting from a discrete invariance condition, that encodes the precise circumstances in which two naive asymptotic expansions (centered at different, nearby initial times) represent the same solution, a continuous, or infinitesimal, invariance condition is derived in the limit where the initial times approach each other. Woodruff developed his method to treat WKBJ type problems, as well as weakly perturbed systems in which the linear part is slowly varying with matrix $A(\epsilon t)$.

Fundamental analysis of the RG condition, RG_3, and the RG method is presented in [20–23,10]. First, it was shown that RG_3 is an envelope equation in the sense of classical differential geometry. In particular, given a family of curves representing naive approximations for a given solution of a perturbed ODE based at different initial times, the condition RG_3 yields the envelope of this family. In addition to the insight that this approach offers for RG, it has also been used to advance the applications of RG. A notable example is presented in [22], where the asymptotic wave forms are found for the Schrödinger equation of the quantum anharmonic oscillator. In addition, it is shown in [10] that the RG method is a technique of reduction for systems possessing an attracting slow manifold. For equations in which the unperturbed system is assumed to have an attracting slow manifold, the renormalization of the initial condition (step RG_2) is equivalent to choosing initial data on the slow manifold. All of the fast components of the initial data are eliminated, and a natural reduction is achieved, along the lines of the results of [24], as the authors note. The analysis is carried out to second order, and extensions are presented for problems with autonomous perturbations in which there are some zero eigenvalues.

Fundamental asymptotic analysis of the RG method has also been presented in [30–32] for systems subject to small-amplitude, time-periodic perturbations and for weakly nonlinear, autonomous perturbations of planar oscillators. In these works, a simplified version of the RG method is presented. A central new feature is a multiple-time scale ansatz in which a slow time $\tau = \epsilon t$ is explicitly introduced and in which the initial data is replaced by a slowly varying amplitude. This work has been recently generalized in [38], with an emphasis on the relationship to the methods of averaging and multiple time scales.

In [44], the energy preserving and dissipation preserving properties of RG are studied. It is shown that for dissipative problems where the eigenvalues of the matrix A all have negative real part, the renormalized equations are also dissipative. Moreover, it is shown that the size of the attracting ball depends in a nontrivial manner on the order of truncation as well as on ϵ. In [40], the RG method is investigated, with special emphasis on the distinctions between the Wilson RG approach and the Gell-Mann and Low formulation, and examples are given for which the RG method fails due to slow modulation of the perturbation term. A number of examples are also studied in [37], and a proto-RG method is introduced that simplifies the sometimes onerous task of finding naive perturbation expansions. In the context of Hamiltonian systems subject to small-amplitude Hamiltonian perturbations, it has been shown [47,48] that the RG method yields results equivalent to those obtained from canonical Hamiltonian perturbation theory, up to and including $\mathcal{O}(\epsilon^2)$. Finally, for completeness, we note that RG has also been applied to derive reduced or amplitude equations for certain nonlinear partial differential equations; see [4–6,12,13,10,24,28].

1.5. Outline of the article

The article is organized as follows: In Section 2 we review the renormalization group method of Chen, Goldenfeld and Oono [5, 6], which we will refer to as CGO-RG for the remainder of the article. In Section 3, we apply an equivalent RG procedure to the autonomous systems (1.1) to derive asymptotic expansions of solutions up to and including $\mathcal{O}(\epsilon)$ and show the equivalence to the classical Poincaré–Birkhoff NF theory in this context. In Section 4, we extend the analysis of the RG method to second order for equations of the form (1.1), and prove that the RG and NF methods yield the same normal form equations to second order. In Section 5, we use Rayleigh's equation as an example to illustrate the general analytical results for autonomous systems. Then, in Section 6, we turn our attention to equations with nonautonomous perturbations of the form (1.2) and apply the RG method to first order. In Section 7, we show that the reduced equation derived by applying the RG method is equivalent to the time-asymptotic normal form for these systems, based on a KBM average. In Section 8, we illustrate the equivalence between the RG method for nonautonomous equation (1.2) and the time-asymptotic normal form theory on the example of the Mathieu equation. Finally, in Section 9 we state a theorem about how well solutions of the reduced equations derived by RG and NF theory approximate solutions of the original systems (1.1) and (1.2).

2. The CGO-RG method

In this section, we describe the CGO-RG method and review how it is implemented on autonomous initial value problems of the form

$$\begin{aligned} x &= Ax + \epsilon f(x), \\ x(T_0) &= w(T_0), \end{aligned} \tag{2.1}$$

where $f(x) = \sum_{\alpha,i} C_{\alpha,i} x^\alpha e_i$, α is a multi-index, i runs from 1 to n, e_i is the standard Euclidean basis vector, $\epsilon \ll 1$, the sum is finite, and T_0 denotes the initial time. We further assume that $x \in \mathbb{C}$ and that the matrix A is diagonal. The goal is to derive asymptotic expansions of solutions of this differential equation on time scales of $\mathcal{O}(1/\epsilon)$. The CGO-RG method consists of the following five steps:

1. Derive a naive perturbation expansion for the solution of the given differential equation.
2. Make a preparatory change of variables to remove all instances of the initial condition. An exception is made for secular terms in which a factor of T_0 exists explicitly.

3. Introduce an arbitrary time τ in between t and T_0.
4. Renormalize the solution to remove those terms involving $(\tau - T_0)$.
5. Apply the RG condition

$$\left.\frac{\mathrm{d}x}{\mathrm{d}\tau}\right|_{\tau=t} = 0 \tag{2.2}$$

to the renormalized solution, since the solution of the differential equation should be independent of the arbitrary parameter τ.

One begins by supposing a naive perturbation expansion for the solution to (2.1),

$$x(t) = x_0(t) + \epsilon x_1(t) + \epsilon^2 x_2(t) + \cdots \tag{2.3}$$

and by substituting this expansion into the differential equation (2.1). Equating like powers of ϵ, one obtains the following sequence of differential equations:

$$\begin{aligned} \dot{x_0} &= Ax_0 \\ \dot{x_1} &= Ax_1 + f(x_0) \\ \dot{x_2} &= Ax_2 + D_x f(x_0)x_1, \end{aligned} \tag{2.4}$$

etc. For the time being, we are only interested in the solutions up to first order. The solutions are

$$\begin{aligned} x_0(t) &= \mathrm{e}^{A(t-T_0)}w(T_0) \\ x_1(t) &= \mathrm{e}^{A(t-T_0)}\int_{T_0}^{t}\mathrm{e}^{-A(s-T_0)}f(\mathrm{e}^{A(s-T_0)}w(T_0))\mathrm{d}s. \\ &= \mathrm{e}^{A(t-T_0)}\left((t-T_0)\sum_{\Lambda_{\alpha,i}=0}C_{\alpha,i}w(T_0)^{\alpha}e_i + \sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}(\mathrm{e}^{\Lambda_{\alpha,i}(t-T_0)}-1)w(T_0)^{\alpha}e_i\right), \end{aligned}$$

where $\Lambda_{\alpha,i} = \sum_{k=1}^{n}\alpha_k\lambda_k - \lambda_i$. Thus, the naive expansion to first order is

$$x(t) = \mathrm{e}^{A(t-T_0)}\left(w(T_0) + \epsilon(t-T_0)\sum_{\Lambda_{\alpha,i}=0}C_{\alpha,i}w(T_0)^{\alpha}e_i + \epsilon\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}(\mathrm{e}^{\Lambda_{\alpha,i}(t-T_0)}-1)w(T_0)^{\alpha}e_i\right).$$

Notice that those terms with $\Lambda_{\alpha,i} = 0$ in the naive expansion are secular terms that, due to their unbounded nature, cause the asymptotic property of (2.3) to be lost on time scales of $\mathcal{O}(1/\epsilon)$. The CGO-RG procedure was created to treat such terms.

With the naive solution in hand, one now proceeds to make a preparatory change of variables. This change of variables is aimed at absorbing all instances of the initial condition into our integration constant. The exception occurs within the secularity $(t - T_0)$. For these terms, one does not absorb the initial condition into an integration constant. Explicitly, this change of variables is

$$v(T_0) = \mathrm{e}^{-AT_0}w(T_0) - \epsilon\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}\mathrm{e}^{\Lambda_{\alpha,i}T_0}(\mathrm{e}^{-AT_0}w(T_0))^{\alpha}e_i + \mathcal{O}(\epsilon^2).$$

After some calculations, the naive expansion becomes

$$x(t) = \mathrm{e}^{At}v(T_0) + \epsilon(t-T_0)\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}=0}C_{\alpha,i}v(T_0)^{\alpha}e_i + \epsilon\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}\mathrm{e}^{\Lambda_{\alpha,i}t}v(T_0)^{\alpha}e_i + \mathcal{O}(\epsilon^2). \tag{2.5}$$

Next, one introduces an arbitrary time τ into the secular term between t and T_0,

$$x(t) = \mathrm{e}^{At}v(T_0) + \epsilon(t-\tau+\tau-T_0)\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}=0}C_{\alpha,i}v(T_0)^{\alpha}e_i + \epsilon\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}\mathrm{e}^{\Lambda_{\alpha,i}t}v(T_0)^{\alpha}e_i + \mathcal{O}(\epsilon^2).$$

The idea is to absorb those terms with $(\tau - T_0)$ that are secular into the integration constant $v(T_0)$. To do this, one renormalizes the constant of integration by introducing a new integration constant depending on τ. Explicitly,

$$V(\tau) = v(T_0) + \epsilon(t-T_0)\sum_{\Lambda_{\alpha,i}=0}C_{\alpha,i}v(T_0)^{\alpha}e_i + \mathcal{O}(\epsilon^2).$$

The renormalized expansion is now

$$x(t) = \mathrm{e}^{At}V(\tau) + \epsilon(t-\tau)\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}=0}C_{\alpha,i}V(\tau)^{\alpha}e_i + \epsilon\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}\mathrm{e}^{\Lambda_{\alpha,i}t}V(\tau)^{\alpha}e_i + \mathcal{O}(\epsilon^2). \tag{2.6}$$

Finally, one applies the RG condition. In particular, one differentiates the renormalized expansion with respect to τ,

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \mathrm{e}^{At}\frac{\mathrm{d}V}{\mathrm{d}\tau} - \epsilon\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i}V(\tau)^{\alpha}e_i + \epsilon(t-\tau)\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i}\frac{\mathrm{d}}{\mathrm{d}\tau}V(\tau)^{\alpha}e_i + \epsilon\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}\mathrm{e}^{\Lambda_{\alpha,i}t}\frac{\mathrm{d}}{\mathrm{d}\tau}V(\tau)^{\alpha}e_i + \mathcal{O}(\epsilon^2),$$

and then evaluates this derivative at $\tau = t$ which removes the third term in the right of the above equation. Then the resulting expression is set equal to zero to yield

$$\left.\frac{\mathrm{d}x}{\mathrm{d}\tau}\right|_{\tau=t} = \mathrm{e}^{At}\frac{\mathrm{d}V}{\mathrm{d}\tau} - \epsilon\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i}V(\tau)^{\alpha}e_i + \epsilon\mathrm{e}^{At}\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}\mathrm{e}^{\Lambda_{\alpha,i}t}\frac{\mathrm{d}}{\mathrm{d}\tau}V(\tau)^{\alpha}e_i + \mathcal{O}(\epsilon^2) = 0.$$

Finally, multiplying by e^{-At} and noting that $\frac{\mathrm{d}V}{\mathrm{d}\tau} = \mathcal{O}(\epsilon)$, one absorbs the final term into $\mathcal{O}(\epsilon^2)$ and hence obtains

$$\frac{\mathrm{d}V}{\mathrm{d}\tau} = \epsilon\sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i}V^{\alpha}e_i + \mathcal{O}(\epsilon^2). \tag{2.7}$$

This is precisely the first-order amplitude equation of (2.1), and it governs the solutions, free of secularities, at this order on time scales up to and including $\mathcal{O}(1/\epsilon)$.

3. An equivalent RG method

In this section, we introduce an equivalent, simplified version of the procedure discussed in Section 2. The purpose of this simplification is to highlight the mathematical underpinnings of the CGO-RG method. The focus of the remainder of the article will be the application and analysis of this method.

We make the following observations concerning the CGO-RG method of Section 2. First, we note that Eqs. (2.5) and (2.6) are equivalent, except with T_0 replaced with τ and $v(T_0)$ replaced with $V(\tau)$. Thus, effectively in steps 2–4 of the CGO-RG method the given initial time and initial condition, T_0 and $w(T_0)$, are replaced by an arbitrary initial time τ and integration constant $V(\tau)$, respectively.

The second observation concerns the RG condition (2.2). We note that evaluation of the derivative in the RG condition (2.2) at $\tau = t$ is unnecessary to obtaining the final equation, since $\frac{\mathrm{d}V}{\mathrm{d}\tau} = \mathcal{O}(\epsilon)$ in (2.7). In this case, evaluation at $\tau = t$ serves to rename the independent variable in the final equation (2.7).

