Bifurcation to coherent structures in nonlocally coupled systems

Arnd Scheel and Tianyu Tao

University of Minnesota, School of Mathematics, 206 Church St. S.E., Minneapolis, MN 55455, USA

Abstract

We show bifurcation of localized spike solutions from spatially constant states in systems of nonlocally coupled equations in the whole space. The main assumptions are a generic bifurcation of saddle-node or transcritical type for spatially constant profiles, and a symmetry and second moment condition on the convolutions kernel. The results extend well known results for spots, spikes, and fronts, in locally coupled systems on the real line and for radially symmetric profiles in higher space dimensions. Rather than relying on center manifolds, we pursue a more direct approach, deriving leading order asymptotics and Newton corrections for error terms. The key ingredient is smoothness of Fourier multipliers arising from discrepancies between nonlocal operators and their local long-wavelength approximations.

1 Introduction

Much of the success of modeling has been based on infinitesimal descriptions of physical laws, leading to differential and partial differential equations as models for physical processes. However, many physical interactions are inherently nonlocal, at least at a coarser modeling level, leading to nonlocal spatial coupling, as well as dependence of time evolution on history. From a dynamical systems point of view, a natural question in this context is in how far nonlocally coupled equations may behave qualitatively differently from locally coupled problems. As usual, one can approach this question from several vantage points, striving in particular to point to phenomena where nonlocality generates new phenomena, or delineate situations where nonlocal and local equations behave in analogous fashions. Our contribution here is mainly towards the latter aspect, showing that local bifurcations in nonlocal systems produce coherent structures completely analogous to local systems. We will however also comment on phenomena that are qualitatively different as a result of nonlocality.

We will focus on a fairly simple model problem, stationary solutions to nonlocally ocupled systems of the form

$$U + \mathcal{K} * U = \mathcal{N}(U;\mu), \tag{1.1}$$

where $U = U(x) \in \mathbb{R}^k$, $x \in \mathbb{R}^n$, $\mathcal{K} * U$ stands for matrix convolution,

$$(\mathcal{K} * U(x))_i = \sum_{j=1}^{m} \int_{\mathbb{R}^n} \mathcal{K}_{i,j}(x-y) U_j(y) dy, \qquad 1 \le i \le k,$$

and $\mathcal{N}(U;\mu)$ encodes nonlinear terms which depend on a parameter $\mu \sim 0$.

We are interested in the existence of spatially localized solutions $U(x) \to U_\infty$, $|x| \to \infty$ in the prototypical setting of a transcritical bifurcation of spatially constant solutions $U(x) \equiv m$.

In the remainder of this introduction, we shall first provide some background and motivation for this kind of question, Section 1.1, and then give precise assumptions and results in Sections 1.2 and 1.3. We collect some standard notation used throughout at the end of the introduction.

1.1 Motivation

Applications with nonlocal coupling. Nonlocal coupling has been suggested as a more appropriate modling assumptions in fields all across the sciences. Prominent examples are nonlocal dispersal of plant seeds in ecology [9], fractional powers of the Laplacian as limits of random walks with Levy jmumps [19], models for neural fields [7], nonlocal interactions in models for Bose-Einstein condensates [22], kinetic equations for interacting particles [21], shallow water-wave models [8], or material science [4]. In many of these models, one is interested in spatially localized or front-like solutions, stationary, periodic, or propagating in time, which we will here refer to collectively as coherent structures. Such solutions are usually found through a traveling-wave ansatz, thus eliminating or compactifying time dependence. In some special situations, Many of the above examples can thereby be reduced to problems of the type (1.1), and we will give details for some cases in Section 4.

Arguably, the most powerful results for existence and stability of coherent structures rely on a formulating the existence problem as an ordinary differential equation, a method sometimes referred to as "spatial dynamics" [24] — for locally coupled equations, in essentially one-dimensional geometries such as the real line, cyliners $\mathbb{R} \times \Omega$, or radially symmetry. We will briefly overview such results from our perspective here, discuss in how far they generalize to nonlocally coupled equations, before discussing our contribution here in more detail in Sections 1.2–1.3.

Local coupling — results. Replacing nonlocal coupling by diffusive coupling, say $D\Delta U$, with D positive, diagonal, we can consider the elliptic system

$$\Delta U + \mathcal{N}(U; \mu) = 0, \tag{1.2}$$

with $x \in \mathbb{R}^n$. In the simplest case of $n = 1$, one can study the resulting ordinary differential equation

$$U_x = V, \qquad V_x = -\mathcal{N}(U; \mu),$$

using dynamical systems methods such as center-manifold reduction and normal form transformations, thus leading to nearly universal predictions for bifurcations of coherent structures. In addition to existence results, such methods also allow one to state rather general uniqueness and non-existence results. In higher space dimensions, $n > 1$, such reduction techniques Are not known to be applicable, except for the context of radial symmetry, which allows one to formulate the existence problem as a dynamical system in the radial variable r,

$$U_r = V, \qquad V_r = -\frac{n-1}{r} V - \mathcal{N}(U; \mu).$$

Slightly adapted center manifold and normal form theory again leads to near-universal classifications of small-amplitude coherent structures; see [25] for a comprehensive exposition of these techniques and [16] for backgroujnd on center manifolds and normal forms.

in the simplest case of a transcritical bifurcation in the nonlinearity, with suitable additional assumptions on the linear part, one finds a reduced equation on the center manifold of the form

$$\begin{aligned} u_x &= v + \mathrm{O}\left(|\mu|(|u| + |v|) + u^2 + v^2\right) \\ v_x &= \mu u - u^2 + \mathrm{O}\left(v^2 + (|u| + |v|)(\mu^2 + u^2 + v^2)\right), \end{aligned}$$

which at leading order, after scaling, reduces to

$$u_{xx} - u + u^2 = 0, \tag{1.3}$$

with an explicit localized solution $u(x) = \ldots$. Using revsibility $x \mapsto -x$ one then readily shows persistence of this solution at higher order. Exploiting the characterization of center manifolds, one also obtains non-existence of other localized solutions and, in fact, a complete characterization of small bounded solutions.

