UNIVERSITY OF BATH

Citation for published version:
Matthies, K & Wayne, CE 2006, 'Wave pinning in strips', Proceedings of the Royal Society of Edinburgh Section A - Mathematics, vol. 136, no. 5, pp. 971-995. https://doi.org/10.1017/S0308210500004832

DOI:
10.1017/S0308210500004832

Publication date:
2006

Document Version
Publisher's PDF, also known as Version of record

Link to publication

University of Bath

Wave pinning in strips

Karsten Matthies[*]
Institut für Mathematik I, Freie Universität Berlin,
Arnimallee 2–6, 14195 Berlin, Germany

C. E. Wayne
Department of Mathematics and Statistics and
Center for Biodynamics, Boston University,
111 Cummington Street, Boston, MA 02215, USA
(`cew@math.bu.edu`)

We consider the existence of stationary or pinned waves of reaction–diffusion equations in heterogeneous media. By combining averaging, homogenization and dynamical-systems techniques we prove under mild non-degeneracy conditions that if the heterogeneity is periodic with period ε, pinned solutions persist at most for intervals in parameter space whose length is $\mathcal{O}(\mathrm{e}^{-c/\sqrt{\varepsilon}})$.

1. Introduction

Travelling waves in nonlinear partial differential equations (PDEs) have been used in a variety of different circumstances to model, among other things, combustion, population dynamics and the properties of electrical signals in nerve tissue. Most early work in this area assumed that the material properties of the system were constant in space and time, i.e. that the coefficients in the PDE did not depend on the independent variables in the problem. More recently, there has been a great deal of study of the more realistic situation in which the coefficients are space or time dependent. For a recent review of the literature in this area we refer to the reader to [9].

In this paper we wish to focus on a particular phenomenon, known as 'pinning', which may occur in such heterogeneous media. It has been observed that there are PDEs that exhibit travelling-wave solutions when the coefficients are homogeneous, but if the coefficients are made heterogeneous there can be a failure of propagation and the wave becomes 'pinned' to the heterogeneity, that is, one obtains a stationary solution with non-trivial spatial structure. Typically, this structure looks much like the travelling-wave profile before it was pinned.

Such pinned solutions may be important because their presence can inhibit the transfer of information or material from one part of the system to another. A more positive role for such solutions is that such localized regions of activity could then possibly be used to store information in some location. For a discussion of pinned solutions in biological models, see [3].

[*]Present address: Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK (`k.matthies@maths.bath.ac.uk`).

In order to get an idea of how important such solutions are for realistic applications, we investigate their sensitivity to changes in the system parameters and we find that they can exist in only *very* small regions of parameter space. Prior work [1,3] has shown that homogenization theory and averaging theory can be used in the construction of pinned solutions. Here we combine these two approaches to prove under mild non-degeneracy conditions that if the heterogeneity is periodic with period ε, pinned solutions persist at most for intervals in parameter space whose length is $\mathcal{O}(\mathrm{e}^{-c/\sqrt{\varepsilon}})$.

Our approach is based on a formulation of related elliptic equations as (ill-posed) spatial dynamical systems, where an unbounded space direction is considered as a time direction. This is based on an idea of Kirchgässner [4]. Another important ingredient is homogenization theory, which gives effective descriptions for heterogeneous elliptic problems. Here the heterogeneities, in the form of rapidly varying coefficients, are removed, so that the resulting approximate equation is homogeneous, i.e. the equation is homogenized (for a general reference see, for example, [2]). We note in passing that our approach is not limited to small heterogeneities. The homogenization step can also deal with large, periodic, inhomogeneous coefficients (at least those of the form described below).

To show the exponential smallness of certain effects, we then apply an averaging transformation to the homogenized equation which further reduces the size of the inhomogeneity up to exponentially small errors. Averaging up to exponentially small remainder terms was developed by Neishtadt [7] for analytic ordinary differential equations. Extensions to PDEs were first developed for the time-averaging of parabolic equations [5]. Here we use and extend a version for the homogenization of certain elliptic problems on strips [6].

We are looking for travelling waves for parabolic equations on strips $(y, x) \in \mathbb{R} \times \Omega$, where $\Omega = [0, 2\pi]^d$. As a main example, consider a scalar equation, where the diffusion coefficients and the nonlinearity both depend periodically on the unbounded space direction y:

$$\partial_\tau u = d_0\left(\frac{y}{\varepsilon}\right)\partial_y^2 u + d_1\left(\frac{y}{\varepsilon}\right)\Delta_x u + f\left(u, x, \frac{y}{\varepsilon}; \lambda, \varepsilon\right). \tag{1.1}$$

We impose periodic boundary conditions on $\partial\Omega$. The diffusion coefficients are assumed to be periodic in y/ε, and be positive and continuously differentiable functions. The nonlinearity is assumed to be an entire function of u and x and continuous in y/ε and ε. Assume furthermore that f and all its derivatives with respect to u and x are differentiable with respect to the parameter $\lambda \in \Lambda$, where Λ is a bounded, open interval of the real line.

2. Spatial dynamics and the main result

We are interested in pinned or standing waves, i.e. in solutions to the static problem

$$d_0\left(\frac{y}{\varepsilon}\right)\partial_y^2 u + d_1\left(\frac{y}{\varepsilon}\right)\Delta_x u + f\left(u, x, \frac{y}{\varepsilon}; \lambda, \varepsilon\right) = 0, \tag{2.1}$$

where $u(y, \cdot)$ converges to the y-independent equilibria u^1 and u^2 for $y \to \pm\infty$. As we will consider only this static problem, we can assume without restriction that

$d_0 \equiv 1$ after dividing by $d_0(y/\varepsilon)$, and, thus, for the remainder of the paper we consider only the equation

$$\partial_y^2 u + d\left(\frac{y}{\varepsilon}\right)\Delta_x u + f\left(u, x, \frac{y}{\varepsilon}; \lambda, \varepsilon\right) = 0, \tag{2.2}$$

where $d(\xi) = d_1(\xi)/d_0(\xi)$. Note that we do not assume that the variations in $d(\xi)$ are small: this is not a perturbation around the constant coefficient case.

Associated with the heterogeneous problem (2.2), we consider a homogenized problem

$$\partial_y^2 u + \bar{d}\Delta_x u + \bar{f}(u, x; \lambda, \varepsilon) = 0, \tag{2.3}$$

with $\bar{d}$ given by the homogenization formula

$$\bar{d} = \int_0^1 \frac{d_1(\xi)}{d_0(\xi)}\,\mathrm{d}\xi, \tag{2.4}$$

and the nonlinearity is averaged over the 'fast' variable:

$$\bar{f}(u, x; \lambda, \varepsilon) = \int_0^1 f(u, x, s; \lambda, \varepsilon)\,\mathrm{d}s. \tag{2.5}$$

We assume the existence of a standing wave $\bar{u}$ in the homogenized problem (2.3) connecting two y-independent equilibria, u^1 and u^2. More precisely, we assume that the following hypotheses hold.

(H1) $\bar{u}$, u^1, and u^2 solve (2.3) for $\lambda = \lambda_0 \in \Lambda$, $\varepsilon = 0$. Furthermore, $\partial_y u^{1,2} \equiv 0$.

(H2) $\lim_{y\to\infty} \bar{u}(\cdot, y) = u^1(\cdot)$ and $\lim_{y\to-\infty} \bar{u}(\cdot, y) = u^2(\cdot)$.

(H3) $-\bar{d}_\Omega \Delta_x u^i - D\bar{f}(u^i, x; \lambda_0, 0)$, $i = 1, 2$, are both strictly positive operators.

REMARK 2.1. In the case of a zero-dimensional cross-section (i.e. Ω is a point), these hypotheses are satisfied for bistable nonlinearities. For examples of travelling pulses in this case, see [9].

We will use additional transversality assumptions. These conditions will allow us to follow the evolution of the pinned solutions as we vary ε and λ. This transversality condition is given in terms of the linearization of (2.3) along $\bar{u}$. To formulate this, we rewrite (2.3) as a 'dynamic' equation in the unbounded spatial direction y, which from now on we will call t. By letting

$$U = \begin{pmatrix} u \\ u_t \end{pmatrix}, \tag{2.6}$$

we have

$$U_t = \bar{A}U + \bar{F}(U; \lambda, \varepsilon) \tag{2.7}$$

with

$$\bar{A}U = \begin{pmatrix} 0 & 1 \\ -\bar{d}\Delta_x & 0 \end{pmatrix}\begin{pmatrix} u \\ u_t \end{pmatrix},$$

$$\bar{F}(U; \lambda, \varepsilon)(x) = \begin{pmatrix} 0 \\ -\bar{f}(u, x; \lambda, \varepsilon) \end{pmatrix}.$$

In this notation, the standing-wave equation of the heterogeneous problem (2.2) has the form

$$U_t = A\left(\frac{t}{\varepsilon}\right)U + F\left(U, \frac{t}{\varepsilon}; \lambda, \varepsilon\right) \tag{2.8}$$

with

$$A\left(\frac{t}{\varepsilon}\right)U = \begin{pmatrix} 0 & 1 \\ -d\left(\frac{t}{\varepsilon}\right)\Delta_x & 0 \end{pmatrix} \begin{pmatrix} u \\ u_t \end{pmatrix},$$

$$F\left(U, \frac{t}{\varepsilon}; \lambda, \varepsilon\right)(x) = \begin{pmatrix} 0 \\ -f\left(u, x, \frac{t}{\varepsilon}; \lambda, \varepsilon\right) \end{pmatrix}.$$

REMARK 2.2. The hypotheses (H1)–(H3) on the stationary solution of the averaged problem translate immediately into analogous statements about the solution of (2.7) (which we will also refer to as (H1)–(H3)).

We consider (2.7) and (2.8) to be infinite-dimensional dynamical systems on a phase space X. The space X is given as a function space on the cross-section Ω of the cylindrical domain. In particular we choose $H^s(\Omega) \times H^{s-1}(\Omega)$ with $s > \max(1, \frac{1}{2}d)$ intersected with the periodic boundary conditions on $\partial\Omega$ on the cross-section $\Omega = [0, 2\pi]^d$ of the cylindrical domain. If we equip the domain $X^1 = D(A) \subset H^{s+1} \times H^s$ of the unbounded operator A with the norm $|U|_{X^1} = |U|_X + |AU|_X$, then X^1 is a Banach space. A solution of (2.7) (respectively, (2.8)) is defined to be:

(i) $U(\cdot) \in BC^0(\mathbb{R}, X^1) \cap BC^1(\mathbb{R}, X)$;

(ii) $U(\cdot)$ satisfies equations (2.7) (respectively, (2.8)) on $\mathbb{R}$ with values in X.

