On averaging methods for partial differential equations*

Ferdinand Verhulst
Mathematisch Instituut
University of Utrecht
PO Box 80.010, 3508 TA Utrecht
The Netherlands

1999

Contents

*To be published in "SPT98-Symmetry and Perturbation Theory II", A.Degasperis and G.Gaeta eds., World Scientific (1999)

1 Introduction

The analysis of weakly nonlinear partial differential equations both qualitatively and quantitatively, is emerging as an exciting field of investigation. In this report we consider specific results, related to averaging but we do not aim at completeness. The sections 3.1 and 3.4 contain important material which is not easily accessible in the literature.
Of the literature which we will not discuss in detail we should mention: Formal approximation methods which have been nicely presented by Cole and Kevorkian [1981]. A number of formal methods for nonlinear hyperbolic equations on unbounded domains have been analysed with respect to the question of asymptotic validity by van der Burgh [1979].
An adaptation of the Poincaré-Lindstedt method for periodic solutions of weakly nonlinear hyperbolic equations was given by Hale [1967]; note that this is a rigorous method, based on the implicit function theorem. An early version of the Galerkin-averaging method, presented in section 3.2, can be found in Rafel [1983] who considers vibrations of bars.
In the sequel ε will always be a small, positive parameter.

1.1 Qualitative aspects

The analysis of asymptotic approximations with proofs of validity rests firmly on the qualitative theory of weakly nonlinear partial differential equations. Existence and uniqueness results are avalailable which involve typically contraction, other fixed point methods and maximum principles; we will also use projection methods in Hilbert spaces.
Some of our examples will concern conservative systems (sections 3.3-4). In the theory of finite-dimensional Hamiltonian systems we have for nearly-integrable systems the celebrated KAM-theorem which, under certain non-degeneracy conditions, guarantees the persistence of many tori in the non-integrable system. For infinite-dimensional, conservative systems we now have the KKAM-theorems developed by Kuksin [1991]. Finite-dimensional invariant manifolds obtained in this way are densely filled with quasi-periodic orbits; these are the kind of solutions we obtain by our approximation methods. It is stressed however, that identification of approximate solutions with solutions covering invariant manifolds is only possible if the validity of the approximation has been demonstrated.

1.2 Averaging for ode's

Averaging is concerned with equations which have been put into the Lagrange standard form:

$$\dot{x} = \varepsilon f(t, x), \quad x(0) = x_0 \quad , \tag{1}$$

where $x \in R^n$, $f(t, x)$ is T-periodic in t and x_0 is the initial value of $x(t)$. We introduce the average:

$$f^0(x) = \frac{1}{T} \int_0^T f(t, x) dt$$

and consider the associated system:

$$\dot{y} = \varepsilon f^0(y), \quad y(0) = x_0 \quad . \tag{2}$$

Solving (2) we have found an approximation of $x(t)$, as under rather general conditions we can prove:

$$x(t) - y(t) = O(\varepsilon) \quad \text{for} \quad 0 \leq \varepsilon t \leq C$$

with O the standard order symbol, C a positive constant independent of ε. Sometimes this is also expressed verbally as: $y(t)$ is an order ε-approximation of $x(t)$ on the time-scale $1/\varepsilon$. Note that as ε is a small parameter, this is a long time-scale. Proofs can be found in the monographs Sanders and Verhulst [1985]; see also Bogoliubov and Mitropolsky [1961] and Verhulst [1996] for an introduction.
Four remarks should be added.

1. First, $f(t, x)$ might not be T-periodic but may have an average in the sense that:
$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(t, x) dt = f^0(x)$$
exists. The averaging procedure then also produces an approximation of $x(t)$ on the time-scale $1/\varepsilon$; in this case the error may be somewhat larger than ε, for instance $\sqrt{\varepsilon}$ or $\varepsilon \ln \varepsilon$. If $f(t, x)$ is quasi-periodic, i.e. f can be expressed as a finite sum of periodic functions with periods independent over the reals, then the error is again $O(\varepsilon)$.

2. If equation (2) contains an equilibrium solution c, so $f^0(c) = 0$, while c is a hyperbolic fixed point (there are no eigenvalues with real part zero) then there exists a T-periodic solution $\phi(t)$ in an ε-neighbourhood of $x_0 = c$. We have $\phi(t) = c + O(\varepsilon)$ for all time. Moreover, if $y(t) = c$ is an asymptotically stable solution of (2), then any solution $x(t)$ of (1) which starts in an interior subset of the domain of attraction of c, is approximated by the solution of the corresponding averaged equation for all time. For a precise formulation and proof see Sanders and Verhulst [1985].

3. To derive the standard form (1) is not altogether trivial, especially if we start with a nonlinear, unperturbed problem. See Verhulst [1996] and for nonlinear, unperturbed problems Sanders and Verhulst [1985], Rand [1990], Coppola and Rand [1991].

