Exponential Averaging under Rapid Quasiperiodic Forcing

Karsten Matthies

Department of Mathematical Sciences,University of Bath, Claverton Down, Bath, BA2 7AY, UK
K.Matthies@maths.bath.ac.uk

Abstract

We derive estimates on the magnitude of the interaction between a wide class of analytic partial differential equations and a high-frequency quasiperiodic oscillator. Assuming high regularity of initial conditions, the equations are transformed to an uncoupled system of an infinite dimensional dynamical system and a linear quasiperiodic flow on a torus; up to coupling terms which are exponentially small in the smallest frequency of the oscillator. The main technique is based on a careful balance of similar results for ordinary differential equations by Simó, Galerkin approximations and high regularity of the initial conditions. Similar finite order estimates assuming less regularity are also provided. Examples include reaction-diffusion and nonlinear Schrödinger equations.

MSC Subjects 37L65 (34C29, 35B40 35Q55)

Short Title Exponential Averaging

1 Introduction

Averaging and homogenisation principles are widely used to derive effective models for systems containing fast scales in their original description. These descriptions are usually ordinary or partial differential equations with explicit dependencies on fast variables; examples range from celestial mechanics to elasticity and nonlinear optics. There are two main approaches for the averaging of multiscale problems. The first one is based on weak convergence methods (see e.g. [Tar79, Tar86, Bor98, JKO94]); but these methods do not provide directly a quantitative error estimate for the approximation by the effective system. The other method is based on formal or rigorous asymptotic expansions in small parameters, e.g. the period of a rapid external excitation or the length scale of a microstructure. Quantitative estimates can be obtained but the class of problems is more restrictive than for weak convergence methods: the estimates are often limited to fast periodic forcing and are still only to a finite order in the small parameter, see e.g. [AKN97, BLP84, BP89, JKO94, LM88].

When iterating finite order asymptotic procedures to obtain an expansion or a series, one cannot expect convergence in general – even for analytic ordinary differential equations. But still beyond a finite asymptotic analysis there are estimates of exponential order giving upper estimates on the divergence effects for analytic ordinary differential equations. The first results of this type were by Nekhoroshev for the perturbation of fully integrable Hamiltonian equations [Nek79] and by Neishtadt for the periodic forcing of analytic differential equations [Nei84]. The approach by Neishtadt is to transform the differential equation by some periodic coordinate change such that the remainder term is small, i.e. an equation like

$$\dot{u} = f(u, t/\varepsilon) \tag{1.1}$$

is transformed to

$$\dot{v} = \bar{f}(v;\varepsilon) + r(v, t/\varepsilon, \varepsilon), \tag{1.2}$$

where the remainder is exponentially small in ε on bounded domains for all t:

$$|r(v, t/\varepsilon, \varepsilon)| \le C\exp(-c/\varepsilon). \tag{1.3}$$

Simó [Sim94] extended this to rapid quasiperiodic forcing

$$\begin{aligned} \dot{u} &= f(u, \theta) \\ \dot{\theta} &= \frac{1}{\varepsilon}\omega, \end{aligned} \tag{1.4}$$

where $\theta \in \mathbb{T} = (\mathbb{R}/\mathbb{Z})^p$ and where the frequency vector fulfils some Diophantine conditions to control small denominators that appear in the analysis: There exist some $C > 0$ and $\tau > p-1$, such that $|(\omega, m)| > C|m|^{-\tau}$ for all $m \in \mathbb{Z}^p \setminus \{0\}$. Diophantine conditions appear in KAM theory and various other contexts in the dynamical systems literature, for references see e.g. [AKN97]. Then again one can construct for analytic right-hand sides a coordinate change, such that the remainder term is exponentially small

$$\begin{aligned} \dot{v} &= \bar{f}(v;\varepsilon) + r(v, \theta, \varepsilon) \\ \dot{\theta} &= \frac{1}{\varepsilon}\omega, \end{aligned} \tag{1.5}$$

with $|r(v, \theta, \varepsilon)| \le C\exp(-c\varepsilon^{-1/(\tau+1)})$.

Extending such results to partial differential equations poses several obstacles. Consider as an example some scalar reaction-diffusion equation

$$u_t = u_{xx} + f(u, t/\varepsilon) \tag{1.6}$$

with periodic boundary condition in $[0, 1]$. Firstly the right-hand side is not continuous or even bounded on any sensible phase space, whereas analyticity is crucial in all the above results. Furthermore the class of allowed transformation is much smaller, when we want to preserve the semilinear structure of the equation. Using the regularisation property of (1.6) it was possible to extend the Neishtadt results to a class of periodically forced parabolic partial differential equations in [Mat01], who obtained a different exponential estimate of order $\mathrm{O}(\exp(-c/\varepsilon^{1/3}))$. In [MS03], we extended the results to a wider class of evolution equations with a rapid one-degree-of-freedom forcing. There the attention is restricted to highly regular initial data, instead of using the regularisation properties of (1.6). Again we obtain different exponential estimates and we also provide lower estimates for an example, showing the different exponential behaviour of ordinary and partial differential equations. Similar methods of combining averaging techniques with exponentially good Galerkin approximations were used in different contexts in [Bam05, Mat05, TW07]. For results in the direction of Nekhoroshev estimates, see e.g. [Bam99, Poe99].

In this paper we give finite order and exponential order averaging results for partial differential equations under temporal quasi-periodic forcing both for highly regular initial data and for general initial data of regularising equations. This extends Simó's result for ordinary differential equations [Sim94]. Crucial assumptions are Diophantine conditions on the frequency vectors. Similar Diophantine conditions were also used in the spatial homogenisation of semilinear parabolic equation in [FV01] and for some low order temporal averaging results in [EZ03]. In this paper we use methods different from [Mat01, MS03] to remove and estimate periodic and quasiperiodic forcings. We also provide estimates in examples where the rapid dependencies are in the coefficients of the unbounded differential operators of the partial differential equations. Furthermore we give an explicit example showing the spatial nonlocality of averaging in partial differential equations.

The paper is organised in the following way. In section 2, we state our main theorems on finite order and exponential averaging. In section 3 we first explain the general procedure underlying both proofs before giving the details of each proof separately. A number of examples including regularising reaction-diffusion equations and nonlinear Schrödinger equations are discussed in section 4.

2 Main results

We state our main results on higher-order and exponential averaging of abstract infinite dimensional evolution equations with fast quasiperiodic forcing. First we introduce the notation used throughout, then we state the averaging theorems in the case (A) of finite regularity, and case (B) in the analytic setting.

We consider abstract evolution equations. Let X be a real Banach space, let A be a closed densely defined, possibly unbounded, operator with domain $D(A)$ generating a strongly continuous semigroup. We denote by $\mathbb{T} = (\mathbb{R}/\mathbb{Z})^p$ the p-dimensional torus, on which the quasiperiodic motion is given by a parallel flow. On the phase space $X \times \mathbb{T}$ we define

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}u &= Au + f(u) + g(u,\theta,\varepsilon) \\ \frac{d}{dt}\theta &= \frac{1}{\varepsilon}\omega \end{aligned} \tag{2.1}$$

with initial conditions $u(0) = u_0; \theta(0) = \theta_0$ and $u \in X$, $\theta \in \mathbb{T}$ and $\varepsilon > 0$ the small real parameter. The regularity assumptions on $f : X \to X$ and $g : X \times \mathbb{T} \times \mathbb{R} \to X$ will be stated below.

A key assumption is that we can approximate the full equation by (not necessarily finite-dimensional) ordinary differential equations.

Hypothesis 2.1 *We assume that there exists a sequence of (Galerkin) projections $(P_N)_{N\in\mathbb{N}}$ which satisfy the following requirements:*

(i) the sequence of projections converges strongly to the identity on X

$$\lim_{N\to\infty} P_N u = u \text{ in } X \text{ for all } u \in X; \tag{2.2}$$

(ii) the projections P_N commute with A on its domain of definition

$$P_N A u = A P_N u \text{ for all } u \in D(A); \tag{2.3}$$

(iii) the operator A is bounded on range P_N

$$|AP_N u|_X \le N|P_N u|_X \text{ for all } u \in X. \tag{2.4}$$

Remark 2.2 *In examples like $A = \Delta$ or $A = i\Delta$ with periodic boundary conditions we choose P_N as a projection to Fourier modes.*

The next assumption is needed to control small denominators and resonances.

Hypothesis 2.3 (Diophantine condition) *Let $\omega \in \mathbb{R}^p$ be such that there exist constants $\gamma > 0$ and $\tau > p - 1$ such that for all $m \in \mathbb{Z}^p \setminus \{0\}$*

$$|(m, \omega)| \ge \gamma |m|^{-\tau} \tag{2.5}$$

where $(\ ,\ )$ is the inner product on $\mathbb{R}^p$ and $|m| = \sum_{j=1}^p |m_j|$ is the norm of m. In the periodic case $p = 1$, (2.5) is equivalent to $\omega \ne 0$.

Remark 2.4 *The Diophantine condition (2.5) is a generic property for $\omega \in \mathbb{R}^p$.*

Hypothesis 2.5 (Zero mean) *For all ε and $u \in X$ the quasi-periodic term has zero mean*

$$\int_{\theta\in\mathbb{T}} g(u, \theta, \varepsilon) \mathrm{d}\theta = 0. \tag{2.6}$$

As we will not require fixed smoothing assumptions on the equations, we consider spaces of regular initial data. We could define them as the domains of $D(|A|^\alpha)$ or Gevrey spaces like $D(\exp(\sigma|A|^\nu))$, but it may require some work to define $|A|$, so we assume the existence of maps Λ_α or $\Gamma_{\sigma,\nu}$, which essentially behave like $|A|^\alpha$ and $\exp(\sigma|A|^\nu)$ respectively:

Hypothesis 2.6 (Higher order approximation) *Assume that there exists a closed, densely defined, boundedly invertible operator Λ_α for $\alpha \geq 1$ with domain of definition*

$$\mathcal{Y}_\alpha := D(\Lambda_\alpha) \subset D(A), \tag{2.7}$$

such that range $\mathrm{Rg}\,(P_N) \subset \mathcal{Y}_\alpha$, $\Lambda_\alpha(\mathrm{Rg}\,(P_N)) = \mathrm{Rg}\,(P_N)$ for all N, and $\Lambda_\alpha A u = A\Lambda_\alpha u$, for all $u \in \mathrm{Rg}\,(P_N)$. We equip the spaces $\mathcal{Y}_\alpha$ with the graph norm

$$|u|_{\mathcal{Y}_\alpha} = |u|_X + |\Lambda_\alpha u|_X. \tag{2.8}$$

We assume that the functions in $\mathcal{Y}_\alpha$ have higher algebraic order approximations by the Galerkin projections P_N

$$|\Lambda_\alpha^{-1}(id - P_N)|_{L(X,X)} \leq C_0 N^{-\alpha}, \tag{2.9}$$

for some N-independent constant $C_0(\alpha)$.

Hypothesis 2.7 (Exponential approximation) *Assume that there exists a closed, densely defined, boundedly invertible operator $\Gamma_{\sigma,\nu}$ with domain of definition*

$$\mathcal{G}_{\sigma,\nu} := D(\Gamma_{\sigma,\nu}) \subset D(A), \tag{2.10}$$

such that $\mathrm{Rg}\,P_N \subset \mathcal{G}_{\sigma,\nu}$, $\Gamma_{\sigma,\nu}(\mathrm{Rg}\,P_N) = \mathrm{Rg}\,P_N$ for all N, and $\Gamma_{\sigma,\nu}Au = A\Gamma_{\sigma,\nu}u$, for all $u \in \mathrm{Rg}\,P_N$. We equip the Gevrey spaces $\mathcal{G}_{\sigma,\nu}$ with the graph norm

$$|u|_{\mathcal{G}_{\sigma,\nu}} = |u|_X + |\Gamma_{\sigma,\nu}u|_X. \tag{2.11}$$

We assume that Gevrey-smooth functions in $G_{\sigma,\nu}$ are exponentially well approximated by the Galerkin projections P_N

$$|\Gamma_{\sigma,\nu}^{-1}(id - P_N)|_{L(X,X)} \leq C_0 \exp(-c_0/N^\nu), \tag{2.12}$$

for N-independent constants $C_0(\sigma,\nu)$ and $c_0(\sigma,\nu)$.