The results in [25] extend this machinery to radially symmetric solutions in higher space dimension, leading to a leading order equation of the form

$$u_{rr} + \frac{n-1}{r} u_r - u + u^2 = 0.$$

This type of results is also available for Turing and Hopf bifurcations [25]. It is however not immediately applicable to the type of nonlocally coupled equations described above, with recent progress that we shall describe next.

Nonlocal coupling — center manifolds. Going back to (1.1), we in general not find an obvious formulation as a dynamical systems, with the notable exception of convolutions kernels with a rational Fourier transform. Consider for instance $\mathcal{K}(x) = \frac{1}{2}\mathrm{e}^{-|x|}$, $x \in \mathbb{R}$, with Fourier transform $\hat{\mathcal{K}}(\xi) = (1+\xi^2)^{-1}$, for which we can formally write (1.1) as $(\,\mathrm{id}\, - \Delta)^{-1}U + \mathcal{N}(U;\mu) = 0$, which in turn is equivalent to

$$\begin{aligned} -U + W + \mathcal{N}(U;\mu) &= 0, \\ W - \Delta W &= U. \end{aligned}$$

Under suitable assumptions, equivalent to stability assumptions made in [25], one can solve the first equation for $U = \Psi(W;\mu)$ and insert into the second equation, thus obtaining a local equation for W,

$$\Delta W - W + \Psi(W;\mu) = 0,$$

which is amenable to the methods from [25].

The restriction to kernels with rational Fourier transform is clearly restrictive, excluding for instance Gaussians, and one naturally wonders if similar results hold outside of this class. The more recent results in [14] answer this question in the affirmative, for $x \in \mathbb{R}$ and $\mathcal{K}, \mathcal{K}'$ exponentially localized, such that $\hat{\mathcal{K}}$ is analytic in a strip of the complex plane $|\,\mathrm{Re}\,\xi| < \eta$.

Some smoothness of $\mathcal{K}$ appears to be necessary, as the counter example of $\mathcal{K} = \frac{1}{2}(\delta_{-1} + \delta_1)$ shows, which produces a simple iteration

$$\frac{1}{2}\left(U(x+1) - 2U(x) + U(x-1)\right) + \mathcal{N}(U;\mu) = 0,$$

with completely uncorrelated solutions on lattices $x \in x_0 + \mathbb{Z}$. A finite-dimensional reduction to an ordinary differential equation here does not seem possible.

On the other hand, exponential localization appears to be necessary in the sense that it solutions to reduced differential equations would typically converge exponentially. Nevertheless, our results remove the assumption of exponential localization, at the expense of a lacking uniqueness argument. We still reduce to the simple ordinary differential equation (1.3) or its higher-dimensional analogue, with exponentially localized solutions, and find weaker far-field decay only at higher order in the bifurcation parameter μ.

1.2 Setup — linear nonlocal diffusive coupling and local bifurcations of spatially constant states

Within the context described in the previous section, we are now ready to state our main assumption and results. We start with assumptions on the linear part, keeping in mind that the nonlinearity will be assumed to be of quadratic order in U, μ, then turn to assumptions on the nonlinearity, before stating our main result.

Linear diffusive coupling. Let I_k denote the identity matrix of size k and consider the linearized operator $I_k + \mathcal{K}*$.

Hypothesis (L) *We assume that $\mathcal{K}$satisfies the following properties:*

(i) localization: *$\mathcal{K}$ has finite moments of order less than 2, that is, $\mathcal{K}(x)$, $|x|^2\mathcal{K}(x) \in L^1(\mathbb{R}^n, \mathbb{R}^{k\times k})$;*

(ii) symmetry: $\mathcal{K}(x) = \mathcal{K}(\gamma x)$ *for all* $x \in \mathbb{R}^n$ *and all* $\gamma \in \Gamma \subset \mathbf{O}(n)$*, a subgroup of the orthogonal matrices with*

$$d\mathrm{Fix}\,\Gamma = \{x \mid \gamma x = x, \text{ for all } \gamma \in \Gamma\} = \{0\};$$

(iii) minimal nullspace: $\ker\left(I_k + \int \mathcal{K}\right) = \operatorname{span} \mathcal{E}_1$ *for some* $0 \neq \mathcal{E}_1 \in \mathbb{R}^k$*; we then choose* $\mathcal{E}_1^*$ *such that* $\ker\left(I_k + \int \mathcal{K}^T\right) = \operatorname{span} \mathcal{E}_1^*$*;*

(iv) nondegenerate second moments: *the matrix of projected second moments* S *with entries*

$$S_{ij} = \int x_i x_j \langle \mathcal{E}_1^*, \mathcal{K}(x)\mathcal{E}\rangle \mathrm{d}x,$$

is positive definite;

(v) invertibility for nonzero wavenumbers: $I_k + \int \mathrm{e}^{\mathrm{i}\langle \xi, x\rangle}\mathcal{K}(x)\mathrm{d}x$ *is invertible for all* $\xi \neq 0$.

The assumption on positive definiteness can be readily replaced by negative definiteness, simply multiplying the equation by -1. Note that first moments, $\int x\mathcal{K}(x)\mathrm{d}x$ vanish, since

$$\int x\mathcal{K}(x)\mathrm{d}x = \int xK(\gamma x)\mathrm{d}x = \gamma^{-1}\int y\mathcal{K}(y)\mathrm{d}y,$$

for all $\gamma \in \Gamma$, hence $\int x\mathcal{K}(x)\mathrm{d}x = 0$ since Γ fixes the origin, only. In fact, this is the primary reason for us to require the symmetry mentioned here. Nonvanishing first moments correspond to an effective directional transport which would need to be compensated by a drift term $c\cdot\nabla U$, say, in order to find coherent structures.

Typical examples of symmetry groups Γ are $\Gamma = \mathbf{O}(n)$, $\Gamma = \{\mathrm{id}, -\mathrm{id}\}$ or the group generated by reflections at hyperplanes, $x_j \mapsto -x_j$.