The linearization of (2.7) along the solution $\bar{U}$ is then given by

$$\left.\begin{array}{l} \mathcal{L}: BC^0(\mathbb{R}, X^1) \cap BC^1(\mathbb{R}, X) \to BC^0(\mathbb{R}, X), \\ \mathcal{L}V(t) = V_t(t) - AV(t) - D_U\bar{F}(\bar{U}(t); \lambda_0, 0)V(t). \end{array}\right\} \tag{2.9}$$

We assume the following generic property:

$$\mathrm{Ker}(\mathcal{L}) = \mathrm{span}(\dot{\bar{U}}(\cdot)). \tag{2.10}$$

Physically, this condition just means that the only zero eigenvalue of the linearized problem is the one that comes from the translation invariance of the averaged equation. The transversality assumption is

$$\mathrm{Rg}\,\mathcal{L} \oplus D_\lambda\bar{F}(\bar{U}(\cdot); \lambda, 0)|_{\lambda=\lambda_0} = BC^0(\mathbb{R}, X), \tag{2.11}$$

i.e. the rank defect can be removed when varying the parameter. More intuitively, this condition will guarantee (in a later section) that, when we reduce the question of existence of fronts to a question about zeros of a certain function, the derivative of this function (with respect to λ) will be non-zero and the implicit function theorem can be applied to construct the fronts.

Solutions of elliptic equations like (2.7) with f an entire function are analytic in x. Here we express such classes of highly regular functions as Gevrey spaces. We define these spaces as follows. For each $k \in \mathbb{Z}^d \setminus \{0\}$, the operator $\bar{A}$ has eigenfunctions (in X) given by

$$v_k(x) = \begin{pmatrix} \exp(\mathrm{i}k \cdot x) \\ \sqrt{\bar{d}}|k| \exp(\mathrm{i}k \cdot x) \end{pmatrix}$$

with positive eigenvalue $\sqrt{\bar{d}}|k|$ and

$$w_k(x) = \begin{pmatrix} \exp(\mathrm{i}k \cdot x) \\ -\sqrt{\bar{d}}|k| \exp(\mathrm{i}k \cdot x) \end{pmatrix}$$

with negative eigenvalue $-\sqrt{\bar{d}}|k|$. The zero eigenspace is one dimensional and spanned by $v_0 = (1,0)$. We add to these eigenvectors the vector $w_0 = (0,1)$ in order to form a basis with respect to which we can expand any vector $U \in X$ as $U = \sum_k (a_k v_k + b_k w_k)$. The X-norm is equivalent to

$$|U|_X^2 = \sum_k \{(1 + |k|)^{2s}(|a_k|^2 + |b_k|^2)\}.$$

We define the Gevrey space $\mathcal{G}_\sigma$ to be those elements of X for which the norm

$$|U|_{\mathcal{G}_\sigma}^2 = \sum_k \{(1 + |k|)^{2s}(|a_k|^2 + |b_k|^2)\mathrm{e}^{2\sigma|k|}\}$$

is finite.

REMARK 2.3. Note that $\mathcal{G}_0 = X$.

REMARK 2.4. If the nonlinear term, f, is entire (as we assume) and we fix $\sigma_0 > 0$, then every bounded solution, $u(t)$, of (2.2) lies in all $\mathcal{G}_\sigma$, $0 \leqslant \sigma \leqslant \sigma_0$, for each $t \in \mathbb{R}$ (see [6]).

THEOREM 2.5. *In (2.2) let d_0 and d be 1-periodic and differentiable in y/ε and assume that f is an entire function in x and u and continuous in y/ε and ε. Let f and all its derivatives with respect to u and x be differentiable in $\lambda \in \Lambda$. Assume the existence of a standing wave $\bar{U} = \begin{pmatrix} \bar{u} \\ \bar{u}_t \end{pmatrix}$ of (2.3) satisfying (H1)–(H3), (2.10) and (2.11). Standing waves U_ε of (2.8) can then only exist in an exponentially small parameter interval, i.e. there exist positive constants c_1, c_2, ε_0, B and b and a function $\lambda_1(\varepsilon)$ such that if $(U_\varepsilon, \tilde{\lambda}_\varepsilon)$ is a standing wave of (2.8), for $0 < \varepsilon \leqslant \varepsilon_0$, that satisfies*

$$\|V_\varepsilon - \bar{U}\|_{BC^0(\mathbb{R}, X^1) \cap BC^1(\mathbb{R}, X)} \leqslant c_1,$$
$$|\tilde{\lambda}_\varepsilon - \lambda_1(\varepsilon)| \leqslant c_2,$$

then

$$|\tilde{\lambda}_\varepsilon - \lambda_1(\varepsilon)| \leqslant B \exp\left(-\frac{b}{\varepsilon^{1/2}}\right). \tag{2.12}$$

REMARK 2.6. The analysis applies directly to systems with the same diffusion coefficients for each component: with $u \in H^s(\Omega, \mathbb{R}^n)$ it is possible to analyse pinned waves of

$$u_t = d_0\left(\frac{x}{\varepsilon}\right)\partial_y^2 u + d\left(\frac{y}{\varepsilon}\right)\Delta u + f\left(u, x, \frac{y}{\varepsilon}; \lambda, \varepsilon\right).$$

It also applies to systems with diffusion coefficients that are independent of y/ε, but possibly different for each component:

$$u_t = \mathrm{diag}(d_0^1, \dots, d_0^n)\partial_y^2 u + \mathrm{diag}(d^1, \dots, d^n)\Delta u + f\left(u, x, \frac{y}{\varepsilon}; \lambda, \varepsilon\right).$$

3. Proof of theorem 2.5

We will use an implicit function argument starting with $\bar{u}$ to show the persistence of pinned solutions under bounded perturbations. In particular, we will set up a Lyapunov–Schmidt reduction to deal with the rank defect of $\mathcal{L}$. For this we use methods by Peterhof *et al.* [8] on exponential dichotomies for equations of the form

$$U_t = \bar{A}U + B(t)U. \tag{3.1}$$

To show the exponentially small pinning properties, we employ a homogenization method similar to a result in [6]. To apply both of these results, we will first transform (2.8) in order to remove the dependence on t/ε in the main part. This transformation will create a rapidly varying, unbounded perturbation, which is formally small and will be estimated later.

3.1. Time changes and linear transformations

The goal of this subsection is to make a change of variables that eliminates the 'fast' time dependence in the linear part of (2.8). By a linear transformation and time-rescaling we will move the fast variable in the main part of the equation,

$$U_t = A\left(\frac{t}{\varepsilon}\right)U, \tag{3.2}$$

to a higher order in ε.

First we rescale the time as

$$t = \tau + \varepsilon\Theta\left(\frac{\tau}{\varepsilon}\right)$$

with $\Theta(\cdot) = \Theta(\cdot + 1)$. We then transform the dependent variable according to

$$U = \Phi\left(\frac{\tau}{\varepsilon}\right)V = \begin{pmatrix} 1 & \varepsilon\phi\left(\dfrac{\tau}{\varepsilon}\right) \\ 0 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{3.3}$$

with $\phi(\cdot) = \phi(\cdot + 1)$. If we re-express equation (2.8) in the new variables we find that

$$V_\tau = \begin{pmatrix} v_2 + \left(\Theta'\left(\dfrac{\tau}{\varepsilon}\right) - \phi'\left(\dfrac{\tau}{\varepsilon}\right) \right) v_2 + \varepsilon \Lambda_1 v_1 + \varepsilon^2 \Lambda_2 v_2 \\[4mm] -\left(1 + \Theta'\left(\dfrac{\tau}{\varepsilon}\right) \right) d\left(\dfrac{\tau}{\varepsilon} + \Theta\left(\dfrac{\tau}{\varepsilon}\right) \right) \Delta_x v_1 - \varepsilon \Lambda_1 v_2 \end{pmatrix}$$
$$+ \left(1 + \Theta'\left(\dfrac{\tau}{\varepsilon}\right) \right) \Phi^{-1}\left(\dfrac{\tau}{\varepsilon}\right) F\left(\Phi\left(\dfrac{\tau}{\varepsilon}\right) V, \dfrac{\tau}{\varepsilon} + \Theta\left(\dfrac{\tau}{\varepsilon}\right); \lambda, \varepsilon \right), \quad (3.4)$$

where

$$\Lambda_1 v = \left(1 + \Theta'\left(\dfrac{\tau}{\varepsilon}\right) \right) d\left(\dfrac{\tau}{\varepsilon} + \Theta\left(\dfrac{\tau}{\varepsilon}\right) \right) \phi\left(\dfrac{\tau}{\varepsilon}\right) \Delta_x v$$

and

$$\Lambda_2 v = \left(1 + \Theta'\left(\dfrac{\tau}{\varepsilon}\right) \right) d\left(\dfrac{\tau}{\varepsilon} + \Theta\left(\dfrac{\tau}{\varepsilon}\right) \right) \left(\phi\left(\dfrac{\tau}{\varepsilon}\right) \right)^2 \Delta_x v.$$

The leading-order term in (3.4) depending on the 'fast time' τ/ε is then in the second component:

$$-d\left(\dfrac{\tau}{\varepsilon} + \Theta\left(\dfrac{\tau}{\varepsilon}\right) \right) \left(1 + \Theta'\left(\dfrac{\tau}{\varepsilon}\right) \right) \Delta_x v_1.$$

In order to eliminate the fast time dependence of this term we choose Θ to satisfy

$$-d(s + \Theta(s))(1 + \Theta'(s)) = -\bar{d}. \qquad (3.5)$$

The properties of Θ are given in the following lemma.

LEMMA 3.1. *There exist $\bar{d} > 0$ and a differentiable, periodic function Θ such that (3.5) is satisfied. Furthermore, one can choose Θ such that one also has*

$$\int_0^1 \Theta(\theta) \, d\theta = 0.$$

Proof. Let Θ be the solution of the ordinary differential equation

$$\left. \begin{aligned} \dot{\Theta}(s) &= \frac{\bar{d}}{d(s + \Theta(s))} - 1, \\ \Theta(s_0) &= 0. \end{aligned} \right\} \qquad (3.6)$$

The periodicity of Θ, i.e. $\Theta(\cdot) = \Theta(\cdot + 1)$, can be ensured by an appropriate choice of $\bar{d} > 0$. Indeed, the solution $\Theta(s_0 + 1)$ depends continuously on the parameter $\bar{d}$, and $\Theta(s_0 + 1) < \Theta(s_0)$ for $\bar{d} = 0$ and $\Theta(s_0 + 1) > \Theta(s_0)$ for $\bar{d}$ large. Hence, there exists a $\bar{d} > 0$, such that $\Theta(s_0 + 1) = \Theta(s_0)$, which implies that $\Theta(\cdot) = \Theta(\cdot + 1)$ by periodicity of the differential equation.