Remark 2.8 *For sectorial operators like the Laplacian $A = \Delta$ on regular domains, the higher regularity spaces $\mathcal{Y}_\alpha$ can be defined as a fractional power space, and the Gevrey space by letting $\Gamma_{\sigma,\nu} = \exp(\sigma(-A)^\nu)$. In some cases it is then possible to prove high regularity after a short transient time for a wide class of initial data; see section 4.1. For the nonlinear Schrödinger equation we can use $\Gamma_{\sigma,\nu} = \exp(\sigma|A|^\nu)$ with $A = \mathrm{i}\Delta$ on $H^s(\mathbb{R},\mathbb{C})$.*

Next we state precisely the regularity assumptions on the nonlinearities with Y denoting a general Banach space and $B_R(Y)$ the ball of radius R around 0 in Y.

Hypothesis 2.9 (Finite regularity) *The nonlinearities $f : Y \to Y, g : Y \times \mathbb{T} \times \mathbb{R} \to Y$ are C^ℓ with bounded C^ℓ-norm on balls $B_R(Y)$ of radius R independent of ε: $\|f\|_{C^\ell(B_R(Y),Y)} \leq C(R,\ell)$ and $\|g\|_{C^\ell(B_R(Y)\times\mathbb{T}\times[0,\varepsilon_0],Y)} \leq C(R,\ell)$ for both $Y = X$ and $Y = \mathcal{Y}_\alpha$.*

To make analyticity precise we extend our function spaces to complex Banach spaces in the standard way. For a general Banach space Y we denote by

$$Y_{\mathbb{C}} = Y \times Y \tag{2.13}$$

the standard complexification $u = u_1 + \mathrm{i}u_2$. If Y is a Hilbert space, the norm is given by $|(u_1,u_2)|_{Y_{\mathbb{C}}}^2 = |u_1|_Y^2 + |u_2|_Y^2$. For general Banach spaces we use the slightly more complicated $|(u_1,u_2)|_{Y_{\mathbb{C}}} = \max_{z_1^2+z_2^2=1} |z_1 u_1 + z_2 u_2|$ to ensure $|\lambda(u_1,u_2)|_{Y_{\mathbb{C}}} = |\lambda||(u_1,u_2)|_{Y_{\mathbb{C}}}$ for $\lambda \in \mathbb{C}$. Complexification of the torus is understood as the quotient of the complexification of the real space $\mathbb{R}^p$ to $\mathbb{C}^p$ by real integer translations $\mathbb{T}_{\mathbb{C}} := (\mathbb{C}/\mathbb{Z})^p$.

Analyticity properties of functions are quantified by the help of extensions to the complexified domain. We therefore introduce a short notation. For any open subset U of a Banach space Y, we define the open complex δ-extension $U + \delta$ for any $\delta > 0$ by

$$U + \delta := \{y \in Y_{\mathbb{C}} | \inf_{u \in U} |u - y|_Y < \delta\}. \tag{2.14}$$

Smoothness properties of the nonlinearities can now be made precise.

Hypothesis 2.10 (Analyticity of nonlinearities) *There is a Gevrey class $Y = \mathcal{G}_{\sigma,\nu}$ with $\sigma, \nu > 0$ and a constant $\delta > 0$ for the size of the complex extension such that the following properties of the nonlinearities hold.*

The nonlinearities $f : (Y + \delta) \to Y_{\mathbb{C}}; g : (Y + \delta) \times (\mathbb{T} + \delta) \times \mathbb{R} \to Y_{\mathbb{C}}$ are analytic and bounded on bounded subsets when considered on Gevrey spaces, extended in the complex direction. In addition, all of the above statements are assumed to hold when the space of Gevrey regularity $Y = \mathcal{G}_{\sigma,p}$ is replaced by $Y = X$.

The hypothesis above are required to derive estimates about transformations of equation (2.1). To deduce results about the semigroup of solutions of equation (2.1) we will assume well-posedness.

Hypothesis 2.11 (Well-posedness on $\mathcal{Y}_\alpha$) *The operator A generates a strongly continuous semigroup both on X and $\mathcal{Y}_\alpha$.*

Hypothesis 2.12 (Well-posedness on $\mathcal{G}_{\sigma,\nu}$) *The operator A generates a strongly continuous semigroup both on X and $\mathcal{G}_{\sigma,\nu}$.*

Theorem A (Finite regularity) *Assume Hypothesis 2.1 on the existence of bounded approximations, Hypothesis 2.3 on the Diophantine conditions on the frequency vector ω with constants γ and τ, Hypothesis 2.5 on the zero mean, Hypothesis 2.6 on the higher order approximation and Hypothesis 2.9 on the regularity for some fixed space $\mathcal{Y}_\alpha$ and some fixed differentiability level ℓ.*

Then, for any ball of radius R in $\mathcal{Y}_\alpha$ there exists an $\varepsilon_0 > 0$, such that for $0 < \varepsilon < \varepsilon_0$ the following holds.

There exists a C^1 transformation $u = v + \varepsilon w(v, \theta, \varepsilon)$, which is near identity both in X and $\mathcal{Y}_\alpha$, such that the transformed equation has the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} v &= Av + f(v) + \bar{g}(v, \varepsilon) + r(v, \theta, \varepsilon) \\ \frac{d}{dt}\theta &= \frac{1}{\varepsilon}\omega \end{aligned} \tag{2.15}$$

with initial conditions $v(0) = u_0; \theta(0) = \theta_0$ and with $\bar{g}$ and r both bounded on balls in X; furthermore

$$|\bar{g}(v, \varepsilon)|_X \leq C_{\alpha,l}(|v|_X)\varepsilon^{\alpha/(q(\ell)+\alpha)} \tag{2.16}$$

$$|r(v, \theta, \varepsilon)|_X \leq C_\alpha(|v|_{\mathcal{Y}_\alpha})\varepsilon^{q(\ell)\alpha/(q(\ell)+\alpha)}, \tag{2.17}$$

with the exponent given by $q(\ell) = \left\lfloor \frac{\ell}{\lceil \tau+p+1 \rceil} \right\rfloor$ (where $\lfloor\,\rfloor$ is the integer part and $\lceil \tau + p + 1 \rceil$ is the smallest integer larger than $\tau + p + 1$).

Assume also Hypothesis 2.11 on well-posedness, then the solutions of the truncated equation

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \bar{v} &= A\bar{v} + f(\bar{v}) + \bar{g}(\bar{v}, \varepsilon) \\ \frac{d}{dt}\theta &= \frac{1}{\varepsilon}\omega \end{aligned} \tag{2.18}$$

with initial conditions $\bar{v}(0) = u_0; \theta(0) = \theta_0$, remain close to the solutions of the transformed equation (2.15): Suppose one of the solutions $v(t)$ or $\bar{v}(t)$ of (2.15) resp. (2.18) remains inside the ball $B_R(\mathcal{Y}_\alpha)$ for $0 \leq t \leq T$, then there exists a constant $C(T, R)$ such that

$$|v(t) - \bar{v}(t)|_X \leq C(T, R)\varepsilon^{q(\ell)\alpha/(q(\ell)+\alpha)}. \tag{2.19}$$

Theorem B (Analyticity) *Assume Hypothesis 2.1 on the existence of bounded approximations, Hypothesis 2.3 on the Diophantine conditions on the frequency vector ω with constants γ and τ, Hypothesis 2.5 on the zero mean, Hypothesis 2.7 on the exponential approximation and Hypothesis 2.10 on the analytic regularity for some fixed space $\mathcal{G}_{\sigma,\nu}$.*

Then, for any ball of radius R in $\mathcal{G}_{\sigma,\nu}$ there exists an $\varepsilon_0 > 0$, such that for $0 < \varepsilon < \varepsilon_0$ the following holds.

There exists a C^1 near identity transformation $u = v + \varepsilon w(v, \theta, \varepsilon)$ of both $\mathcal{G}_{\sigma,\nu}$ and X such that the transformed equation has the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} v &= Av + f(v) + \bar{g}(v, \varepsilon) + r(v, \theta, \varepsilon) \\ \frac{d}{dt}\theta &= \frac{1}{\varepsilon}\omega \end{aligned} \tag{2.20}$$

with initial conditions $v(0) = u_0; \theta(0) = \theta_0$ *and with* $\bar{g}$ *and* r *both bounded on balls in* X*; furthermore the remainder term is exponentially small on balls of the Gevrey space,*

$$|\bar{g}(v, \varepsilon)|_X \leq C(|v|_X)\varepsilon^{(\tau+1)/(\tau+1+1/\nu)} \tag{2.21}$$

$$|r(v, \theta, \varepsilon)|_X \leq C(|v|_{\mathcal{G}_{\sigma,\nu}}) \exp(-c/\varepsilon^{1/(\tau+1+1/\nu)}). \tag{2.22}$$

Assume also Hypothesis 2.12 on well-posedness in X *and in the Gevrey class* $\mathcal{G}_{\sigma,\nu}$*, then the solutions of the truncated equation*

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\bar{v} &= A\bar{v} + f(\bar{v}) + \bar{g}(\bar{v}, \varepsilon) \\ \frac{d}{dt}\theta &= \frac{1}{\varepsilon}\omega \end{aligned} \tag{2.23}$$

with initial condition $\bar{v}(0) = u_0; \theta(0) = \theta_0$ *remain close to the solutions of the transformed equation (2.20): Suppose one of the solutions* $v(t)$ *or* $\bar{v}(t)$ *of (2.20), resp. (2.23), remains inside the ball* $B_R(\mathcal{G}_{\sigma,\nu})$ *for* $0 \leq t \leq T$*, then there exists a constant* $C(T, R)$ *such that*

$$|v(t) - \bar{v}(t)|_X \leq C(T, R) \exp(-c/\varepsilon^{1/(\tau+1+1/\nu)}). \tag{2.24}$$

Remark 2.13 *(i) In partial differential equations with local nonlinearities of the form* $f(u)(x) = f(u(x))$*, both the corrector* $\bar{g}$ *and the remainder are in general nonlocal operators, mapping the phase spaces* X*,* $\mathcal{Y}_\alpha$ *and* $\mathcal{G}_{\sigma,\nu}$ *into themselves. We show that the use of nonlocal transformations is already necessary in an example of low order averaging in section 4.3.*

(ii) For periodic forcing, $p = 1$*, there is no small denominator problem and it is possible to choose* $\tau = 0$ *and still obtain the same estimate. The forcing* g *then only needs to be continuous in* θ*; see [MS03].*

(iii) The transformation $id + \varepsilon w$ *is defined iteratively and for fixed* $\varepsilon > 0$*, it transforms only a Galerkin approximation space of finite index. Each transformation step in this space corresponds to the formal removing of the* θ*-dependent term that is of the lowest order in* ε*.*

(iv) When considering an equation with bounded A*, this gives an ordinary differential equation on a possibly infinite dimensional space. Then it is possible with a small variation of the proof to obtain the results with formally* $\alpha = \infty$ *(resp.* $\nu = \infty$*). In the analytic case we recover Simó's result [Sim94] on exponential averaging of quasiperiodic ODE with a remainder estimate of order* $\exp(-c/\varepsilon^{1/(\tau+1)})$*. In the finite regularity case we obtain* $|r| \leq C\varepsilon^{q(\ell)}$ *and* $|\bar{g}| \leq C\varepsilon$*.*

(v) Lower exponential estimates are an intriguing problem. Some lower estimates in the periodic case $p = 1$ *can be found in [MS03].*

(vi) The remainder term r *might not have zero mean; for another variant see remark 3.3.*

3 Approximation and Averaging

The general procedure to prove both theorems is similar. Using Hypothesis 2.1 we will derive approximations by ordinary differential equations on $P_N X$. In these equations we remove the quasiperiodic part by a number of near-identity transformations of the approximation space. Formally we increase the order in ε of the remaining quasiperiodic terms by one in each transformation, but in general this process diverges even in the analytic case. Rigorous estimates depend on the regularity of the right hand side. The magnitude of the contribution by the complement of $P_N X$ is estimated using the higher regularity that we assume on the initial conditions in Hypothesis 2.6 and 2.7. The number of transformations and the approximation space $P_N X$ will be chosen depending on the regularity class (Theorem A; resp. B) and on ε to simultaneously minimise the two contributions of the quasiperiodic remainder term:

- the remainder of the averaging procedure of the ordinary differential equation
- the remainder due to the Galerkin approximation.