Remark 1.1 (normalizing second moments) . *There exists a coordinate change* $x = Ty$ *such that* $\tilde{\mathcal{K}}(y) := |\det T|\mathcal{K}(Ty)$ *satisfies*

$$\tilde{S}_{ij} = \int x_i x_j \langle \mathcal{E}_1^*, \tilde{\mathcal{K}}(x)\mathcal{E}\rangle \mathrm{d}x = 2\delta_{ij}.$$

Indeed, let λ_i*,* $i = 1, \ldots, k$ *be the eigenvalues of* S *and* P_i *be the associated spectral projections. Define*

$$T^{-1} = \sqrt{2}\sum_i \lambda_i^{-1/2} P_i, \qquad M(y) = \langle \mathcal{E}_1^*, \tilde{\mathcal{K}}(y)\mathcal{E}_1\rangle,$$

then

$$\begin{aligned}
\tilde{S}_{ij} &= \int y_i y_j M(y)\mathrm{d}y = \int (T^{-1}x)_i (T^{-1}x)_j \langle \mathcal{E}_1^*, \mathcal{K}(x)\mathcal{E}_1\rangle \mathrm{d}x \\
&= \sum_{k,\ell} T^{-1}_{ik} T^{-1}_{j\ell} S_{k\ell} = T^{-1}S(T^{-1})^* = T^{-1}ST^{-1} \\
&= 2\left(\sum_i \lambda_i^{-1/2} P_i\right)\left(\sum_j \lambda_i P_j\right)\left(\sum_m \lambda_i^{-1/2} P_m\right) \\
&= 2\delta_{ij}
\end{aligned}$$

by the orthogonality of P_i*. Moreover, by differentiating the identity* $\mathcal{K}(\gamma x) = \mathcal{K}(x)$ *in* x *twice, we find* S *commutes with every* $\gamma \in \Gamma$*. As a result, the spectral projections, and hence* T_0*, commute with every* $\gamma \in \Gamma$*.*

The assumptions can also be stated in terms of the associated Fourier determinant $\mathcal{D}$,

$$\mathcal{D}(\xi) := \det(I_k + \widehat{\mathcal{K}}(\xi)).$$

One readily finds that $\mathcal{D}$ is of class $\mathscr{C}^2$ by the assumption on second moments, and

$$\mathcal{D}(0) = 0, \qquad \mathcal{D}''(0) \neq 0.$$

The characteristic function $\mathcal{D}$ was also used in [14], identifying zeros of $\mathcal{D}$ on the real axis with bounded solutions to the linear equation, and, more generally, multiplicities of zeros adding up to the dimension of a reduced center manifold as algebraic multiplicities of a center subspace. In the setup there, $\mathcal{D}$ was analytic, allowing readily for characterizing the multiplicity of roots. We here assume just enough regularity, $\mathscr{C}^2$ to make sense of a "double root" of $\mathcal{D}$.

Remark 1.2 (generalizing linear assumptions) *Most examples of nonlinear problems would involve a nontrivial pointwise linear part, say, $AU + \mathcal{K} * U$. One quickly sees that there and more general terms should be viewed as less smooth, namely Dirac-δ contributions to the matrix kernel. Whenever this principle part, say, the matrix A, is invertible, the system can be readily put into our form by applying A^{-1}. On the other hand, when this principal part possesses a kernel, our assumption of invertibility for nonzero wavenumbers would be violated asymptotically, for $\ell \to \infty$. Bifurcation solutions in such situations are not necessarily smooth.*

Yet a different interpretation would refer to the spectrum of the linear part $I_k + \mathcal{K}*$, given by the set of λ for which $I_k + \mathcal{K}* - \lambda I_k$ is not invertible, or, equivalently, the closure of the set of λ such that $\det(I_k + \widehat{\mathcal{K}}(\xi) - \lambda I_k) = 0$ for some $\xi \in \mathbb{R}$. Clearly, $\lambda = 0$ is in the spectrum, choosing $\xi = 0$. Also, $\lambda = 0$ is minimal in multiplicity in the sense that it is an eigenvalue only for $\xi = 0$, and it's geometric multiplicity at $\xi = 0$ is minimal. Assuming that, in addition, $\lambda = 0$ is algebraically simple, one redadily finds that the continuation of λ as an eigenvalue in ξ is quadratic, $\lambda = \langle \mathcal{S}\xi, \xi\rangle + \ldots$, with definite symmetric matrix $\mathcal{S}$. Such spectral information can in general be converted into heat decay estimates for $U_t = -U + \mathcal{K} * U$ [3].

Transcritical bifurcation for spatially constant solutions. As mentioned, we assume the presence of a simple transcritical bifurcation in the nonlinearity $\mathcal{N}$. The assumptions that follow are generic and necessary for a non-degenerate bifurcation scenario; see for example [11]. They single out relevant terms in systems of equations that lead to bifurcations as in the simple scalar example $\mathcal{N}(u;\mu) = \mu u - u^2$.

Hypothesis (TC) *We assume that $\mathcal{N} = \mathcal{N}(U;\mu) : \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}^k$ satisfies the following conditions:*

(i) smoothness: $\mathscr{C}^K(\mathbb{R}^k \times \mathbb{R}; \mathbb{R}^k)$, $K = 1 + \ell + 2$;

(ii) trivial solution: $\mathcal{N}(0;\mu) = 0$ *for all* μ;

(iii) criticality: $D_U\mathcal{N}(0;0) = 0$;

(iv) generic linear unfolding:

$$\alpha := \langle D_{\mu,U}\mathcal{N}(0;0)\mathcal{E}_1, \mathcal{E}_1^*\rangle \neq 0; \tag{1.4}$$

(v) generic nonlinearity:

$$\beta := \frac{1}{2}\langle D_{U,U}\mathcal{N}(0;0)[\mathcal{E}_1, \mathcal{E}_1], \mathcal{E}_1^*\rangle \neq 0;. \tag{1.5}$$

Smoothness assumptions ensures that, the superposition operator $U(\cdot) \mapsto \mathcal{N}(U(\cdot);\mu)$ defined by $\mathcal{N}$ is of class $\mathscr{C}^1$ as a map on $H^\ell(\mathbb{R}^n, \mathbb{R}^k)$, see [23] for instance.