On the basis of these observations, we distill the CGO-RG method into the following three steps:

1. Derive a naive perturbation expansion for the solution of the given differential equation with an arbitrary initial time t_0 and initial condition $w(t_0)$.
2. Renormalize the initial condition by absorbing those terms in the naive expansion that are time independent and bounded (or that grow sublinearly) into $w(t_0)$.
3. Apply the RG condition

$$\frac{\mathrm{d}x}{\mathrm{d}t_0} = 0. \tag{3.1}$$

We henceforth refer to this procedure as the RG method. Note that the three steps correspond to steps RG_1, RG_2, and RG_3 in the commutative diagram (1.3).

The main result of this section is that the RG method yields the NF equations for (2.1) up to and including $\mathcal{O}(\epsilon)$. We show this in Section 3.1. Later in Section 3.2, we discuss the role that the RG condition (3.1) plays in the RG procedure.

3.1. RG yields the Normal form equation up to and including $\mathcal{O}(\epsilon)$

In this section, we apply the simplified three-step RG method to the system given in (2.1). As in the CGO-RG method, the first step is to derive a naive perturbation expansion of the solution to the differential equation. In this case, however, we will solve the initial value problem for an arbitrary initial time, t_0, and initial condition $w(t_0)$. We find

$$x(t) = \mathrm{e}^{A(t-t_0)}w(t_0) + \epsilon(t-t_0)\mathrm{e}^{A(t-t_0)}\sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i}w(t_0)^{\alpha}e_i + \epsilon\mathrm{e}^{A(t-t_0)}\sum_{\Lambda_{\alpha,i}\neq 0}\frac{C_{\alpha,i}}{\Lambda_{\alpha,i}}(\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} - 1)w(t_0)^{\alpha}e_i. \tag{3.2}$$

Next, we renormalize the solution to isolate the resonant terms. The object to be renormalized is $w(t_0)$, which is replaced by an integration constant

$$w(t_0) = W(t_0) + \sum_{k=1}^{\infty} a_k(t_0, W(t_0))\epsilon^k, \tag{3.3}$$

where a_k is an n-dimensional vector. We are free to choose $a_k : \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ as we please provided that the resulting series is an asymptotic series (see Remark 2). Thus, with the choice

$$a_1 = \sum_{\Lambda_{\alpha,i} \neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} W(t_0)^\alpha e_i, \tag{3.4}$$

we push the autonomous part of the nonresonant term in (3.2) to higher order. This leaves us with the first-order renormalized expansion

$$x(t) = \mathrm{e}^{A(t-t_0)} W + \epsilon (t-t_0) \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i + \epsilon \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i} \neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} W^\alpha e_i. \tag{3.5}$$

For ease of notation we are writing W for $W(t_0)$.

We now apply the RG condition (3.1) to produce an evolution equation for $W(t_0)$. Differentiating (3.5) with respect to t_0, we find

$$\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t_0} &= -A\mathrm{e}^{A(t-t_0)} W + \mathrm{e}^{A(t-t_0)} \frac{\mathrm{d}W}{\mathrm{d}t_0} - \epsilon \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i \\
&\quad - \epsilon (t-t_0) A \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha \mathrm{e}_i + \epsilon (t-t_0) \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} \sum_j \alpha_j \frac{W^\alpha}{W^{(j)}} \frac{\mathrm{d}W^{(j)}}{\mathrm{d}t_0} e_i \\
&\quad - \epsilon A \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i} \neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} W^\alpha \mathrm{e}_i - \epsilon \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i} \neq 0} C_{\alpha,i} \mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} W^\alpha e_i \\
&\quad + \epsilon \mathrm{e}^{A(t-t_0)} \sum_{\Lambda_{\alpha,i} \neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} \sum_j \alpha_j \frac{W^\alpha}{W^{(j)}} \frac{\mathrm{d}W^{(j)}}{\mathrm{d}t_0} e_i.
\end{aligned}$$

Setting this quantity equal to zero and clearing the exponentials, we find to $\mathcal{O}(1)$, $\frac{\mathrm{d}W}{\mathrm{d}t_0} = AW$. Therefore, we may substitute $\lambda_j W^{(j)}$ for $\frac{\mathrm{d}W^{(j)}}{\mathrm{d}t_0}$. Also, we observe that the matrix A in the fourth and sixth terms may be pulled inside the sum to obtain $Ae_i = \lambda_i e_i$. Thus,

$$\begin{aligned}
\frac{\mathrm{d}W}{\mathrm{d}t_0} &= AW + \epsilon \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i - \epsilon (t-t_0) \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} \left(\sum_j \alpha_j \lambda_j - \lambda_i \right) W^\alpha e_i \\
&\quad + \epsilon \sum_{\Lambda_{\alpha,i} \neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \left(\Lambda_{\alpha,i} + \lambda_i - \sum_j \alpha_j \lambda_j \right) W^\alpha e_i + \mathcal{O}(\epsilon^2),
\end{aligned}$$

which simplifies to

$$\frac{\mathrm{d}W}{\mathrm{d}t_0} = AW + \epsilon \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i + \mathcal{O}(\epsilon)^2,$$

since the sum is over terms for which $\Lambda_{\alpha,i}$ in the third term, and since the fourth term vanishes by the definition of $\Lambda_{\alpha,i}$.

Therefore, to first order the RG equation is

$$\frac{\mathrm{d}W}{\mathrm{d}t_0} = AW + \epsilon \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i, \tag{3.6}$$

$W(T_0) =$ inverse of (3.3).

The RG equation (3.6) is precisely the NF equation of (2.1) up to and including $\mathcal{O}(\epsilon)$; see (A.4) in Appendix A. Furthermore, the change of coordinates executed in (3.3) is related to the change of variables performed in NF theory. Recall from (3.3) and (3.4) that the change of coordinates in RG is

$$w(t_0) = W(t_0) + \epsilon \sum_{\Lambda_{\alpha,i} \neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} W(t_0)^\alpha e_i.$$

Comparing this with the NF change of coordinates given in (A.3),

$$x = y + \epsilon \sum_{\Lambda_{\alpha,i} \neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} y^\alpha e_i,$$

we see that the two changes of coordinates have exactly the same form. The only difference is that the RG method makes the change of variables on the initial condition, whereas in NF theory it is made on the dependent variables of the original equation. Hence, we have demonstrated the main point of this section: that the RG method and NF theory are equivalent up to and including $\mathcal{O}(\epsilon)$.

For completeness, we refer the reader to Section 9 where a precise theorem is stated about how well solutions of the RG equation, or equivalently the normal form equation, approximate solutions of the original differential equation.

Remark 2. We renormalize $w(t_0)$ in (3.3) to leading order as $W(t_0)$ and not $e^{-At_0}W(t_0)$ as was done in Section 2. This difference is not of great importance as our approach produces the NF equations while the CGO-RG approach produces the amplitude equations. We make the choice not to absorb the exponential e^{-At_0} because we prefer to think of RG as an operator that fixes the linear part of the vector field. Of course, in some applications, see for example Section 8, the amplitude equations are preferable. Also, the use of the word renormalize in step 2 is slightly different from that in the CGO method.

3.2. The RG condition

In this section, we discuss the RG condition (3.1) and its role in the RG method. Our interpretation of the RG condition is similar to that of [46] and [10] among others, but our interpretation of its role in the RG method differs. We will discuss the impact of the RG condition on both solutions of differential equations and their asymptotic approximations.

Consider, for a moment, the general ODE

$$\frac{dx}{dt} = F(x), \tag{3.7}$$

where $x \in \mathbb{R}^n$ and F is globally Lipschitz. Let $w(t_0) : \mathbb{R} \to \mathbb{R}^n$ be an arbitrary function. We denote the solution of (3.7) by $x(t, t_0, w(t_0))$, where the solution is evaluated at time t, with initial time t_0 and initial condition $w(t_0)$. We are interested in the conditions on $w(\cdot)$ that insure the invariance of the solution. In other words, we are interested in when two solutions of (3.7) are equal,

$$x(t, t_0, w(t_0)) = x(t, t_0', w(t_0')), \tag{3.8}$$

where t_0 and t_0' are two different initial times.

The unsurprising answer, due to existence and uniqueness, is that the solutions are equal if and only if $w(t_0)$ and $w(t_0')$ lie on the same solution curve, or equivalently, if the evolution of $w(t_0)$ is given by (3.7). Note that the right hand side of (3.8) does not depend on t_0. This suggests that – for solutions of (3.7) – the invariance condition is equivalent to the RG condition $\frac{dx}{dt_0} = 0$. To show this, we write the solution in integral form as

$$x(t, t_0, w(t_0)) = w(t_0) + \int_{t_0}^{t} F(x(s, t_0, w(t_0)))ds. \tag{3.9}$$

Applying the RG condition, we find

$$0 = \frac{dx}{dt_0} = \frac{dw}{dt_0} - F(x(t_0, t_0, w(t_0))) + \int_{t_0}^{t} DF(x(s, t_0, w(t_0))) \cdot \frac{dx}{dt_0}(s, t_0, w(t_0))ds.$$

Since $\frac{dx}{dt_0} = 0$ for all s, we obtain

$$\frac{dw}{dt_0} = F(x(t_0, t_0, w(t_0))) = F(w(t_0)),$$

recovering the original ODE, (3.7). Thus, applying the RG condition to a family of solution curves parameterized by their initial data simply recovers the ODE which generates the family.

What we have just seen is that if we are given a solution curve, then the RG condition gives back the corresponding differential equation. In the current RG method, the situation is slightly different as we deal with asymptotic approximations of solution curves and not the curves themselves. As a result, we are confronted with two additional considerations. First, these approximations are truncated at a finite order and therefore do not exactly solve any ODE as they are only flows up to the order of truncation. Second, in general, these naive approximations will contain secular terms that limit the domain of validity to bounded time intervals.

The first consideration is readily addressed by observing that if we substitute a naive perturbation series into (3.7) and asymptotically expand the right hand side, we are left with a recursive sequence of differential equations to solve. Therefore, up to any finite order (i.e. the order of truncation), we can apply the same analysis as we did above and reproduce the sequence of differential equations up to the order of truncation.

Secondly, if the naive approximation contains secular terms then the approximation is only valid locally, i.e. there exists a $C > 0$ such that the approximation is valid for $|t - t_0| < C$. Hence, the RG condition is only applied locally as well.

The RG procedure makes good use of the above simple observation in the following manner. Consider an arbitrary change of coordinates,

$$w = \phi(W),$$

where $W : \mathbb{R} \to \mathbb{R}^n$ and $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism. The physical interpretation given to ϕ is that it is the relationship between the initial condition $w(t_0)$ and an integration constant $W(t_0)$. Using again the integral form of the solution (3.9) we find

$$x(t) = \phi(W(t_0)) + \int_{t_0}^{t} F(x(s, t_0, \phi(W(t_0))))\mathrm{d}s.$$

Applying the RG condition again, we find

$$0 = \frac{\mathrm{d}x}{\mathrm{d}t_0} = (D\phi)\frac{\mathrm{d}W}{\mathrm{d}t_0} - F(x(t_0, t_0, \phi(W(t_0)))) + \int_{t_0}^{t} DF(x(s, t_0, \phi(W(t_0)))) \cdot \frac{\mathrm{d}x}{\mathrm{d}t_0}(s, t_0, \phi(W(t_0)))\mathrm{d}s.$$

This reduces to

$$\frac{\mathrm{d}W}{\mathrm{d}t_0} = (D\phi)^{-1}F(\phi(W)).$$

Thus, the RG procedure produces an evolution equation for the integration constant $W(t_0)$. Moreover, if the same change of coordinates is applied to the original dependent variables, i.e.

$$x = \phi(X),$$

then we find that X satisfies the same differential equation as $W(t_0)$,

$$\frac{\mathrm{d}X}{\mathrm{d}t} = (D\phi)^{-1}F(\phi(X)).$$

The principal advantage to using asymptotic expansions is that in many cases we can write them down in closed form provided that the leading order problem is solvable. In turn, this makes selection of a change of coordinates like ϕ straightforward because we only have to collect instances of the initial condition into an integration constant. This change of coordinates can be applied also in the original vector field, but without solving for the naive approximation we may not be able to guess the form of this transformation a priori. It is in this sense that RG_2 may be viewed as the essential reductive step of the RG method.

Remark 3. At first glance, the RG condition appears to be a needlessly complicated way to produce a differential equation from the corresponding solution curve. A much more straightforward method would be to simply differentiate the solution with respect to time. However, in this case the solution curve $x(t, t_0, w(t_0))$ is a function of three variables and hence differentiation with respect to time produces an equation given strictly in terms of t, t_0, and $w(t_0)$ without any explicit dependence on x. We must then invert the relationship between x and t, t_0 and $w(t_0)$ to produce the differential equation. In general, such a computation will not be trivial.