So we will first describe the formal averaging procedure in the approximation space and its extension to the full space which are common in both proofs. The estimates and the choices of index N and the number of averaging steps j^* are then given separately below.

So consider equation (2.1) for $(u_N, \theta) \in P_N X \times \mathbb{T}$

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} u_N &= A u_N + P_N f(u_N) + P_N g(u_N, \theta, \varepsilon) \\ \frac{d}{dt}\theta &= \frac{1}{\varepsilon}\omega \end{aligned} \tag{3.1}$$

with projected initial conditions $u_N(0) = P_N u_0; \theta(0) = \theta_0$. We perform repeated averaging steps, indexed by j. For $j = 0$ we let $\bar{g}_0(u_N, \varepsilon) = \int_{\mathbb{T}} g(u_N, \theta, \varepsilon) \mathrm{d}\theta = 0$ and $\tilde{g} = P_N g$ using Hypothesis 2.5.

Suppose that we already performed j transformation steps and we have an equation of the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} u_N &= A u_N + P_N f(u_N) + \bar{g}_j(u_N, \varepsilon) + \tilde{g}_j(u_N, \theta, \varepsilon) \\ \frac{d}{dt}\theta &= \frac{1}{\varepsilon}\omega, \end{aligned} \tag{3.2}$$

where $\int_{\mathbb{T}} \tilde{g}_j(u_N, \theta, \varepsilon) \mathrm{d}\theta = 0$. Then the next transformation is given in terms of the temporal Fourier expansion of the remainder term. Letting

$$\tilde{g}_j(u_N, \theta, \varepsilon) = \sum_{m \in \mathbb{Z}^p} \tilde{g}_{jm}(u_N, \varepsilon) \exp(\mathrm{i} 2\pi (m, \theta)) \tag{3.3}$$

with $\tilde{g}_{j0}(u_N, \varepsilon) = 0$ then we transform $u_N = v_N + \varepsilon w_j(v_N, \theta, \varepsilon)$ with

$$w_j(v_N, \theta, \varepsilon) = \sum_{m \in \mathbb{Z}^p, m \neq 0} \frac{\tilde{g}_{jm}(v_N, \varepsilon)}{2\pi \mathrm{i}(m, \omega)} \exp(\mathrm{i} 2\pi (m, \theta)) \tag{3.4}$$

For notational briefness, we usually suppress arguments, $w_j = w_j(v_N, \theta, \varepsilon)$. The transformed equation is given by

$$\begin{aligned}
\left(id + \varepsilon \frac{\partial}{\partial v_N} w_j\right) \dot{v}_N + \varepsilon \frac{\partial}{\partial \theta} w_j \dot{\theta} &= \frac{\mathrm{d}}{\mathrm{d}t} u_N \\
&= A v_N + P_N f(v_N) + \bar{g}_j(v_N, \varepsilon) + \tilde{g}_j(v_N, \theta, \varepsilon) \\
&\quad + \varepsilon A w_j + P_N f(v_N + \varepsilon w_j) - P_N f(v_N) \\
&\quad + \bar{g}_j(v_N + \varepsilon w_j, \varepsilon) - \bar{g}_j(v_N, \varepsilon) \\
&\quad + \tilde{g}_j(v_N + \varepsilon w_j, \theta, \varepsilon) - \tilde{g}_j(v_N, \theta, \varepsilon). \qquad (3.5)
\end{aligned}$$

Hence on a formal level we remove the largest quasiperiodic term $\tilde{g}_j(v_N, \theta, \varepsilon)$. The transformed equation is

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} v_N &= A v_N + P_N f(v_N) + \bar{g}_j(v_N, \varepsilon) + r_j(v_N, \theta, \varepsilon) \\
\frac{\mathrm{d}}{\mathrm{d}t} \theta &= \frac{1}{\varepsilon} \omega, \qquad (3.6)
\end{aligned}$$

where

$$\begin{aligned}
r_j = (id + \varepsilon \frac{\partial}{\partial v_N} w_j)^{-1} \Big\{ &-\varepsilon \frac{\partial}{\partial v_N} w_j (A v_N + P f(v_N) + \bar{g}_j(v_N, \varepsilon)) \\
&+ \varepsilon A w_j + P_N f(v_N + \varepsilon w_j) - P_N f(v_N) \\
&+ \bar{g}_j(v_N + \varepsilon w_j, \varepsilon) - \bar{g}_j(v_N, \varepsilon) \\
&+ \tilde{g}_j(v_N + \varepsilon w_j, \theta, \varepsilon) - \tilde{g}_j(v_N, \theta, \varepsilon) \Big\}. \qquad (3.7)
\end{aligned}$$

To derive this we used the identity $(id + \varepsilon \frac{\partial}{\partial v_N} w_j)^{-1} = id - (id + \varepsilon \frac{\partial}{\partial v_N} w_j)^{-1} \varepsilon \frac{\partial}{\partial v_N} w_j$. Then (3.6) can be written in the form

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} v_N &= A v_N + P_N f(v_N) + \bar{g}_{j+1}(v_N, \varepsilon) + \tilde{g}_{j+1}(v_N, \theta, \varepsilon) \\
\frac{\mathrm{d}}{\mathrm{d}t} \theta &= \frac{1}{\varepsilon} \omega \qquad (3.8)
\end{aligned}$$

with

$$\bar{g}_{j+1}(v_N) = \bar{g}_j(v_N) + \int_{\mathbb{T}} r_j(v_N, \theta, \varepsilon) \mathrm{d}\theta \qquad (3.9)$$

$$\tilde{g}_{j+1}(v_N) = r_j(v_N, \theta, \varepsilon) - \int_{\mathbb{T}} r_j(v_N, \theta, \varepsilon) \mathrm{d}\theta, \qquad (3.10)$$

such that $\int_{\mathbb{T}} \tilde{g}_{j+1}(v_N, \theta, \varepsilon) d\theta = 0$ for all v_N and ε.

As an asymptotic expansion for all $j \in \mathbb{N}$ will not converge in general, we will stop after a certain number of steps. So after performing j^* transformations

$$u_N = (id + \varepsilon w_0) \circ (id + \varepsilon w_1) \circ \ldots \circ (id + \varepsilon w_{j^*-1}) v_N = (id + \varepsilon w) v_N, \qquad (3.11)$$

we obtain a transformation of equation (3.1) to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}v_N &= Av_N + P_N f(v_N) + \bar{g}_{j^*}(v_N,\varepsilon) + \tilde{g}_{j^*}(v_N,\theta,\varepsilon)\\ \frac{\mathrm{d}}{\mathrm{d}t}\theta &= \frac{1}{\varepsilon}\omega\end{aligned}\tag{3.12}$$

still with projected initial data $v_N(0) = P_N u_0$; $\theta(0) = \theta_0$. We extend this to the full space by

$$u = v + \varepsilon w(P_N v),\tag{3.13}$$

and noting that the range of w is then still in $P_N X$. Then we obtain

$$v = (P_N. + \varepsilon w(P_N.,\theta,\varepsilon))^{-1}(P_N u) + (id - P_N)u,\tag{3.14}$$

such that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}v &= (id - P_N)\{Av + f(v + \varepsilon w(P_N v)) + g(v + \varepsilon w(P_N v),\theta,\varepsilon)\}\\ &\quad +(id + \varepsilon\frac{\partial}{\partial v_N}w(P_N v))^{-1}P_N\{A(v + \varepsilon w(P_N v)) + f(v + \varepsilon w(P_N v)) + g(v + \varepsilon w(P_N v),\theta,\varepsilon)\}\\ &= Av + f(v) + \bar{g}_{j^*}(P_N v,\varepsilon) + \tilde{g}_{j^*}(P_N v,\theta,\varepsilon) + r^*(v,\theta,\varepsilon) \qquad (3.15)\\ \frac{\mathrm{d}}{\mathrm{d}t}\theta &= \frac{1}{\varepsilon}\omega\end{aligned}$$

with initial data $v(0) = u_0; \theta(0) = \theta_0$, and the additional remainder term collects all terms due to the error of the Galerkin approximation

$$\begin{aligned}r^*(v,\theta,\varepsilon) &= (id - P_N)\Big\{f(v + \varepsilon w(P_N v)) - f(v) + g(v + \varepsilon w(P_N v),\theta,\varepsilon)\Big\}\\ &\quad +(id + \varepsilon\frac{\partial}{\partial v_N}w(P_N v))^{-1}P_N\Big\{f(v + \varepsilon w(P_N v)) - f(P_N v + \varepsilon w(P_N v))\\ &\qquad +g(v + \varepsilon w(P_N v),\theta,\varepsilon) - g(P_N v + \varepsilon w(P_N v),\theta,\varepsilon)\Big\}.\end{aligned}\tag{3.16}$$

Its X-norm can be estimated by expressions of the form $C|\bar{v} - P_N\bar{v}|_X$ for some bounded $\bar{v} \in \mathcal{G}_{\sigma,\nu}$ or $\mathcal{Y}_\alpha$ using the boundedness and the differentiability of f and g on balls. The expression $C|\bar{v} - P_N\bar{v}|_X$ can be estimated using the higher regularity in Hypothesis 2.6 and 2.7.

Estimating the right-hand side of (3.15), we can then prove the estimates on $v(t)$ by a Gronwall argument. The difference $y(t) = v(t) - \bar{v}(t)$ satisfies the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}y(t) = A(t)y + h(t)\tag{3.17}$$

with $A(t) = A + \int_0^1(\partial_v(f + \bar{g})(s\bar{v} + (1-s)v)\mathrm{d}s$ and $h(t) = r(v(t)) - r(v(\bar{t}))$. Then estimate (2.19) follows by the Gronwall lemma, since f, $\bar{g}$ and r possess bounded derivatives on bounded sets of $\mathcal{G}_{\sigma,\nu}$ or $\mathcal{Y}_\alpha$. We apply the Gronwall lemma in X, because the estimates on r hold in the X-topology. They do not hold in general in $\mathcal{G}_{\sigma,\nu}$ or $\mathcal{Y}_\alpha$.

The general outline so far leads to proofs of theorems A and B by giving rigorous estimates on the remainder terms $\tilde{g}_{j^*}(v,\theta,\varepsilon)$ and $r^*(v,\theta,\varepsilon)$.