The choice of a transcritical setup here is for convenience and other elementary bifurcations can be treated in a similar fashion. A saddle-node bifurcation, where $\langle D_\mu\mathcal{N}(0;0), \mathcal{E}_1^*\rangle \neq 0$, $\langle D_{U,U}\mathcal{N}(0;0)[\mathcal{E}_1, \mathcal{E}_1], \mathcal{E}_1^*\rangle \neq 0$, can be transformed into a transcritical bifurcation after subtracting one of the branches, $\tilde{U} = U - U_-(\mu)$ and reparameterizing, say, $\mu = \tilde{\mu}^2$. Also, more general nonlocal dependence of $\mathcal{N}$ on U can be allowed. In such a case, Hypothesis (TC) applies to $\mathcal{N}$ acting on spatially constant solutions U.

1.3 Bifurcation of spikes — main result

We are now ready to state our main result. As suggested above, we would like to compare our problem to the local problem

$$\Delta u + \alpha\mu u + \beta u^2 = 0, \tag{1.6}$$

when $U \sim u\mathcal{E}_1$. For $\alpha\mu > 0$, this equation possesses localized ground states of the form

$$u_{\mathrm{c}}(x;\mu) = -\alpha\beta^{-1}\mu u_*(\sqrt{\alpha\mu}x), \tag{1.7}$$

where $u_*(y)$ is the (positive) ground state to

$$\Delta u - u + u^2 = 0; \tag{1.8}$$

see [27] for background information on existence of such ground states and their properties.

Theorem 1 (bifurcation of spikes) *Fix $n < 6$ and $\ell > n/2$. Assume Hypotheses (L) and (TC), and recall the definition of T_0 from Remark 1.1 and α, β from* (1.4) *and* (1.5).

There then exists a constant $\mu_0 > 0$ such that for all $0 < \alpha\mu < \mu_0$, the nonlocally coupled system (1.1) *possesses a family of nontrivial solutions $U_* = U_*(\cdot;\mu) \in H^\ell(\mathbb{R}^n;\mathbb{R}^k)$. Moreover, $U_*(x;\mu)$ is given to leading order through*

$$U_*(x;\mu) = (\alpha\beta^{-1}\mu u_*(\sqrt{\alpha\mu}T_0x) + w(\sqrt{\mu}x;\mu)\mathcal{E}_1 + u_\perp(x;\mu) \tag{1.9}$$

where

- *u_* from* (1.7)*is the (scalar) radially symmetric ground state to* (1.8)*;*
- *the corrector $w(\sqrt{\mu}x;\mu)$ satisfies*[1] *$\|w(\cdot;\mu)\|_{H^\ell} \to 0$ as $\mu \to 0$;*
- *$\langle u_\perp(x;\mu), \mathcal{E}_1\rangle = 0$, and $\|u_\perp\|_{H^\ell} = \mathrm{O}(\mu^2)$.*

Moreover, $U_*(x;\mu) = U_*(\gamma x;\mu)$ for all $\gamma \in \Gamma$.

We comment briefly on the scope of this result and outline the main idea of proof.

We believe that our assumptions are to some extent sharp. Second moments are necessary to define diffusive behavior and obtain the limiting ground state problem. One would suspect that weaker localization, $\hat{\mathcal{K}} \sim |\xi|^2\alpha$ for $\xi \sim 0$ would lead to reduced problems based on the fractional Laplacian $(-\Delta)^\alpha$, with somewhat analogous results. Symmetry of the kernel is necessary to some extent to prevent drift of the resulting spikes. In other words, we find at leading order a manifold of translates of a ground state as solutions. The symmetry condition guarantees that all these solutions persist to higher order as non-degenerate fixed points in the fixed point subspace of the action of Γ on profiles. We suspect that alternate assumptions could involve a variation structure in the problem. We will comment on possible drift and the technical ramifications in the discussion section.

The proof is eventually based on a rather direct contraction mapping principle. We prepare the equation by performing linear transformations singling out the neutral direction $\mathcal{E}_1$, followed by scaling and solving the equation in the complement $\mathcal{E}_1^\perp$. Finally, an ansatz substituting u_* and a suitable corrector yields a contraction mapping for the corrector in a small neighborhood of the origin.

Key to the argument is precise formulation of the convergence of $-I_k + \mathcal{K}$ to Δ, which we accomplish by carefully preconditioning our system; see equation (3.2).

Outline. We perform coordinate changes and rescalings in Section 2, preparing for the proof through Lyapunov-Schmidt reduction and contraction mappings in Section 3. Section 4 shows some more concrete applications of our result, and we conclude with a discussion in Section 5.

[1] should w and $u_\perp$ be depend on T_0x rather than x?

Notation. For a vector $u = (u_1, \ldots, u_k) \in \mathbb{R}^k$, we write $|u|$ to denote its usual Euclidean norm $|u| = \sum_{i=1}^k u_i^2$. We also use the standard multi-index notation in $\mathbb{R}^n$, that is we have $\alpha = (\alpha_1, \cdots, \alpha_n)$ with $\alpha_i \in \{0, 1, \ldots\}$ and $\alpha! = \alpha_1! \cdots \alpha_n!$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$. So that $D^\alpha u = \dfrac{\partial^{|\alpha|} u}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}}$.