4. If necessary, we may calculate higher order approximations of the solutions of (1). This yields an improvement of the error estimate but the validity of this result is usually still on the time scale $1/\varepsilon$.

2 Operators with a continuous spectrum

Various forms of averaging techniques are being used in the literature. They are sometimes indicated by terms like 'homogenization' or 'regularization' methods and their main purpose is to stabilize numerical integration schemes for partial differential equations. However, apart from numerical improvements we are also interested in asymptotic estimates of validity and in qualitative characteristics of the solutions. This will be the subject of the subsequent sections.

2.1 Averaging of operators

A typical problem formulation would be to consider the Gauchy problem (or later an initial-boundary value problem) for equations like

$$u_t + Lu = \varepsilon f(u), \quad t > 0, u(0) = u_0. \tag{3}$$

L is a linear operator, $f(u)$ represents the nonlinear terms.
To obtain a standard form (1) suitable for averaging in the case of a partial

differential equation can already pose a formidable technical problem, even in the case of simple geometries. However it is reasonable to suppose that one can solve the 'unperturbed' ($\varepsilon = 0$) problem in sufficient explicit form before proceeding to the nonlinear equation.

A number of authors, in particular in the former Sowjet-Union, have addressed problem (3). For a survey of results see Mitropolskii, Khoma and Gromyak [1997]; see also Shtaras [1987].

There still doesnot exist a unified mathematical theory with a satisfactory approach to higher order approximations (normalisation to arbitrary order) and enough convincing examples.

Here we shall follow the theory developed by Krol [1991] which has some interesting applications. Consider the problem (3) with two spatial variables x, y and time t; assume that, after solving the unperturbed problem, by a variation of constants procedure we can write the problem in the form

$$\frac{\partial F}{\partial t} = \varepsilon L(t) F, \quad F(x, y, 0) = \gamma(x, y). \tag{4}$$

We have

$$L(t) = L_2(t) + L_1(t) \tag{5}$$

where

$$L_2(t) = b_1(x, y, t)\frac{\partial^2}{\partial x^2} + b_2(x, y, t)\frac{\partial^2}{\partial x \partial y} + b_3(x, y, t)\frac{\partial^2}{\partial y^2}, \tag{6}$$

$$L_1(t) = a_1(x, y, t)\frac{\partial}{\partial x} + a_2(x, y, t)\frac{\partial}{\partial y} \tag{7}$$

in which $L_2(t)$ is a uniformly elliptic operator on the domain, L_1, L_2 and so L are T-periodic in t; the coefficients a_i, b_i and γ are C^∞ and bounded with bounded derivatives.

We average the operator L by averaging the coefficients a_i, b_i over t:

$$\bar{a}_i(x, y) = \frac{1}{T}\int_0^T a_i(x, y, t)dt, \quad \bar{b}_i(x, y) = \frac{1}{T}\int_0^T b_i(x, y, t)dt, \tag{8}$$

producing the averaged operator $\bar{L}$. As an approximating problem for (4) we now take

$$\frac{\partial \bar{F}}{\partial t} = \varepsilon \bar{L}\bar{F}, \quad \bar{F}(x, y, 0) = \gamma(x, y). \tag{9}$$

A rather straightforward analysis shows existence and uniqueness of the solutions of problems (4) and (9) on the time-scale $1/\varepsilon$. Krol [1991] proves the following theorem:
Let F be the solution of initial value problem (4) and $\bar{F}$ the solution of initial value problem (9), then we have the estimate $\|F - \bar{F}\| = O(\varepsilon)$ on the time-scale $1/\varepsilon$. The norm $\|.\|$ is the supnorm on the spatial domain and on the time-scale $1/\varepsilon$.
The classical approach to prove such a theorem would be to transform equation (4) by a near-identity transformation to an averaged equation which satisfies equation (4) to a certain order in ε. In this approach we meet in our estimates fourth-order derivatives of F; this puts serious restrictions on the method. Instead Krol [1991] applies a near-identity transformation to $\bar{F}$ which is autonomous and about which we have explicit information.The actual proof involves barrier functions and the Phragmèn-Lindelöf principle (see for instance Protter and Weinberger [1967]).

2.2 Application to a time-periodic advection-diffusion problem

As an application Krol [1991] considers the transport of material (chemicals or sediment) by advection and diffusion in a tidal basin. In this case the advective flow is nearly periodic and diffusive effects are small. The problem can be formulated as

$$\frac{\partial C}{\partial t} + \nabla.(uC) - \varepsilon \Delta C = 0, \quad C(x, y, 0) = \gamma(x, y), \tag{10}$$

where $C(x, y, t)$ is the concentration of the transported material, the flow $u = u_0(x, y, t) + \varepsilon u_1(x, y)$ is given; u_0 is T-periodic in time and represents the tidal flow, εu_1 is a small reststream. As the diffusion process is slow, we are interested in a long time-scale approximation.
If the flow is divergence-free the unperturbed ($\varepsilon = 0$) problem is given by