3.1 Proof of Theorem A

Before we can estimate the remainder terms, we have to ensure that the formal transformations are well-defined, i.e. that we can control the small denominator in (3.4) for $1 \le j \le j^*$. We will choose $j^* \le \frac{\ell}{\lceil \tau+p \rceil}$ independent of ε and N, where $\lceil y \rceil$ denotes the smallest integer greater than y. For integer tuples $m \in \mathbb{Z}^p$, we denote $|m| := \sum_{k=1}^p |m_k|$. Consider the first averaging of (3.1) then

$$w_0(v,\theta,\varepsilon) = \sum_{m\in\mathbb{Z}^p} \frac{g_m(v,\varepsilon)}{2\pi \mathrm{i}(m,\omega)} \exp(2\pi\mathrm{i}(m,\theta)) \tag{3.18}$$

with $g_m(v,\varepsilon) = \int_{\mathbb{T}} g(v,\theta,\varepsilon)\exp(2\pi\mathrm{i}(m,\theta))\mathrm{d}\theta$. By integration by parts we obtain

$$g_m(v,\varepsilon) = -\int_{\mathbb{T}} (2\pi m_1)^{-1}\partial_{\theta_1} g(v,\theta,\varepsilon)\exp(2\pi\mathrm{i}(m,\theta))\mathrm{d}\theta. \tag{3.19}$$

Choose k such that $\max\{|m_1|,\dots,|m_p|\} \le |m_k|$. Repeating the integration by parts ℓ times with the component θ_k of $\theta = (\theta_1,\dots,\theta_p)$ we obtain

$$|g_m(v,\varepsilon)| \le \frac{\|g\|_{C^\ell}}{(2\pi)^\ell}|m|_\infty^{-\ell} \le \frac{\|g\|_{C^\ell}}{(2\pi)^\ell} p^\ell |m|^{-\ell} \tag{3.20}$$

Similarly we can estimate the m-th Fourier mode of any ν^{th} order partial derivative, denoted as $\left(\left(\frac{\partial}{\partial\theta}\right)^\nu g\right)_m$ with respect to the components of θ:

$$\left|\left(\left(\frac{\partial}{\partial\theta}\right)^\nu g\right)_m (v)\right| \le \frac{\|g\|_{C^\ell}}{(2\pi)^{\ell-\nu}}|m|_\infty^{-\ell+\nu} \le \frac{p^{\ell-\nu}}{(2\pi)^{\ell-\nu}}\|g\|_{C^\ell}|m|^{-\ell+\nu} \tag{3.21}$$

Using the Diophantine condition of Hypothesis 2.3 we see for any ν^{th} derivative of w_0 with respect to θ:

$$\left|\left(\frac{\partial}{\partial\theta}\right)^\nu w_0(v,\theta,\varepsilon)\right| \le \frac{p^{\ell-\nu}}{\gamma(2\pi)^{\ell-\nu+1}} \sum_{m\in\mathbb{Z}^p} \|g\|_{C^\ell}|m|^{-\ell+\nu+\tau} \le M\|g\|_{C^\ell} < \infty \tag{3.22}$$

for $-\ell+\tau+\nu < -p$,with $M \ge 1$ only depending on τ. Thus w_0 is $(\ell - \lceil \tau+p \rceil)$ times differentiable with respect to θ, if $\tau \notin \mathbb{N}$ and it is also $\ell - (\tau+p+1) = \ell - \lceil \tau+p \rceil$ times differentiable with respect to θ, if $\tau \in \mathbb{N}$. Then $\tilde g_1$ inherits the differentiability from w_0 after a further differentiation, see (3.7)-(3.10). Hence, both for $\tau \notin \mathbb{N}$ and $\tau \in \mathbb{N}$, $\tilde g_1$ is $\ell - \lceil \tau+p+1 \rceil$ times differentiable and inductively w_j is $\ell - j\lceil \tau+p+1 \rceil$-times differentiable.

Assuming boundedness of g in the C^ℓ norms as in Hypothesis 2.9 by a constant $C(R)$ on both spaces $Y = X$ and $Y = \mathcal{Y}_\alpha$, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ the first transformation is well-defined. To estimate remainder term and its derivatives up to differentiability order $\nu = \ell - \lceil \tau+p+1 \rceil$, we first use (3.22) and a direct Taylor estimate of (3.7) on $P_N Y$ for $j = 1$

$$\begin{aligned}
\|r_1\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} &\le 2\varepsilon\|\frac{\partial}{\partial v_N} w_1 A v_N + A w_1\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} \\
&\quad +\|P_N f(v_N+\varepsilon w_1) - P_N f(v_N) + \varepsilon\frac{\partial}{\partial v_N} w_1 P_N f(v_N)\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} \\
&\quad +\|\tilde g_1(v_N+\varepsilon w_1,\theta,\varepsilon) - \tilde g_1(v_N,\theta,\varepsilon)\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} \\
&\le 4\varepsilon NMC(R) + 4\varepsilon MC(R)\|P_N f\|_{BC^{\nu+1}(B_R(Y),Y)} + 2\varepsilon MC(R)^2
\end{aligned} \tag{3.23}$$

This yields estimates of the remainder terms in (3.9,3.10)

$$
\begin{aligned}
\|\tilde{g}_1\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} &\leq \|r_1\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} \\
&\leq 4\varepsilon NMC(R) + \varepsilon(4\|P_N f\|_{BC^{\nu+1}(B_R(Y),Y)} + 2C(R))MC(R) \quad (3.24) \\
\|\bar{g}_1\|_{BC^\nu(B_R(Y),Y)} &\leq 2\|r_1\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} \\
&\leq 8\varepsilon NMC(R) + \varepsilon(8\|P_N f\|_{BC^{\nu+1}(B_R(Y),Y)} + 4C(R))MC(R) \quad (3.25)
\end{aligned}
$$

We then proceed inductively for N large and for εN small compared to $C(R)$ and for a fixed C independent of ε and N. Then we obtain after j steps with $\nu = \ell - j\lceil \tau + p + 1\rceil$

$$
\begin{aligned}
\|\tilde{g}_j\|_{BC^\nu(B_R(Y)\times\mathbb{T},Y)} &\leq \varepsilon^j(4NM + C)^j C(R) \\
\|\bar{g}_j\|_{BC^\nu(B_R(Y),Y)} &\leq 2\|\bar{g}_1\|_{BC^{\ell-\lceil\tau+p+1\rceil}(B_R(Y),Y)} \leq 17\varepsilon NMC(R). \quad (3.26)
\end{aligned}
$$

So finally we arrive for the remainder term in the Galerkin approximation space with $q(\ell) = \left\lfloor \frac{\ell}{\lceil \tau+p+1\rceil} \right\rfloor$

$$\|\tilde{g}_{j^*}\|_{BC^1(B_R(Y)\times\mathbb{T},Y)} \leq \varepsilon^{q(\ell)}(8MN + C)^{q(\ell)}C(R) \quad (3.27)$$

$$\|\bar{g}_{j^*}\|_{BC^1(B_R(Y),Y)} \leq 17\varepsilon NMC(R). \quad (3.28)$$

As explained above the remainder term due to the Galerkin approximation can be controlled in the X-norm using hypothesis 2.6 by a term of order $N(\varepsilon)^{-\alpha}$.

Choosing a coupling of ε and N, such that we simultaneously minimise the remainder terms $\tilde{g}_{j^*}$ and $N(\varepsilon)^{-\alpha}$, gives

$$N(\varepsilon) = \varepsilon^{-q(\ell)/(q(\ell)+\alpha)}. \quad (3.29)$$

Then both remainder terms are of order $\varepsilon^{\alpha q(\ell)/(q(\ell)+\alpha)}$ on bounded sets in $\mathcal{Y}_\alpha$ with constants uniform in ε and $N(\varepsilon)$ but depending on the size of the bounded set. By (3.26), the estimate on $\bar{g}_{j^*}$ holds already for $u \in X$. This proves the estimates for the terms in the transformed equation. With the general Gronwall argument in (3.17), we also obtain the estimates on the solutions. □

3.2 Proof of Theorem B

Following the argument above, any fixed number of transformation steps is well-defined in the analytic setting. The number of steps that we use will depend on ε. To estimate the remainder term we use the complex extension as in Hypothesis 2.10. In each transformation step we will estimate $\bar{g}_j$ and $\tilde{g}_j$ on $D_j = (B_R(Y) + \delta_j) \times (\mathbb{T} + \delta_j)$. The transformations are then constructed such that $(id + \varepsilon w_j)^{-1} : D_{j+1} \to D_j$. We set $\delta_j = \delta - j\eta(\varepsilon)$, where the step size $\eta(\varepsilon)$ in the reduction of the complex extension is chosen later in (3.52). We also introduce an intermediate domain $D_{j,\frac{1}{2}} = (B_R(Y) + \delta_{j+\frac{1}{2}}) \times (\mathbb{T} + \delta_{j+1})$. For notational convenience we denote for a function h and any complex domain D for some given ε

$$|h|_D = \sup_{(u,\theta)\in D} |h(u,\theta,\varepsilon)|_Y \tag{3.30}$$

and for derivatives $\frac{\partial h}{\partial u}(u_0,\theta,\varepsilon) \in L(Y,Y)$ we consider their operator norms

$$\left\| \frac{\partial h}{\partial u} \right\|_D = \sup_{(u_0,\theta)\in D} \sup_{v\in Y, |v|_Y=1} \left| \frac{\partial h}{\partial u}(u_0,\theta,\varepsilon) v \right|_Y . \tag{3.31}$$

We assume inductively

$$|\bar{g}_j|_{D_j} \le \bar{b}_j \text{ and } |\tilde{g}_j|_{D_j} \le \tilde{b}_j. \tag{3.32}$$

The bound on the nonlinearity f is denoted as $|f|_{D_0} =: b$. Then we estimate w_j given in (3.4) in the intermediate domain $D_{j,\frac{1}{2}}$. Fixing a direction in the complex extension of the torus $\mathbb{T}$ with $\theta = (\theta_1, \ldots, \theta_p) \in \mathbb{T}$, analyticity gives estimates on the Fourier expansion of $\tilde{g}_j$:

$$\begin{aligned} & \int_0^1 \tilde{g}_j(v,\theta_1,\ldots,\theta_p) \exp(-2\pi i(m,\theta)) d\theta_1 \qquad (3.33) \\ = & \int_0^{i\pm\delta_j} \tilde{g}_j(v,\theta_1,\ldots,\theta_p) \exp(-2\pi i(m,\theta)) d\theta_1 + \int_{\pm i\delta_j}^{1+i\pm\delta_j} \tilde{g}_j(v,\theta_1,\ldots,\theta_p) \exp(-2\pi i(m,\theta)) d\theta_1 \\ & + \int_{1+i\pm\delta_j}^1 \tilde{g}_j(v,\theta_1,\ldots,\theta_p) \exp(-2\pi i(m,\theta)) d\theta_1. \end{aligned}$$

The first and last integral cancel due to periodicity whereas the middle integral yields for the right choice of sign an exponential factor $\exp(-2\pi|m_1|\delta_j)$. By repeating this in the other directions on the torus $\mathbb{T}$ we obtain for the Fourier coefficients of index m:

$$|\tilde{g}_{jm}|_{D_{j,\frac{1}{2}}} \le \frac{\tilde{b}_j}{2\pi} \mathrm{e}^{-2\pi\delta_j|m|}. \tag{3.34}$$

Then we have on the intermediate domain $D_{j,\frac{1}{2}}$ for the next transformation

$$\begin{aligned} |w_j|_{D_{j,\frac{1}{2}}} & \le \frac{\tilde{b}_j}{4\pi^2} \sum_{m\in\mathbb{Z}^p, m\neq 0} \frac{\mathrm{e}^{-2\pi\delta_j|m|}}{|(m,\omega)|} \mathrm{e}^{2\pi\delta_{j+1}|m|} \\ & = \frac{\tilde{b}_j}{4\pi^2} \sum_{m\in\mathbb{Z}^p, m\neq 0} \frac{\mathrm{e}^{-2\pi\eta(\varepsilon)|m|}}{|(m,\omega)|}. \end{aligned} \tag{3.35}$$

Using the Diophantine condition hypothesis 2.3 one can prove the next lemma.