4. The RG method to second order

In this section, we extend the RG analysis of Section 3.1 up to and including $\mathcal{O}(\epsilon^2)$. We begin by finding the naive approximation at second order. As shown in (2.4), the second-order differential equation is

$$\dot{x}_2 = Ax_2 + (D_x f(x_0))x_1. \tag{4.1}$$

The essential step is to obtain a computable expression for the second term on the right hand side of (4.1). The matrix $D_x f$ is an $n \times n$ matrix whose (i, j)th component is

$$(D_x f(x))_{i,j} = \sum_{\alpha} C_{\alpha,i}\alpha_j \frac{x^{\alpha}}{x^{(j)}}.$$

Evaluating x at $x_0(t) = \mathrm{e}^{A(t-t_0)}w(t_0)$, we find

$$(D_x f(x_0))_{i,j} = \sum_{\alpha} C_{\alpha,i}\alpha_j \mathrm{e}^{\Lambda_{\alpha,j}(t-t_0)} \frac{w(t_0)^{\alpha}}{w(t_0)^{(j)}}.$$

Next, multiplying the matrix $D_x f(x_0)$ by the vector $x_1(t)$, we obtain the vector

$$(D_x f(x_0))x_1 = \sum_{\alpha,i} \sum_{\Lambda_{\beta,j}=0} (t-t_0) C_{\alpha,i} C_{\beta,j} \alpha_j \mathrm{e}^{(\Lambda_{\alpha,j}+\lambda_j)(t-t_0)} \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} e_i + \sum_{\alpha,i} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \mathrm{e}^{(\Lambda_{\alpha,j}+\lambda_j)(t-t_0)} (\mathrm{e}^{\Lambda_{\beta,j}(t-t_0)} - 1) \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} e_i. \tag{4.2}$$

Now that we have an expression for $(D_x f(x_0))x_1$, we proceed to solve (4.1) with the aid of an integrating factor to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{-A(t-t_0)} x_2) = \mathrm{e}^{-A(t-t_0)} (D_x f(x_0))x_1.$$

The impact of the exponential on the right hand side of the previous equation is to replace $\mathrm{e}^{(\Lambda_{\alpha,j}+\lambda_j)(t-t_0)}$ by $\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)}$ in $(D_x f(x_0))x_1$; see (4.2). We now solve the differential equation, noting that $x_2(t_0) = 0$ by assumption. For ease of integration, we split the two double sums in $(D_x f(x_0))x_1$ into six double sums according to whether or not the arguments of the exponentials vanish. Thus, the corresponding solution is

$$x_2(t) = \mathrm{e}^{A(t-t_0)} \int_{t_0}^{t} (\mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI}) \mathrm{d}s, \tag{4.3}$$

where

$$\mathrm{I} = \sum_{\Lambda_{\alpha,i}=0} \sum_{\Lambda_{\beta,j}=0} C_{\alpha,i} C_{\beta,j} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} (s-t_0) e_i,$$

$$\mathrm{II} = \sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}=0} C_{\alpha,i} C_{\beta,j} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} (s-t_0) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} e_i,$$

$$\mathrm{III} = \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} (\mathrm{e}^{(\Lambda_{\alpha,i}+\Lambda_{\beta,j})(s-t_0)}) e_i,$$

$$\mathrm{IV} = \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}=0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} e_i,$$

$$\mathrm{V} = -\sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} (\mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)}) e_i,$$

$$\mathrm{VI} = -\sum_{\Lambda_{\alpha,i}=0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} e_i.$$

Upon integrating each of these terms, we obtain the following quantities:

$$\mathsf{A} = \int_{t_0}^{t} \mathrm{I}\, \mathrm{d}s = \sum_{\Lambda_{\alpha,i}=0} \sum_{\Lambda_{\beta,j}=0} C_{\alpha,i} C_{\beta,j} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} \frac{(t-t_0)^2}{2} e_i$$

$$\mathsf{B} = \int_{t_0}^{t} \mathrm{II}\, \mathrm{d}s = \sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}=0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} C_{\beta,j} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} \left((t-t_0)\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} - \frac{\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)}}{\Lambda_{\alpha,i}} + \frac{1}{\Lambda_{\alpha,i}} \right) e_i$$

$$\mathsf{C} = \int_{t_0}^{t} \mathrm{III}\, \mathrm{d}s = \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} \left(\frac{\mathrm{e}^{(\Lambda_{\alpha,i}+\Lambda_{\beta,j})(t-t_0)} - 1}{\Lambda_{\alpha,i} + \Lambda_{\beta,j}} \right) e_i$$

$$\mathsf{D} = \int_{t_0}^{t} \mathrm{IV}\, \mathrm{d}s = \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}=0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} (t-t_0) e_i$$

$$\mathsf{E} = \int_{t_0}^{t} \mathrm{V}\, \mathrm{d}s = -\sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} (\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} - 1) e_i$$

$$\mathsf{F} = \int_{t_0}^{t} \mathrm{VI}\, \mathrm{d}s = -\sum_{\Lambda_{\alpha,i}=0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{w(t_0)^{\alpha+\beta}}{w(t_0)^{(j)}} (t-t_0) e_i.$$

The terms A through F comprise the coefficients on ϵ^2 in the naive expansion.

Next, we renormalize the naive expansion, which up to and including second order is given by (3.2) and (4.3). The renormalization of the $\mathcal{O}(\epsilon)$ terms follows exactly what was done in Section 3.1. However, this renormalization at first order introduces two terms at second order that arise after expanding $(W+\epsilon a_1)^\alpha$ in (3.2). We will call these terms R_2 (resp. N_2) corresponding to the second-order terms that come from expanding the first-order resonant (resp. nonresonant) terms. Performing this expansion, we find

$$(W+\epsilon a_1)^\alpha = W^\alpha + \epsilon \sum_{\Lambda_{\beta,j}\neq 0} \alpha_j \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \frac{W^{\alpha+\beta}}{W^{(j)}} + \mathcal{O}(\epsilon^2).$$

Thus,

$$\mathsf{R}_2 = \sum_{\Lambda_{\alpha,i}=0} \sum_{\Lambda_{\beta,j}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\beta,j}} C_{\beta,j}\alpha_j (t-t_0) \frac{W^{\alpha+\beta}}{W^{(j)}} e_i$$

and

$$\mathsf{N}_2 = \sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} \frac{C_{\alpha,i}C_{\beta,j}}{\Lambda_{\alpha,i}\Lambda_{\beta,j}} \alpha_j \left(\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} - 1\right) \frac{W^{\alpha+\beta}}{W^{(j)}} e_i.$$

Notice that R_2 and N_2 are the same as F and E up to $\mathcal{O}(\epsilon^3)$ respectively, but with opposite signs. Therefore, their difference is $\mathcal{O}(\epsilon^3)$. Hence, the second-order expansion, renormalized up to and including $\mathcal{O}(\epsilon)$ but not yet to $\mathcal{O}(\epsilon^2)$, is

$$x(t) = \mathrm{e}^{A(t-t_0)} \left[W + \epsilon(t-t_0) \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i + \epsilon \sum_{\Lambda_{\alpha,i}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} W^\alpha e_i + \epsilon^2 (a_2 + \mathsf{A} + \mathsf{B} + \mathsf{C} + \mathsf{D}) \right] + \mathcal{O}(\epsilon^3),$$

where terms A through D are all evaluated at $w(t_0) = W(t_0) + \mathcal{O}(\epsilon)$.

Next, we choose a_2 so as to absorb all the constant homogeneous terms at second order into a single integration constant. We choose

$$a_2 = -\sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}=0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}^2} C_{\beta,j}\alpha_j \frac{W(t_0)^{\alpha+\beta}}{W(t_0)^{(j)}} e_i + \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{W(t_0)^{\alpha+\beta}}{W(t_0)^{(j)}} \frac{1}{\Lambda_{\alpha,i}+\Lambda_{\beta,j}} e_i. \tag{4.4}$$

This leaves us with the second-order renormalized expansion

$$\begin{aligned} x(t) = \mathrm{e}^{A(t-t_0)} \Bigg[& W + \epsilon(t-t0) \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i + \epsilon \sum_{\Lambda_{\alpha,i}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} W^\alpha e_i \\ & + \epsilon^2 \sum_{\Lambda_{\alpha,i}=0} \sum_{\Lambda_{\beta,j}=0} C_{\alpha,i} C_{\beta,j} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} \frac{(t-t_0)^2}{2} e_i \\ & + \epsilon^2 \sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}=0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} C_{\beta,j}\alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} \left((t-t_0)\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)} - \frac{\mathrm{e}^{\Lambda_{\alpha,i}(t-t_0)}}{\Lambda_{\alpha,i}} \right) e_i \\ & + \epsilon^2 \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} \frac{\mathrm{e}^{(\Lambda_{\alpha,i}+\Lambda_{\beta,j})(t-t_0)}}{\Lambda_{\alpha,i}+\Lambda_{\beta,j}} e_i \\ & + \epsilon^2 \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}=0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} (t-t_0) e_i \Bigg]. \end{aligned} \tag{4.5}$$

The final step is to apply the RG condition (3.1). We make the following observation: if one pulls the exponential $\mathrm{e}^{A(t-t_0)}$ into the sums and double sums in (4.5) one gets $\mathrm{e}^{\lambda_i(t-t_0)}$ and therefore to leading order

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t_0} \left(\alpha_j \mathrm{e}^{\lambda_i(t-t_0)} \frac{W^{\alpha+\beta}}{W^{(j)}} \right) &= -\alpha_j\lambda_i \mathrm{e}^{\lambda_i(t-t_0)} \frac{W^{\alpha+\beta}}{W^{(j)}} + \alpha_j \mathrm{e}^{\lambda_i(t-t_0)} \sum_m (\alpha_m + \beta_m - \delta_{jm}) \frac{W^{\alpha+\beta}}{W^{(j)}W^{(m)}} \frac{\mathrm{d}W^{(m)}}{\mathrm{d}t_0} \\ &= \alpha_j \mathrm{e}^{\lambda_i(t-t_0)} (\Lambda_{\alpha,i} + \Lambda_{\beta,j}) \frac{W^{\alpha+\beta}}{W^{(j)}}, \end{aligned}$$

where we have used the fact that to leading order $\frac{dW}{dt_0} = AW$. Applying the RG condition, clearing exponentials, and recalling that the $\mathcal{O}(\epsilon)$ terms were computed earlier, we find

$$\begin{aligned}
\frac{dx}{dt_0} = {} & \frac{dW}{dt_0} - AW - \epsilon \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i + \epsilon^2 (t-t_0) \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} \sum_j \alpha_j \frac{W^\alpha}{W^{(j)}} \sum_{\Lambda_{\beta,j}=0} C_{\beta,j} W^\beta e_i \\
& + \epsilon^2 \sum_{\Lambda_{\alpha,i}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} e^{\Lambda_{\alpha,i}(t-t_0)} \sum_j \alpha_j \frac{W^\alpha}{W^{(j)}} \sum_{\Lambda_{\beta,j}=0} C_{\beta,j} W^\beta e_i - \epsilon^2 \sum_{\Lambda_{\alpha,i}=0} \sum_{\Lambda_{\beta,j}=0} C_{\alpha,i} C_{\beta,j} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} (t-t_0) e_i \\
& + \epsilon^2 \sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}=0} C_{\alpha,i} C_{\beta,j} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} \left((t-t_0) e^{\Lambda_{\alpha,i}(t-t_0)} - \frac{e^{\Lambda_{\alpha,i}(t-t_0)}}{\Lambda_{\alpha,i}} \right) e_i \\
& + \epsilon^2 \sum_{\Lambda_{\alpha,i}\neq 0} \sum_{\Lambda_{\beta,j}=0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} C_{\beta,j} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} \left(-e^{\Lambda_{\alpha,i}(t-t_0)} - \Lambda_{\alpha,i}(t-t_0) e^{\Lambda_{\alpha,i}(t-t_0)} + e^{\Lambda_{\alpha,i}(t-t_0)} \right) e_i \\
& + \epsilon^2 \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} e^{(\Lambda_{\alpha,i}+\Lambda_{\beta,j})(t-t_0)} e_i \\
& - \epsilon^2 \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}\neq 0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} e^{(\Lambda_{\alpha,i}+\Lambda_{\beta,j})(t-t_0)} e_i - \epsilon^2 \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}=0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} e_i .
\end{aligned}$$

Cancelling five pairs of terms and setting the above expression equal to zero, we find that the RG equation to second order is

$$\begin{aligned}
& \frac{dW}{dt_0} = AW + \epsilon \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} W^\alpha e_i + \epsilon^2 \sum_{\Lambda_{\alpha,i}+\Lambda_{\beta,j}=0} \sum_{\Lambda_{\beta,j}\neq 0} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \alpha_j \frac{W^{\alpha+\beta}}{W^{(j)}} e_i , \qquad (4.6) \\
& W(T_0) = \text{inverse of (3.3)}.
\end{aligned}$$

This differential equation is equivalent to the NF equation for (2.1) which we review in Appendix A. In particular, see Eq. (A.8). Therefore, we have shown that for the autonomous vector field given by (2.1) the RG procedure produces equivalent results to NF theory at second order. In addition, we recall that the second-order coordinate change in the RG procedure was given in (4.4). Likewise, the nonresonant terms at second order in the NF procedure are given in (A.6) and (A.7) after removing the resonant terms (A.8). A short calculation reveals that a coordinate change equivalent to (3.3) with (4.4) removes the nonresonant terms at second order. We have therefore shown that the normal form and RG equations are equivalent up to and including $\mathcal{O}(\epsilon^2)$.