$$\frac{\partial C_0}{\partial t} + u_0 \nabla C_0 = 0, \quad C_0(x, y, 0) = \gamma(x, y), \tag{11}$$

a first-order equation which can be integrated along the characteristics with solution $C_0 = \gamma(Q(t)(x, y))$. In the spirit of variation of constants we introduce the change of variables

$$C(x, y, t) = F(Q(t)(x, y), t) \tag{12}$$

We expect F to be slowly time-dependent when introducing (12) into the original equation (10). Using again the technical assumption that the flow $u_0 + \varepsilon u_1$ is divergence-free we find a slowly varying equation of the form (4). Note that the assumption of divergence-free flow is not essential, it only facilitates the calculations.
Krol [1991] presents some extensions of the theory and explicit examples where the slowly varying equation is averaged to obtain a time-independent parabolic problem. Quite often the latter problem still has to be solved numerically and one may wonder what then the use is of this technique. The answer is that one needs solutions on a long time-scale and that numerical integration of an equation where the fast periodic oscillations have been eliminated is a much safer procedure.
In the analysis presented thus far we have considered unbounded domains. To study the equation on spatially bounded domains, adding boundary conditions, doesnot present serious obstacles to the techniques and the proofs. An example is given below.

2.3 Boundary conditions and sources

An extension of the advection-diffusion problem has been obtained by Heijnekamp et al. [1995]. They considered the problem with initial and boundary values on the two-dimensional domain $\Omega, 0 \leq t < \infty$

$$\frac{\partial C}{\partial t} + \nabla.(uC) - \varepsilon \Delta C + \varepsilon f(C) = \varepsilon B(x, y, t), \tag{13}$$

$$C(x, y, 0) = \gamma(x, y), (x, y) \in \Omega \tag{14}$$

$$C(x, y, t) = 0, (x, y) \in \partial\Omega \times [0, \infty). \tag{15}$$

The flow u is expressed as above, the term $f(C)$ is a small reaction-term representing for instance the reactions of material with itself or the settling down of sediment; $B(x, y, t)$ is a T-periodic source term, for instance representing dumping of material.
Note that we chose the Dirichlet problem; the Neumann problem would be more realistic but it presents some problems, boundary layer corrections and complications in the proof of asymptotic validity which we avoid here.
The next step is to obtain the standard form (4) by the variation of constants procedure (12) which yields

$$U_t = \varepsilon L(t)U - \varepsilon f(U) + \varepsilon D(x, y, t) \tag{16}$$

where $L(t)$ is a uniform elliptic T-periodic operator generated by the (unperturbed) time t flow operator as before, $D(x,y,t)$ is produced by the inhomogeneous term B. Averaging over time t produces the averaged equation

$$\bar{U}_t = \varepsilon \bar{L}\bar{U} - \varepsilon \bar{f}(x,y,\bar{U}) + \varepsilon \bar{D}(x,y) \tag{17}$$

with appropriate initial-boundary values.
Krol's [1991] theorem formulated above produces that $U(t) - \bar{U}(t) = O(\varepsilon)$ on the time-scale $1/\varepsilon$. It is interesting that we can obtain a stronger result in this case. Using sub- and supersolutions in the spirit of maximum principles (Protter and Weinberger, 1967) we can show that the $O(\varepsilon)$ estimate is *valid for all time.*
Another interesting aspect is that the presence of the source term triggers off the existence of a unique periodic solution which is attracting the flow. In the theory of averaging in the case of ordinary differential equations the existence of a periodic solution is derived from the implicit function theorem. In the case of averaging of this parabolic initial-boundary value problem one has to use a topological fixed point theorem.
The paper by Heijnekamp et al. [1995] contains an explicit example for a circular domain with reaction-term $f(C) = aC^2$ and for the source term B Dirac-delta functions.

3 Operators with a discrete spectrum

In this section we shall be concerned with weakly nonlinear hyperbolic equations of the form

$$u_{tt} + Au = \varepsilon g(u, u_t, t, \varepsilon) \tag{18}$$

where A is a positive, selfadjoint linear differential operator on a separable real Hilbert space. Equation (18) can be studied in various ways. First we shall discuss theorems by Buitelaar [1993] who considers more general semilinear wave equations with a discrete spectrum to prove asymptotic estimates on the $1/\varepsilon$ time-scale.
The procedure involves solving an equation corresponding with an infinite number of ordinary differential equations. In most interesting cases resonance will make this virtually impossible and we have to take recourse to truncation techniques; we discuss results by Krol [1989] on the asymptotic

validity of truncation methods which at the same time yield information on the time-scale of interaction of modes.
Another fruitful approach of weakly nonlinear wave equations like (18) is by the multiple times-scales method. In the discussion and the examples we shall compare some of the methods.