We shall use the standard Sobolev spaces on $\mathbb{R}^n$ with values in $\mathbb{R}^k$, which are denoted by $W^{\ell,p}(\mathbb{R}^n, \mathbb{R}^p)$ or simply $W^{\ell,p}(\mathbb{R}^n)$ when $k = 1$ or even $W^{\ell,p}$ whenever it is convenient to do so and does not cause confusions. For $\ell \geq 0$ and $1 \leq p \leq \infty$

$$W^{\ell,p}(\mathbb{R}^n, \mathbb{R}^k) := \{u \in L^p(\mathbb{R}^n, \mathbb{R}^k) : \partial^\alpha u \in L^p(\mathbb{R}^n, \mathbb{R}^k), 1 \leq |\alpha| \leq \ell\},$$

with norm

$$\|u\|_{W^{\ell,p}(\mathbb{R}^n,\mathbb{R}^k)} = \begin{cases} \left(\sum_{1 \leq |\alpha| \leq \ell} \|\partial^\alpha u\|_{L^p(\mathbb{R}^n;\mathbb{R}^k)}\right)^{1/p}, & 1 \leq p < \infty \\ \max_{1 \leq |\alpha| \leq \ell} \|\partial^\alpha u\|_{L^\infty(\mathbb{R}^n;\mathbb{R}^k)}, & p = \infty. \end{cases}$$

We use $H^\ell(\mathbb{R}^n, \mathbb{R}^k)$ to denote the space $W^{\ell,2}(\mathbb{R}^n, \mathbb{R}^k)$, we will also use $\mathscr{C}_b^\ell(\mathbb{R}^n, \mathbb{R}^k)$ to denote the space of $\ell-$times bounded continuously differentiable functions for $\ell = 0, 1, \ldots, \infty$.

Finally, we use the usual Fourier transform on $\mathbb{R}^n$,

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) \mathrm{e}^{-\mathrm{i}\langle \xi, x\rangle} \mathrm{d}x$$

for a Schwartz function f, which extends by isometry to all $f \in L^2(\mathbb{R}^n, \mathbb{R}^k)$.

2 Normal forms and scalings

In this section we perform the necessary coordinate change and rescaling that simplifies (1.1).

2.1 Normal forms on the linear part

Recall that by converting to the variable $y = T_0^{-1} x$, the second moment matrix S corresponding to the modified convolution kernel $\tilde{K}(y) = |\det T_0| \mathcal{K}(T_0 y)$ becomes $2I_k$. Upon setting $\tilde{U}(y) = U(T_0 y)$, the linear part $U(x) + \mathcal{K} * U(x)$ of (1.1) now equals $\tilde{U}(y) + \tilde{\mathcal{K}} * \tilde{U}(y)$. Define $\mathcal{T} := I_k + \tilde{\mathcal{K}} *$ and its symbol $\widehat{\mathcal{T}}(\xi) = I_k + \widehat{\tilde{\mathcal{K}}}(\xi)$. We shall transform the matrix $\widehat{\mathcal{T}}(\xi)$ to a "normal form" via the following lemma.

Lemma 2.1 *There exist invertible $k \times k$ matrices P, Q, and a multiplier operator L whose symbol $\widehat{L}(\xi) \in L^\infty(\mathbb{R}^n, \mathbb{R}^{k \times k})$ such that*

$$\widehat{L}(\xi) P \widehat{\mathcal{T}}(\xi) Q = \mathrm{diag}\{\frac{|\xi|^2}{1 + |\xi|^2}, I_{k-1}\}.$$

Moreover, with respect to the usual basis, we can write the L and its symbol in matrix form

$$L = \begin{pmatrix} L_{cc} & L_{ch} \\ L_{hc} & L_{hh} \end{pmatrix}, \qquad \widehat{L}(\xi) = \begin{pmatrix} \widehat{L}_{cc}(\xi) & \widehat{L}_{ch}(\xi) \\ \widehat{L}_{hc}(\xi) & \widehat{L}_{hh}(\xi) \end{pmatrix},$$

where terms with subscript cc denote a scalar, terms with subscript ch a $(k-1)$ dimensional row vector, terms with subscript hc a $(k-1)$ dimensional column vector, terms with subscript hh a $(k-1) \times (k-1)$ matrix. We then have $\widehat{L}_{cc}, \widehat{L}_{ch}, \widehat{L}_{hc}$ are continuous and bounded functions in ξ and $\widehat{L}_{hc} \in L^\infty(\mathbb{R}^n, \mathbb{R}^k)$.

Proof. we divide our construction in 2 steps.

Step 1. First observe that $\int \mathcal{K}(x)\mathrm{d}x = \int \mathcal{K}(T_0 y)|\det T_0|\mathrm{d}y = \int \tilde{\mathcal{K}}(y)\mathrm{d}y$, therefore we still choose $\mathcal{E}_1$ and $\mathcal{E}_1^*$ which spans $\ker(I_k + \int \tilde{\mathcal{K}}) = \ker \widehat{\mathcal{T}}(0)$ and $\ker(I_k + \int \tilde{\mathcal{K}}^T) = \ker \widehat{\mathcal{T}}(0)^T$ by (iii) of Hypothesis (L).

Possessing only a one dimensional kernel, the rank of $\widehat{\mathcal{T}}(0)$ is equal to $k-1$. It is then standard that there exist invertible matrices P and Q such that

$$P\widehat{\mathcal{T}}(0)Q = \mathrm{diag}\{0, I_{k-1}\}.$$

Next, note that the first moment of $\tilde{\mathcal{K}}$ vanishes, $\int y\tilde{\mathcal{K}}(y) = |\det T_0|T_0^{-1}\int x\mathcal{K}(x)dx = 0$ if $x = T_0 y$. By the assumption on finite second moment, the entries of $\widehat{\mathcal{T}}(\xi)$ are $\mathscr{C}^2$ functions in ξ. Taylor expand near $\xi = 0$, we find that

$$P\widehat{\mathcal{T}}(\xi)Q = \begin{pmatrix} \widehat{\mathcal{T}}_{cc}(\xi) & \widehat{\mathcal{T}}_{ch}(\xi) \\ \widehat{\mathcal{T}}_{hc}(\xi) & \widehat{\mathcal{T}}_{hh}(\xi) \end{pmatrix},$$

where $\widehat{\mathcal{T}}_{cc} = |\xi|^2 + o(|\xi|^2)$ exactly due to the normalization condition $\int x_i x_j \langle \mathcal{E}_1^*, \tilde{\mathcal{K}}(x)\mathcal{E}\rangle dx = 2\delta_{ij}$. The off-diagonal components $\widehat{\mathcal{T}}_{ch}(\xi), \widehat{\mathcal{T}}_{hc}(\xi)$ are $k-1$ sized row, column vectors respectively, both are of order $\mathrm{O}(|\xi|^2)$. $\widehat{\mathcal{T}}_{hh}(\xi)$ is a $k \times k$ matrix which has the expansion $I_{k-1} + \mathrm{O}(|\xi|^2)$.