The RG equation (4.6) is an evolution equation for the integration constant $W(t_0)$. The solution of this evolution (or normal form) equation may then be used to obtain an approximation of the solution $w(t_0)$ of the original problem, (2.1), that is valid up to and including $\mathcal{O}(\epsilon^2)$. In particular, we plug $W(t_0)$ into (3.3) with a_1 and a_2 specified as in (3.4) and (4.4).

5. Example

In this section, we illustrate the RG method of Sections 3 and 4 on the Rayleigh oscillator, given by the following differential equation:

$$\frac{d^2y}{dt^2} + y = \epsilon \left\{ \frac{dy}{dt} - \frac{1}{3} \left(\frac{dy}{dt} \right)^3 \right\}. \qquad (5.1)$$

One may convert (5.1) into a system of the form (1.1). This is done using the complex coordinate $z = x + iy$ and its complex conjugate as new variables so that the linear part of the new system is diagonal with eigenvalues $-i$ and i. We elect to, equivalently, work directly with the second-order scalar equation (5.1).

Substituting the naive expansion

$$y = y_0 + \epsilon y_1 + \epsilon^2 y_2 + \cdots$$

into the differential equation, we find at each order

$$\begin{aligned}
& \mathcal{O}(1): \quad \ddot{y}_0 + y_0 = 0, \\
& \mathcal{O}(\epsilon): \quad \ddot{y}_1 + y_1 = \dot{y}_0 - \frac{1}{3}\dot{y}_0^3, \\
& \mathcal{O}(\epsilon^2): \quad \ddot{y}_2 + y_2 = \dot{y}_1 - \dot{y}_0^2 \dot{y}_1 .
\end{aligned}$$

The solutions are

$$
\begin{aligned}
y_0(t) &= A\mathrm{e}^{\mathrm{i}(t-t_0)} + c.c.\\
y_1(t) &= \frac{\mathrm{i}}{24}A^3\mathrm{e}^{\mathrm{i}(t-t_0)} + \frac{1}{2}A(1-A\bar{A})(t-t_0)\mathrm{e}^{\mathrm{i}(t-t_0)} - \frac{\mathrm{i}}{24}A^3\mathrm{e}^{3\mathrm{i}(t-t_0)} + c.c.\\
y_2(t) &= \left(\frac{A^3}{32}\left(1-\frac{3}{2}A\bar{A}-\frac{A^2}{6}\right)+\frac{1}{192}A^5\right)\mathrm{e}^{\mathrm{i}(t-t_0)} - \frac{\mathrm{i}}{8}A\left(1-\frac{A^2\bar{A}^2}{2}-\frac{A^2}{6}-\frac{A\bar{A}^{-3}}{6}+\frac{A^3\bar{A}}{3}\right)(t-t_0)\mathrm{e}^{\mathrm{i}(t-t_0)}\\
&\quad + \frac{A}{8}(1-4A\bar{A}+3A^2\bar{A}^2)(t-t_0)^2\mathrm{e}^{\mathrm{i}(t-t_0)} - \frac{\mathrm{i}}{16}A^3(1-A\bar{A})(t-t_0)\mathrm{e}^{3\mathrm{i}(t-t_0)}\\
&\quad - \frac{A^3}{32}\left(1-\frac{3}{2}A\bar{A}-\frac{A^2}{6}\right)\mathrm{e}^{3\mathrm{i}(t-t_0)} - \frac{1}{192}A^5\mathrm{e}^{5\mathrm{i}(t-t_0)} + c.c.
\end{aligned}
$$

Here we have chosen the homogeneous parts of the solutions to y_1 and y_2 so that the solutions vanish at that initial time, i.e. $y_1(t_0) = y_2(t_0) = 0$. We next renormalize the integration constant A, absorbing the homogeneous parts of the solution into it and creating a new integration constant $\mathcal{A} = \mathcal{A}(t_0)$. We begin at first order by requiring

$$A = \mathcal{A} - \epsilon\frac{\mathrm{i}\mathcal{A}^3}{24} + \epsilon^2 a_2 + \mathcal{O}(\epsilon^3).$$

Applying this change of variables we are left with

$$
\begin{aligned}
y_0(t) &= \mathcal{A}\mathrm{e}^{\mathrm{i}(t-t_0)} + c.c.\\
y_1(t) &= \frac{1}{2}\mathcal{A}(1-\mathcal{A}\bar{\mathcal{A}})(t-t_0)\mathrm{e}^{\mathrm{i}(t-t_0)} - \frac{\mathrm{i}}{24}\mathcal{A}^3\mathrm{e}^{3\mathrm{i}(t-t_0)} + c.c.\\
y_2(t) &= \left(\frac{\mathcal{A}^3}{32}\left(1-\frac{3}{2}\mathcal{A}\bar{\mathcal{A}}-\frac{\mathcal{A}^2}{6}\right)+a_2\right)\mathrm{e}^{\mathrm{i}(t-t_0)} - \frac{\mathrm{i}}{8}\mathcal{A}\left(1-\frac{\mathcal{A}^2\bar{\mathcal{A}}^2}{2}\right)(t-t_0)\mathrm{e}^{\mathrm{i}(t-t_0)}\\
&\quad + \frac{\mathcal{A}}{8}(1-4\mathcal{A}\bar{\mathcal{A}}+3\mathcal{A}^2\bar{\mathcal{A}}^2)(t-t_0)^2\mathrm{e}^{\mathrm{i}(t-t_0)} - \frac{\mathrm{i}}{16}\mathcal{A}^3(1-\mathcal{A}\bar{\mathcal{A}})(t-t_0)\mathrm{e}^{3\mathrm{i}(t-t_0)}\\
&\quad - \frac{\mathcal{A}^3}{32}\left(1-\frac{3}{2}\mathcal{A}\bar{\mathcal{A}}-\frac{\mathcal{A}^2}{6}\right)\mathrm{e}^{3\mathrm{i}(t-t_0)} - \frac{1}{192}\mathcal{A}^5\mathrm{e}^{5\mathrm{i}(t-t_0)} - \frac{1}{192}\mathcal{A}^5\mathrm{e}^{3\mathrm{i}(t-t_0)} + c.c.
\end{aligned}
$$

In turn, we select a_2 so to remove the homogeneous terms at second order so that the total renormalization transformation is

$$A = \mathcal{A} - \epsilon\frac{\mathrm{i}\mathcal{A}^3}{24} - \epsilon^2\frac{\mathcal{A}^3}{32}\left(1-\frac{3}{2}\mathcal{A}\bar{\mathcal{A}}-\frac{\mathcal{A}^2}{6}\right) + \mathcal{O}(\epsilon^3).$$

We now apply the RG condition, which isolates the resonant terms at second order to leave the RG equation, correct to $\mathcal{O}(\epsilon^3)$,

$$\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}t_0} = \mathrm{i}\mathcal{A} + \epsilon\frac{\mathcal{A}}{2}(1-\mathcal{A}\bar{\mathcal{A}}) - \epsilon^2\frac{\mathrm{i}}{8}\mathcal{A}\left(1-\frac{\mathcal{A}^2\bar{\mathcal{A}}^2}{2}\right). \tag{5.2}$$

This equation is exactly the normal form of (5.1). If we let $\mathcal{A} = \frac{R}{2}\mathrm{e}^{\mathrm{i}\theta}$ and substitute into (5.2) then we get the following system of amplitude and phase equations:

$$\frac{\mathrm{d}R}{\mathrm{d}t_0} = \frac{\epsilon}{2}R\left(1-\frac{R^2}{4}\right) \tag{5.3}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t_0} = 1 - \frac{\epsilon^2}{8}\left(1-\frac{R^4}{32}\right). \tag{5.4}$$

Since the fixed point $R^* = 2$ of the truncated system (5.3) is hyperbolic, the untruncated equation has a hyperbolic limit cycle which deviates at most by $\mathcal{O}(\epsilon^2)$ from a circle of radius 2. Standard techniques show that (5.3) gives a valid approximation of the radial variable for all time; however, Eq. (5.4) can only be expected to be valid on time scales of $\mathcal{O}(1/\epsilon^2)$.

6. RG applied to nonautonomous equations (1.2)

We now apply the RG methodology to nonautonomous systems (1.2), to which classical NF theory does not apply. In Section 7.1, we use the RG method to guide us in developing a normal form theory for these same vector fields, and then we show that the two methods produce identical results to first order.

We revisit the problem posed in (1.2),

$$\begin{aligned} \dot{x} &= Ax + \epsilon \sum_{\alpha,i} f_{\alpha,i}(t) x^\alpha e_i, \\ x(t_0) &= w(t_0), \end{aligned} \tag{6.1}$$

where $x \in \mathbb{C}^n$, A is an $n \times n$ constant diagonal matrix with purely imaginary eigenvalues, e_i is the ith unit vector, and the sum is finite. Also, $\alpha \in \mathbb{N}^n$ is a multi-index so that $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and we will assume that the f's satisfy the KBM condition defined in (6.3).

Substituting a naive perturbation expansion again produces a sequence of differential equations as in (2.4). The naive expansion to first order is therefore

$$x(t) = \mathrm{e}^{A(t-t_0)} w(t_0) + \epsilon \mathrm{e}^{A(t-t_0)} \sum_{\alpha,i} \int_{t_0}^{t} \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} f_{\alpha,i}(s) w(t_0)^\alpha e_i \mathrm{d}s. \tag{6.2}$$

In the autonomous case, a term in the expansion was considered resonant if it grew like $(t - t_0)$. We carry the same definition of resonance over to the nonautonomous case, formalized by the notion of a KBM_λ average. The KBM_λ average of a function is defined as

$$f^{(\lambda)} = \lim_{(T-T_0)\to\infty} \frac{1}{(T - T_0)} \int_{T_0}^{T} \mathrm{e}^{\lambda t} f(t) \mathrm{d}t \tag{6.3}$$

for $\lambda \in \mathbb{C}$. A function $f(t)$ is said to be KBM_λ if the KBM_λ average converges for all choices of T_0. The notion of KBM vector fields was introduced in [3] and developed in [27,41], where KBM_0 was used. We follow [41] in calling this the Krylov–Bogolyubov–Mitropolskii average. We will assume that all $f_{\alpha,i}$ in (6.1) are $\mathsf{KBM}_{\Lambda_{\alpha,i}}$. Given this definition, we consider a term $f_{\alpha,i}$ to be resonant if it has nonzero $\mathsf{KBM}_{\Lambda_{\alpha,i}}$ average. We can therefore split $f_{\alpha,i}$ into resonant and nonresonant parts as

$$f_{\alpha,i}(t) = f_{\alpha,i}^{\mathsf{R}}(t) + f_{\alpha,i}^{\mathsf{NR}}(t),$$

where

$$f_{\alpha,i}^{\mathsf{R}}(t) = \mathrm{e}^{-\Lambda_{\alpha,i}t} f^{(\Lambda_{\alpha,i})} = \mathrm{e}^{-\Lambda_{\alpha,i}t} \left(\lim_{(T-T_0)\to\infty} \frac{1}{T - T_0} \int_{T_0}^{T} \mathrm{e}^{\Lambda_{\alpha,i}t} f_{\alpha,i}(t) \mathrm{d}t \right).$$

With this in mind, we split the $\mathcal{O}(\epsilon)$ term in the expansion (6.2) into two integrals based upon whether the term $f_{\alpha,i}$ is resonant or not. Thus (6.2) becomes

$$x(t) = \mathrm{e}^{A(t-t_0)} \left(w(t_0) + \epsilon \int_{t_0}^{t} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{R}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} w(t_0)^\alpha \mathrm{e}_i \mathrm{d}s + \epsilon \int_{t_0}^{t} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{NR}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} w(t_0)^\alpha \mathrm{e}_i \mathrm{d}s \right). \tag{6.4}$$

After integration the term involving f^{NR} will contain terms that grow slower than $(t - t_0)$. With this definition of nonresonance, we make the same renormalization as we did in the autonomous case. Namely, we renormalize the initial conditions as in (3.3) to remove the autonomous part of the nonresonant integral above, or the lower limit of integration. This quantity can only be specified up to a constant, so we choose to split the integral at an arbitrary fixed time T_0 and then absorb the resulting autonomous integral with the following choice of a_1 at $\mathcal{O}(\epsilon)$:

$$a_1 = -\int_{t_0}^{T_0} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{NR}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} W^\alpha \mathrm{e}_i \mathrm{d}s. \tag{6.5}$$

Our choice of a_1 removes the lower bound of integration in the nonresonant part of (6.4) and leaves the renormalized expansion as

$$x(t) = \mathrm{e}^{A(t-t_0)} W + \epsilon \mathrm{e}^{A(t-t_0)} \int_{t_0}^{t} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{R}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} W^\alpha \mathrm{e}_i \mathrm{d}s + \epsilon \mathrm{e}^{A(t-t_0)} \int_{T_0}^{t} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{NR}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} W^\alpha \mathrm{e}_i \mathrm{d}s.$$

We now apply the RG condition, (3.1). Differentiating with respect to t_0, we find