3.1 Averaging results by Buitelaar

Consider the semilinear initial value problem

$$\frac{dw}{dt} + \mathcal{A}w = \varepsilon f(w, t, \varepsilon), \quad w(0) = w_0 \tag{19}$$

where $-\mathcal{A}$ generates a uniformly bounded C_0-group $H(t), -\infty < t < +\infty$, on the separable Hilbert space X (in fact the original formulation is on a Banach space but here we focus on Hilbert spaces), f satisfies certain regularity conditions and can be expanded with respect to ε in a Taylorseries, at least to some order. A generalized solution is defined as a solution of the integral equation

$$w(t) = H(t)w_0 + \varepsilon \int_0^t H(t-s) f(w(s), s, \varepsilon) ds. \tag{20}$$

Using the variation of constants transformation $w(t) = H(t)z(t)$ we find the integral equation corresponding with the standard form

$$z(t) = w_0 + \varepsilon \int_0^t F(z(s), s, \varepsilon) ds, \quad F(z, s, \varepsilon) = H(-s) f(H(s)z, s, \varepsilon). \tag{21}$$

Introduce the average F^0 of F by

$$F^0(z) = \lim_{T \to \infty} \frac{1}{T} \int_0^T F(z, s, 0) ds \tag{22}$$

and the averaging approximation $\bar{z}(t)$ of $z(t)$ by

$$\bar{z}(t) = w_0 + \varepsilon \int_0^t F_0(\bar{z}(s)) ds. \tag{23}$$

We mention that:

- f has to be Lipschitz-continuous and uniformly bounded on $\bar{D}\times[0,\infty)\times[0,\varepsilon_0]$ where D is an open, bounded set in the Hilbertspace X.

- F is Lipschitz-continuous in D, uniformly in t and ε.

Under these rather general conditions Buitelaar [1993] proves that $z(t)-\bar{z}(t)=o(1)$ on the time-scale $1/\varepsilon$.
In the case that $F(z,t,\varepsilon)$ is T-periodic in t we have the estimate $z(t)-\bar{z}(t)=O(\varepsilon)$ on the time-scale $1/\varepsilon$.

It turns out that in the right frame-work we can use again the methods of proof as they were developed for averaging in ordinary differential equations. Note that, assuming that X is a separable Hilbert space and that $-i\mathcal{A}$ is self-adjoint and generates a denumerable, complete orthonormal set of eigenfunctions, we have that if $f(w,t,\varepsilon)$ is almost-periodic, $F(z,t,\varepsilon)$ is almost-periodic in t. This means that in this case the limit $F_0(z)$ exists.
A straightforward application is then to consider semilinear initial value problems of hyperbolic type

$$u_{tt}+Au=\varepsilon g(u,u_t,t,\varepsilon),\quad u(0)=u_0,u_t(0)=v_0 \tag{24}$$

where A is a positive self-adjoint linear operator on a separable Hilbert space. An example is the wave equation

$$u_{tt}-u_{xx}=\varepsilon f(u,u_x,u_t,t,x,\varepsilon),\quad t\geq 0,0<x<1 \tag{25}$$

where

$$u(0,t)=u(1,t)=0,u(x,0)=\phi(x),u_t(x,0)=\psi(x),0\leq x\leq 1;$$

also the Klein-Gordon equation

$$u_{tt}-u_{xx}+a^2u=\varepsilon u^3,\quad t\geq 0,0<x<\pi,a>0. \tag{26}$$

Buitelaar presents extensive applications to rod and beam equations. A rod problem with extension and torsion produces two linear and nonlinearly coupled Klein-Gordon equations which is a system with many resonances. A number of cases are explored.

3.2 Galerkin-averaging results by Krol

The averaging result by Buitelaar is of importance in its generality, in many interesting cases however the resulting averaged system is still difficult to analyse and we need additional theorems. One of the most important techniques involves truncation which has been discussed by Krol [1989].
Consider again the initial-boundary value problem for the nonlinear wave equation (25). The normalized eigenfunctions of the unperturbed ($\varepsilon = 0$) problem are $v_n(x) = \sqrt{2}\sin(n\pi x), n = 1, 2, \cdots$ and we propose to expand the solution of the initial-boundary value problem for equation (25) in a Fourier series with respect to these eigenfunctions of the form

$$u(t,x) = \sum_{n=1}^{\infty} u_n(t)v_n(x). \tag{27}$$

By taking inner products this yields an infinite system of ordinary differential equations which is equivalent to the original problem. The next step is then to truncate this infinite dimensional system and apply averaging to the truncated system. The truncation is known as Galerkin's method and one has to estimate the combined error of truncation and averaging.