We conclude the proof by showing that, if we choose s_0 appropriately, we can also obtain

$$\int_0^1 \Theta(\theta) \, d\theta = 0.$$

We begin by integrating by parts:

$$I(s_0) = \int_{s_0}^{s_0+1} \Theta(s)\,\mathrm{d}s = -\int_{s_0}^{s_0+1} s\dot{\Theta}(s)\,\mathrm{d}s = \int_{s_0}^{s_0+1} s\left(\frac{\bar{d}}{d(s+\Theta(s))} - 1\right)\mathrm{d}s,$$

where the first equality uses the fact that $\Theta(s_0) = 0$ to eliminate the boundary terms. We now use the identity

$$\bar{d}^{-1} = \int_{s_0}^{s_0+1} \frac{1}{d(s+\Theta(s))}\,\mathrm{d}s = \int_0^1 \frac{1}{d(s+\Theta(s))}\,\mathrm{d}s$$

to rewrite the last integral as

$$\bar{d}\int_{s_0}^{s_0+1} s\left(\frac{1}{d(s+\Theta(s))} - \int_{s_0}^{s_0+1} \frac{1}{d(\theta+\Theta(\theta))}\,\mathrm{d}\theta\right)\mathrm{d}s$$

$$= \bar{d}\int_{s_0}^{s_0+1} s\left(\frac{1}{d(s+\Theta(s))} - \int_0^1 \frac{1}{d(\theta+\Theta(\theta))}\,\mathrm{d}\theta\right)\mathrm{d}s$$

$$= \bar{d}\int_0^1 (s+s_0)\left(\frac{1}{d(s+s_0+\Theta(s+s_0))} - \int_0^1 \frac{1}{d(\theta+\Theta(\theta))}\,\mathrm{d}\theta\right)\mathrm{d}s$$

$$= \bar{d}\int_0^1 s\left(\frac{1}{d(s+s_0+\Theta(s+s_0))} - \int_0^1 \frac{1}{d(\theta+\Theta(\theta))}\,\mathrm{d}\theta\right)\mathrm{d}s.$$

Now integrate $I(s_0)$ with respect to s_0:

$$\int_0^1 I(s_0)\,\mathrm{d}s_0 = \bar{d}\int_0^1 s\int_0^1 \left(\frac{1}{d(s+s_0+\Theta(s+s_0))} - \int_0^1 \frac{1}{d(\theta+\Theta(\theta))}\,\mathrm{d}\theta\right)\mathrm{d}s_0\,\mathrm{d}s$$

$$= 0.$$

The last equality follows from the periodicity of d and Θ. Hence, there is a choice of s_0 such that $I(s_0) = 0$ and, consequently, the corresponding solution of (3.6) has

$$\int_0^1 \Theta(s)\,\mathrm{d}s = 0.$$

$\square$

REMARK 3.2. Once we know of the existence of a function Θ satisfying the properties described in the lemma, it is easy to compute its value. If we integrate both sides of (3.5) from 0 to 1 and make the change of variables $\sigma = s + \Theta(s)$, we have

$$\bar{d} = \int_0^1 d(s+\Theta(s))(1+\Theta'(s))\,\mathrm{d}s = \int_0^1 d(\sigma)\,\mathrm{d}\sigma. \qquad (3.7)$$

Thus, $\bar{d}$ is the average of $d(\sigma)$ over one period of the inhomogeneity. Recalling the definition of d in terms of the coefficients of the original equation (1.1), we see that

$$\bar{d} = \int_0^1 \frac{d_1(\sigma)}{d_0(\sigma)}\,\mathrm{d}\sigma. \qquad (3.8)$$

Having fixed Θ in this way, we next eliminate the term

$$\left(\Theta'\left(\frac{\tau}{\varepsilon}\right) - \phi'\left(\frac{\tau}{\varepsilon}\right)\right)v_2$$

in the first component by choosing $\phi(s) = \Theta(s)$.

If V is called U again, we see that, in the new variables, (2.8) takes the form

$$U_\tau = \bar{A}U + B\left(\frac{\tau}{\varepsilon}\right)U + \bar{F}(U; \lambda, \varepsilon) + \tilde{F}\left(U, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right) \tag{3.9}$$

with $\bar{A}$ as in (2.7) and by using (3.6),

$$B\left(\frac{\tau}{\varepsilon}; \varepsilon\right)U = \varepsilon \begin{pmatrix} \Theta\left(\dfrac{\tau}{\varepsilon}\right)\bar{d}\Delta_x u_1 + \varepsilon\Theta^2\left(\dfrac{\tau}{\varepsilon}\right)\bar{d}\Delta_x u_2 \\[2mm] -\Theta\left(\dfrac{\tau}{\varepsilon}\right)\bar{d}\Delta_x u_2 \end{pmatrix}$$

and

$$\tilde{F}\left(U, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right)(x)$$

$$= \left(1 + \Theta'\left(\frac{\tau}{\varepsilon}\right)\right)F\left(\Phi\left(\frac{\tau}{\varepsilon}\right)U, \frac{\tau}{\varepsilon} + \Theta\left(\frac{\tau}{\varepsilon}\right); \lambda, \varepsilon\right)$$

$$- \begin{pmatrix} \varepsilon\Theta\left(\dfrac{\tau}{\varepsilon}\right)f\left(u_1 + \varepsilon\phi\left(\dfrac{\tau}{\varepsilon}\right)u_2, \dfrac{\tau}{\varepsilon} + \Theta(\tau + \varepsilon); \lambda, \varepsilon\right) \\[2mm] 0 \end{pmatrix} - \bar{F}(U; \lambda, \varepsilon).$$

The averages over a period are denoted by

$$\langle g(\cdot)\rangle = \int_0^1 g(\theta)\, \mathrm{d}\theta$$

and can be computed for B using our choice of Θ:

$$\langle B(\cdot)U\rangle = \varepsilon^2 \begin{pmatrix} \displaystyle\int_0^1 \Theta^2(\theta)\, \mathrm{d}\theta\, \bar{d}\Delta_x u_2 \\[2mm] 0 \end{pmatrix}. \tag{3.10}$$

3.2. Persistence

The transformations of the previous section result in a homogenized equation, (2.7), perturbed by rapidly varying terms $B(\tau/\varepsilon)$ and $\tilde{F}(U, \tau/\varepsilon; \lambda, \varepsilon)$. In §4 we will apply an averaging method to show that the effects of $\tilde{F}$ can be made exponentially small. With this in mind, the present section is devoted to studying the persistence of standing waves in the averaged equation when it is subjected to small perturbations. Using a Lyapunov–Schmidt argument based on the theory of exponential dichotomies (see [8]), we show that, under the non-degeneracy and transversality hypotheses we stated in §1, standing waves in the averaged model

persist under small, bounded, homogeneous and heterogeneous perturbations. Furthermore, we derive some necessary conditions for the persistence of standing waves under unbounded, but formally small, perturbations like the term $B(\tau/\varepsilon)U$ in (3.9).

Consider an equation of the form

$$U_\tau = \bar{A}U + \bar{F}(U; \lambda, \varepsilon) + \mathcal{P}\left(U, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right) \tag{3.11}$$

(we suppress the dependence of $\bar{F}$ and $\mathcal{P}$ on x, since it plays little role in this section).

Our principal interest is in the case considered in the previous subsection where $\mathcal{P} = BU + \tilde{F}$, but it will be of use in the following sections to consider this slightly more general case.

The persistence proof is based the functional analytic framework of exponential dichotomies as in [8], which we follow very closely wherever possible. To prove the existence of a pinned wave U_ε, we again introduce a new variable V by letting $U(\tau) = \bar{U}(\tau) + V(\tau)$. Using the fact that $\bar{U}$ satisfies (2.7) with $\lambda = \lambda_0$ and $\varepsilon = 0$ we find that

$$V_\tau = \bar{A}V + D\bar{F}(\bar{U}(\tau); \lambda_0, 0)V + G\left(V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right) \tag{3.12}$$

with

$$G\left(V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right) = \bar{F}(\bar{U} + V; \lambda, \varepsilon) - \bar{F}(\bar{U}; \lambda_0, 0) - D\bar{F}(\bar{U}; \lambda_0, 0)V + \mathcal{P}\left(\bar{U} + V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right).$$

Again, with reference to the material of the previous section note that, for an equation of the form (3.9), G will have the more explicit form:

$$G\left(V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right) = \bar{F}(\bar{U} + V; \lambda, \varepsilon) + \tilde{F}\left(\bar{U} + V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right)$$

$$+ \varepsilon B\left(\frac{\tau}{\varepsilon}\right)(\bar{U} + V) - \bar{F}(\bar{U}; \lambda_0, 0) - D\bar{F}(\bar{U}; \lambda_0, 0)V. \tag{3.13}$$

Next we will show that the linear part of (3.12) has an exponential dichotomy on $\mathbb{R}^+$ and $\mathbb{R}^-$. First, we consider the linearized equations around equilibria

$$U^i = \begin{pmatrix} u^i \\ 0 \end{pmatrix}, \quad i = 1, 2,$$

at $\lambda = \lambda_0$ and $\varepsilon = 0$:

$$U_t = \bar{A}U + D\bar{F}(U^i; \lambda_0, 0)U.$$

The respective projections on stable and unstable subspaces are given by

$$P_{i,\pm} = \frac{1}{2} \begin{pmatrix} I & \pm(L_i)^{-1/2} \\ \pm(L_i)^{1/2} & I \end{pmatrix}, \tag{3.14}$$

where $L_i = -\bar{d}\Delta_x - D\bar{f}(u^i, x; \lambda_0, 0)$ is positive and self-adjoint. For the existence of these projections, the assumption (H3) that $-\bar{d}\Delta_x - D\bar{f}(u^i, x; \lambda_0, 0)$ is a positive operator in (2.3) is crucial, as it implies that the equilibria U^1 and U^2 are hyperbolic

equilibria of the spatial dynamics equation. Using [8, theorem 1 and corollary 1], we see that the linearization along the standing wave,

$$U_\tau(\tau) = \bar{A}U(\tau) + D\bar{F}(\bar{U}(\tau); \lambda_0, 0)U(\tau), \qquad (3.15)$$

has exponential dichotomies on $\mathbb{R}^+$ and $\mathbb{R}^-$ denoted by $P(\tau)$ and $Q(\tau)$, respectively. The intersection $\mathrm{Rg}(P(0)) \cap \mathrm{Rg}(Q(0))$ of solutions bounded in forward and backward time is then one dimensional by assumption (2.10).