Lemma 3.1 ([Sim94, Lemma 1]) *Let $\omega \in \mathbb{R}^p$ satisfy Hypothesis 2.3 with $\tau > p-1$. Then*

$$s_k = \sum_{m\in\mathbb{Z}^p\setminus\{0\}, |m|\le k} |(m,\omega)|^{-1} \le G k^\tau, \tag{3.36}$$

where G is a positive constant depending on ω, but independent of k.

The lemma was stated in [Sim94] and a proof was given in [Sim99]. For the convenience of the reader, a proof is also given in the appendix. Then we use the Γ-function

$$\sum_{k\in\mathbb{N}} \mathrm{e}^{-2\pi\eta(\varepsilon)k} k^{\tau} \le 2\int_0^{\infty} \mathrm{e}^{-2\pi\eta(\varepsilon)k} k^{\tau}\mathrm{d}k = \frac{2\Gamma(\tau+1)}{(2\pi\eta(\varepsilon))^{\tau+1}}, \tag{3.37}$$

to obtain

$$\begin{aligned} |w_j|_{D_{j,\frac{1}{2}}} &\le \frac{\tilde{b}_j}{2\pi}\sum_{k\in\mathbb{N}} \mathrm{e}^{-2\pi\eta(\varepsilon)k}(s_k - s_{k-1}) = \frac{\tilde{b}_j}{2\pi}\sum_{k\in\mathbb{N}} s_k(\mathrm{e}^{-2\pi\eta(\varepsilon)k} - \mathrm{e}^{-2\pi\eta(\varepsilon)(k+1)}) \\ &\le \frac{\tilde{b}_j}{2\pi}(1-\mathrm{e}^{-2\pi\eta(\varepsilon)})G\sum_{k\in\mathbb{N}} \mathrm{e}^{-2\pi\eta(\varepsilon)k}k^{\tau} \le \tilde{b}_j\frac{K}{\eta(\varepsilon)^{\tau}}, \end{aligned} \tag{3.38}$$

where $K \ge 1$ depends on τ, but it is independent of $\tilde{b}_j$ and $\eta(\varepsilon)$. The transformation $id + \varepsilon w_j$ is well-defined, when we can ensure that if $v \in D_{j,\frac{1}{2}}$, then $u \in D_j$. For this, it is enough to have $\varepsilon|w_j|_{D_{j,\frac{1}{2}}} < \eta/2$, i.e.

$$\varepsilon\tilde{b}_j\frac{K}{\eta(\varepsilon)^{\tau}} < \frac{\eta(\varepsilon)}{2} \tag{3.39}$$

by a suitable choice of $\eta(\varepsilon)$ below.

Before giving the crucial inductive estimates on $|\tilde{g}_{j+1}|_{D_{j+1}}$ and $|\bar{g}_{j+1}|_{D_{j+1}}$ we collect some preliminary estimates. A key ingredient is the Cauchy estimate, which gives estimates of derivatives of a function by uniform estimates on the function on larger complex domains.

Lemma 3.2 (Cauchy estimate) *Let $f : \Omega \subset Y \to Y$ be analytic on a complex Banach space Y. Then $\|\frac{\partial f}{\partial u}\|_{\Omega-\eta} \le \frac{|f|_{\Omega}}{\eta}$ in the notation of (3.30) and (3.31).*

A similar lemma was also used in [Mat01]. For the convenience of the reader, a new and clearer proof is given in the appendix.

Hence we obtain

$$\left\|\frac{\partial w_j}{\partial v_N}\right\|_{D_{j+1}} \le \frac{2|w_j|_{D_{j,\frac{1}{2}}}}{\eta} \le \tilde{b}_j\frac{2K}{\eta(\varepsilon)^{\tau+1}} \tag{3.40}$$

Then $\left\|\varepsilon\frac{\partial w_j}{\partial v_N}\right\|_{D_{j+1}} < 1$ if (3.39) holds and

$$\left\|\left(id + \varepsilon\frac{\partial w_j}{\partial v_N}\right)^{-1}\right\|_{D_{j+1}} < 1 - \tilde{b}_j\frac{2K\varepsilon}{\eta(\varepsilon)^{\tau+1}}. \tag{3.41}$$

Also using the Cauchy estimate and the mean value theorem we obtain

$$|\tilde{g}_j(v_N + \varepsilon w_j) - \tilde{g}_j(v_N)|_{D_{j+1}} \le \varepsilon\left\|\frac{\partial}{\partial v_N}\tilde{g}_j\right\|_{D_{j+1}}|w_j|_{D_{j,\frac{1}{2}}} \le \varepsilon\frac{2|\tilde{g}_j|_{D_{j,\frac{1}{2}}}}{\eta}|w_j|_{D_{j,\frac{1}{2}}} \le \tilde{b}_j\tilde{b}_j\frac{2K\varepsilon}{\eta(\varepsilon)^{\tau+1}} \tag{3.42}$$

and

$$|(P_Nf+\bar{g}_j)(v_N+\varepsilon w_j)-(P_Nf+\bar{g}_j)(v_N)|_{D_{j+1}} \leq \varepsilon \left\| \frac{\partial}{\partial v_N}(P_Nf+\bar{g}_j) \right\|_{D_{j+1}} |w_j|_{D_{j,\frac{1}{2}}} \leq (b+\bar{b}_j)\frac{2K\varepsilon}{\eta(\varepsilon)^{\tau+1}}\tilde{b}_j. \tag{3.43}$$

Now we are in the position to estimate $|\tilde{g}_{j+1}|_{D_{j+1}}$ and $|\bar{g}_{j+1}|_{D_{j+1}}$. We estimate the different parts in (3.7). To simplify notation we let $\psi = \frac{2K\varepsilon}{\eta(\varepsilon)^{\tau+1}}$. Then

$$\begin{aligned} |r_j|_{D_{j+1}} &\leq (1-\psi\tilde{b}_j)^{-1}((N+b+\bar{b}_j)\tilde{b}_j\psi + N\tilde{b}_j\psi + b\tilde{b}_j\psi + \bar{b}_j\tilde{b}_j\psi + \tilde{b}_j\tilde{b}_j\psi), \\ &= (1-\psi\tilde{b}_j)^{-1}(2(N+b+\bar{b}_j)+\tilde{b}_j)\tilde{b}_j\psi. \end{aligned} \tag{3.44}$$

then we obtain $|\tilde{g}_{j+1}|_{D_{j+1}} \leq \tilde{b}_{j+1}$ and $|\bar{g}_{j+1}|_{D_{j+1}} \leq \bar{b}_{j+1}$ with

$$\bar{b}_{j+1} = \bar{b}_j + (1-\psi\tilde{b}_j)^{-1}(2(N+b+\bar{b}_j)+\tilde{b}_j)\tilde{b}_j\psi \tag{3.45}$$

$$\tilde{b}_{j+1} = 2(1-\psi\tilde{b}_j)^{-1}(2(N+b+\bar{b}_j)+\tilde{b}_j)\tilde{b}_j\psi \tag{3.46}$$

Now we will choose ε_0, the index of the Galerkin approximation $N(\varepsilon)$, the number of averaging steps j^* and the step size $\eta(\varepsilon)$ for the Cauchy estimates such that

$$\bar{b}_j \leq 2\bar{b}_0 \text{ and } \tilde{b}_{j+1} \leq \tilde{b}_j \leq 2\bar{b}_0. \tag{3.47}$$

For $j=0$ these inequalities hold by definition. We first perform a single averaging step to obtain $\bar{b}_0 = C\varepsilon N(\varepsilon)$ as in the proof of theorem A and decrease the domain by $\delta/4$. If (3.47) holds, all transformations are well-defined by (3.39), if we can ensure

$$\psi\bar{b}_0 < \frac{1}{2} \text{ for } 0<\varepsilon<\varepsilon_0. \tag{3.48}$$

Our choice of parameter will be such that

$$2(1-\psi\tilde{b}_j)^{-1}(2(N+b+\bar{b}_j)+\tilde{b}_j)\psi < \frac{1}{\mathrm{e}}. \tag{3.49}$$

Then we obtain from (3.45) and (3.46), that

$$\tilde{b}_j \leq \left(\frac{1}{\mathrm{e}}\right)^j 2\bar{b}_0 \tag{3.50}$$

$$\begin{aligned} \bar{b}_{j+1} &< \bar{b}_j + \frac{1}{2\mathrm{e}}\tilde{b}_j \leq \bar{b}_0 + \frac{1}{2\mathrm{e}}(\tilde{b}_0+\ldots+\tilde{b}_j) \leq \bar{b}_0 + \frac{1}{2\mathrm{e}}\sum_{k=0}^{j}\frac{1}{\mathrm{e}^k}2\bar{b}_0 \\ &= \bar{b}_0(1+\frac{1}{\mathrm{e}-1}) < 2\bar{b}_0, \end{aligned} \tag{3.51}$$

i.e. (3.47) will hold. We choose

$$\eta(\varepsilon) = \bar{K}(\varepsilon N(\varepsilon))^{1/(1+\tau)}; \quad \bar{K} = (64\mathrm{e}K(\bar{b}_0+1+b))^{1/(1+\tau)}. \tag{3.52}$$

Thus (3.48) holds as then $\psi\bar{b}_0 = \frac{\bar{b}_0}{N(\varepsilon)32(\bar{b}_0+1+b)} < \frac{1}{2}$ for all $0<\varepsilon<\varepsilon_0$ as long $N(\varepsilon)\geq 1$ and $\eta(\varepsilon)<\delta/2$. Furthermore (3.49) follows from (3.47), as

$$\begin{aligned} 2(1-\psi\tilde{b}_j)^{-1}(2(N(\varepsilon)+b+\bar{b}_j)+\tilde{b}_j)\psi &\leq 4\frac{2(N(\varepsilon)+b)+4\bar{b}_0}{32\mathrm{e}N(\varepsilon)(\bar{b}_0+1+b)} \\ &< \frac{8N(\varepsilon)}{32\mathrm{e}N(\varepsilon)}+\frac{8b}{32\mathrm{e}b}+\frac{16\bar{b}_0}{32\mathrm{e}\bar{b}_0}=\frac{1}{\mathrm{e}}. \end{aligned} \tag{3.53}$$

Then the iterative steps are well-defined; all iterative estimates hold and the number of averaging steps j^* is chosen such that the complex domain $(B_R(Y)+\delta/2)\times(\mathbb{T}+\delta/2)$ is contained in D_{j^*}, then

$$j^* = \left\lfloor \frac{\delta}{4\eta(\varepsilon)} \right\rfloor > \frac{\delta}{4\bar{K}(\varepsilon N(\varepsilon))^{1/(1+\tau)}} - 1. \tag{3.54}$$