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t_0} &= -A\mathrm{e}^{A(t-t_0)} W + \mathrm{e}^{A(t-t_0)} \frac{\mathrm{d}W}{\mathrm{d}t_0} - \epsilon \mathrm{e}^{A(t-t_0)} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{R}}(t_0) W^\alpha \mathrm{e}_i \mathrm{d}s - \epsilon A \mathrm{e}^{A(t-t_0)} \int_{t_0}^{t} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{R}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} W^\alpha \mathrm{e}_i \mathrm{d}s \\ &\quad + \epsilon \mathrm{e}^{A(t-t_0)} \int_{t_0}^{t} \sum_{\alpha,i} f_{\alpha,i}^{\mathsf{R}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} \sum_j \alpha_j \frac{W^\alpha}{W^{(j)}} \frac{\mathrm{d}W^{(j)}}{\mathrm{d}t_0} e_i \mathrm{d}s - \epsilon \mathrm{e}^{A(t-t_0)} \int_{t_0}^{t} \sum_{\alpha,i} \Lambda_{\alpha,i} f_{\alpha,i}^{\mathsf{R}}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} W^\alpha \mathrm{e}_i \mathrm{d}s \end{aligned}$$

$$-\epsilon A e^{A(t-t_0)}\int_{T_0}^{t}\sum_{\alpha,i} f^{\mathsf{NR}}_{\alpha,i}(s)e^{\Lambda_{\alpha,i}(s-t_0)}W^{\alpha}e_i\,ds+\epsilon e^{A(t-t_0)}\int_{T_0}^{t}\sum_{\alpha,i} f^{\mathsf{NR}}_{\alpha,i}(s)e^{\Lambda_{\alpha,i}(s-t_0)}\sum_j \alpha_j\frac{W^{\alpha}}{W^{(j)}}\frac{dW^{(j)}}{dt_0}e_i\,ds$$
$$-\epsilon e^{A(t-t_0)}\int_{T_0}^{t}\sum_{\alpha,i}\Lambda_{\alpha,i}f^{\mathsf{NR}}_{\alpha,i}(s)e^{\Lambda_{\alpha,i}(s-t_0)}W^{\alpha}e_i\,ds.$$

Setting this expression equal to zero and clearing the exponentials, we find as in the autonomous case that the fourth, fifth, and sixth terms cancel by the definition of $\Lambda_{\alpha,i}$, because $dW^{(j)}/dt_0 = AW^{(j)}$ to leading order. Likewise, the terms in the last two lines also sum to zero exactly. Therefore, we find the following nonautonomous RG equation, truncated to $\mathcal{O}(\epsilon)$:

$$\frac{dW}{dt_0} = AW+\epsilon\sum_{\alpha,i} f^{\mathsf{R}}_{\alpha,i}(t_0)W^{\alpha}e_i. \tag{6.6}$$

In the next section we will derive NF equations to the system in (6.1) that are identical to this equation. We will also prove in Section 9 that solutions to this NF equation stay close to solutions of the original equation thus justifying the RG method provided here. As a final point, we note that the vector field in (6.6) is nonautonomous and therefore somewhat problematic to solve directly. However, by making the change of variables $y = e^{At_0}W$, we reduce Eq. (6.6) to the autonomous equation

$$\dot{y} = \epsilon\sum_{\alpha,i} f^{(\Lambda_{\alpha,i})}y^{\alpha}e_i. \tag{6.7}$$

Eq. (6.7) is equivalent to the equation that one obtains by first transforming (6.1) into rotating coordinates and then averaging. In this example, the amplitude equations are more convenient than the NF equations for practical calculations.

7. Equivalence of RG theory to the NF theory—nonautonomous perturbations

7.1. Nonautonomous normal form theory

In this section, we develop a NF theory for nonautonomous systems (1.2) based on Krylov–Bogolyubov–Mitropolsky averages. This nonautonomous NF theory is a natural extension of Poincaré–Birkhoff NF theory (please see Appendix A for details). See also [43] for another extension to nonautonomous systems.

We introduce a near-identity change of variables

$$x = y+\epsilon g(y,t), \qquad g:\mathbb{C}^n\times\mathbb{R}\to\mathbb{C}^n, \tag{7.1}$$

with the goal of removing as many nonlinear terms as possible in (6.1). In the new variables, (6.1) becomes

$$\dot{y} = Ay+\epsilon\left(Ag(y,t)-Dg(y,t)Ay-\frac{\partial g}{\partial t}(y,t)+\sum_{\alpha,i} f_{\alpha,i}(t)y^{\alpha}e_i\right)+\mathcal{O}(\epsilon^2). \tag{7.2}$$

Let $[Ay,g](y,t) = Dg(y,t)Ay - Ag(y,t)$. To remove the nonlinear terms at $\mathcal{O}(\epsilon)$ in (7.2), we want to solve the PDE

$$\frac{\partial g}{\partial t}(y,t)+[Ay,g](y,t) = \sum_{\alpha,i} f_{\alpha,i}(t)y^{\alpha}e_i. \tag{7.3}$$

This equation is linear in g, and thus it is sufficient to solve separately the equations

$$\frac{\partial g_{\alpha,i}}{\partial t}(y,t)+[Ay,g_{\alpha,i}](y,t) = f_{\alpha,i}(t)y^{\alpha}e_i.$$

Choosing $g_{\alpha,i}(y,t) = h_{\alpha,i}(t)y^{\alpha}e_i$, we obtain the following ODE for $h_{\alpha,i}$:

$$\dot{h}_{\alpha,i}(t)+\Lambda_{\alpha,i}h_{\alpha,i}(t) = f_{\alpha,i}(t),$$

whose solution satisfying $h(T_0) = 0$ is

$$h_{\alpha,i}(t) = e^{-\Lambda_{\alpha,i}t}\int_{T_0}^{t} e^{\Lambda_{\alpha,i}\tau}f_{\alpha,i}(\tau)d\tau.$$

In contrast to the situation in Appendix A, here the change of coordinates (7.1) can be formally defined for all $f_{\alpha,i}$ under consideration, including resonant terms. However, for $f^{\mathsf{R}}_{\alpha,i}$ the resulting solution causes ϵg to become $\mathcal{O}(1)$ for $t = \mathcal{O}(1/\epsilon)$ time scales. Thus, on these time scales (7.1) no longer defines a near-identity change of coordinates, and the asymptotic expansions undertaken to produce (7.2) are not valid. Accordingly, we choose $h_{\alpha,i}(t)$ so that

$$\dot{h}_{\alpha,i}(t)+\Lambda_{\alpha,i}h_{\alpha,i}(t) = f^{\mathsf{NR}}_{\alpha,i}(t),$$

or

$$h_{\alpha,i}(t) = \mathrm{e}^{-\Lambda_{\alpha,i}t}\int_{T_0}^{t} \mathrm{e}^{\Lambda_{\alpha,i}\tau} f^{\mathsf{NR}}_{\alpha,i}(\tau)\mathrm{d}\tau. \tag{7.4}$$

Since Eq. (7.3) for g is linear, it is clear that if we define

$$g(y,t) = \sum_{\alpha,i} h_{\alpha,i}(t) y^{\alpha} e_i, \tag{7.5}$$

then substituting (7.5) into (7.2) and truncating at $\mathcal{O}(\epsilon)$ leaves the first-order NF equation

$$\dot{y} = Ay + \epsilon \sum_{\alpha,i} f^{\mathsf{R}}_{\alpha,i}(t) y^{\alpha} e_i. \tag{7.6}$$

This is the same equation as was produced by the RG method; see (6.6).

7.2. The relationship between the RG method and NF theory

For the nonautonomous vector fields (6.1) there is a clear connection between RG and NF theory, just as there was in the autonomous case, as shown in Sections 3 and 4. We have shown in Sections 6 and 7.1 that the two methods produce identical results. In this subsection, we also highlight that the mechanics by which they produce these results are equivalent by comparing the change of coordinates used in the two methods.

The RG procedure renormalizes arbitrary initial conditions. In the nonautonomous case, this renormalization was given in (6.5) by

$$w(t_0) = W(t_0) + \epsilon \int_{T_0}^{t_0} \sum_{\alpha,i} f^{\mathsf{NR}}_{\alpha,i}(s) \mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)} W(t_0)^{\alpha} e_i \mathrm{d}s.$$

On the other hand, the NF transformation was given in (7.1), (7.4) and (7.5) by

$$x = y + \epsilon \int_{T_0}^{t_0} \sum_{\alpha,i} f^{\mathsf{NR}}_{\alpha,i}(\tau) \mathrm{e}^{\Lambda_{\alpha,i}(\tau-t)} y^{\alpha} e_i \mathrm{d}\tau.$$

The only difference between the two transformations is that t_0 is replaced with t in the NF case. This difference corresponds to the fact that the RG method solves the backwards problem by finding evolution equations for integration constants, while NF theory works in forward time with the solution itself.

8. Nonautonomous example

In this section, we consider the Mathieu equation [18]. It is a second-order nonautonomous differential equation given by

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + (a + 2\epsilon(\cos t))y = 0. \tag{8.1}$$

Here ϵ is taken to be a small, positive parameter, and a is a real parameter. We are interested in the stability of (8.1) as ϵ and a vary. For small ϵ, the (ϵ, a) plane is filled with stable solutions except for tongues emanating from the points $a = n^2/4$ for n a positive integer. We focus on the case $n = 1$ and attempt to find an asymptotic expansion for the boundary of the stability region above $a = 1/4$, i.e. we suppose $a = 1/4 + \epsilon a_1 + \epsilon^2 a_2 + \cdots$, where we note that the use of a_i in this expansion is traditional, and we do not expect them to be confused with the coefficients a_i in the near-identity coordinate change used in the RG method.

We begin by letting $\dot{y} = x/2$ and transforming (8.1) into a system of first-order nonautonomous differential equations given by

$$\begin{aligned} \dot{x} &= -\frac{y}{2} - 4\epsilon(\cos t)y - \epsilon 2a_1 y + \mathcal{O}(\epsilon^2) \\ \dot{y} &= \frac{x}{2}. \end{aligned} \tag{8.2}$$

This system is better studied in complex notation, so we make the following invertible change of coordinates:

$$\begin{aligned} z &= x + \mathrm{i}y \\ \bar{z} &= x - \mathrm{i}y \end{aligned} \iff \begin{aligned} x &= \frac{1}{2}(z + \bar{z}) \\ y &= -\frac{\mathrm{i}}{2}(z - \bar{z}) \end{aligned}$$

which diagonalizes the linear part of (8.2). This yields the following equation (and its complex conjugate) to study:

$$\dot{z} = \frac{\mathrm{i}}{2}z + \epsilon(2\mathrm{i}(\cos t)(z-\bar{z}) + a_1\mathrm{i}(z-\bar{z})) + \mathcal{O}(\epsilon^2), \tag{8.3}$$

which is precisely of the form (1.2).

Before applying the RG method, we pause to compute the $\mathsf{KBM}_{\Lambda_{\alpha,i}}$ average of each of the terms in (8.3). Most of the averages are straightforward to compute. The $2\mathrm{i}(\cos t)\bar{z}$ term is the least trivial, and we compute its $\mathsf{KBM}_{\Lambda_{\alpha,i}}$ average explicitly:

$$f^{(\Lambda_{(0,1),1})} = \lim_{t-t_0\to\infty}\frac{1}{t-t_0}\int_{t_0}^{t} 2\mathrm{i}(\cos s)\mathrm{e}^{-\mathrm{i}s}\mathrm{d}s = \lim_{t-t_0\to\infty}\frac{2\mathrm{i}}{t-t_0}\int_{t_0}^{t}\frac{\mathrm{e}^{\mathrm{i}s}+\mathrm{e}^{-\mathrm{i}s}}{2}\mathrm{e}^{-\mathrm{i}s}\mathrm{d}s = \mathrm{i}.$$

The remaining terms are listed in the following table.