The solution operators of the exponential dichotomies are denoted as

$$\phi^s_{1,2}(\tau, \theta)z, \quad \tau \geqslant \theta \text{ for the stable part,}$$
$$\phi^u_{1,2}(\tau, \theta)z, \quad \tau \leqslant \theta \text{ for the unstable part.}$$

As in [8, theorem 4], we decompose

$$\mathrm{Rg}(\phi^s_1(0,0)) = Y_1 \oplus \mathrm{span}(\dot{U}(0)),$$
$$\mathrm{Rg}(\phi^u_2(0,0)) = Y_2 \oplus \mathrm{span}(\dot{U}(0)).$$

We are looking for solutions of the full perturbed equation (3.12) that are bounded on $\mathbb{R}^+$ and $\mathbb{R}^-$, respectively. This is equivalent to the existence of $\xi_1 \in Y_1$ and $\xi_2 \in Y_2$, such that

$$V_1(\tau) = \phi^s_1(\tau, 0)\xi_1 + \int_0^\tau \phi^s_1(\tau, \theta)G\left(V_1(\theta), \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta$$
$$- \int_\tau^\infty \phi^u_1(\tau, \theta)G\left(V_1(\theta), \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta \quad \text{for } \tau > 0,$$

$$V_2(\tau) = \phi^u_2(\tau, 0)\xi_2 + \int_0^\tau \phi^u_2(\tau, \theta)G\left(V_2(\theta), \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta$$
$$+ \int_{-\infty}^\tau \phi^s_2(\tau, \theta)G\left(V_2(\theta), \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta \quad \text{for } \tau < 0.$$

Hence, for all $\xi_1 \in Y_1$ and $\xi_2 \in Y_2$ small, we get bounded solutions $V_1(\tau; \xi_1, \lambda, \varepsilon)$ and $V_2(\tau; \xi_2, \lambda, \varepsilon)$ by the implicit function theorem. For small $\varepsilon > 0$, we now try to find $\xi = \xi_1 + \xi_2 \in Y_1 \oplus Y_2$ and λ such that $V_1(0; \xi, \lambda, \varepsilon) = V_2(0; \xi, \lambda, \varepsilon)$. For bounded perturbations G, this holds if and only if there is a solution to

$$(\phi^s_1(0,0) - \phi^u_2(0,0))\xi = \int_{-\infty}^0 \phi^s_2(0,\theta)G\left(V_2(\theta), \theta, \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta$$
$$+ \int_0^\infty \phi^u_1(0,\theta)G\left(V_1(\theta), \theta, \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta. \quad (3.16)$$

On the other hand, if G is unbounded and we have a bounded solution $V(\tau) \in X$, such that $G(V(\tau), \tau, \tau/\varepsilon; \lambda, \varepsilon)$ is bounded in X, then we can also write

$$V_1(\tau) = \phi^s_1(\tau, 0)\xi_1 + \int_0^\tau \phi^s_1(\tau, \theta)G\left(V_1(\theta), \theta, \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta$$
$$- \int_\tau^\infty \phi^u_1(\tau, \theta)G\left(V_1(\theta), \theta, \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta \quad \text{for } \tau > 0,$$

$$V_2(\tau) = \phi_2^u(\tau, 0)\xi_1 + \int_0^\tau \phi_2^u(\tau, \theta)G\left(V_2(\theta), \theta, \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta$$

$$+ \int_{-\infty}^\tau \phi_2^s(\tau, \theta)G\left(V_2(\theta), \theta, \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta \quad \text{for } \tau < 0,$$

with $V_1(0) = V_2(0)$. Thus, we can proceed as for bounded G to obtain (3.16), noting that the results are necessary conditions only.

Next we show that the left-hand side, $L = \phi_1^s(0, 0) - \phi_2^u(0, 0)$, of (3.16) is a Fredholm operator of index 0. By [8, theorem 1],

$$\phi_1^s(0, 0) = P_{1,+} + P_{1,-}(S_1 + K_1)$$

and

$$\phi_2^u(0, 0) = P_{2,-} + P_{2,+}(S_2 + K_2),$$

with S_1 and S_2 arbitrarily small and K_1 and K_2 compact in $L(X)$. Thus,

$$P(0) - Q(0) = \phi_1^s(0, 0) - \phi_2^u(0, 0)$$
$$= P_{1,+} - P_{1,-} + (P_{1,-} - P_{2,-}) + P_{1,-}(S_1 + K_1) - P_{2,+}(S_2 + K_2)$$

is a Fredholm operator of index 0, since $P_{1,+} - P_{1,-}$ is invertible in $L(X)$ and thus Fredholm of index 0. This persists under the small perturbations $P_{1,-}S_1 - P_{2,+}S_2$ and the compact perturbations $P_{1,-}K_1 - P_{2,+}K_2$, as well as under the compact perturbation $P_{1,-} - P_{2,-}$. Its compactness follows from a direct calculation using (3.14) and the Sobolev embedding theorem.

The null space is $\text{Ker}(L) = \text{span}(\dot{U}(0))$ by (2.10). Furthermore $\text{Rg}(L) = \{\eta \in X, \langle \psi(0), \eta \rangle = 0\}$, where ψ is the unique bounded solution of the adjoint of (3.15), up to a multiple. Thus, the Lyapunov–Schmidt reduction method implies that if G is bounded, and (3.16) has a solution $(\xi(\varepsilon), \lambda(\varepsilon))$ for $\varepsilon > 0$, then

$$\Gamma(V; \lambda, \varepsilon) = \left(\psi(0), \left\{\int_{-\infty}^0 \phi_2^s(0, \theta)G\left(V_2(\theta; \xi(\varepsilon)), \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta\right.\right.$$
$$\left.\left. + \int_0^\infty \phi_1^u(0, \theta)G\left(V_1(\theta; \xi(\varepsilon)), \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right) d\theta\right\}\right) = 0. \quad (3.17)$$

Here, the notation $V_{1,2}(\theta; \xi(\varepsilon))$ is used simply to remind the reader of the dependence of $V_{1,2}$ on the initial condition ξ, and we define $V : \mathbb{R} \to X$ to be the function that is equal to $V_2(t)$ for t negative and $V_1(t)$ for t positive.

REMARK 3.3. Note that by applying the inner product inside the integrals in (3.17) and taking the adjoint of $\phi_{1,2}^{u,s}$, this expression can be rewritten as

$$\Gamma(V; \lambda, \varepsilon) = \int_{-\infty}^\infty \left(\psi(\theta), G\left(V(\theta; \xi(\varepsilon)), \frac{\theta}{\varepsilon}; \lambda, \varepsilon\right)\right) d\theta = 0. \quad (3.18)$$

We assume that Γ is a continuous function of ε for $\varepsilon > 0$. Suppose now that

$$\lim_{\varepsilon \to 0} \Gamma(0; \lambda_0, \varepsilon) \equiv \Gamma(0; \lambda_0, 0) = 0. \quad (3.19)$$

Suppose furthermore that

$$D_\lambda \Gamma(0; \lambda, 0)|_{\lambda=\lambda_0} \equiv \lim_{\varepsilon \to 0^+} D_\lambda \Gamma(0; \lambda, \varepsilon)|_{\lambda=\lambda_0} \neq 0. \qquad (3.20)$$

The Lyapunov–Schmidt method then guarantees that, for ε sufficiently small and positive, there exists $(V_\varepsilon(\tau), \lambda_\varepsilon)$ such that (3.16) holds.

Thus, we can make the following proposition.

PROPOSITION 3.4. *Assume that $\bar{U}$ is a standing wave of (2.7) satisfying (H1)–(H3), (2.10) and (2.11). Let Γ be defined by (3.17). If (3.19) and (3.20) are satisfied, there exists $\varepsilon_0 > 0$ such that for every $0 < \varepsilon < \varepsilon_0$ there exists $\lambda_\varepsilon \in \Lambda$ and $V_\varepsilon \in X$ such that $\bar{U} + V_\varepsilon$ is a standing wave solution of (3.11) with $\lambda = \lambda_\varepsilon$. Furthermore, we can make $|V_\varepsilon|_X$ and $|\lambda_0 - \lambda_\varepsilon|$ arbitrarily small by choosing ε_0 small enough.*

Proof. The proof of existence was discussed just prior to the statement of the above proposition. The fact that $|V_\varepsilon|_X$ and $|\lambda_0 - \lambda_\varepsilon|$ can be made arbitrarily small follows from the fact that Γ is smooth with respect to V and λ and continuous with respect to ε. $\qquad \square$

REMARK 3.5. The solution $\bar{U}(\tau) + V_\varepsilon(\tau)$ constructed in proposition 3.4 will actually lie not just in X but in $\mathcal{G}_\sigma$ for each τ by the results of [6].

REMARK 3.6. The same analysis holds for $U_{\varepsilon,\beta}$ near the shifted $\bar{U}(\cdot + \beta)$. Then, typically,

$$\int_{-\infty}^{\infty} \left(\psi(\tau), D_\beta G\left(U_{\varepsilon,\beta}(\theta), \theta, \frac{\theta}{\varepsilon}; \lambda_\varepsilon, \varepsilon \right) \right) d\theta \neq 0,$$

and the standing wave persists under changing λ. However, as we demonstrate in the next section, it can persist over only a very small range of λ.

3.3. Exponentially small pinning

To estimate the parameter interval in which pinning can occur, we again consider equation (3.9). If $B + \tilde{F} + \bar{F}$ were bounded on bounded sets and analytic on both X and Gevrey spaces $\mathcal{G}_\sigma$, we could apply the results of [6, theorem 2] to equation (3.9) in order to construct a time-periodic transformation of the phase space X, such that the transformed equation would have the form

$$V_\tau = AV + \bar{F}_{\exp}(V; \lambda, \varepsilon) + R\left(V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon \right) \qquad (3.21)$$

with an estimate on the remainder

$$\left| R\left(V(\tau), \frac{\tau}{\varepsilon}; \lambda, \varepsilon \right) \right|_X \leqslant C \exp\left(-\frac{c}{\varepsilon^{1/2}} \right).$$

Unfortunately, the unboundedness of the term B prevents us from applying these results directly. However, if we take advantage of the *a priori* information we have about the smoothness of solutions of (3.9), we can obtain results very similar to those of [6], which will be sufficient for our needs.