The first step is that (25) with its initial-boundary values has exactly one solution in a suitably chosen Hilbert space $\mathcal{H}_k = H_0^k \times H_0^{k-1}$ where H_0^k are the well-known Sobolev spaces consisting of functions u with derivatives $U^{(k)} \in L_2[0,1]$ and $u^{(2l)}$ zero on the boundary whenever $2l < k$. It is not trivial but rather standard to establish existence and uniqueness of solutions *on the time-scale* $1/\varepsilon$ under certain mild conditions on f; examples are righthandsides f like $u^3, uu_t^2, \sin u, \sinh u_t$ etc. Moreover we note that:

1. If $k \geq 3, u$ is a classical solution of equation (25).

2. If $f = f(u)$ is an odd function of u, one can find an even energy integral. If such an integral represents a positive definite energy integral, we are able to prove existence and uniqueness for all time; see also Reed [1976].

In Galerkin's truncation method one considers only the first N modes of the expansion (27) which we shall call the projection u_N of the solution u on a N-dimensional space. To find u_N we have to solve a $2N$-dimensional system of ordinary differential equations for the expansion coefficients $u_n(t)$ with appropriate (projected) initial values. The estimates for the error $\|u - u_N\|$ depend strongly on the smoothness of the righthandside f of equation (25) and the initial values $\phi(x), \psi(x)$ but, remarkably enough, not on ε. Krol [1989] finds supnorm estimates on the time-scale $1/\varepsilon$ and as $N \to \infty$ of the form

$$\begin{aligned} \|u - u_N\|_\infty &= O(N^{\frac{1}{2}-k}) &(28)\\ \|u_t - u_{Nt}\|_\infty &= O(N^{\frac{3}{2}-k}). &(29) \end{aligned}$$

We shall return later to estimates in the analytic case.

As mentioned before the truncated system is in general difficult to solve. Averaging as described in section 1 for the periodic case, produces an approximation $\bar{u}_N$ of u_N and finally the following

Galerkin-averaging theorem (Krol [1989])
Consider the initial-boundary value problem

$$u_{tt} - u_{xx} = \varepsilon f(u, u_x, u_t, t, x, \varepsilon), \quad t \geq 0, 0 < x < 1 \tag{30}$$

where

$$u(0,t) = u(1,t) = 0, u(x,0) = \phi(x), u_t(x,0) = \psi(x), 0 \leq x \leq 1.$$

Suppose that f is k-times continuously differentiable and satisfies the existence and uniqueness conditions on the time-scale $1/\varepsilon, (\phi, \psi) \in \mathcal{H}_k$; if the solution of the initial-boundary problem is (u, u_t) and the approximation obtained by the Galerkin-averaging procedure $(\bar{u}_N, \bar{u}_{Nt})$ we have on the time-scale $1/\varepsilon$

$$\begin{aligned} \|u - \bar{u}_N\|_\infty &= O(N^{\frac{1}{2}-k}) + O(\varepsilon), \quad N \to \infty, \varepsilon \to 0 &(31)\\ \|u_t - \bar{u}_{Nt}\|_\infty &= O(N^{\frac{3}{2}-k}) + O(\varepsilon), \quad N \to \infty, \varepsilon \to 0. &(32) \end{aligned}$$

There are a number of remarks:

- Taking $N = O(\varepsilon^{-\frac{2}{2k-1}})$ we obtain an $O(\varepsilon)$-approximation on the time-scale $1/\varepsilon$. So, the required number of modes decreases when the regularity of the data and the order up to which they satisfy the boundary conditions, increases.

- However, this decrease of the number of required modes is not uniform in k. So it is not obvious for which choice of k the estimates are optimal at a given value of ε.

- An interesting case arises if the nonlinearity f satisfies the regularity conditions for all k. This happens for instance if f *is an odd polynomial in* u and with analytic initial values. In such cases the results can be improved by introducing Hilbert spaces of analytic functions (so-called Gevrey classes). The estimates by Krol [1989] for the approximations on the time-scale $1/\varepsilon$ obtained by the Galerkin-averaging procedure become in this case

$$\begin{aligned} \|u - \bar{u}_N\|_\infty &= O(N^{-1}a^{-N}) + O(\varepsilon), \quad N \to \infty, \varepsilon \to 0 && (33) \\ \|u_t - \bar{u}_{Nt}\|_\infty &= O(a^{-N}) + O(\varepsilon), \quad N \to \infty, \varepsilon \to 0, && (34) \end{aligned}$$

  where the constant a arises from the bound one has to impose on the size of the strip around the real axis on which analytic continuation is permitted in the initial-boundary value problem.
  The important implication is that, because of the a^{-N}-term we need only $N = O(|log\varepsilon|)$ terms to obtain an $O(\varepsilon)$ approximation on the time-scale $1/\varepsilon$.
  We shall return to this important point in section 3.4.

- Here and in the sequel we have chosen Dirichlet boundary conditions. It is stressed that this is by way of example and not a restriction. We can also use the method for Neumann conditions, periodic boundary conditions etc.

- It is possible to generalize these results to higher dimensional (spatial) problems; see Krol [1989] for remarks and Pals [1996] for an analysis of a two-dimensional nonlinear Klein-Gordon equation with Dirichlet boundary conditions. Also it is possible to include dispersion although not without some additional difficulties; see section 3.3.