The final transformed nonlinearities on $P_{N(\varepsilon)}Y$ are $P_{N(\varepsilon)}f(v)+\bar{g}_{j^*(\varepsilon)}(v,\varepsilon)+\tilde{g}_{j^*(\varepsilon)}(v,\theta,\varepsilon)$ with

$$|P_{N(\varepsilon)}f(.)+\bar{g}_{j^*(\varepsilon)}(.,\varepsilon)|_{D_{j^*}} < b+2\bar{b}_0 \tag{3.55}$$

and the first desired exponential estimate

$$|\tilde{g}_{j^*(\varepsilon)}(.,.,\varepsilon)|_{D_{j^*}} \leq \left(\frac{1}{\mathrm{e}}\right)^{j^*} 2\bar{b}_0 \leq C\exp\left(-c_2(\varepsilon N(\varepsilon))^{1/(1+\tau)}\right), \tag{3.56}$$

all estimates up to here hold both for $Y=X$ and $Y=\mathcal{G}_{\sigma,\nu}$. Another exponential estimate is obtained for the remainder terms r^* in (3.16) using the exponential approximation hypothesis 2.7 for $v\in\mathcal{G}_{\sigma,\nu}$

$$|r^*(v,\theta,\varepsilon)|_X \leq C(|v|_{\mathcal{G}_{\sigma,\nu}})\exp(-c_0/(N(\varepsilon))^\nu). \tag{3.57}$$

Balancing both exponential estimates yields an optimal choice for $(\varepsilon N(\varepsilon))^{1/(1+\tau)}=(N(\varepsilon))^{-\nu}$, i.e. $N(\varepsilon)=\varepsilon^{-1/(1+\nu(\tau+1))}$. Then we finally have the exponential estimate for the nonautonomous remainder $r=r^*+\tilde{g}_{j^*(\varepsilon)}$ as in (2.22)

$$|r(v,\theta,\varepsilon)|_X \leq C(|v|_{\mathcal{G}_{\sigma,\nu}})\exp(-c/\varepsilon^{1/(\tau+1+1/\nu)}). \tag{3.58}$$

For the autonomous correction $\bar{g}=\bar{g}_{j^*}$we obtain as in the proof of theorem A

$$|\bar{g}|_X \leq C\varepsilon N(\varepsilon)+C(|v|_{\mathcal{G}_{\sigma,\nu}})\exp(-c/\varepsilon^{1/(\tau+1+1/\nu)}) \leq C\varepsilon^{(\tau+1)/(\tau+1+1/\nu)}. \tag{3.59}$$

We complete the proof of theorem B by using the Gronwall argument to extend the estimates from the equation to the solutions. □

Remark 3.3 *To enforce the condition* $\int_{\mathbb{T}} r = 0$*, we let*

$$\begin{aligned} \bar{g} &= \bar{g}_{j^*} + \int_{\mathbb{T}} r^* \\ r &= r^* + \tilde{g}_{j^*(\varepsilon)} - \int_{\mathbb{T}} r^* \end{aligned}$$

Then we obtain similar estimates as in theorems A and B, except that we have $|\bar{g}(v,\varepsilon)|_X \leq C(|v|_Y)\varepsilon^{(\tau+1)/(\tau+1+1/\nu)}$ *with* $Y=\mathcal{Y}_\alpha$ *or* $Y=\mathcal{G}_{\sigma,\nu}$ *instead of (2.16) or (2.21).*

4 Examples

In this section we give examples of equations that can be written in the form (2.1), such that theorems A and B can be applied.

4.1 Reaction-Diffusion equation

In reaction-diffusion equations describing a light sensitive Belousov Zhabotinsky reaction, an external forcing can be introduced in the reaction terms for example through light changes, see e.g. [SaScWu99] and references therein. Such equations with a fast quasiperiodic external forcing and with species $u_1, \dots, u_n$ have then the form for $u = (u_1, \dots u_n)$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} u &= \operatorname{diag}(d_1, \dots, d_n)\Delta u + f(u, \theta) = \operatorname{diag}(d_1, \dots, d_n)\Delta u + \bar{f}(u) + g(u, \theta) &(4.1)\\
\dot{\theta} &= \frac{1}{\varepsilon}\omega &(4.2)\\
(u(0), \theta(0)) &= (u_0, \theta_0) \in H^s(\Omega, \mathbb{R}^n) \times \mathbb{T}
\end{aligned}$$

with periodic boundary conditions on $\Omega = [0, L]^d$. If $f : \mathbb{R}^n \times \mathbb{T} \to \mathbb{R}^n$ is an entire function in u and real analytic in θ, then (4.2) has highly regular solutions. Starting with initial data $u_0 \in X = H^s_{per}(\Omega, \mathbb{R}^n)$, with possibly non-integer $s > d/2$, the solutions are bounded in Gevrey spaces after any finite transient. If $|u(t)|_X < M$ for $0 \le t \le T$, then solutions become highly regular

$$\begin{aligned}
|u(t)|_{\mathcal{G}_{t,1/2}} &\le 2M \text{ for } 0 \le t \le t^* \\
|u(t)|_{\mathcal{G}_{t^*,1/2}} &\le 2M \text{ for } t^* \le t \le T,
\end{aligned}$$

for some $t^* > 0$ and with $|u|_{\mathcal{G}_{\sigma,1/2}} = |u|_X + |\exp(\sigma(-\Delta)^{1/2})u|_X$; for proofs and details see e.g. [FeTi98, Mat01]. Using a spatial Fourier decomposition of u, the hypothesis 2.1 is fulfilled, the assumptions 2.3 and 2.5 hold. The exponential approximation property 2.7 holds in $\mathcal{G}_{\sigma,1/2}$ with $c_0 = \sigma$ and for general initial data after a transient, since the Gevrey norm can be expressed in spatial Fourier modes. For $u(x) = \sum_{k\in\mathbb{Z}^d} u_k \exp(2\pi \mathrm{i}(k, x)) \in \mathcal{G}_{\sigma,1/2}$ with $u_k \in \mathbb{C}^n$, we have $|u|^2_{\mathcal{G}_{\sigma,1/2}} = \sum_{k\in\mathbb{Z}^d} |u_k|^2(1 + \exp(\sigma|k|))^2$. Then we let $P_N u = \sum_{k\in\mathbb{Z}^d, 4\pi^2|k|^2 \le N} u_k \exp(2\pi \mathrm{i}(k, x))$ and then

$$|(id - P_N)u|_X = |\sum_{k\in\mathbb{Z}^d, |k|>\sqrt{N}} u^k \exp(2\pi \mathrm{i}(k, x))|_X \le \exp(-\sigma N^{1/2})|u|_{\mathcal{G}_{\sigma,1/2}}. \tag{4.3}$$

The assumption, that f is entire in u and real analytic in θ, gives analyticity of the nonlinearities in $\mathcal{G}_{\sigma,1/2}$ and X; see [Mat01]. Using standard semigroup arguments, $D\Delta$ generates a strongly continuous semigroup in X. The same arguments also apply to $\mathcal{G}_{\sigma,1/2}$.

The nonautonomous remainder can then be estimated for Gevrey initial data

$$|r(u, \theta, \varepsilon)|_X \le C(|u|_{\mathcal{G}_{\sigma,1/2}}) \exp(-c/\varepsilon^{1/(\tau+3)}) \tag{4.4}$$

and the correction term is bounded by $C\varepsilon^{(\tau+1)/(\tau+3)}$. For general initial data $u_0 \in X$ the exponent c increases in time as the regularity of the solutions increases

$$|r(u(t),\theta(t),\varepsilon)|_X \leq C(|u_0|_X)\exp(-c(t)/\varepsilon^{1/(\tau+3)}) \tag{4.5}$$

with $c(t) = \min(t, t^*, c)$, where t^* is the maximal Gevrey exponent. For H^s initial data the estimates on $\bar{g}$ have to be modified, in $(id - P_N)X$, we obtain for r^* in general only boundedness.

If the nonlinearity f possesses only finite differentiability, only finite Sobolev regularity of solutions can be assumed. Theorem A can be applied with $\mathcal{Y}_\alpha = H^{s+2\alpha}_{per}$ as $|u|^2_{H^{s+2\alpha}} = \sum_{k\in\mathbb{Z}^d} |u^k|^2(1 + |k|^{s+2\alpha})^2$ and

$$|(id - P_N)u|_X \leq CN^{-\alpha}|u|_{H^{s+2\alpha}} \tag{4.6}$$

for some α depending on the differentiability ℓ of f. Again a version holds, starting with H^s initial data and using regularising to higher Sobolev spaces, as well.

For both the analytic and the finite differentiability case, the correction term is given by the iterative procedure. A good approximation is the correction in the first step

$$\bar{g}(u,\varepsilon) = \int_{\mathbb{T}} D_u g(u,\theta,\varepsilon) w_0(P_{N(\varepsilon)}v,\theta,\varepsilon)\mathrm{d}\theta \tag{4.7}$$

where w_0 is the first transformation as defined in (3.18) and $N(\varepsilon)$ was chosen in the proofs as for the finite regularity case $N(\varepsilon) = \varepsilon^{q(\ell)/(q(\ell)+\alpha)}$ and in the analytic setting $N(\varepsilon) = \varepsilon^{-1/(1+\nu(\tau+1))}$.

Another way to create a periodic forcing in the nonlinearity is a time-periodic change of the diffusion coefficient. As an example we consider a scalar reaction-diffusion equation

$$\frac{\mathrm{d}}{\mathrm{d}t}u = d(t\omega_1/\varepsilon)\Delta u + f(u). \tag{4.8}$$

with $d(s) = \bar{d} + \tilde{d}(s)$ and $\int_0^1 \tilde{d}(s)\mathrm{d}s = 0$. We can transform (4.8) into the general form of the theorems with a rescaling of time $\tau(t)$. Let $\tau(t) = \int_0^t d(s\omega_1/\varepsilon)/\bar{d}\,\mathrm{d}s$, then $\tau(l\varepsilon/\omega_1) = l\varepsilon/\omega_1$ for $l \in \mathbb{Z}$ and the derivative $\tau'(t) = d(t\omega_1/\varepsilon)/\bar{d}$ is periodic, such that this is a near-identity reparametrisation of time. In the new time the equation is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tau}u &= \frac{\mathrm{d}u}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{1}{\tau'(t(\tau))}\frac{\mathrm{d}u}{\mathrm{d}t}\\ &= \frac{\bar{d}}{d(t(\tau)\omega_1/\varepsilon)}\left(d(t(\tau)\omega_1/\varepsilon)\Delta u + f(u)\right) = \bar{d}\Delta u + \frac{\bar{d}}{d(t(\tau)\omega_1/\varepsilon)}f(u)\end{aligned}$$

The nonlinearity can be understood as a periodic forcing: let $L(T) = \int_0^T d(s)/\bar{d}\mathrm{d}s$, which is clearly periodic. Then

$$\frac{\bar{d}}{d(t(\tau)\omega_1/\varepsilon)}f(u) = \frac{\bar{d}}{d(L(\tau\omega_1/\varepsilon))}f(u), \tag{4.9}$$

as

$$L(\tau\omega_1/\varepsilon) = \int_0^{\omega_1\tau/\varepsilon} d(s)/\bar{d}\mathrm{d}s = \frac{\omega_1}{\varepsilon}\int_0^\tau d(s\omega_1/\varepsilon)/\bar{d}\mathrm{d}s = t(\tau)\frac{\omega_1}{\varepsilon}. \tag{4.10}$$

On the one-dimensional torus we add the equation

$$\dot{\theta} = \frac{1}{\varepsilon}\omega_1, \tag{4.11}$$

such that both equations together have the desired form. Then as above the theory of this paper or alternatively the theorems of [Mat01] are applicable and there exists an autonomous equation that is exponentially close to the original system in the analytic framework.