PROPOSITION 3.7. *Fix $r > 0$. Under the assumptions of theorem 2.5, there exist $\varepsilon_0, \sigma_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$ there is a change of coordinates $U = \Phi(V) = V + \varepsilon W(V, \tau/\varepsilon; \lambda, \varepsilon)$ which is analytic and invertible on the ball of radius r in both X and the Gevrey space $\mathcal{G}_\sigma$, $0 \leqslant \sigma < \sigma_0$. If U evolves according to (2.8), then $V = \Phi^{-1}(U)$ satisfies*

$$V_\tau = \bar{A}V + \bar{F}_{\exp}(V; \lambda, \varepsilon) + R\left(V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right). \tag{3.22}$$

Furthermore, there exist $C_r, c > 0$ (depending on σ) such that, for every V with $|V|_{\mathcal{G}_\sigma} < r$ and $\lambda \in \Lambda$,

$$\left| R\left(V, \frac{\tau}{\varepsilon}; \lambda, \varepsilon\right) \right|_X \leqslant C_r \exp\left(-\frac{c}{\varepsilon^{1/2}}\right). \tag{3.23}$$

The homogenized term, $\bar{F}_{\exp}$ is close to $\bar{F}$ in the sense that there exists $C_F(\sigma) > 0$ such that, for every V with $|V|_{\mathcal{G}_\sigma} < r$,

$$|\bar{F}(U; \lambda, \varepsilon) - \bar{F}_{\exp}(U; \lambda, \varepsilon)|_X \leqslant C_F(\sigma)\varepsilon^{1/2}, \tag{3.24}$$

$$|D_\lambda \bar{F}(U; \lambda, \varepsilon) - D_\lambda \bar{F}_{\exp}(U; \lambda, \varepsilon)|_X \leqslant C_F(\sigma)\varepsilon^{1/2}. \tag{3.25}$$

Finally, the transformation, Φ, is near to the identity in the sense that W is uniformly bounded on the ball of radius r in either X or $\mathcal{G}_\sigma$.

REMARK 3.8. Note that the estimates (3.23) and (3.24) are somewhat unusual in that they require the argument of the function to lie in the Gevrey space $\mathcal{G}_\sigma$, while giving an estimate of the functions only in the much weaker X-norm. Thus, they cannot be used effectively to bound the evolution of (3.22), for example. However, if one has some additional, *a priori* estimates of the solution (as one does in the present instance), such estimates can nonetheless be quite useful.

The proposition is proved in the next section. We complete the proof of the main theorem first.

Proof of theorem 2.5. Begin by noting that, if we set $\mathcal{P} = F_{\exp} - \bar{F}$ in (3.11), then, since we know that $\bar{U} \in \mathcal{G}_\sigma$, (3.24) implies that (3.19) holds, while (2.11) ensures that (3.20) is satisfied. Thus, proposition 3.4 implies that there exists $\varepsilon_0 > 0$ such that, for every $0 < \varepsilon < \varepsilon_0$, there exists a standing wave $(\bar{U}_\varepsilon, \lambda_1(\varepsilon))$ for

$$U_t = \bar{A}U + \bar{F}_{\exp}(U; \lambda, \varepsilon) \tag{3.26}$$

close to $(\bar{U}, \lambda_0)$. Let $\tilde{U}_\varepsilon \in \mathcal{G}_\sigma, \tilde{\lambda}_\varepsilon$ be another standing wave solution of (2.8) close to $\bar{U}$, i.e. $|\tilde{U}_\varepsilon - \bar{U}|_X < c_1$, where c_1 will be specified below. If $\tilde{V}_\varepsilon \equiv \tilde{U}_\varepsilon - \bar{U}$, and G is given by (3.13), then (3.17) implies that

$$\int_{\mathbb{R}} \left(\Psi(\tau), G\left(\tilde{V}_\varepsilon(\tau), \frac{\tau}{\varepsilon}; \tilde{\lambda}_\varepsilon, \varepsilon \right) \right) d\tau = 0.$$

Applying the transformation Φ to $\tilde{U}_\varepsilon$ (and still denoting the transformed variable by $\tilde{U}_\varepsilon$), we see that

$$\tilde{U}_{\varepsilon,\tau} = \bar{A}\tilde{U}_\varepsilon + \bar{F}(\tilde{U}_\varepsilon; \tilde{\lambda}_\varepsilon, \varepsilon) + \{\bar{F}_{\exp}(\tilde{U}_\varepsilon; \tilde{\lambda}_\varepsilon, \varepsilon) - \bar{F}(\tilde{U}_\varepsilon; \tilde{\lambda}_\varepsilon, \varepsilon) + R(\tilde{U}_\varepsilon; \tilde{\lambda}_\varepsilon, \varepsilon)\}.$$

Then (3.19) implies that

$$I_1(\tilde{V}_\varepsilon, \tilde{\lambda}_\varepsilon, \varepsilon) + I_2(\tilde{V}_\varepsilon, \tilde{\lambda}_\varepsilon, \varepsilon) = 0$$

with

$$I_1(\tilde{V}_\varepsilon, \lambda, \varepsilon) = \int_{\mathbb{R}} \left(\Psi(\theta), \bar{F}_{\exp}\left(\tilde{U}_\varepsilon(\theta), \frac{\theta}{\varepsilon}; \lambda, \varepsilon \right) \right.$$

$$\left. - \bar{F}(\bar{U}(\theta); \lambda_0, 0) - D\bar{F}(\bar{U}(\theta); \lambda_0, 0)\tilde{V}_\varepsilon \right) d\theta,$$

$$I_2(\tilde{V}_\varepsilon, \lambda, \varepsilon) = \int_{\mathbb{R}} \left(\Psi(\theta), R\left(\tilde{U}_\varepsilon(\theta), \frac{\theta}{\varepsilon}; \lambda, \varepsilon \right) \right) d\theta.$$

Recalling that all bounded solutions of (2.2) lie in $\mathcal{G}_\sigma$ for all t, we see that there exist positive constants C_2 and c_2 such that, for any bounded solution $\tilde{U}_\varepsilon$ of (2.2), any $0 < \varepsilon < \varepsilon_0$ and any $\lambda \in \Lambda$, we have

$$I_2(\tilde{V}_\varepsilon, \lambda, \varepsilon) \leqslant C_2 \exp\left(-\frac{c_2}{\varepsilon^{1/2}} \right).$$

On the other hand, since $\bar{U}_\varepsilon$ is a solution of (3.26), $I_1(\bar{V}_\varepsilon; \lambda_1(\varepsilon), \varepsilon) = 0$, where $\bar{V}_\varepsilon = \bar{U}_\varepsilon - \bar{U}$. But, by (2.11) and (3.25),

$$\left| \left. \left(\int_{\mathbb{R}} \left(\Psi(\theta), D_\lambda \bar{F}_{\exp}\left(U(\theta), \frac{\theta}{\varepsilon}; \lambda, 0 \right) \right) \right|_{\lambda=\lambda_0} \right) d\theta \right| > C \qquad (3.27)$$

for all U close to $\bar{U}$. We now choose the constant c_1 which bounds $|\tilde{U}_\varepsilon - \bar{U}|_X$ small enough that $\tilde{U}_\varepsilon$ lies in the neighbourhood where (3.27) holds. Then, there exists $C_{\exp}$ such that for $|\lambda - \lambda_1(\varepsilon)| \geqslant C_{\exp} \exp(-c_2/\varepsilon^{1/2})$ we obtain $|I_1(V, \lambda, \varepsilon)| > |I_2(V, \lambda, \varepsilon)|$. Hence, in order to have $I_1 + I_2 = 0$ we must have

$$|\lambda - \lambda_1(\varepsilon)| < C_{\exp} \exp\left(-\frac{c_2}{\varepsilon^{1/2}} \right),$$

which completes the proof of theorem 2.5. $\qquad \qquad \square$

4. Exponential homogenization

The proof of proposition 3.7 is based on a combination of Galerkin approximation, an iterative averaging procedure and the use of Gevrey class regularity. A similar problem was treated in [5, 6], but the presence of the unbounded term $\varepsilon B(\tau/\varepsilon)V$ in our problem causes additional difficulties in the construction of the averaging transformation.

More precisely, we proceed as follows.

We begin in step 1 by considering (2.7) restricted to a finite-dimensional subspace of X. (The dimension of this subspace will be chosen in an ε-dependent fashion.) For this finite-dimensional system we show that we can make a change of variables (an 'averaging' transformation) such that, in terms of the new variables, the dependence on the 'fast' time τ/ε occurs only in terms that are exponentially small in ε.

In step 2 we then relate the solutions of the finite-dimensional model to the evolution of the full system (3.9). We prove that if the standing wave solution of

the averaged system is very smooth, then the effects of the 'fast' time dependence are also exponentially small in the full system.

REMARK 4.1. Throughout the proof we suppress the dependence on the parameter λ, because the transformations are smooth in λ and the estimates on the averaging transformations are all uniform in λ. The closeness of the λ-derivatives of $\bar{F}_{\exp}$ and $\bar{F}$ is a direct consequence of this. For similar reasons we also suppress the dependence of the various functions on x.

STEP 1 (finite-dimensional transformation). As in [6] we use an ε-dependent Galerkin approximation with approximation space

$$H_N = \mathrm{span}\left\{ \begin{pmatrix} \exp(\mathrm{i}k \cdot x) \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \exp(\mathrm{i}k \cdot x) \end{pmatrix} \text{ with } k \in \mathbb{Z}^d,\ |k|^2 = \sum_{j=1}^{d} k_i^2 \leqslant N^2 \right\}$$

with orthogonal projection P_N in X, note that $\|P^N \bar{A}\| = N$ on X and the Gevrey spaces $\mathcal{G}_\sigma$. Then, noting that $\bar{A}P^N = P^N \bar{A}$ and $BP^N = P^N B$, we define the evolution of $U_N = P^N U$ by

$$\dot{U}_N = \bar{A}U_N + \varepsilon B\left(\frac{\tau}{\varepsilon}\right)U_N + P^N \bar{F}(U_N; \varepsilon) + P^N \tilde{F}\left(U_N, \frac{\tau}{\varepsilon}; \varepsilon\right), \qquad (4.1)$$

where we choose N depending on ε through the formula

$$N(\varepsilon) = [\varepsilon^{-1/2}]. \qquad (4.2)$$

We now transform (4.1) by adapting the proof of Neishtadt's theorem [7] about the averaging of exponential order for finite-dimensional ODEs. We make successive coordinate changes, such that the non-autonomous terms are formally of higher order in ε in the transformed equation:

$$\dot{V}_N = \bar{A}V_N + P^N \bar{F}(V_N; \varepsilon) + \bar{F}_j(V_N; \varepsilon) + R_j\left(V_N, \frac{\tau}{\varepsilon}; \varepsilon\right). \qquad (4.3)$$

We will then give estimates on the transformed equation that are uniform in ε and $N(\varepsilon)$. In the transformed equation, the non-autonomous terms are exponentially small in ε.