3.3 A nonlinear Klein-Gordon equation

As a prototype of a nonlinear wave equation with dispersion consider the nonlinear Klein-Gordon equation

$$u_{tt} - u_{xx} + u = \varepsilon u^3, \quad t \geq 0, 0 < x < \pi \qquad (35)$$

with boundary conditions $u(0,t) = u(\pi,t) = 0$ and initial values $u(x,0) = \phi(x), u_t(x,0) = \psi(x)$ which are supposed to be sufficiently smooth.
The problem has been studied by many authors, for an introduction to formal approximation procedures see Kevorkian and Cole [1981].
What do we know qualitatively? It follows from Krol's [1989] analysis that we have existence and uniqueness of solutions on the time-scale $1/\varepsilon$ and for all time if we add a minus sign on the righthand side. Kuksin [1991], Bobenko and Kuksin [1995] consider Klein-Gordon equations as a perturbation of the (integrable) sine-Gordon equation and prove, in an infinite-dimensional version of KAM-theory, the persistence of most finite-dimensional invariant manifolds in system (35). See also the subsequent discussion of results by Bourgain [1996] and Bambusi [1998].
We start with the eigenfunction expansion (27) where we have

$$v_n(x) = \sin(nx), \lambda_n^2 = n^2 + 1, n = 1, 2, \cdots$$

for the eigenfunctions and eigenvalues. Substituting the expansion in the equation (35) and taking the L_2 inner product with $v_n(x)$ for $n = 1, 2, \cdots$ produces an infinite number of coupled ordinary differential equations. As the spectrum is nonresonant we can average (or normalize) to any truncation number N. The result is that the actions are constant to this order of approximation, the angles are varying slowly as a function of the energy level of the modes.
With regards to the asymptotic character of the estimates we can make the following observations:

- Stroucken and Verhulst [1987] prove that, depending on the smoothness of the initial values (ϕ, ψ) we need $N = O(\varepsilon^{-\beta})$ modes (β a positive constant) to obtain an $O(\varepsilon^{\alpha})$ approximation ($0 < \alpha \leq 1$) on the time-scale $1/\varepsilon$.

- Note that according to Buitelaar [1993], section 3.2, we have the case of averaging of an almost periodic infinite dimensional vector field which yields an $o(1)$ approximation on the time-scale $1/\varepsilon$ in the case of general, smooth initial values.

- It is not difficult to improve the result in the case of finite-mode initial values, i.e. the initial values can be expressed in a finite number of eigenfunctions $v_n(x)$. In this case the error becomes $O(\varepsilon)$ on the time-scale $1/\varepsilon$ if N is taken large enough.

- Using the method of two time-scales van Horssen and van der Burgh [1988] construct an asymptotic approximation of the same form with estimate $O(\varepsilon)$ on the time-scale $1/\sqrt{\varepsilon}$. Van Horssen [1992] develops a method to prove an $O(\varepsilon)$ approximation on the time-scale $1/\varepsilon$ which is applied to the nonlinear Klein-Gordon equation with a quadratic nonlinearity $(-\varepsilon u^2)$.

- Stroucken and Verhulst [1987] also have constructed a second-order approximation in the small parameter ε. It turns out that there exists a small interaction between modes with number m and number $3m$ which probably involves much longer time-scales than $1/\varepsilon$. This is still an open problem.

- Bourgain [1996] considers the nonlinear Klein-Gordon equation (35) in the rather general form

$$u_{tt} - u_{xx} + V(x)u = \varepsilon f(u), \quad t \geq 0, 0 < x < \pi \tag{36}$$

  with V a periodic, even function and $f(u)$ an odd polynomial in u. Assuming rapid decrease of the amplitudes in the eigenfunction expansion (27) and diophantine (non-resonance) conditions on the spectrum, it is proved that *infinite*-dimensional invariant tori persist in the nonlinear wave equation (36) corresponding with almost-periodic solutions. The proof involves a perturbation expansion which is valid on a time-scale $1/\varepsilon^M$ with $M > 0$ a fixed number.

- Bambusi [1998] considers the nonlinear Klein-Gordon equation (35) in the general form

$$u_{tt} - u_{xx} + mu = \varepsilon\phi(x, u), \quad t \geq 0, 0 < x < \pi \tag{37}$$

  and the same boundary conditions. The function $\phi(x, u)$ is polynomial in u, entire analytic and periodic in x and odd in the sense that $\phi(x, u) = -\phi(-x, -u)$.
  Under a certain non-resonance condition on the spectrum Bambusi

[1998] shows that the solutions remain close to finite-dimensional invariant tori corresponding with quasi-periodic motion on time-scales longer than $1/\varepsilon$.
The results of Bourgain [1996] and Bambusi [1998] add to the understanding and interpretation of the averaging results and, as we are describing phenomena which are really there, it raises the question of how to obtain longer time-scale approximations.