4.2 Nonlinear Schrödinger equations

In fibre optics communications in long cables there are periodically changing dispersion coefficients and periodically distributed amplification sites. This can be modelled by a nonlinear Schrödinger equation,

$$\mathrm{i}u_z + D(\omega_1 z/\varepsilon)u_{tt} + C(\omega_2 z/\varepsilon)|u|^2 u = 0. \tag{4.12}$$

The direction of the evolution is z. The equation describes how a temporal profile $u(z,.)$ – given as a function of t – is transported along the cable described by the z-direction, see e.g. [NM92]. We use the phase space $X = H^1(\mathbb{R},\mathbb{C}) = H^1(\mathbb{R},\mathbb{R}^2)$ for $u(z,.)$. The corresponding Gevrey classes are defined as

$$\mathcal{G}_{\sigma,1/2} = \left\{ u \in H^1(\mathbb{R},\mathbb{R}^2) | |u|_{H^1} + |\exp(\sigma|\partial_{tt}|^{1/2})u|_{H^1} < \infty \right\}. \tag{4.13}$$

Note that the nonlinearity is entire in $u_1 = \operatorname{Re} u$ and $u_2 = \operatorname{Im} u$. We first construct a linear Floquet-type transformation to transform (4.12) into the general form of the theorems. The linear equation can be solved by a spatial Fourier transform. The complex Fourier modes $u_k \exp(\mathrm{i}kt)$ with $k \in \mathbb{R}$ and $u_k \in \mathbb{C}$ are eigenfunctions for the linear evolution

$$\frac{\mathrm{d}}{\mathrm{d}z} u_k(z) = \mathrm{i}D(\omega_1 z/\varepsilon)k^2 u_k(z) \tag{4.14}$$

Then the evolution is given by

$$\begin{aligned} u_k(z) &= u_k(0)\exp\left(\int_0^z \mathrm{i}k^2 D(s\omega_1/\varepsilon)\mathrm{d}s\right) \\ &= u_k(0)\exp\left(\int_0^z \mathrm{i}k^2 \tilde{D}(s\omega_1/\varepsilon)\mathrm{d}s\right)\exp(\mathrm{i}k^2\bar{d}z) \end{aligned}$$

with $D(s) = \bar{D} + \tilde{D}(s)$ and $\int_0^1 \tilde{D}(s)\mathrm{d}s = 0$.

We apply the periodic Floquet-type transformation, $Q(z,\varepsilon)$, defined by

$$v_k = Q_k(z)u_k = \exp\left(-\int_0^z \mathrm{i}k^2\tilde{D}(s\omega_1/\varepsilon)\mathrm{d}s\right)u_k \tag{4.15}$$

on each Fourier mode separately. Then we obtain the quasiperiodic equation, where $\hat{}$ denotes the Fourier transform

$$\frac{\mathrm{d}}{\mathrm{d}z}v(z) = \frac{\mathrm{d}}{\mathrm{d}z}\left(\int_{k\in\mathbb{R}} Q_k(z)u_k(z)\mathrm{d}k\right)$$

$$
\begin{aligned}
&= \int_{k\in\mathbb{R}} \left(-\mathrm{i}k^2 \tilde{D}(z\omega_1/\varepsilon) Q_k(z) u_k(z) + Q_k(z) \frac{\mathrm{d}}{\mathrm{d}z} u_k(z) \right) \mathrm{d}k \\
&= \int_{k\in\mathbb{R}} \Big[-\mathrm{i}k^2 \tilde{D}(z\omega_1/\varepsilon) Q_k(z) u_k(z) + \mathrm{i}k^2 (\tilde{D}(z\omega_1/\varepsilon) + \bar{D}) Q_k(z) u_k(z) \\
&\quad + Q_k(z) \left(\widehat{(C(\omega_2 z/\varepsilon)} |u|^2 u) \right) \Big] \mathrm{d}k \\
&= \mathrm{i}\bar{D} v_{tt} + Q(z,\varepsilon) \left(C(\omega_2 z/\varepsilon) |Q^{-1}(z,\varepsilon) v|^2 Q^{-1}(z,\varepsilon) v \right). \qquad (4.16)
\end{aligned}
$$

The transformation $Q(.,\varepsilon)$ has period ε/ω_1. So rewriting $\tilde{Q}(z\omega_1/\varepsilon,\varepsilon) = Q(z,\varepsilon)$, where $\tilde{Q}$ is then 1-periodic in its first argument, we obtain the desired form of (2.1):

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}z} v(z) &= \mathrm{i}\bar{D} v_{tt} + \tilde{Q}(\theta_1,\varepsilon) C(\theta_2) |\tilde{Q}^{-1}(\theta_1,\varepsilon) v|^2 \tilde{Q}^{-1}(\theta_1,\varepsilon) v \\
\dot{\theta} &= \frac{1}{\varepsilon} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}. \qquad (4.17)
\end{aligned}
$$

The transformations Q and $\tilde{Q}$ leave each complex Fourier mode invariant and is an isometry on each mode. Thus $\tilde{Q}$ maps the spaces X and $\mathcal{G}_{\sigma,1/2}$ to themselves and furthermore

$$
\tilde{Q}(\theta_1,\varepsilon)(C(\theta_2)|\tilde{Q}^{-1}(\theta_1,\varepsilon).|^2 \tilde{Q}^{-1}(\theta_1,\varepsilon).) : Y \to Y \qquad (4.18)
$$

for $Y = X$ and $Y = \mathcal{G}_{\sigma,1/2}$. Then the theorem is applicable when we use a truncation of the Fourier transform as a Galerkin approximation. The Diophantine conditions on (ω_1,ω_2) still has to be assumed. Then an effective autonomous description can be given for regular initial data.

4.3 Nonlocality

We give an example of a linear parabolic partial differential equation with periodic forcing, where we can see that the transformation has to be nonlocal to remove higher order terms. Consider

$$
u_t(x,t) = u_{xx}(x,t) + f(t/\varepsilon,x) u(x,t), \qquad (4.19)
$$

with $u(0,t) = u(1,t)$ and $x \in [0,1]$. A specific choice of f will be made later. As the equation is linear, we can restrict out attention to linear transformations – nonlinear ones will create a larger error. A general linear, local and near-identity transformation has the form

$$
u(x,t) = v(x,t) + \varepsilon w(t/\varepsilon,x) v(x,t), \qquad (4.20)
$$

where $w(\tau,x)$ is 1-periodic in τ and in a certain regularity class in X to ensure that $\mathcal{G}_{\sigma,\nu}$ resp. $\mathcal{Y}_\alpha$ is mapped to itself. Then the transformed equation is

$$
v_t = v_{xx} + \frac{1}{1+\varepsilon w(t/\varepsilon,x)} \left[\varepsilon \partial_{xx} w(t/\varepsilon,x) v + 2\varepsilon \partial_x w(t/\varepsilon,x) v_x + f(t/\varepsilon,x)(v + \varepsilon w(t/\varepsilon,x) v) - \partial_\tau w(t/\varepsilon,x) v \right] \qquad (4.21)
$$

Using the periodic boundary conditions and the fact that w is periodic in $\tau = t/\varepsilon$, we expand w in a Fourier series simultaneously in x and τ.

$$w(\tau, x) = \sum_{l\in\mathbb{Z}, m\in\mathbb{Z}} w_{lm} \exp(\mathrm{i}2\pi(l\tau + mx)) \tag{4.22}$$

We obtain a relation between the Fourier expansion of f and w, which we have to fulfil to make the nonautonomous terms small in (4.21). Denoting t/ε by τ, this relation is

$$\Big|\frac{1}{1+\varepsilon w(\tau,x)}\left[\varepsilon\partial_{xx}w(\tau,x)v + 2\varepsilon\partial_x w(\tau,x)v_x + f(\tau,x)(v+\varepsilon w(\tau,x)v) - \partial_\tau w(\tau,x)v\right]$$
$$-\int_0^1 \frac{1}{1+\varepsilon w(\theta,x)}\left[\varepsilon\partial_{xx}w(\theta,x)v + 2\varepsilon\partial_x w(\theta,x)v_x + f(\theta,x)(v+\theta w(\tau,x)v) - \partial_\theta w(\theta,x)v\right]\mathrm{d}\theta\Big|_{H^s} \in \mathrm{O}(\varepsilon^\nu)$$

with w bounded in H^s. In particular as $H^s(S^1\times S^1)$ is an algebra for any (possibly non-integer) $s > 1$, this is equivalent to

$$|\left[\varepsilon\partial_{xx}w(\tau,x)v + 2\varepsilon\partial_x w(\tau,x)v_x + f(\tau,x)(v+\varepsilon w(\tau,x)v) - \partial_\tau w(\tau,x)v\right]$$
$$-\int_0^1 \left[\varepsilon\partial_{xx}w(\theta,x)v + 2\varepsilon\partial_x w(\theta,x)v_x + f(\theta,x)(v+\theta w(\tau,x)v) - \partial_\tau w(t/\varepsilon,x)v\right]\mathrm{d}\theta|_{H^s} \in \mathrm{O}(\varepsilon^\nu).$$

Now letting $v = 1$, we need in particular in terms of Fourier-modes for $l \neq 0$,

$$w_{lm}(-\mathrm{i}2\pi l - \varepsilon 4\pi^2 m^2) + f_{lm} + \varepsilon \sum_{l_1\in\mathbb{Z}, m_1\in\mathbb{Z}} w_{l_1,m_1} f_{l-l_1,m-m_1} \in \mathrm{O}(\varepsilon^\nu). \tag{4.23}$$

To achieve $= 0$ there is a unique solution for ε small by the implicit function theorem for f and w in H^s with $s > 1$. Then the nonlinearity wf is differentiable again as H^s is an algebra. Thus w is defined by this relation up to an error $\mathrm{O}(\varepsilon^\nu)$ in ε of some order ν, that we want to achieve. In particular we have

$$\left|w_{lm} - \frac{1}{\mathrm{i}2\pi l + \varepsilon 4\pi^2 m^2} f_{lm}\right| \le \varepsilon\frac{|f||w|}{2\pi} + \mathrm{O}(\varepsilon^\nu) \le \varepsilon\frac{|f|^2}{2\pi^2} + \mathrm{O}(\varepsilon^\nu) \tag{4.24}$$

as $|w| \le |f|/\pi$ for ε small enough. If we now derive in the same way another relation for w using $v = v_k \frac{1}{k^s}\exp(i2\pi kx)$, we obtain instead

$$w_{lm}(-\mathrm{i}2\pi l - \varepsilon 4\pi^2 m^2 + \varepsilon 4\pi^2 mk) + f_{lm} + \varepsilon \sum_{l_1\in\mathbb{Z}, m_1\in\mathbb{Z}} w_{l_1,m_1} f_{l-l_1,m-m_1} \in \mathrm{O}(\varepsilon^\nu), \tag{4.25}$$

i.e. again by the the implicit function theorem

$$\left|w_{lm} - \frac{1}{\mathrm{i}2\pi l + \varepsilon 4\pi^2 m^2 - \varepsilon 4\pi^2 mk} f_{lm}\right| \le \varepsilon\frac{|f|^2}{2\pi^2} + \mathrm{O}(\varepsilon^\nu) \tag{4.26}$$

Now we consider the forcing $f(\tau, x) = \exp(2\pi\mathrm{i}\tau)\exp(2\pi\mathrm{i}x)$. Then we compare the two values for w_{11} given by (4.24) and (4.26), which differ for ε small by more than

$$\left|\frac{1}{\mathrm{i}2\pi + \varepsilon 4\pi^2} - \frac{1}{\mathrm{i}2\pi + \varepsilon 4\pi^2 + \varepsilon 4\pi^2 k}\right| \ge \frac{\varepsilon k}{2}. \tag{4.27}$$

This yields a direct contradiction for some fixed $k > \frac{|f|^2}{\pi^2}$ and any choice of $\nu \ge 1$, independent of ε and the choice of function spaces $\mathcal{G}_{\sigma,\nu}$ resp. $\mathcal{Y}_\alpha$. Hence we cannot achieve a higher order averaging result using a local transformation $id + \varepsilon w$.