   We describe the formal coordinate changes needed to remove non-autonomous terms. For a moment we suppress the dependence of V on N to simplify the notation. We work on the complex extension of the projected ball in $\mathcal{G}_\sigma$. More precisely, if $B_{\mathcal{G}_\sigma}$ is the ball of radius $\mathcal{R}$ in $\mathcal{G}_\sigma$, we define $D_{N,\sigma} = P^N(B_{\mathcal{G}_\sigma})$, and

$$\mathcal{N}_\delta^\sigma(D_{N,\sigma}) := \left\{ U \in H_N \ \Big|\ \inf_{V \in D_{N,\sigma}} |U - V|_{\mathcal{G}_\sigma} < \delta \right\}. \qquad (4.4)$$

REMARK 4.2. We could choose the radius, $\mathcal{R}$, of the ball on which we define our changes of variables to be as large as we want, though obviously the constants that appear in the proof would depend on $\mathcal{R}$. For convenience, we fix $\mathcal{R}$ once and for all to be twice the max of $\bar{U}(t)$ for $\bar{U}$, the standing wave solution of (2.7).

We define the supremum norm on these neighbourhoods by

$$\|f\|_{\mathcal{N}_\delta^\sigma(D)} = \sup_{U \in \mathcal{N}_\delta^\sigma(D)} |f(U)|_{\mathcal{G}_\sigma}$$

Since σ will be fixed throughout this section we will typically denote this norm in the much simpler fashion:

$$\|f\|_{\mathcal{N}_\delta^\sigma(D)} = \|f\|_\delta,$$

if there is no possibility of confusion.

We construct the averaging transformation inductively and at each step of the induction we shrink the size of the complex neighbourhood slightly. There will be r steps in our induction (r will be fixed below) and we define a sequence of constants $\delta = \delta_0 > \delta_1 > \delta_2 > \cdots > \delta_r > \frac{1}{2}\delta$.

After j inductive steps, the equations will have the form

$$\dot{U} = AU + P^N \bar{F}(U;\varepsilon) + \bar{F}_j(U;\varepsilon) + R_j\left(U, \frac{t}{\varepsilon}; \varepsilon\right). \tag{4.5}$$

We will estimate the size of the various terms on the complex domain $\mathcal{N}_{\delta_j}^\sigma(D_{N,\sigma})$. The 'remainder' term R_j will be chosen to have zero average,

$$\langle R_j \rangle(U, \varepsilon) = \int_0^1 R_j(U, \theta, \varepsilon) \, d\theta = 0,$$

and to start the induction we will take

$$\bar{F}_0 = \langle P^N \tilde{F} + P^N B \rangle \quad \text{and} \quad R_0 = (P^N \tilde{F} + P^N B) - \langle P^N \tilde{F} + P^N B \rangle.$$

Starting with (4.5), we consider a coordinate change of the form

$$U = V + \varepsilon W_j\left(V, \frac{t}{\varepsilon}; \varepsilon\right) \tag{4.6}$$

with $W(\cdot, t/\varepsilon; \varepsilon) : P^N X \to P^N X$, periodic in t/ε with period 1. Substitution into (4.5) yields

$$\dot{V} + \varepsilon \frac{\partial}{\partial V} W_j(V, \tau; \varepsilon) \dot{V} + \frac{\partial}{\partial \tau} W_j(V, \tau; \varepsilon)$$
$$= \bar{A}(V + \varepsilon W_j(V, \tau; \varepsilon)) + P^N \bar{F}(V + \varepsilon W_j(V, \tau; \varepsilon); \varepsilon)$$
$$+ \bar{F}_j(V + \varepsilon W_j(V, \tau; \varepsilon); \varepsilon) + R_j(V + \varepsilon W_j(V, \tau; \varepsilon), \tau; \varepsilon).$$

We will choose W_j to 'kill' the (formally) largest term that depends on the 'fast' time. In the present case, this is R_j, and to remove it we pick

$$W_j(V, \tau; \varepsilon) = \int_0^\tau R_j(V, \theta; \varepsilon) \, d\theta \tag{4.7}$$

and obtain

$$\dot{V} = \bar{A}V + P^N \bar{F}(V; \varepsilon) + \bar{F}_j(V; \varepsilon) + a\left(V, \frac{\tau}{\varepsilon}, \varepsilon\right), \tag{4.8}$$

where

$$a(V, \theta; \varepsilon) = \left(1 + \varepsilon \frac{\partial}{\partial V} W_j\right)^{-1}$$
$$\times \{\varepsilon P^N \bar{A} W_j + P^N (\bar{F}(V + \varepsilon W_j; \varepsilon) - \bar{F}(V; \varepsilon))$$
$$+ P^N (\bar{F}_j(V + \varepsilon W_j; \varepsilon) - \bar{F}_j(V; \varepsilon))$$
$$+ P^N (R_j(V + \varepsilon W_j, \theta; \varepsilon) - R_j(V, \theta; \varepsilon))\}$$
$$+ \{(\mathbf{1} + \varepsilon D_V W_j)^{-1} - \mathbf{1}\}(\bar{A}V + P^N \bar{F}(V; \varepsilon) + \bar{F}_j(V; \varepsilon)). \quad (4.9)$$

(In this expression we have suppressed the dependence of many of the functions on variables like ε and λ to prevent the formula from becoming too unwieldy.) We then choose

$$\bar{F}_{j+1}(V, \varepsilon) = \bar{F}_j(V) + \langle a(V, \cdot, \varepsilon)\rangle, \qquad R_{j+1}(V, \tau, \varepsilon) = a(V, \tau, \varepsilon) - \langle a(V, \cdot, \varepsilon)\rangle \quad (4.10)$$

and we see that (4.8) has the form needed for our inductive argument.

We now begin the task of rigorously estimating the various terms that result from the above procedure. We begin by bounding the terms $\bar{F}_0$ and $\bar{R}_0$ that appear in our original equation (4.1).

By construction we have $\|\bar{F}_0(\cdot; \varepsilon)\|_\delta \leqslant \|\langle P^N \tilde{F}(\cdot, \varepsilon)\rangle\|_\delta + \|\langle P^N B \cdot\rangle\|_\delta$ With the aid of (3.10) the second term on the right-hand side can be bounded by $C\varepsilon^2 N^2$. To bound the first term we note that $\langle P^N \tilde{F}(\cdot, \varepsilon)\rangle = P^N \langle \tilde{F}(\cdot, \varepsilon)\rangle$ and write

$$\langle \tilde{F}\rangle = \left\langle (1 + \Theta'(\theta)) \begin{pmatrix} 0 \\ -f(v_1 + \varepsilon\phi(\cdot)v_2, \cdot + \Theta(\cdot); \varepsilon) \end{pmatrix}\right\rangle$$
$$- \left\langle \begin{pmatrix} 0 \\ -f(v_1, \cdot; \varepsilon) \end{pmatrix}\right\rangle - \left\langle \begin{pmatrix} \varepsilon\Theta(\cdot)f(v_1 + \varepsilon\phi(\cdot)v_2, \cdot + \Theta(\cdot); \varepsilon) \\ 0 \end{pmatrix}\right\rangle.$$

The smoothness of f allows us to bound the $\|\cdot\|_\delta$ norm of the last term by $C\varepsilon$. In order to bound the difference of the first two terms, note that, by the change-of-variables formula,

$$\langle f(v_1, \cdot, \varepsilon)\rangle = \langle (1 + \Theta'(\cdot))f(v_1, \cdot + \Theta(\cdot); \varepsilon)\rangle.$$

With this identity and the smoothness of f we can bound the difference of the first two terms in the expression for $\langle \tilde{F}\rangle$ by $C\varepsilon N$.

By applying analogous estimates to R_0, we obtain the following lemma.

LEMMA 4.3. *There exist $\varepsilon_0 > 0$ and $C > 0$ such that if $0 < \varepsilon < \varepsilon_0$,*

$$\|\bar{F}_0(\cdot, \varepsilon)\|_\delta \leqslant C(\varepsilon N + \varepsilon^2 N^2) \leqslant C\varepsilon^{1/2},$$
$$\sup_{\psi \in [0,1]} \|R_0(\cdot, \psi, \varepsilon)\|_\delta \leqslant C.$$

REMARK 4.4. The fact that the norm of R_0 is only bounded by a constant rather than $C\varepsilon^{1/2}$ is due to the presence of the term

$$\varepsilon\Theta(\psi)\bar{d}\Delta_x u_1,$$

in BU. This term has an average of zero and hence does not contribute to F_0, but it is present in R_0 and can only be bounded by $C\varepsilon N^2 \leqslant C$.

We now construct inductively a sequence of changes of variables which successively transforms (4.1) into (4.5). Begin by defining the change of variables

$$U = \Phi_0(V) \equiv V + \varepsilon W_0\left(V, \frac{t}{\varepsilon}, \varepsilon\right), \tag{4.11}$$

with

$$W_0(V, \xi, \varepsilon) = \int_0^\xi R_0(V, s, \varepsilon)\,\mathrm{d}s.$$

From the bound on R_0 in lemma 4.3, we have

$$\sup_{\xi \in [0,1]} \|W_0(\cdot, \xi, \varepsilon)\|_\delta \leqslant M_0$$

for some $M_0 > 0$. Furthermore, recalling that f is an entire function we see that R_0, and hence W_0 is analytic on $\mathcal{N}_\delta^\sigma(D_{N,\sigma})$ and thus we can use Cauchy estimates to bound its derivatives (see [5, lemma 7]):

$$\left\|\varepsilon \frac{\partial W_0}{\partial V}\right\|_{\delta_j - K_0} \leqslant \frac{\varepsilon\|W_0\|_\delta}{K_0} \leqslant \frac{M_0\varepsilon}{K_0}. \tag{4.12}$$

These estimates, plus the inverse function theorem, imply that if we choose $K_0 = \frac{1}{4}\delta$ and set $\delta_1 = \frac{3}{4}\delta$ we have the following lemma.