Step 2. Set $H(\xi) = \mathrm{diag}\{\frac{1+|\xi|^2}{|\xi|^2}, I_{k-1}\}$ for $\xi \neq 0$. We find

$$\begin{aligned} P\widehat{\mathcal{T}}(\xi)QH(\xi) &= \begin{pmatrix} \widehat{\mathcal{T}}_{cc}(\xi) & \widehat{\mathcal{T}}_{ch}(\xi) \\ \widehat{\mathcal{T}}_{hc}(\xi) & \widehat{\mathcal{T}}_{hh}(\xi) \end{pmatrix} \begin{pmatrix} \frac{1+|\xi|^2}{|\xi|^2} & \\ & I_{k-1} \end{pmatrix} \\ &= \begin{pmatrix} \widehat{\mathcal{T}}_{cc}(\xi)\frac{1+|\xi|^2}{|\xi|^2} & \widehat{\mathcal{T}}_{ch}(\xi) \\ \widehat{\mathcal{T}}_{hc}(\xi)\frac{1+|\xi|^2}{|\xi|^2} & \widehat{\mathcal{T}}_{hh}(\xi) \end{pmatrix}. \end{aligned}$$

The fact that $\widehat{\mathcal{T}}_{cc}(0) = D_\xi \widehat{\mathcal{T}}_{cc}(0) = 0$ and the normalization assumption on the second moment matrix implies that $\widehat{\mathcal{T}}_{cc}(\xi) = |\xi|^2 T(\xi)$ for some continuous function $T(\xi)$. On the other hand, we have $|\widehat{\mathcal{T}}_{hc}(\xi)|/|\xi|^2 \leq C$ for some constant C near $\xi = 0$. Therefore, for $|\xi| \neq 0$ small, $P\widehat{\mathcal{T}}(\xi)QH(\xi)$ is of the form $\begin{pmatrix} 1 & \cdots 0 \cdots \\ \mathrm{O}(1) & I_{k-1} \end{pmatrix} + o(|\xi|)$. It follows that $P\widehat{\mathcal{T}}(\xi)QH(\xi)$ is invertible with uniform bounds on the inverse near $\xi = 0$, and its inverse is also of the form $\begin{pmatrix} 1 & \cdots 0 \cdots \\ \mathrm{O}(1) & I_{k-1} \end{pmatrix} + o(|\xi|)$, for $\xi \neq 0$ small.

For any other $\xi \neq 0$ finite, it follows from (v) of Hypothesis (L), invertibility for nonzero wavenumbers, that $P\widehat{\mathcal{T}}(\xi)QH(\xi)$ is invertible for each ξ. Moreover, as $|\xi| \to \infty$, $\widehat{\mathcal{T}}(\xi) \to I_k$ by Riemann-Lebesgue and $H(\xi) \to I_k$. We conclude therefore that $P\widehat{\mathcal{T}}(\xi)QH(\xi)$ is uniformly invertible on $\mathbb{R}^n \backslash \{0\}$.

We then define the multiplier L by setting its symbol $\widehat{L}(\xi)$ equal to $[P\widehat{\mathcal{T}}(\xi)QH(\xi)]^{-1}$. Then $\widehat{L}(\xi) \in L^\infty$ and it follows that

$$L(\xi)P\widehat{\mathcal{T}}(\xi)Q = H(\xi)^{-1}[P\widehat{\mathcal{T}}(\xi)Q]^{-1}P\widehat{\mathcal{T}}(\xi)Q = H(\xi)^{-1} = \mathrm{diag}\{\frac{|\xi|^2}{1+|\xi|^2}, I_{k-1}\},$$

which is as stated in the lemma.

If we write

$$L = \begin{pmatrix} L_{cc} & L_{ch} \\ L_{hc} & L_{hh} \end{pmatrix}, \quad \widehat{L}(\xi) = \begin{pmatrix} \widehat{L}_{cc}(\xi) & \widehat{L}_{ch}(\xi) \\ \widehat{L}_{hc}(\xi) & \widehat{L}_{hh}(\xi) \end{pmatrix},$$

then from above computations, it follows that the entries $\widehat{L}_{cc}, \widehat{L}_{ch}, \widehat{L}_{hh}$ are continuous at $\xi = 0$ and $\mathscr{C}^2$ for $\xi \neq 0$, while $\widehat{L}_{hc}$ is only bounded in ξ. This verifies the last claim in the lemma and concludes the proof. ∎

Since $\widehat{L} \in L^\infty$, we know the multiplier L maps H^ℓ in H^ℓ and is bounded. Denote by M the multiplier operator with symbol $|\xi|^2/(1+|\xi|^2) := m(\xi)$, set $V(y) = Q^{-1}U(y)$, with standard coordinates $V(y) = (v_c(y), v_h(y))^T$.

Then after precondition (1.1) with LP we obtain an equivalent equation

$$\operatorname{diag}(M, I_{n-1})V(y) = LP\mathcal{N}(QV(y);\mu). \tag{2.1}$$

In particular, the linear part of the system decouples. We will preform a series of rescalings to further simplify this equation in the following sections.

2.2 Taylor expansion and rescaling

Taylor jets of the nonlinearity. Write $v_h = (v_2, \ldots, v_k)^T$ in the standard coordinate in $\mathbb{R}^{k-1}$. Set $\mathcal{H}(V;\mu) := P\mathcal{N}(QV;\mu)$. Then with respect to the standard basis in $\mathbb{R}^k$, we denote by $\mathcal{H}_c(V;\mu)$ the first component of the nonlinearity $\mathcal{H}$ and $\mathcal{H}_h(V;\mu)$ the remaining $k-1$ components.