Acknowledgement

The author would like to thank Prof. C. Simó for explaining the proof of lemma 3.1. Furthermore, the careful reading and comments by the referees were very helpful. Financial support of the Deutsche Forschungsgemeinschaft (DFG) in the Schwerpunktprogramm 'Modellierung, Analysis und effektive Simulation von Mehrskalenproblemen' is gratefully acknowledged.

A Appendix

We prove lemma 3.1. For $\tau > p-1$, we can restrict our attention just to points in $\mathbb{Z}^p$, which are close to the hyper-plane $\omega^\perp$, as

$$\sum_{m\in\mathbb{Z}^p,|m|\le k,\mathrm{dist}(m,\omega^\perp)\ge 1} \frac{1}{|(m,\omega)|} \le Ck^{p-1}\log k. \tag{A.1}$$

This can be estimated by $(G/2)k^\tau$ uniformly in k.

To discuss those points in $\mathbb{Z}^p$, which are close to $\omega^\perp$, first assume that the constant γ in the estimate $|(m,\omega)| \ge \gamma|m|^{-\tau}$ is sharp, i.e.

$$\gamma = \inf_{m\in\mathbb{Z}^p\setminus 0}\{|(m,\omega)||m|^\tau\} \tag{A.2}$$

The key idea is now to classify the $m \in \mathbb{Z}^p, |m| \le k$ in how close they are to this worst case estimate for the small divisor $1/(m,\omega)$. In each case we estimate the number of m and their influence in the sum. Now fix some $\rho > 1$ and let

$$B^s = \{m \in \mathbb{Z}^p | \gamma\rho^s \le |(m,\omega)||m|^\tau < \gamma\rho^{s+1}\} \tag{A.3}$$

First consider all those m in B^0 with $2k/3 \le |m| \le k$. Then we obtain lower estimates on the distance of $m_1, m_2 \in B^0, m_1 \ne m_2$ in the following way: As

$$|(m_1-m_2,\omega)|k^\tau \le |(m_1,\omega)|k^\tau + |(m_2,\omega)|k^\tau \le (3/2)^\tau|(m_1,\omega)||m_1|^\tau + |(m_2,\omega)||m_2|^\tau \le 2(3/2)^\tau\gamma\rho \tag{A.4}$$

and using the Diophantine estimate on $m_1 - m_2$, which could be in B^0, we also have

$$|(m_1-m_2,\omega)||m_1-m_2|^\tau > \gamma. \tag{A.5}$$

Combining these two inequalities yields

$$\begin{aligned} |m_1-m_2|^\tau &> \frac{2^\tau}{23^\tau\rho}k^\tau \text{ or} \\ |m_1-m_2| &> \frac{2}{3\sqrt[\tau]{2\rho}}k \end{aligned}$$

Then the number $m \in B^0$ with $2k/3 < |m| < k$ is bounded by the number of balls of radius $2k/(3\sqrt[\tau]{2\rho})$ that can be packed into annulus $2k/3 \le |m| \le k$ close to the hyper-plane $\omega^\perp$. This is bounded by a constant of the form $C\rho^{(p-1)/\tau}$, where C is independent of k, p, ω, τ and γ. Then the contribution of B^0 to the main sum in the lemma can be estimated by

$$\sum_{2k/3\le|m|<k, m\in B^0} \frac{1}{|(m,\omega)|} \le C\rho^{(p-1)/\tau}\gamma k^\tau. \tag{A.6}$$

Such an analysis is repeated for B^s with $s>0$. As above we obtain estimates on the number of m in B^s with $2k/3 \le |m| < k$, first we have the lower estimate on their distance

$$|m_1 - m_2| > \frac{2}{3\sqrt[\tau]{2\rho^{s+1}}}k, \tag{A.7}$$

such that the number of $m \in B^s$ is bounded by $C\rho^{(s+1)(p-1)/\tau}$ and their overall contribution is bounded by

$$C\rho^{(s+1)(p-1)/\tau}\rho^{-s}\gamma k^\tau \tag{A.8}$$

Hence the sum of all interesting B_s, i.e. $\rho^s \le m^\tau/\gamma$, is bounded as $p-1<\tau$:

$$C\rho^{(p-1)/\tau}\gamma k^\tau \sum_{s=0}^{\infty}\left(\rho^{\frac{p-1}{\tau}-1}\right)^s = C\rho^{(p-1)/\tau}\gamma k^\tau(1-\rho^{\frac{p-1}{\tau}-1})^{-1} \tag{A.9}$$

Then this can be extended from the annulus $2k/3 \le |m| \le k$ to the ball $|m| \le k$:

$$\sum_{|m|<k}\frac{1}{|(m,\omega)|} \le C\rho^{(p-1)/\tau}\gamma(1-\rho^{\frac{p-1}{\tau}-1})^{-1}k^\tau(1+(2/3)^\tau+(2/3)^{2\tau}+(2/3)^{3\tau}+\ldots) \le (G/2)k^\tau \tag{A.10}$$

Together with the contribution of the points with distance larger than 1 to the hyper-plane $\omega^\perp$, we obtain the desired estimate.□

Now we prove the Cauchy estimate (lemma 3.2). The lemma is a direct consequence of the Cauchy formula for vector-valued holomorphic functions of a single complex variable. We want to differentiate in the direction v, letting without restriction $\|v\|_Y = 1$. For any $u \in \Omega - \eta$ we take a circle in the complex plane defined by $u + zv, z \in \mathbb{C}, |z| = \eta$. Then we have

$$f(u) = \frac{1}{2\pi i}\oint_{u+zv,z\in\mathbb{C},|z|=\eta}\frac{f(u+zv)}{-z}\mathrm{d}z \tag{A.11}$$

So taking a partial derivative in the direction v, we first consider for $h \in \mathbb{C}$

$$\begin{aligned} f(u+hv) &= \frac{1}{2\pi i}\oint_{u+hv+zv,z\in\mathbb{C},|z|=\eta}\frac{f(u+hv+zv)}{-z}\mathrm{d}z = \frac{1}{2\pi i}\oint_{u+\tilde{z}v,\tilde{z}\in\mathbb{C},|\tilde{z}-h|=\eta}\frac{f(u+\tilde{z}v)}{-\tilde{z}+h}\mathrm{d}\tilde{z} \\ &= \frac{1}{2\pi i}\oint_{u+\tilde{z}v,\tilde{z}\in\mathbb{C},|\tilde{z}|=\eta}\frac{f(u+\tilde{z}v)}{-\tilde{z}+h}\mathrm{d}\tilde{z}, \end{aligned} \tag{A.12}$$

where the last equality holds by the Cauchy theorem in $\mathbb{C}$. Thus for the directional derivative in direction v, we have

$$\begin{aligned} D_u f(u)v &= \lim_{h\to 0}\frac{1}{h}\left(f(u+hv)-f(u)\right) = \lim_{h\to 0}\frac{1}{2\pi i}\oint_{u+zv,z\in\mathbb{C},|z|=\eta} f(u+zv)\frac{1}{h}\left(\frac{1}{-z+h}-\frac{1}{-z}\right)\mathrm{d}z \\ &= \frac{1}{2\pi i}\oint_{u+zv,z\in\mathbb{C},|z|=\eta}\frac{f(u+zv)}{z^2}. \end{aligned} \tag{A.13}$$

Thus
$$\|D_u f\|_{\Omega-\eta} \leq \frac{1}{2\pi} 2\pi\eta \frac{|f|_\Omega}{\eta^2} = \frac{|f|_\Omega}{\eta} \tag{A.14}$$
and the lemma is proved. □

References

[AKN97] V.Arnold, V.Kozlov, and A. Neishtadt, Mathematical Aspects of Classical and Celestial Mechanics, Springer Verlag, Berlin (1997).

[BP89] N. Bachvalov, and G. Panasenko, Homogenisation: averaging processes in periodic media, Nauka, Moscow (1984); Kluwer, Dordrecht (1989).

[Bam99] D. Bambusi, Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Math. Z.* **230** (1999), 345–387.

[Bam05] D.Bambusi, Galerkin averaging method and Poincaré normal form for some quasilinear PDEs, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **4** (2005), 669–702

[BLP84] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North Holland, Amsterdam (1984).

[BoMi61] N. Bogoliubov, and Y. Mitroploskii, *Asymptotic Methods in the theory of nonlinear oscillations*, Gordon and Breach, New York (1961).

[Bor98] F. Bornemann, *Homogenization in time of singularly perturbed mechanical systems*, Lecture Notes in Mathematics **1687**, Springer-Verlag, Berlin (1998).

[EZ03] M. Efendiev, and S. Zelik, The regular attractor for the reaction-diffusion system with a nonlinearity rapidly oscillating in time and its averaging, *Adv. Differ. Equ.* **8** (2003), 673-732.

[FeTi98] A. Ferrari, and E. Titi, Gevrey regularity for nonlinear analytic parabolic equations, *Comm. Part. Diff. Eq.* **23** (1998), 1-16.

[FV01] B. Fiedler, and M. Vishik, Quantitative homogenization of analytic semigroups and reaction diffusion equations with diophantine spatial frequencies. *Adv. Differ. Equ.* **6** (2001), 1377–1408.

[JKO94] V. Jikov, S. Kozlov, and O. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, New-York (1994).

[LM88] P.Lochak, C. Meunier, *Multiphase Averaging for Classical Systems*, Appl. Math. Sc. **72**, Springer-Verlag, New York (1988).

[Mat01] K. Matthies, Time-averaging under fast periodic forcing of parabolic partial differential equations: Exponential estimates, *J. Differential Equations* **174** (2001), 133-180.

[MS03] K. Matthies, and A. Scheel, Exponential Averaging of Hamiltonian Evolution Equations, *Trans. Amer. Math. Soc.* **355** (2003), 747-773.

[Mat05] K. Matthies, Homogenisation of Exponential Order for Elliptic Systems in Infinite Cylinders, *Asymptot. Anal.* **43** (2005), 205-232.

[Nei84] A. Neishtadt, The separation of motions in systems with rapidly rotating phase, *J. Appl. Math. Mech.* **48** (1984), 133-139.

[Nek79] N.N. Nekhorošev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems, (Russian) *Uspehi Mat. Nauk* **32** (1977), no. 6 (198), 5–66.

[NM92] A. Newell, and J. Moloney, *Nonlinear Optics*, Advanced Topics in the Interdisciplinary Sciences, Addison-Wesley, Redwood City (1992).

[Poe99] J. Pöschel, On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem by Bambusi, *Nonlinearity* **12** (1999), 1587–1600.

[SaScWu99] B. Sandstede, A. Scheel, and C. Wulff, Bifurcations and dynamics of spiral waves. *J. Nonlinear Science* **9** (1999), 439–478.

[Sim94] C. Simó, Averaging under fast quasiperiodic forcing, in *Hamiltonian Mechanics* (Torun, 1993), *NATO Adv. Sci. Inst.Ser. B Phys.* **331** (1994), 13-34.

[Sim99] C. Simó, Lectures Institut de Mec. Celes., Paris (1999).

[Tar79] L. Tartar, Compensated compactness and applications to partial differential equations, in: Nonlinear analysis and mechanics: Heriot-Watt Symp., Vol. 4, Edinburgh 1979, Res. Notes Math. **39** (1979), 136-212.

[Tar86] L. Tartar, Oscillations in nonlinear partial differential equations: compensated compactness and homogenization, in: Nolinear systems of partial differential equations in applied mathematics, *Lectures in Applied Mathematics* **74**, AMS, Providence (1986), 243-266.

[TW07] R. Temam, and D. Wirosoetisno, Exponentially accurate approximations for the primitive equations of the ocean, *Discr. Cont. Dyn. Sys. B* **7** (2007), 425-440.