LEMMA 4.5. *There exist $\varepsilon_0 > 0$ such that for every $0 < \varepsilon < \varepsilon_0$, (4.11) defines an analytic and invertible change of variables from $\mathcal{N}_{\delta_1}^\sigma(D_{N,\sigma})$ into $\mathcal{N}_{\delta_0}^\sigma(D_{N,\sigma})$.*

Writing out the differential equation satisfied by V we have

$$\dot{V} = \bar{A}V + \bar{F}(V;\varepsilon) + \bar{F}_1(V;\varepsilon) + R_1\left(V, \frac{t}{\varepsilon}, x, \varepsilon\right),$$

where $\bar{F}_1$ and R_1 are given in terms of $\bar{F}_0$ and R_0 by (4.10). Applying the Cauchy estimates as above, we see that we can bound terms like

$$\|R_0(\cdot + \varepsilon W_0(\cdot), \theta; \varepsilon) - R_0(\cdot, \theta, \varepsilon)\|_{\delta_1} \leqslant \varepsilon \frac{\|R_0(\cdot, \theta, \varepsilon)\|_{\delta_0}}{(\delta_0 - \delta_1)}\|W_0\|_{\delta_1} \leqslant \frac{C\varepsilon}{\delta}.$$

All other terms in $\bar{F}_1$ and R_1 can be bounded in a similar fashion, the most difficult being the term $\|\varepsilon P^N \bar{A} W_0\|_{\delta_1}$, which is bounded by $C\varepsilon N \leqslant C\varepsilon^{1/2}$. The details are left to the reader as an exercise, and we find that we can apply the following lemma.

LEMMA 4.6. *There exist $\varepsilon_0 > 0$ and positive constants $\tilde{B}_1$ and $\tilde{B}_2$, such that*

$$\|\bar{F}_1(\cdot, \varepsilon)\|_{\delta_1} \leqslant \tilde{B}_1 \varepsilon^{1/2},$$

$$\sup_{\psi \in [0,1]} \|R_1(\cdot, \psi, \varepsilon)\|_{\delta_1} \leqslant \tilde{B}_2 \varepsilon^{1/2}.$$

We now proceed inductively. Define $K(\varepsilon) = \tilde{K}\varepsilon^{1/2}$ for $\tilde{K}$ fixed and set $\delta_{j+1} = \delta_j - K(\varepsilon)$ for $j = 1, 2, \ldots, r$. We require $\delta_r \geqslant \frac{1}{2}\delta$ and so we choose

$$r = \left[\frac{\delta}{2K(\varepsilon)}\right] = \left[\frac{\delta}{2\tilde{K}\varepsilon^{1/2}}\right]$$

We then state the following proposition.

PROPOSITION 4.7. *There exist $\varepsilon_0 > 0$ and positive constants B_1 and B_2, such that, for $j = 1, 2, \ldots, r$, there exists an analytic and invertible change of variables*

$$U = \Phi_j(V) = V + \varepsilon W_j\left(V, \frac{\tau}{\varepsilon}, \varepsilon\right)$$

from $\mathcal{N}^\sigma_{\delta_{j+1}}(D_{N,\sigma})$ into $\mathcal{N}^\sigma_{\delta_j}(D_{N,\sigma})$ and such that V satisfies the differential equation

$$\dot{V}_N = \bar{A}V_N + P^N\bar{F}(V_N; \varepsilon) + \bar{F}_{j+1}(V_N; \varepsilon) + R_{j+1}\left(V_N, \frac{\tau}{\varepsilon}; \varepsilon\right), \tag{4.13}$$

with

$$\|\bar{F}_{j+1}(\cdot, \varepsilon)\|_{\delta_{j+1}} \leqslant B_1\varepsilon^{1/2} \tag{4.14}$$

and

$$\sup_{\psi \in [0,1]} \|R_{j+1}(\cdot, \psi, \varepsilon)\|_{\delta_{j+1}} \leqslant M_{j+1} \equiv 2^{-(j+1)}B_2\varepsilon^{1/2}. \tag{4.15}$$

Proof. Assume that

$$\|\bar{F}_j\|_{\delta_j} \leqslant B_1\varepsilon^{1/2}, \tag{4.16}$$

$$\|R_j\|_{\delta_j} \leqslant M_j \quad \text{with } M_j = 2^{-j}B_2\varepsilon^{1/2}, \tag{4.17}$$

for the constants B_1, B_2 chosen below. Lemma 4.6 implies that these estimates hold for $j = 1$ and we will prove that (4.14) and (4.15) follow from them, completing the induction.

REMARK 4.8. To simplify notation, we suppress the arguments of W and the dependence on time and parameters N and ε in the functions.

Define the $(j+1)$st change of variables by

$$U = \Phi_j(V) = V + \varepsilon W_j\left(V, \frac{t}{\varepsilon}, \varepsilon\right),$$

where W_j is defined by (4.7). The inductive estimates on R_j, plus the Cauchy estimates imply that

$$\|\varepsilon W_j\|_{\delta_j} \leqslant M_j\varepsilon \tag{4.18}$$

and

$$\left\|\varepsilon\frac{\partial W_j}{\partial V}\right\|_{\delta_j - K(\varepsilon)} \leqslant \frac{M_j\varepsilon}{K(\varepsilon)}, \tag{4.19}$$

from which we see (via the analytic inverse function theorem) that Φ_j defines an analytic and invertible change of variables from $\mathcal{N}^\sigma_{\delta_{j+1}}(D_{N,\sigma})$ into $\mathcal{N}^\sigma_{\delta_j}(D_{N,\sigma})$.

We can estimate the correction term a in (4.6):

$$\|a\|_{\delta_j - K(\varepsilon)}$$
$$\leqslant \|(D_V\Phi_j)^{-1}\{\varepsilon\bar{A}W_j + P^N\bar{F}(V + \varepsilon W_j) - P^N F(V)$$
$$+ \bar{F}_j(V + \varepsilon W_j) - \bar{F}_j(V) + R_j(V + \varepsilon W_j) - R_j(V)$$
$$+ \{(\mathbf{1} + \varepsilon D_V W_j)^{-1} - \mathbf{1}\}W_j[\bar{A}V + P^N\bar{F}(V) + \bar{F}_j(V)]\}\|_{\delta_j - K(\varepsilon)}.$$

(Again, we have suppressed the dependence of the various functions on all variables except those important for the inductive estimates.)

For $V \in \mathcal{N}^\sigma_{\delta_{j+1}}(D_{N,\sigma})$, $\|AV\|_{\delta_{j+1}}$ is bounded by $N\|V\|_{\delta_{j+1}}$. Using the inverse function theorem and the mean value theorem to estimate $(D_V\Phi_j)^{-1}$ and $(D_V\Phi_j)^{-1} - \mathbf{1}$ we have

$$\|a\|_{\delta_j - K(\varepsilon)}$$
$$\leqslant \sum_{k=0}^{\infty} \left\|\varepsilon\frac{\partial}{\partial V}W_j\right\|^k_{\delta_j - K(\varepsilon)}$$
$$\times \left\{\|\varepsilon W_j\|_{\delta_j}\left[N + \left\|\frac{\partial}{\partial V}P^N F\right\|_{\delta_j - K(\varepsilon)} + \left\|\frac{\partial}{\partial V}\bar{F}_j\right\|_{\delta_j - K(\varepsilon)} + \left\|\frac{\partial}{\partial V}R_j\right\|_{\delta_j - K(\varepsilon)}\right] \right.$$
$$\left. + C\left\|\varepsilon\frac{\partial}{\partial V}W_j\right\|_{\delta_j - K(\varepsilon)}[N\|V\|_{\delta_j} + \|P^N\bar{F}\|_{\delta_j} + \|\bar{F}_j\|_{\delta_j}]\right\}.$$

Applying the Cauchy estimate again gives

$$\|a\|_{\delta_j - K(\varepsilon)} \leqslant 2\left\{\varepsilon M_j\left[N + \frac{M_j}{K(\varepsilon)} + \frac{3B_1}{K(\varepsilon)}\right] + \frac{M_j\varepsilon}{K(\varepsilon)}[NM + B_1 + 2B_1]\right\},$$

where M is the radius of D_N (see (4.4)). Recalling the definition of $K(\varepsilon)$, we obtain

$$\|a\|_{\delta_j - K(\varepsilon)} \leqslant M_j\left[2\varepsilon^{1/2} + 6\frac{B_1}{\tilde{K}}\varepsilon^{1/2} + 2\frac{M}{\tilde{K}} + 6\frac{B_1}{\tilde{K}}\varepsilon^{1/2}\right] \leqslant \tfrac{1}{4}M_j$$

for sufficiently large $\tilde{K}$ uniformly in $0 < \varepsilon < \varepsilon_0$ and $N(\varepsilon)$. Therefore,

$$\|R_{j+1}\|_{\delta_{j+1}} < \tfrac{1}{2}M_j = M_{j+1} \quad \text{and} \quad \|\bar{F}_{j+1} - \bar{F}_j\|_{\delta_{j+1}} < \tfrac{1}{4}M_j.$$

Hence,

$$\|\bar{F}_{j+1}\|_{\delta_{j+1}} \leqslant \|\bar{F}_0\|_{\delta_{j+1}} + \sum_{k=0}^{j}\|\bar{F}_{k+1} - \bar{F}_k\|_{\delta_{j+1}}$$
$$\leqslant B_2\varepsilon^{1/2} + \frac{1}{4}\sum_{k=0}^{j}M_k \leqslant B_2\varepsilon^{1/2} + \tfrac{1}{2}B_2\varepsilon^{1/2} \leqslant B_1\varepsilon^{1/2}. \qquad (4.20)$$

Thus, the inductive statements (4.16) and (4.17) are satisfied for $j+1$ for $0 < \varepsilon < \varepsilon_0$.

□

We now define $\Phi_* = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_r$. The inductive estimates, plus the definitions of δ_j and r, then imply that Φ_* is an analytic and invertible change of variables from $\mathcal{N}_{\delta/2}^\sigma(D_{N,\sigma})$ into $\mathcal{N}_\delta^\sigma(D_{N,\sigma})$. Furthermore, this change of variables transforms (4.1) into

$$\dot{V} = \bar{A}V + \bar{F}(V;\varepsilon) + \bar{F}_*(V;\varepsilon) + R_*\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right). \tag{4.21}$$

From the inductive estimates we obtain estimates on the norms of F_* and R_*, uniform in $\varepsilon \to 0$ and $N(\varepsilon) \to \infty$, on $\mathcal{N}_{\delta/2}^\sigma(D_{N,\sigma}) \subset \mathcal{N}_{\delta_r}^\sigma(D_{N,\sigma})$. More precisely, we obtain the following corollary.