Term	α	$\Lambda_{\alpha,1}$	$f^{(\Lambda_{\alpha,i})}$
$2\mathrm{i}\cos t$	$(1,0)$	0	0
$-2\mathrm{i}\cos t$	$(0,1)$	$-\mathrm{i}$	i
$a_1\mathrm{i}$	$(1,0)$	0	$a_1\mathrm{i}$
$-a_1\mathrm{i}$	$(0,1)$	$-\mathrm{i}$	0

We now split each term into its resonant and nonresonant parts,

$$f_{\alpha,i}(t) = f^{\mathsf{R}}_{\alpha,i}(t) + f^{\mathsf{NR}}_{\alpha,i}(t).$$

This splitting is straightforward except in the case when $\alpha, i = (0,1), 1$, as we demonstrated above. Recalling our definition of $f^{\mathsf{R}}_{\alpha,i}$ in Section 6 we note

$$\begin{aligned} f^{\mathsf{R}}_{(1,0),1}(t) &= a_1\mathrm{i} \\ f^{\mathsf{R}}_{(0,1),1}(t) &= \mathrm{i}\mathrm{e}^{\mathrm{i}t}. \end{aligned} \tag{8.4}$$

We now apply the RG procedure to (8.3) by first supposing a naive perturbation expansion to the solution to first order

$$z(t) = z_0(t) + \epsilon z_1(t) + \cdots.$$

Plugging this series into (8.3) and solving order by order, we find the solution (following (6.4))

$$z(t) = \mathrm{e}^{\frac{\mathrm{i}}{2}(t-t_0)}\left(z(t_0) + \epsilon a_1\mathrm{i}(t-t_0)z - \epsilon\mathrm{i}\mathrm{e}^{\mathrm{i}t_0}(t-t_0)\bar{z} + \epsilon\int_{t_0}^{t}\sum_{\alpha,i} f^{\mathsf{NR}}_{\alpha,i}(s)\mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)}z^{\alpha}e_i\mathrm{d}s\right).$$

Since the integral on the right hand side is a nonresonant term, it grows more slowly than $(t-t_0)$, and hence we absorb it into the initial conditions by selecting

$$z(t_0) = Z(t_0) - \epsilon\int_{t_0}^{T_0}\sum_{\alpha,i} f^{\mathsf{NR}}_{\alpha,i}(s)\mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)}Z^{\alpha}e_i\mathrm{d}s.$$

This leaves us with the renormalized expansion

$$z(t) = \mathrm{e}^{\frac{\mathrm{i}}{2}(t-t_0)}\left(Z + \epsilon a_1\mathrm{i}(t-t_0)Z - \epsilon\mathrm{i}(t-t_0)\mathrm{e}^{\mathrm{i}t_0}\bar{Z} + \epsilon\int_{T_0}^{t}\sum_{\alpha,i} f^{\mathsf{NR}}_{\alpha,i}(s)\mathrm{e}^{\Lambda_{\alpha,i}(s-t_0)}Z^{\alpha}e_i\mathrm{d}s\right) + \mathcal{O}(\epsilon^2).$$

Applying the RG condition (3.1), we find the following differential equation:

$$\frac{\mathrm{d}Z}{\mathrm{d}t_0} = \frac{\mathrm{i}}{2}Z + \epsilon a_1\mathrm{i}Z + \epsilon\mathrm{i}\mathrm{e}^{\mathrm{i}t_0}\bar{Z} + \mathcal{O}(\epsilon^2),$$

which is equivalent to the normal form of (8.3) if one takes into account (8.4). Thus, this example illustrates the main result of Sections 6 and 7 that the RG and NF approaches for nonautonomous perturbations are equivalent.

To determine the stability of the original problem we follow the same procedure as in NF and the method of averaging and convert into rotating coordinates via the transformation $Z = \mathrm{e}^{\frac{\mathrm{i}}{2}t_0}W$. This yields the autonomous amplitude equation

$$\frac{\mathrm{d}W}{\mathrm{d}t_0} = \epsilon\mathrm{i}(a_1W - \bar{W}) + \mathcal{O}(\epsilon^2).$$

Splitting W into its real and imaginary parts, $W = W_R + iW_I$, we convert the complex differential equation into a real planar differential equation

$$\frac{dW_R}{dt_0} = -\epsilon(a_1 + 1)W_I$$
$$\frac{dW_I}{dt_0} = -\epsilon(a_1 - 1)W_R.$$

Standard linear analysis shows that the rotational orbit is stable if $\epsilon^2(a_1^2 - 1) > 0$ or equivalently if $|a_1| > 1$.

9. Validity of normal form theory

We have opted to compare the RG method to NF theory not only because of the similarities between the two approaches but also because NF theory can be rigorously justified. In particular we have the following theorem which states that, given a nonlinear system (6.1), there is a canonical "simplest" equation which can be used to approximate the original equation on time scales of $\mathcal{O}(1/\epsilon)$.

Theorem 9.1. *Consider the ODE (6.1) in which A is diagonal with purely imaginary eigenvalues, and consider the first-order normal form (7.6). Let $x(0) = z(0)$. Then there exist a constant $T = T(x(0)) > 0$ and a function $\phi_T(\epsilon)$ such that $\lim_{\epsilon\to 0}\phi_T(\epsilon) = 0$ and $|x - z| = \mathcal{O}(\phi_T(\epsilon))$ for all $t \in [-T/\epsilon, T/\epsilon]$ and for all ϵ sufficiently small.*

The proof of this theorem can be found in Appendix B. The asymptotic order of the function $\phi_T(\epsilon)$ in this lemma depends solely on how fast the limit in (6.3) converges for each α and i. For example, it can be shown that, if the original equation is constant, periodic or quasiperiodic, then $\phi_T(\epsilon) = \mathcal{O}(\epsilon)$. On the other hand, there are functions for which $\epsilon \ll \phi_T(\epsilon) \ll 1$. For instance, in the case of $f(t) = (\text{a periodic function}) + 1/\sqrt{t}$, the limit in (6.3) converges at the rate $1/\sqrt{T}$ as $T \to \infty$, and $\phi_T(\epsilon) = \mathcal{O}(\sqrt{\epsilon})$.

Theorem 9.1 can be applied whenever A has purely imaginary eigenvalues and is conjugate to a diagonal matrix, after applying the linear change of coordinates that diagonalizes A. Generalization of this theorem to higher orders as well as cases in which A is allowed to vary slowly, and has eigenvalues with negative real parts, will be treated in a forthcoming article. Finally, despite the fact that this approach can be extended to obtain higher order approximations, it typically does not provide approximations on time scales longer than $\mathcal{O}(1/\epsilon)$.

Acknowledgments

We thank Peter Kramer, Bob O'Malley, Ferdinand Verhulst, Djoko Wirosoetisno, and Mohammed Ziane for useful comments and discussion. LD was supported in part by the NSF's VIGRE program under NSF grant DMS-9983646. MH was supported in part by NSF grant DMS-0109427 to the Center for BioDynamics. KJ was supported in part by NSF grants DMS-0244529, ATM-0417867, and a GEAR grant from the University of Houston. TK was supported in part by NSF grant DMS-0306523.

Appendix A. Review of Poincaré–Birkhoff normal form theory for autonomous systems (1.1)

In this section, we recall the basic normal form procedure for a system of first-order, autonomous differential equations. The approach that we take here is equivalent to the classical Poincaré–Birkhoff normal form except that we grade the vector field in powers of ϵ and not based upon the degree of the polynomials in the vector field. We shall go up to and including terms of $\mathcal{O}(\epsilon^2)$ so that we may compare the NF equation to the RG equation derived earlier.

We begin with the differential equation given by (2.1) or

$$\dot{x} = Ax + \epsilon\sum_{\alpha,i} C_{\alpha,i}x^\alpha e_i \tag{A.1}$$

and compute to first order. In particular, we suppose a near-identity change of variables of the form $x = y + \epsilon w_1(y)$ for some function w_1 and plug into (A.1). After some manipulation, we find

$$\dot{y} = Ay + \epsilon(-Dw_1(y)Ay + Aw_1(y)) + \epsilon\sum_{\alpha,i} C_{\alpha,i}y^\alpha e_i + \mathcal{O}(\epsilon^2). \tag{A.2}$$

A term in $f(y)$ is nonresonant if it lies in the range of the linear operator $[Ay, w_1] = (Dw_1(y)Ay - Aw_1(y))$. Conversely, a resonant f lies in the complement of the range of this operator. We will take as a basis for the space of possible vector fields those vector fields of homogeneous monomials. Taking a different basis will yield a different normal form. For a general element of this basis, $y^\alpha e_i$, we observe that

$$[Ay, y^\alpha e_i] = \Lambda_{\alpha,i}y^\alpha e_i.$$

Therefore, if $\Lambda_{\alpha,i} = 0$ then $y^\alpha e_i$ lies in the complement to the range of $[Ay, y^\alpha e_i]$ and is resonant. On the other hand, if $\Lambda_{\alpha,i} \neq 0$ then those terms are nonresonant and may be removed.

Returning to the normal form for the specific f given in (2.1), we notice that if $\Lambda_{\alpha,i} \neq 0$, then the terms are nonresonant and may be removed by an appropriate choice of w_1. Here we choose

$$w_1 = \sum_{\Lambda_{\alpha,i}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} y^\alpha e_i. \tag{A.3}$$

Then,

$$[Ay, w_1] = Dw_1 Ay - Aw_1 = \sum_{\Lambda_{\alpha,i}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \sum_{k=1}^{n} (\alpha_k \lambda_k) y^\alpha e_i - \sum_{\Lambda_{\alpha,i}\neq 0} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} \lambda_i y^\alpha e_i = \sum_{\Lambda_{\alpha,i}\neq 0} C_{\alpha,i} y^\alpha e_i.$$

Hence, the nonresonant terms in f lie in the range of the operator $[Ay, w_1]$. Substituting (A.3) into (A.2) we find that the NF equation to first order is

$$\dot{y} = Ay + \epsilon \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} y^\alpha e_i + \mathcal{O}(\epsilon^2). \tag{A.4}$$

Remark 4. We show in Section 9 that solutions to the first-order NF equations stay $\mathcal{O}(\epsilon)$ close to solutions of the original system. This is due to the fact that the change of coordinates given in (A.3) is $\mathcal{O}(\epsilon)$ on compact sets for sufficiently small ϵ and truncating the normal form at first order only introduces $\mathcal{O}(\epsilon^2)$ error.

A second-order normal form can be obtained by introducing a second change of coordinates $y = z + \epsilon^2 w_2(z)$ to remove resonant terms at second order. Instead of explicitly computing the function $w_2(z)$ that removes all nonresonant terms at second order, we opt to simply determine what the resonant terms should be. To do this, we must expand (A.2) out to second order. We let $f^{\mathsf{R}}(z) = Aw_1 - Dw_1 Az + f(z)$ be the resonant part of f at first order and expand $(I + \epsilon w_1)^{-1} = I - \epsilon Dw_1 + \epsilon^2 (Dw_1)^2 + \mathcal{O}(\epsilon^3)$ to get

$$\dot{z} = Az + \epsilon f^{\mathsf{R}}(z) + \epsilon^2((Dw_1)^2 Az - Dw_1 Aw_1 - Dw_1 f(z) + Df(z)w_1) + \mathcal{O}(\epsilon^3).$$

Again using the definition of f^{R}, we further reduce the above equation to

$$\dot{z} = Az + \epsilon f^{\mathsf{R}}(z) + \epsilon^2 (Df(z)w_1 - Dw_1 f^{\mathsf{R}}(z)) + \mathcal{O}(\epsilon^3).$$

We compute the coefficients on ϵ^2, recalling that

$$f^{\mathsf{R}}(z) = \sum_{\Lambda_{\alpha,i}=0} C_{\alpha,i} z^\alpha e_i. \tag{A.5}$$

To illustrate the general argument, assume that f and w_1 are monomials, so that $f(z) = z^\alpha e_i$ and $w_1(z) = z^\beta e_j$. Then for each monomial one has

$$Df(z)w_1(z) = \frac{\alpha_j}{z_j} z^{\alpha+\beta} e_i.$$

Now we return to the general case and sum over all α and i. By linearity and using (A.3), we have

$$Df(z)w_1(z) = \sum_{\substack{\alpha,\beta,i,j \\ \Lambda_{\beta,j}\neq 0}} C_{\alpha,i} \frac{C_{\beta,j}}{\Lambda_{\beta,j}} \frac{\alpha_j}{z_j} z^{\alpha+\beta} e_i; \tag{A.6}$$

and, similarly using (A.5), we find

$$Dg_1(z) f^{\mathsf{R}}(z) = \sum_{\substack{\alpha,\beta,i,j \\ \Lambda_{\beta,j}\neq 0 \\ \Lambda_{\beta,j}=0}} \frac{C_{\alpha,i}}{\Lambda_{\alpha,i}} C_{\beta,j} \frac{\alpha_j}{z_j} z^{\alpha+\beta} e_i. \tag{A.7}$$

The monomials in (A.7) and (A.6) can be written as $z^{\alpha+\beta-e_j}$, where e_j is the vector which is 1 in the jth slot. Therefore, the condition for resonance (or equivalently, the condition which shows that we cannot remove a term) is

$$0 = \langle \lambda, \alpha + \beta - e_j \rangle - \lambda_i = \langle \lambda, \alpha \rangle + \langle \lambda, \beta \rangle - \lambda_i - \lambda_j,$$

which is equivalent to $\Lambda_{\alpha,i} + \Lambda_{\beta,j} = 0$. Notice that it is not possible for any term in $Dg_1(y)f^{\mathsf{R}}(y)$ to satisfy $\Lambda_{\alpha,i} + \Lambda_{\beta,j} = 0$, and we are left with

$$f_2^{\mathsf{R}}(z) = \sum_{\substack{\alpha,\beta,i,j \\ \Lambda_{\beta,j}\neq 0 \\ \Lambda_{\alpha,i}+\Lambda_{\beta,j}=0}} C_{\alpha,i}\frac{C_{\beta,j}}{\Lambda_{\beta,j}}\frac{\alpha_j}{z_j}z^{\alpha+\beta}e_i. \tag{A.8}$$

We have now completed the derivation of the NF equation up to and including $\mathcal{O}(\epsilon^2)$. Moreover, as observed at the end of Section 4, it is equivalent to the RG equation (4.6).

Remark 5. Solutions of the resulting equation will remain $\mathcal{O}(\epsilon^2)$ close to solutions of the first-order normal form for times up to and including $\mathcal{O}(1/\epsilon)$. Of course, to get an $\mathcal{O}(\epsilon^2)$ approximation to the original equation we must transform the solution by the change of coordinates

$$x = y + \epsilon w_1(y) + \epsilon^2 w_2(y).$$

Appendix B. Proof of Theorem 9.1

The proof of Theorem 9.1 is given at the end of the appendix and follows from three lemmas.