3.4 A nonlinear wave equation with infinitely many resonances

In Kevorkian and Cole [1981] and Stroucken and Verhulst [1987] a more exciting and difficult problem is briefly discussed: the inital-boundary value problem

$$u_{tt} - u_{xx} = \varepsilon u^3, \quad t \geq 0, 0 < x < \pi \tag{38}$$

with boundary conditions $u(0,t) = u(\pi,t) = 0$ and initial values $u(x,0) = \phi(x), u_t(x,0) = \psi(x)$ which are supposed to be sufficiently smooth.
Starting with an eigenfunction expansion (27) we have

$$v_n(x) = \sqrt{\frac{2}{\pi}} \sin(nx), \lambda_n^2 = n^2, n = 1, 2, \cdots$$

for the eigenfunctions and eigenvalues. The authors note that as there are an infinite number of resonances, after applying the two time-scales method or averaging, we still have to solve an infinite system of coupled ordinary differential equations. In fact the problem is reminiscent of the famous Fermi-Pasta-Ulam problem (see for references and a discussion for instance Jackson [1978, 1991]) and it displays similar interaction between the modes and recurrence.
Apart from numerical approximation, Galerkin-averaging seems to be a possible approach and we state here the application of Krol's [1989] theory to this problem with the cubic term. Suppose that for the initial values ϕ, ψ we have a finite-mode expansion of M modes only; of course we take $N \geq M$ in the eigenfunction expansion. Now the initial values ϕ, ψ are analytic and Krol [1989] optimizes the way in which the analytic continuation of the initial values takes place. The analysis leads to the estimate for the approximation

$\bar{u}_N$ obtained by Galerkin-averaging:

$$\|u - \bar{u}_N\|_\infty = O(\varepsilon^{\frac{N+1-M}{N+1+2M}}), \quad 0 \leq \varepsilon^{\frac{N+1}{N+1+2M}} t \leq 1. \tag{39}$$

It is clear that if $N \gg M$ the error estimate tends to $O(\varepsilon)$ and the time-scale to $1/\varepsilon$. The result can be interpreted as an upper bound for the speed of energy transfer from the first M modes to higher order modes.

The analysis by van der Aa and Krol [1990]

Consider the coupled system of ordinary differential equations corresponding with problem (38) for arbitrary N; this system is generated by the Hamiltonian H^N. Note that although (38) corresponds with an infinite-dimensional Hamiltonian system, this property does not necessarily carries over to projections.

Important progress has been achieved by van der Aa and Krol [1990] who apply Birkhoff normalisation, which is asymptotically equivalent to averaging, to the Hamiltonian system H^N; the normalized Hamiltonian is indicated by $\bar{H}^N$. Remarkably enough the flow generated by $\bar{H}^N$ for arbitrary N, contains an infinite number of invariant manifolds.

Consider the 'odd' manifold M_1 which is characterized by the fact that only odd-numbered modes are involved in M_1. Inspection of $\bar{H}^N$ reveals that M_1 is an invariant manifold.

In the same way the 'even' manifold M_2 is characterized by the fact that only even-numbered modes are involved; this is again an invariant manifold of $\bar{H}^N$.

In Stroucken and Verhulst [1987] this was noted for $N = 3$ which is rather restricted; moreover it can be extended to manifolds M_m with $m = 2^k q, q$ and odd natural number, k a natural number. It turns out that projections to two modes yield little interaction, so this motivates to look at projections with at least $N = 6$ involving the odd modes $1, 3, 5$ on M_1 and $2, 4, 6$ on M_2.

Van der Aa and Krol [1990] analyse $\bar{H}^6$, in particular the periodic solutions on M_1. For each value of the energy this Hamiltonian produces three normal mode (periodic) solutions which are stable on M_1. Analysing the stability in the full system generated by $\bar{H}^6$ we find again stability.

An open question is if there exist periodic solutions in the flow generated by $\bar{H}^6$ which are not contained in either M_1 or M_2.

What is the relation between the periodic solutions found by averaging and periodic solutions of the original nonlinear wave problem (38)? Van der Aa and Krol [1990] compare with results obtained by Fink et al. [1974] who employ the Poincaré-Lindstedt continuation method to prove existence and to approximate periodic solutions. Related results employing elliptic functions have been derived by Lidskii and Shulman [1988]. It turns out that there is very good agreement but the calculation by the Galerkin-averaging method is technically simpler.

Acknowledgement

The results in this paper reflect for a large part discussions and research carried out at the University of Utrecht and the Technical University of Delft during the last two decades. Comments on the manuscript by Dario Bambusi and Wim van Horssen are gratefully acknowledged.

References

1. Bambusi, D., *On long time stability in Hamiltonian perturbations of nonresonant linear pde's*, preprint Dept. Mathematics,Milano, Italy (1998).

2. Bobenko, A.I. and Kuksin, S., *The nonlinear Klein-Gordon equation on an interval as a perturbed Sine-Gordon equation*, Comment. Meth. Helvetici 70, pp. 63-112 (1995).