COROLLARY 4.9. *There exist $\varepsilon_0 > 0$ and $C, c_1, c_2 > 0$ such that if $0 < \varepsilon < \varepsilon_0$, there exists an analytic and invertible change of variable*

$$U = \Phi_*(V) = V + \varepsilon W_*(V)$$

from $\mathcal{N}_{\delta/2}^\sigma(D_{N,\sigma})$ into $\mathcal{N}_\delta^\sigma(D_{N,\sigma})$, which transforms (4.1) into (4.21). The terms F_*, R_* and W_* can be estimated by

$$\sup_{\xi \in [0,1]} \|R_*(\cdot, \xi; \varepsilon)\|_{\delta/2} < 2^{-r} B_2 \varepsilon^{1/2} < c_2 \exp(-c_1 \varepsilon^{-1/2}), \tag{4.22}$$

$$\|F_*(\cdot; \varepsilon)\|_{\delta/2} < C\varepsilon^{1/2}, \tag{4.23}$$

$$\|W_*(\cdot; \varepsilon)\|_{\delta/2} < C. \tag{4.24}$$

STEP 2 (transformation of the full system). Next we extend our bounds on the finite-dimensional subsystem to the full, infinite-dimensional system. We extend the change of variables Φ_* to a (complex) neighbourhood of the ball in $\mathcal{G}_\sigma$ by defining

$$U = V + \varepsilon W_*\left(P^N V, \frac{\tau}{\varepsilon}; \varepsilon\right).$$

W_* is bounded on bounded sets of X and $\mathcal{G}_\sigma$, as the estimates above were uniform in $\sigma \geqslant 0$ and $X = \mathcal{G}_0$. We write the equation for V as in (3.22):

$$\dot{V} = \bar{A}V + \bar{F}_{\exp}(V;\varepsilon) + R\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right) \tag{4.25}$$

with

$$\bar{F}_{\exp}(V;\varepsilon) + R\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right) = \bar{F}(V;\varepsilon) + \bar{F}_*(P^N V;\varepsilon) + R_*\left(P^N V, \frac{\tau}{\varepsilon}; \varepsilon\right) + b\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right).$$

The additional correction term $b = b_1 + b_2 + b_3$ arises from the infinitely many degrees of freedom that were ignored in (4.1), as well as their interactions with the modes in $P^N X$, and the terms that appear there fall naturally into one of three classes.

The infinite-dimensional corrections b_1 and b_2 come from the modes that were neglected in the transformation Φ_*. From the bounded part of the nonlinearity we have

$$b_1\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right) = (I - P^N)\left\{\bar{F}(V + \varepsilon W_*(P^N V)) - \bar{F}(V) + \tilde{F}\left(V + \varepsilon W_*(P^N V), \frac{\tau}{\varepsilon}; \varepsilon\right)\right\}, \tag{4.26}$$

while the unbounded part contributes

$$b_2\left(V, \frac{\tau}{\varepsilon}\right) = (I - P^N)B\left(\frac{\tau}{\varepsilon}\right)(V + \varepsilon W_*(P^N V)) = (I - P^N)B\left(\frac{\tau}{\varepsilon}\right)V.$$

We get also a correction b_3 in the finite-dimensional Galerkin space, as there is an error in the lower modes due to the neglected influence of the higher Galerkin modes on the lower ones. It is given by

$$b_3\left(V, \frac{t}{\varepsilon}, \varepsilon, N(\varepsilon)\right)$$
$$= (D_{V^N}\Phi_*)^{-1}$$
$$\times P^N\bigg\{\bar{F}(V + \varepsilon W_*(P^N V)) - \bar{F}(P^N V + \varepsilon W_*(P^N V))$$
$$+ \tilde{F}\left(V + \varepsilon W_*(P^N V), \frac{\tau}{\varepsilon}; \varepsilon\right) - \tilde{F}\left(P^N V + \varepsilon W_*(P^N V), \frac{\tau}{\varepsilon}; \varepsilon\right)\bigg\}.$$
$$(4.27)$$

Recall that $V^N = P^N V$.

We define

$$\bar{F}_{\exp}(V; \varepsilon) = \bar{F}(V; \varepsilon) + \bar{F}_*(P^N V; \varepsilon) + \langle b_1(V, \cdot; \varepsilon)\rangle, \qquad (4.28)$$

$$R\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right) = R_*\left(P^N V, \frac{\tau}{\varepsilon}; \varepsilon\right) + b\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right) - \langle b_1(V, \cdot; \varepsilon)\rangle. \qquad (4.29)$$

We start by estimating the size of $\bar{F}_{\exp}(\cdot; \varepsilon) = \bar{F}(\cdot; \varepsilon) + \bar{F}_*(\cdot; \varepsilon) + \langle b_1(\cdot, \cdot; \varepsilon)\rangle$. For the finite-dimensional part we have $|\bar{F}_*(P^N V, \varepsilon, N)|_X \leqslant c_3 \varepsilon^{1/2}$ by (4.23). Estimating $\langle b_1\rangle$, noting $\langle \tilde{F}(V, \cdot; \varepsilon)\rangle = 0$ and $\langle B(\cdot)V\rangle = 0$, for $V \in X = \mathcal{G}_0$ we obtain

$$|\langle b_1(V, \cdot, \varepsilon, N)\rangle|_X$$
$$\leqslant \left|(I - P^N)\bigg\{\bar{F}(V + \varepsilon W_*(P^N V)) - F(V) + \tilde{F}\left(V + \varepsilon P^N W_*(V), \frac{\tau}{\varepsilon}; \varepsilon\right)\bigg\}\right|_X$$
$$\leqslant \sup_{\substack{U = V + rW_*(P^N V), \\ r \in [0, \varepsilon]}} |\bar{F}'(U; \varepsilon)|_{L(X, X)} |\varepsilon W_*(P^N V)|_X$$
$$+ \sup_{\substack{U = V + rW_*(P^N V), \\ r \in [0, \varepsilon]}} |\tilde{F}'(U; \varepsilon)|_{L(X, X)} |\varepsilon W_*(P^N V)|_X$$
$$\leqslant C\varepsilon, \qquad (4.30)$$

because $|\varepsilon W_*(P^N V)|_X \leqslant C\varepsilon$. Thus, we obtain the desired estimate on $\bar{F}_{\exp}$:

$$|F(U; \lambda, \varepsilon) - \bar{F}_{\exp}(U; \lambda, \varepsilon)|_X \leqslant C_1 \varepsilon^{1/2}.$$

The additional terms can then be estimated very easily using the following observation. If $V(t) \in \mathcal{G}_\sigma$, then the difference between V and its projection into H_N is bounded in the X norm by

$$|V(t) - P^N V(t)|_X \leqslant |V(t)|_{\mathcal{G}_\sigma} \exp(-\sigma N). \qquad (4.31)$$

Since we consider Gevrey regular solutions of (4.25), these allow us to estimate the remaining terms:

$$|R|_X = |R_* + b_1 + b_2 + b_3 - \langle b_1 \rangle|_X.$$

For the first part we have, by (4.22),

$$\left| R_*\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right)\right|_X \leqslant c_2 \exp(-c_1 \varepsilon^{-1/2}).$$

For $b_1 - \langle b_1 \rangle$ we use the same analysis as for $\langle b_1 \rangle$ above, except that we also note that $\bar{F}$ and $\tilde{F}$ are bounded in the Gevrey norm $\mathcal{G}_\sigma$. Hence, $b_1(V) \in \mathcal{G}_\sigma$ and, thus,

$$\left| b_1\left(V, \frac{\tau}{\varepsilon}; \varepsilon\right) - \langle b_1(V, \cdot; \varepsilon)\rangle\right|_X$$
$$\leqslant 2\left|(I - P^N)\left\{\bar{F}(V + \varepsilon W_*(P^N V)) - \bar{F}(V) + \tilde{F}\left(V + \varepsilon W_*(P^N V), \frac{\tau}{\varepsilon}; \varepsilon\right)\right\}\right|_X$$
$$\leqslant C(|V|_{\mathcal{G}_\sigma}) \exp(-\sigma N). \tag{4.32}$$

Next we estimate $|b_2(V, \tau/\varepsilon)|_X = |(I - P^N)B(\tau/\varepsilon)V|_X$. Although B is unbounded as an operator on X, it is bounded as a linear operator $\mathcal{G}_\sigma \to \mathcal{G}_{\sigma/2}$. Thus,

$$\left| b_2\left(V, \frac{\tau}{\varepsilon}\right)\right|_X \leqslant C \exp(-(\tfrac{1}{2}\sigma)N). \tag{4.33}$$

Finally, we deal with the term b_3, created by the Galerkin approximation in the finite-dimensional part. For $V \in \mathcal{G}_\sigma$ we define the line segment

$$K_V = \{rV + (1-r)P^N V + \varepsilon W(P^N V) | r \in [0,1]\}.$$

We then obtain

$$|b_3(V, \cdot, \varepsilon, N)|_X \leqslant |(D_{V^N}\Phi_*)^{-1}|_{L(X,X)}$$
$$\times (|D\tilde{F}(V)|_{L(X,X)} + |D\bar{F}(V)|_{L(X,X)})|V - P^N V|_X$$
$$\leqslant C \exp(-\sigma N), \tag{4.34}$$

where we have used (4.31).

We now combine the exponential estimates of steps 1 and 2, using $N(\varepsilon) = [\varepsilon^{-1/2}]$. The estimates (4.32) and (4.34) can be expressed in terms of ε. We then obtain

$$\exp(-c_1 \varepsilon^{-1/2}) \text{ from } (4.22),$$
$$\exp(-\sigma \varepsilon^{-1/2}) \text{ from } (4.32)–(4.34).$$

The estimates are then of order $\exp(-c\varepsilon^{-1/2})$ and the proof is complete.

Acknowledgments

References

1   P. C. Bressloff. Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica* D **155** (2001), 83–100.

2   V. Jikov, S. Kozlov and O. Oleinik. *Homogenization of differential operators and integral functionals* (Springer, 1994).

3   J. P. Keener. Homogenization and propagation in the bistable equation. *Physica* D **136** (2000), 1–17.

4   K. Kirchgässner. Wave solutions of reversible systems and applications. *J. Diff. Eqns* **45** (1982), 113–127.

5   K. Matthies. Time-averaging under fast periodic forcing of parabolic partial differential equations: exponential estimates. *J. Diff. Eqns* **174** (2001), 133–180.

6   K. Matthies. Homogenisation of exponential order for elliptic systems in infinite cylinders. *Asymp. Analysis* **43** (2005), 205–232.

7   A. Neishtadt. On the separation of motions in systems with rapidly rotating phase. *J. Appl. Math. Mech.* **48** (1984), 134–139.

8   D. Peterhof, B. Sandstede and A. Scheel. Exponential dichotomies for solitary wave solutions of semilinear elliptic equations on infinite cylinders. *J. Diff. Eqns* **140** (1997), 266–308.

9   J. Xin. Front propagation in heterogeneous media. *SIAM Rev.* **42** (2000), 161–230.