In this notation, system (1.1) becomes

$$Mv_c + L_{cc}\mathcal{H}_c(v_c, v_h;\mu) + L_{ch}\mathcal{H}_h(v_c, v_h;\mu) = 0, \tag{2.2}$$
$$v_h + L_{hc}\mathcal{H}_c(v_c, v_h;\mu) + L_{hh}\mathcal{H}_h(v_c, v_h;\mu) = 0. \tag{2.3}$$

By part (i) of Hypothesis (TC), we may use Taylor's theorem to write $\mathcal{H}_j$ with $j = c, h$ as

$$\begin{aligned}\mathcal{H}_j(v_c, v_h;\mu) &= \Big(a^j_{101}\mu v_c + a^j_{011}\mu v_h + a^j_{110}v_c v_h + a^j_{200}v_c^2 + a^j_{020}[v_h, v_h]\Big) + \mathcal{R}_j(v_c, v_h;\mu)\\ &:= \mathcal{B}_j(v_c, v_h;\mu) + \mathcal{R}_j(v_c, v_h;\mu),\end{aligned}$$

where for the multiindex $\omega = (l, m, n)$ with $|\omega| = 2$, we denoted $a^j_\omega = \dfrac{1}{\omega!}D^\omega\mathcal{H}_j(0,0;0)$, and the remainder R_j satisfies the pointwise estimate

$$|R_j(v_c, v_h;\mu)| = |R_j(V;\mu)| = \mathrm{O}(\mu^2|V| + \mu|V|^2 + |V|^3) \tag{2.4}$$

for $(V;\mu)$ bounded.

We are in particular interested in the terms μv_c and v_c^2. In equation (2.2), the term μv_c is preconditioned by $L_{cc}a^c_{101} + L_{ch}a^h_{101}$, and the coefficient of v_c^2 is preconditioned by $L_{cc}a^c_{200} + L_{ch}a^h_{200}$. Using Hypothesis (TC), we claim that

$$\alpha = a^c_{101} = \widehat{L}_{cc}(0)a^c_{101} + \widehat{L}_{ch}(0)a^h_{101}, \quad \text{and} \quad \beta = a^c_{200} = \widehat{L}_{cc}(0)a^c_{200} + \widehat{L}_{ch}(0)a^h_{200}.$$

Indeed, to verify the first assertion, use the definition of L in lemma 2.1, we find $\widehat{L}_{cc}(0) = 1$ and $\widehat{L}_{ch}(0) = (0, \ldots, 0)$, thus verifying the second equality $a^c_{101} = \widehat{L}_{cc}(0)a^c_{101} + \widehat{L}_{ch}a^h_{101}$. To verify the first equality, let e_1 denote the standard coordinate vector $(1, 0, \ldots, 0)^T \in \mathbb{R}^k$, then the derivative $a^c_{101} = \dfrac{\partial^2}{\partial\mu\partial v_c}\mathcal{H}_c(0,0;0)$ is given by

$$\langle D_{\mu,V}\mathcal{H}(0;0)e_1, e_1\rangle = \langle D_{\mu,U}P\mathcal{N}(0;0)Qe_1, e_1\rangle = \langle PD_{\mu,U}\mathcal{N}(0;0)\mathcal{E}_1, e_1\rangle = \langle D_{\mu,U}\mathcal{N}(0;0)\mathcal{E}_1, \mathcal{E}_1^*\rangle = \alpha,$$

which verifies the first equality $\alpha = a^c_{101}$. The computations for β is similar.

In the next paragraph, we shall make a series of rescalings to simplify the equation.

Rescaling. Recall we have assumed $\alpha\mu > 0$, set now $\tilde{\mu} = \alpha\mu$ and write $\sqrt{\tilde{\mu}} = \varepsilon$. We then rescale the functions v_c, v_h to $\tilde{v}_c, \tilde{v}_h$ through

$$v_c(\cdot) = \frac{-1}{\beta}\varepsilon^2\tilde{v}_c(\varepsilon\cdot), \quad v_h(\cdot) = \varepsilon^2\tilde{v}_h(\varepsilon\cdot).$$

We substitute these variables in equation (2.2) and (2.3), divide the first equation by $(-1/\beta)\varepsilon^4$, the second by ε^2, and then obtain

$$\varepsilon^{-2}M^{\varepsilon}\tilde{v}_c + \sum_{j=c,h} L^{\varepsilon}_{cj}[\tilde{\mathcal{B}}_j(\tilde{v}_c,\tilde{v}_h) + \varepsilon^{-4}\tilde{\mathcal{R}}_j(\tilde{v}_c,\tilde{v}_h;\varepsilon)], \tag{2.5}$$

$$\tilde{v}_h + \sum_{j=c,h} L^{\varepsilon}_{hj}[\varepsilon^2\tilde{\mathcal{B}}_j(\tilde{v}_c,\tilde{v}_h) + \varepsilon^{-2}\tilde{\mathcal{R}}_j(\tilde{v}_c,\tilde{v}_h;\varepsilon)] = 0. \tag{2.6}$$

Note that here equation (2.5) and (2.6) holds pointwise in $z = \sqrt{\mu}y$. Since y is arbitrary, they hold for all $z \in \mathbb{R}^n$. We rather view them as functional equation in $\tilde{v}_c(\cdot)$ and $\tilde{v}_h(\cdot)$.

By property of the Fourier transform, the rescaled linear operators M^{ε} and L^{ε}_j for $j = cc, ch, hc, hh$ are defined to have symbols $m(\varepsilon\xi)$, $\widehat{L}_j(\varepsilon\xi)$ respectively.