3. Bogoliubov, N.N. and Mitropolsky, Y.A., *Methods in the theory of Nonlinear Oscillations*, Hindustan Publ. Corp./ Gordon and Breach (1961).

4. Bourgain, J., *Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations*, GAFA 6, pp. 201-230 (1996).

5. Buitelaar, R.P., *The method of averaging in Banach spaces*, Thesis University of Utrecht (1993).

6. Coppola, V.T. and Rand, R.H., *Macsyma program to implement averaging using elliptic functions*, in Computer Aided Proofs in Analysis, (K.R.Meyer and D.S.Schmidt, eds.), pp. 71-89, Springer-Verlag (1991).

7. Fink, J.P.,Hall,W.S. and Hausrath,A.R., *A convergent two-time method for periodic differential equations*, J. Diff. Eqs. 15, pp. 459-498 (1974).

8. Hale, J., *Periodic solutions of a class of hyperbolic equations containing a small parameter*, Arch. Rat. Mech. Anal. 23, pp. 380-398 (1967).

9. Heijnekamp,J.J., Krol, M.S. and Verhulst, F., *Averaging in Non-linear Transport Problems*, Math. Methods Applied Sciences 18, pp. 437-448 (1995).

10. Jackson, E.A., *Nonlinearity and irreversibility in lattice dynamics*, Rocky Mountain J. Math. 8, pp. 127-196 (1978).

11. Jackson, E.A., *Perspectives of nonlinear dynamics*, 2 vols., Cambridge University Press (1991).

12. Kevorkian, J. and Cole, J.D. *Perturbation Methods in Applied Mathematics*, Springer-Verlag, Applied Mathematical Sciences 34 (1981).

13. Krol, M.S., *On a Galerkin-averaging Method for weakly Non-linear Wave Equations*, Math. Methods Applied Sciences 11, pp. 649-664 (1989).

14. Krol, M.S., *On the averaging method in nearly time-periodic advection-diffusion problems*, SIAM J. Appl. Math. 51, pp. 1622-1637 (1991).

15. Kuksin, S., *Nearly Integrable Infinite-Dimensional Hamiltonian Systems*, Lecture Notes in Mathematics 1556, Springer-Verlag (1991).

16. Lidskii, B.V. and Shulman, E.I., *Periodic solutions of the equation* $u_{tt} - u_{xx} = u^3$, Functional Anal. Appl. 22, pp. 332-333 (1988).

17. Mitropolsky, Y.A., Khoma, G. and Gromyak, M., *Asymptotic Methods for investigating Quasiwave Equations of Hyperbolic Type*, Kluwer Ac. Publ. (1997).

18. Pals,H., *The Galerkin-averaging method for the Klein-Gordon equation in two space dimensions*, Nonlinear Analysis 27, pp. 841-856 (1996).

19. Protter, M.H. and Weinberger, H.F., *Maximum Principles in Differential Equations*, Prentice-Hall (1967).

20. Rafel, G.G., *Applications of a combined Galerkin-averaging method*, Lecture Notes in Mathematics 985 (F.Verhulst ed.), pp. 349-369, Springer-Verlag (1983).

21. Rand, R., *Using computer algebra to handle elliptic functions in the method of averaging*, in Symbolic Computations and their Impact on Dynamics, American Soc.Mech.Eng. PVP-vol.205, pp. 311-326 (1990).

22. Sanders, J.A. and Verhulst, F., *Averaging methods in nonlinear dynamical systems*, Applied Math. Sciences 59, Springer-Verlag (1985).

23. Shtaras, A.L., *The averaging method for weakly nonlinear operator equations*, Math. USSSR Sbornik 62, pp. 223-242 (1989).

24. Stroucken, A.C.J. and Verhulst, F., *The Galerkin-averaging methof for nonlinear, undamped continous systems*, Math. Meth. in the Appl. Sci. 9, pp. 520-549 (1987).

25. Van der Aa, E. and Krol, M.S., *A weakly nonlinear wave equation with many resonances*, Thesis M.S. Krol, University of Utrecht (1990).

26. Van der Burgh, A.H.P., *On the asymptotic validity of perturbation methods for hyperbolic differential equations* in Asymptotic Analysis, from theory to application (F.Verhulst, ed.) pp. 229-240, Lecture Notes in Mathematics 711, Springer-Verlag (1979).

27. Van Horssen, W.T. and van der Burgh, A.H.P., *On initial boundary value problems for weakly nonlinear telegraph equations. Asymptotic theory and application*, SIAM J. Appl. Math. 48, pp. 719-736 (1988).

28. Van Horssen, W.T., *Asymptotics for a class of semilinear hyperbolic equations with an application to a problem with a quadratic nonlinearity*, nonlinear Analysis TMA 19, pp. 501-530 (1992).

29. Verhulst, F., *Nonlinear differential equations and dynamical systems*, 2^d ed. Springer Verlag (1996).