Lemma B.1. *Pick a $T > 0$, and assume that $y \in \mathbb{C}^n$, the solution of (7.6), stays within some compact set $K \subset \mathbb{C}^n$ for all $t \leq T/\epsilon$. Then for ϵ sufficiently small, the transformation*

$$x = y + \epsilon g(y, t),$$

where g is defined in (7.5), is a diffeomorphism for $|t| \leq T/\epsilon$. In fact, for any K and any $T > 0$ such that $y \in K$ for $t \leq T/\epsilon$, there is a function $\phi_T(\epsilon)$ with $\lim_{\epsilon\to 0}\phi_T(\epsilon) = 0$ such that $|x - y| = \mathcal{O}(\phi_T(\epsilon))$.

Proof. We need consider only one of the nonlinear terms $\epsilon h_{\alpha,i}(t)$ in the change of variables. Recall (7.4), namely

$$h_{\alpha,i}(t) = \mathrm{e}^{-\Lambda_{\alpha,i}t}\int_{T_0}^{t}\mathrm{e}^{\Lambda_{\alpha,i}s}f_{\alpha,i}^{\mathsf{NR}}(s)\mathrm{d}s,$$

where $f_{\alpha,i}^{\mathsf{NR}}(t) = f_{\alpha,i}(t) - f_{\alpha,i}^{\mathsf{R}}(t)$, and $f_{\alpha,i}^{\mathsf{R}}(t)$ is defined in Section 6. Note that

$$\begin{aligned}\lim_{\tau\to\infty}\frac{1}{\tau}\int_{T_0}^{\tau}\mathrm{e}^{\Lambda_{\alpha,i}t}f_{\alpha,i}^{\mathsf{NR}}(t)\mathrm{d}t &= \lim_{\tau\to\infty}\frac{1}{\tau}\int_{T_0}^{\tau}\mathrm{e}^{\Lambda_{\alpha,i}t}(f_{\alpha,i}(t) - f_{\alpha,i}^{\mathsf{R}}(t))\mathrm{d}t \\ &= \lim_{\tau\to\infty}\frac{1}{\tau}\int_{T_0}^{\tau}\mathrm{e}^{\Lambda_{\alpha,i}t}(f_{\alpha,i}(t) - f_{\alpha,i}^{(\Lambda_{\alpha,i})}\mathrm{e}^{-\Lambda_{\alpha,i}t})\mathrm{d}t \\ &= \lim_{\tau\to\infty}\frac{1}{\tau}\left(\int_{T_0}^{\tau}\mathrm{e}^{\Lambda_{\alpha,i}t}f_{\alpha,i}(t)\mathrm{d}t\right) - f_{\alpha,i}^{(\Lambda_{\alpha,i})} = 0.\end{aligned} \tag{B.1}$$

Now, we choose a $T > 0$ and define

$$\phi_{\alpha,i,T}(\epsilon) = \max_{0\leq S\leq T}\left|\frac{\epsilon}{S}\int_{T_0}^{S/\epsilon}\mathrm{e}^{\Lambda_{\alpha,i}t}f_{\alpha,i}^{\mathsf{NR}}(t)\mathrm{d}t\right|.$$

This is a measure of the response of the system to the forcing $f_{\alpha,i}^{\mathsf{NR}}(t)$. The calculation in (B.1) shows that for any KBM function f we have $\phi_{\alpha,i,T}(\epsilon) = o(1)$, as $\epsilon \to 0$.

From the definition of ϕ and (7.4),

$$\max_{0\leq t\leq T/\epsilon}\left|\epsilon h_{\alpha,i}(t)\right| \leq T\phi_{\alpha,i,T}(\epsilon),$$

or

$$\max_{0\leq t\leq T/\epsilon}|\epsilon g(x,t)| \leq \phi_T(\epsilon)\cdot\tilde{g}(x),$$

where $\phi_T(\epsilon) = \max_{\alpha,i} T\phi_{\alpha,i,T}(\epsilon) = o(1)$, as $\epsilon \to 0$, and $\tilde{g}(x)$ is a polynomial in x. ■

Lemma B.2. *Consider the ODE*

$$\dot{z} = Az + \phi_1(\epsilon)f(z,t) \tag{B.2}$$

with a given initial condition $z(T_0) = z_0$ such that $|z_0| \leq R$. Assume that A is diagonal with imaginary eigenvalues, $f(z,t)$ is continuous in z and bounded in t, $\lim_{\epsilon\to 0}\phi_1(\epsilon) = 0$, and $\phi_1(\epsilon) > 0$. Pick $\delta > 1$, and let K be any compact set which properly contains the ball of radius δR. Then there is a $T > 0$ such that $z \in K$ for all $|t| \leq T/\phi_1(\epsilon)$.

Proof. Given any compact K, there is a $C > 0$ such that $f(z,t) \leq C|z|$. Then, we compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}|z|^2 &= \langle z, \dot{z}\rangle + \langle \dot{z}, z\rangle \\ &= \langle Az, z\rangle + \langle z, Az\rangle + \phi_1(\epsilon)(\langle f(z,t), z\rangle + \langle z, f(z,t)\rangle).\end{aligned}$$

Since A is diagonal with imaginary eigenvalues, $\langle Az, z\rangle + \langle z, Az\rangle = 0$. Thus we have

$$\frac{\mathrm{d}}{\mathrm{d}t}|z|^2 \leq 2\phi_1(\epsilon)C|z|^2,$$

and this estimate holds as long as $z(t) \in K$. Applying Gronwall's Inequality, we have

$$|z(t)|^2 \leq |z(T_0)|^2 \exp(2C\phi_1(\epsilon)t).$$

This estimate holds for all t such that $z(t) \in K$. Let $T = (\ln\delta)/2C$. The first possible time that z could leave the set K is given by $\exp(2C\phi_1(\epsilon)t) \geq \delta$, or $t \geq T/\phi_1(\epsilon)$. Therefore, for all $t \leq T/\phi_1(\epsilon)$, the solution $z(t)$ stays in K. ■

Let $T(z_0, R)$ be the largest T for which Lemma B.2 holds given an initial condition z_0 and a compact set $B(R)$, the closed ball of radius R. We will see below that it is not important which compact set we work on, only that there be a compact set on which we can obtain the necessary estimates. $T(z_0, R)$ is a nondecreasing function of R. Define for each initial condition z_0 a new number

$$T^*(z_0) = \lim_{R\to\infty} T(z_0, R).$$

When we have a fixed equation with fixed initial condition, we will abuse notation and drop the dependence on the initial data; thus for an initial value problem we will speak of T^*, which is the longest $\mathcal{O}(1/\epsilon)$ time scale for which a solution of Eq. (B.2) with initial condition z_0 is defined. It should be noted that $T^* < \infty$ for many equations of the form (B.2). This means that there exists a compact set on which the approximation will be valid for all $t \leq T/\epsilon$ only as long as $T < T^*$.

There is one last lemma which relates solutions of the NF equation to its truncated counterpart.

Lemma B.3. *Consider the two equations*

$$\dot{y} = Ay + \phi_1(\epsilon)f(y,t) + \phi_2(\epsilon)g(y,t), \tag{B.3}$$

$$\dot{z} = Az + \phi_1(\epsilon)f(z,t), \tag{B.4}$$

where A is a diagonal matrix with imaginary eigenvalues, $\phi_1(\epsilon)$ and $\phi_2(\epsilon)$ are positive order functions of ϵ with $\phi_2(\epsilon) = o(\phi_1(\epsilon))$, and $f(y,t)$, $g(y,t)$ are bounded in t and continuously differentiable in y. Let T^* denote a time given by Lemma B.2 such that solutions of (B.4) stay in a compact set K for any $0 \leq T < T^*$. If ϵ is chosen sufficiently small, then for all $|t| \leq T/\phi_1(\epsilon)$, we have

$$|y(t) - z(t)| = \mathcal{O}\left(\frac{\phi_2(\epsilon)}{\phi_1(\epsilon)}\right).$$

Proof. We use an argument known as *bootstrapping* or *continuous induction*, following [49]. See also [34] for another example of this argument in a similar context, and [16] for its use in another context in perturbation theory.

Define $\xi = y - z$, and note that the differential equation for ξ is

$$\dot{\xi} = A\xi + \phi_1(\epsilon)(f(y,t) - f(z,t)) + \phi_2(\epsilon)g(y,t).$$

Solving this we find

$$\xi(t) = \phi_1(\epsilon)\mathrm{e}^{At}\int_{T_0}^{t}\mathrm{e}^{-As}(f(y,s) - f(z,s))\mathrm{d}s + \phi_2(\epsilon)\mathrm{e}^{At}\int_{T_0}^{t}\mathrm{e}^{-As}g(y,t)\mathrm{d}s.$$

Since f and g are C^1, we can write

$$\begin{aligned}|f(y,t) - f(z,t)| &\leq \sup_{\beta\in[0,1]}|D_z f((1-\beta)\xi + z, t)||\xi|, \\ |g(y,t)| &\leq |g(z,t)| + \sup_{\beta\in[0,1]}|D_z g((1-\beta)\xi + z, t)||\xi|.\end{aligned}$$

Since $|e^{At}| = 1$ for all t, we have

$$|\xi(t)| \leq \phi_1(\epsilon) \int_{T_0}^{t} \sup_{\beta\in[0,1]} |D_z f((1-\beta)\xi + z, s)||\xi| \mathrm{d}s + \phi_2(\epsilon) \int_{T_0}^{t} |g(z,s)| + \sup_{\beta\in[0,1]} |D_z g((1-\beta)\xi + z, s)||\xi| \mathrm{d}s.$$

We write

$$\eta_\epsilon(z,\xi) = \sup_{\substack{\beta\in[0,1]\\ |t|\leq T/\phi_1(\epsilon)}} |D_z f((1-\beta)\xi + z, t)| + \sup_{\substack{\beta\in[0,1]\\ |t|\leq T/\phi_1(\epsilon)}} \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} |D_z g((1-\beta)\xi + z, t)|.$$

We know that the suprema on the right hand side of this equation exist because of Lemma B.2. We know that z and, thus, ξ stay in some compact set for all $|t| \leq T/\phi_1(\epsilon)$. We then have

$$|\xi(t)| \leq \phi_1(\epsilon) \int_{T_0}^{t} |\eta_\epsilon(z,\xi)||\xi|, \mathrm{d}s + \phi_2(\epsilon) \int_{T_0}^{t} |g(z,s)| \mathrm{d}s.$$

If we write $\chi_\epsilon(z,\xi) = \int_{T_0}^{t} \eta_\epsilon(z,\xi)|\xi| \mathrm{d}s$, then we have

$$\begin{aligned} |\dot{\chi}_\epsilon| &= |\eta_\epsilon||\xi| \leq \phi_1(\epsilon)\eta_\epsilon \int_{T_0}^{t} \eta_\epsilon|\xi| \mathrm{d}s + \phi_2(\epsilon)\eta_\epsilon \int_{T_0}^{t} |g(z,s)| \mathrm{d}s \\ &\leq \phi_1(\epsilon)\eta_\epsilon\chi_\epsilon + \phi_2(\epsilon)\eta_\epsilon \int_{T_0}^{t} |g(z,s)| \mathrm{d}s. \end{aligned}$$

Now, as long as $|t| \leq T/\phi_1(\epsilon)$, we have that $|g(z,s)| \leq C_1$ (see Lemma B.2). Hence,

$$|\dot{\chi}_\epsilon| \leq \phi_1(\epsilon)\eta_\epsilon\chi_\epsilon + \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)}\eta_\epsilon C_1 T,$$

or

$$|\chi_\epsilon(t)| \leq \phi_1(\epsilon) \int_{T_0}^{t} \eta_\epsilon\chi_\epsilon \mathrm{d}s + \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} C_1 T \int_{T_0}^{t} \eta_\epsilon \mathrm{d}s.$$

Applying the integral form of Gronwall's Inequality we get

$$|\chi_\epsilon(t)| \leq \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} C_1 T \int_{T_0}^{t} \eta_\epsilon \left[\exp\left(\int_{s}^{t} \phi_1(\epsilon)\eta_\epsilon(\tau)\mathrm{d}\tau\right)\right] \mathrm{d}s. \tag{B.5}$$

Recall that we are estimating $|\xi(t)|$ for all $|t| \leq T/\phi_1(\epsilon)$. On this time scale, we have

$$\begin{aligned} |\xi(t)| &\leq \phi_1(\epsilon) \int_{T_0}^{t} \eta_\epsilon|\xi| \mathrm{d}s + \phi_2(\epsilon) \int_{T_0}^{t} |(z,s)| \mathrm{d}s \\ &\leq \phi_1(\epsilon)\chi_\epsilon + \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} C_1 T. \end{aligned}$$

Clearly, if we can estimate χ_ϵ on this time domain, we will be done. We now use the bootstrapping argument. It is clear from the definitions that $\xi(T_0) = \chi_\epsilon(T_0) = 0$. Therefore, for ϵ sufficiently small, there is a $U > 0$ such that for $|t| \leq U$, $\xi(t) < 1$. Let us choose $U_1(\epsilon) < 0 < U_2(\epsilon)$ to be the largest interval so that for $t \in [U_1, U_2]$, $|\xi(t)| \leq 1$. We will assume that $|U_1| < T/\phi_1(\epsilon)$ and $|U_2| < T/\phi_1(\epsilon)$ and try to derive a contradiction.

For $t \in [U_1, U_2]$, we know that z and ξ stay in some compact set. Also, since $\eta_\epsilon(z,\xi)$ is a continuous function of its arguments, there is a C_2 so that $|\eta_\epsilon(z,t)| \leq C_2$, and we recall that $|g(z,t)| \leq C_1$. Using these estimates in (B.5), we obtain

$$\begin{aligned} |\chi_\epsilon(t)| &\leq \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} C_1 T C_2 \int_{T_0}^{t} \exp\left[\int_{s}^{t} C_2\phi_1(\epsilon)\mathrm{d}\tau\right] \mathrm{d}s \\ &= \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} C_1 T C_2 \int_{T_0}^{t} \exp\left(C_2\phi_1(\epsilon)(t-s)\right) \mathrm{d}s \\ &= \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} C_1 T C_2 \frac{\exp\left(\phi_1(\epsilon)C_2 t\right)}{\phi_1(\epsilon)C_2} \leq \frac{\phi_2(\epsilon)}{(\phi_1(\epsilon))^2} C_1 T C_2 e^{\phi_1(\epsilon)C_2 U_3}, \end{aligned} \tag{B.6}$$

where $U_3 = \max(|U_1|, |U_2|)$. By assumption, $U_3 < T/\phi_1(\epsilon)$ and we get

$$|\chi_\epsilon(t)| \leq C_3 \frac{\phi_2(\epsilon)}{(\phi_1(\epsilon))^2}. \tag{B.7}$$

Thus we have

$$|\xi(t)| \leq C_4 \frac{\phi_2(\epsilon)}{\phi_1(\epsilon)}. \tag{B.8}$$

Since $\phi_2(\epsilon) = o(\phi_1(\epsilon))$, this is much smaller than 1 for ϵ sufficiently small, contradicting the maximality of U_1 and U_2. Therefore $|\xi(t)| < 1$ for all $|t| \leq T/\phi_1(\epsilon)$. But if this is true, then (B.6) holds for all $|t| \leq T/\phi_1(\epsilon)$, or, in short, (B.7) holds for all $|t| \leq T/\phi_1(\epsilon)$. Applying (B.5) says that (B.8) holds for all $|t| \leq T/\phi_1(\epsilon)$, and we are done. ■

Remark 6. One can see from the mechanics of the proof why the time scale $\mathcal{O}(1/\phi_1(\epsilon))$ is optimal. Notice that in the proof, we see that, roughly,

$$\chi_\epsilon(t) = \mathcal{O}\left(\frac{\phi_2(\epsilon)}{(\phi_1(\epsilon))^2} e^{\phi_1(\epsilon)t}\right) \quad \text{and} \quad \xi(t) = \mathcal{O}\left(\frac{\phi_2(\epsilon)}{\phi_1(\epsilon)} e^{\phi_1(\epsilon)t}\right).$$

It follows that once we go beyond the $1/\phi_1(\epsilon)$ time scale, the exponential growth takes over, and we cannot control this term. In short, it is usually not possible to "trade error for time scale."

On the other hand, we can also see from the mechanics of the proof that we can "trade time scale for error," i.e. if we are willing to shorten our time scale by some amount, then we can improve our estimate by exactly the same amount. In short, if we choose $\phi_3(\epsilon)$ to be any order function such that $\phi_1 = o(\phi_3)$, then we can get an error estimate of $\mathcal{O}(\phi_2\phi_3/\phi_1)$ on a time scale of $\mathcal{O}(\phi_3/\phi_1)$. Simply note that if we multiply the time scale of interest by some function of ϵ, then the Gronwall estimates in the proof will all be multiplied by the same factor.

As a consequence of these three lemmas we now have

Proof of Theorem 9.1. From Lemmas B.1 and B.2, we know that there is a T such that, for $|t| \leq T/\epsilon$, y stays in some compact set K and $x = y + \epsilon g(y, t)$ is a diffeomorphism. In fact, we have that $|x(t) - y(t)| \leq \mathcal{O}(\phi_T(\epsilon))$. Applying Lemma B.3 with $\phi_1(\epsilon) = \epsilon$, and $\phi_2(\epsilon) = \epsilon\phi_T(\epsilon)$ to Eqs. (7.6) and (7.1) we also obtain $|y(t) - z(t)| = \mathcal{O}(\phi_T(\epsilon))$ for all $|t| \leq T/\epsilon$. ■

References

[1] V.I. Arnold, Geometrical Methods in The Theory of Ordinary Differential Equations, second edition, in: Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences by Joseph Szücs, Trans.), vol. 250, Springer-Verlag, New York, 1988 (in Russian).
[2] D.K. Arrowsmith, C.M. Place, An Introduction to Dynamical Systems, Cambridge University Press, Cambridge, 1990.
[3] N.N. Bogolyubov, Y.A. Mitropolskii, Asymptotic Methods in the Theory of Nonlinear Oscillations, Gordon and Break, New York, 1961.
[4] J. Bricmont, A. Kupiainen, G. Lin, Renormalization group and the asymptotics of solutions of nonlinear parabolic equations, Comm. Pure Appl. Math. 47 (1994) 893–922.
[5] L.-Y. Chen, N. Goldenfeld, Y. Oono, Renormalization group and singular perturbations: Multiple scales, boundary layers, and reductive perturbation theory, Phys. Rev. E 543 (1) (1996) 376–394.
[6] L.-Y. Chen, N. Goldenfeld, Y. Oono, Renormalization group theory for global asymptotic analysis, Phys. Rev. Lett. 73 (10) (1994) 1311–1315.
[7] S.-N. Chow, C. Li, D. Wang, Normal Forms and Bifurcations of Planar Vector Fields, Cambridge University Press, Cambridge, UK, 1994.
[8] S. Cohn, Resonance and long time existence for the quadratic semilinear Schrödinger equation, Comm. Pure Appl. Math. 45 (8) (1992) 973–1001.
[9] J.-P. Eckmann, H. Epstein, C.E. Wayne, Normal forms for parabolic partial differential equations, Ann. Inst. H. Poincaré Phys. Theor. 58 (3) (1993) 287–308.
[10] S.-I. Ei, K. Fujii, T. Kunihiro, Renormalization-group method for reduction of evolution equations; invariant manifolds and envelopes, Ann. Phys. 28 (2) (2000) 236–298.
[11] N. Goldenfeld, Lectures on Phase Transitions and the Renormalization Group, Addison-Wesley, 1992.
[12] N. Goldenfeld, O. Martin, Y. Oono, Asymptotics of partial differential equations and the renormalization group, in: H. Segur (Ed.), Asymptotics Beyond all Orders, Plenum, New York, 1991, pp. 375–383.
[13] S. Goto, Y. Masutomi, K. Nozaki, Lie-group approach to perturbative renormalization group method, 1999. arXiv:patt-sol/9905001.
[14] J. Guckenheimer, P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields, Springer-Verlag, New York, 1983.
[15] M.H. Holmes, Introduction to Perturbation Methods, in: Texts in Applied Mathematics, vol. 20, Springer-Verlag, New York, 1995.
[16] C.K.R.T. Jones, T.J. Kaper, N. Kopell, Tracking invariant manifolds up to exponentially small errors, SIAM J. Math. Anal. 27 (2) (1996) 558–577.
[17] P.B. Kahn, Y. Zarmi, Nonlinear Dynamics, in: Wiley Series in Nonlinear Science, John Wiley & Sons Inc., New York, 1998; Exploration through Normal Forms, A Wiley-Interscience Publication.
[18] J. Kevorkian, J.D. Cole, Multiple Scale and Singular Perturbation Methods, in: Applied Mathematical Sciences, vol. 114, Springer-Verlag, New York, 1996.
[19] Y. Kodama, Normal form, symmetry, and infinite-dimensional Lie algebra for systems of ordinary differential equations, PLA 191 (1994) 223–228.
[20] T. Kunihiro, A geometrical formulation of the renormalization group method for global analysis, Progr. Theoret. Phys. 94 (1995) 503–514.
[21] T. Kunihiro, The renormalization group method applied to asymptotic analysis of vector fields, Progr. Theoret. Phys. 97 (1997) 179–200.
[22] T. Kunihiro, Renormalization-group resummation of a divergent series of the perturbative wave functions of the quantum anharmonic oscillator, Phys. Rev. D. 57 (1998) R2025–R2039.
[23] T. Kunihiro, Renormalization-group resummation of a divergent series of the perturbative wave functions of quantum systems, Progr. Theoret. Phys. 131 (Suppl.) (1998) 459–470.
[24] Y. Kuramoto, On the reduction of evolution equations in extended systems, Progr. Theoret. Phys. 99 (Suppl.) (1989) 244–261.
[25] P. Lagerstrom, Matched Asymptotic Expansions: Ideas and Techniques, in: Applied Mathematical Sciences, vol. 76, Springer-Verlag, New York, 1996.
[26] A. Lichtenberg, M. Lieberman, Regular and Stochastic Motion, in: Applied Mathematical Sciences, vol. 38, Springer-Verlag, Heidelberg, 1983.

[27] Y.A. Mitropolsky, A.K. Lopatin, Nonlinear Mechanics, Groups and Symmetry, in: Mathematics and its Applications, vol. 319, Kluwer Academic Publishers Group, Dordrecht, 1995, Translated from the 1988 Russian original by O.S. Limarchenko, A.N. Khruzin and T.P. Boichuk and revised by the authors.
[28] I. Moise, M. Ziane, Renormalization group method. Applications to partial differential equations, J. Dynam. Differential Equations 13 (2) (2001) 275–321.
[29] J.A. Morrison, Comparison of the modified method of averaging and the two variable expansion procedure, SIAM Rev. 8 (1966) 66–85.
[30] B. Mudavanhu, Singular Perturbation techniques: The multiple scales, renormalization group and invariance condition methods. Ph.D. Thesis, University of Washington, 1999.
[31] B. Mudavanhu, R.E. O'Malley, A renormalization group method for nonlinear oscillators, Stud. Appl. Math. 107 (1) (2001) 63–79.
[32] B. Mudavanhu, R.E. O'Malley, A new renormalization method for the asymptotic solution of weakly nonlinear vector systems, SIAM J. Appl. Math. 63 (2) (2002) 373–397.
[33] J. Murdock, Normal Forms and Unfoldings for Local Dynamical Systems, in: Springer Monographs in Mathematics, Springer-Verlag, New York, 2003.
[34] J. Murdock, Perturbations, in: Classics in Applied Mathematics, vol. 27, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999. Theory and Methods, Corrected reprint of the 1991 original.
[35] J. Murray, Asymptotic Analysis, in: Applied Mathematical Sciences, vol. 48, Springer-Verlag, New York, 1984.
[36] A.H. Nayfeh, Method of Normal Forms, in: Wiley Series in Nonlinear Science, John Wiley & Sons Inc., New York, 1993. A Wiley-Interscience Publication.
[37] K. Nozaki, Y. Oono, Renormalization-group theoretical reduction, Phys. Rev. E. 63 (046101) (2001) 18 pages.
[38] R.E. O'Malley, D.B. Williams, Deriving amplitude equations for weakly-nonlinear oscillators and their generalizations, J. Comput. Appl. Math. 190 (1–2) (2006) 3–21.
[39] R.E. O'Malley, Singular Perturbation Methods for Ordinary Differential Equations, in: Appl. Math. Sci. Series, vol. 89, Springer-Verlag, New York, 1991.
[40] G.C. Paquette, Renormalization group analysis of differential equations subject to slowly modulated perturbations, Physica A 276 (1) (2000) 122–163.
[41] J.A. Sanders, F. Verhulst, Averaging Methods in Nonlinear Dynamical Systems, Springer-Verlag, New York, 1985.
[42] J. Shatah, Normal forms and quadratic nonlinear Klein–Gordon equations, Comm. Pure Appl. Math. 38 (5) (1985) 685–696.
[43] S. Siegmund, Normal forms for nonautonomous differential equations, J. Differential Equations 178 (2) (2002) 541–573.
[44] R.M. Temam, D. Wirosoetisno, On the solutions of the renormalized equations at all orders, Adv. Differential Equations 8 (8) (2003) 1005–1024.
[45] S.L. Woodruff, The use of an invariance condition in the solution of multiple-scale singular perturbation problems: Ordinary differential equations, Stud. Appl. Math. 90 (3) (1993) 225–248.
[46] S.L. Woodruff, A uniformly valid asymptotic solution to a matrix system of ordinary differential equations and a proof of its validity, Stud. Appl. Math. 94 (4) (1995) 393–413.
[47] Y.Y. Yamaguchi, Y. Nambu, Renormalization group equations and integrability in Hamiltonian systems, Progr. Theoret. Phys. 100 (1998) 199–204.
[48] Y.Y. Yamaguchi, Y. Nambu, Renormalization group method and canonical perturbation theory, 1999. arXiv:chao-dyn/9902013.
[49] M. Ziane, On a certain renormalization group method, J. Math. Phys. 41 (5) (2000) 3290–3299.