The rescaled nonlinear terms $\tilde{\mathcal{B}}_j, \tilde{\mathcal{R}}_j$ for $j = c, h$ are defined through

$$\tilde{\mathcal{B}}_j(u,v) = \frac{a^j_{101}}{\alpha}u + \frac{a^j_{011}}{\alpha}v + a^j_{110}uv + \frac{a^j_{200}}{-\beta}u^2 + a^j_{020}(-\beta)v^2,$$

$$\tilde{R}_j(u,v;\varepsilon) = \mathcal{R}_j\left(\frac{\varepsilon^2 u}{-\beta}, \varepsilon^2 v; \frac{\varepsilon^2}{\alpha}\right).$$

In particular, the coefficient of the term $\tilde{v}_c$ now equals $a^c_{101}/\alpha = 1$, and the coefficient of $\tilde{v}_c^2$ now equals $a^h_{200}/(-\beta) = \beta/(-\beta) = -1$ by the computation we have done in the previous step. As a consequence, we have

$$\tilde{B}_c(\tilde{v}_c,\tilde{v}_h) = \tilde{v}_c - \tilde{v}_c^2 + \mathrm{O}(|\tilde{v}_h| + |\tilde{v}_h|^2 + \tilde{v}_c|\tilde{v}_h|).$$

From the multiplication algebra property and Sobolev embedding of $H^{\ell}(\mathbb{R}^n,\mathbb{R}^k)$ with $\ell > n/2$, for any $u \in H^{\ell}(\mathbb{R}^n), v \in H^{\ell}(\mathbb{R}^n,\mathbb{R}^{k-1})$, and $j = c, h$ we have the following estimates

$$\|\tilde{\mathcal{B}}_j(u,v)\|_{H^{\ell}} \le C\left(\|u\|_{H^{\ell}} + \|v\|_{H^{\ell}} + \|u\|_{H^{\ell}}\|v\|_{H^{\ell}} + \|u\|^2_{H^{\ell}} + \|v\|^2_{H^{\ell}}\right), \tag{2.7}$$

with some constant C. For the remainder terms $\tilde{\mathcal{R}}_j$, we have

$$\|\tilde{\mathcal{R}}_j(u,v;\varepsilon)\|_{H^{\ell}} = \mathrm{O}(\varepsilon^6) \qquad \text{and } \|D_u\tilde{\mathcal{R}}_j(u,v;\varepsilon)\|_{H^{\ell}\to H^{\ell}} = \mathrm{O}(\varepsilon^6) \tag{2.8}$$

as $\varepsilon \to 0$ from (2.4).

For the rescaled linear operator L^{ε}, using the definition of L in Lemma 2.1 and the Fourier transform characterization of H^{ℓ}, we find the following estimates hold

$$\|(L^{\varepsilon}_{cc} - 1)u\|_{H^{\ell}} \to 0, \quad \|L^{\varepsilon}_{ch}v\|_{H^{\ell}} \to 0, \quad \|(L^{\varepsilon}_{hh} - I_{k-1})w\|_{H^{\ell}} \to 0, \quad \text{and } \|L^{\varepsilon}_{hc}v\|_{H^{\ell}} \le C, \tag{2.9}$$

where $u \in H^{\ell}(\mathbb{R}^n), v \in H^{\ell}(\mathbb{R}^n,\mathbb{R}^{k-1}), w \in H^{\ell}(\mathbb{R}^n,\mathbb{R}^k)$ and C is some constant independent of ε.

Now it remains to understand the behavior of the term $\varepsilon^{-2}M^{\varepsilon}v_c$ as $\varepsilon \to 0$. But before that, we first reduce (2.5) and (2.6) to a scalar equation using a fixed point argument in the next subsection. The main result will be proved shortly after.

To further ease notations, we drop the tildes, and still use $v_j, \mathcal{B}_j, \mathcal{R}_j$ $(j = c, h)$ for the same variables after the rescaling.

3 Lyapunov-Schmidt reduction, leading-order ansatz, and corrections

We first solve (2.6) to obtain v_h as a function of v_c by a fixed point argument. We then substitute this function back into equation (2.5) to obtain a scalar equation for v_c and ε, which will be solved again using a fixed point argument.

3.1 Lyapunov Schmidt reduction

We write the left hand side of (2.6) as $\mathcal{G}(v_h; v_c, \varepsilon)$ with $\mathcal{G}$ defined so that

$$\mathcal{G}(v; u, \varepsilon) = v + \sum_{j=c,h} L^{\varepsilon}_{hj} \left(\varepsilon^2 \mathcal{B}_j(u, v; \varepsilon) + \varepsilon^{-2} \mathcal{R}_j(u, v; \varepsilon)\right),$$

using estimates (2.7) and (2.8), we have $\mathcal{G} : H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1}) \times H^{\ell}(\mathbb{R}^n) \to H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1})$ for each $\varepsilon > 0$. Note that we are treating v_c as an additional (Banach space-valued) parameter. The following lemma accomplishes what we were planning to do.

Lemma 3.1 *Fix $r > 0$ not necessarily small, let B_r denote the ball centered at 0 with radius r in $H^{\ell}(\mathbb{R}^n)$, there then exist $\varepsilon_0 > 0$ sufficiently small and a map $\psi(u, \varepsilon) : B_r \times (0, \varepsilon_0) \to H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1})$ such that $v = \psi(u, \varepsilon)$ solves $\mathcal{G}(v; u, \varepsilon) = 0$. Moreover, the map $u \mapsto \psi(u, \varepsilon)$ is smooth for $u \in B_r$, and we have*

$$\|\psi(u, \varepsilon)\|_{H^{\ell}} = \mathrm{O}(\varepsilon^2), \quad \|D_u\psi(u, \varepsilon)\|_{H^{\ell} \to H^{\ell}} = \mathrm{O}(\varepsilon^2),$$

as $\varepsilon \to 0$, uniformly for $u \in B_r$ where $D_u\psi(u, \varepsilon)$ denotes the Frechet derivative of ψ with respect to u at the point (u, ε).

Proof. We will solve $\mathcal{G}(v; u, \varepsilon) = 0$ using a Newton iteration scheme. For $u \in B_r$ and ε_0 small, we claim the following properties hold for $\mathcal{G}$:

(i) $\|\mathcal{G}(0; u, \varepsilon)\|_{H^{\ell}} = \mathrm{O}(\varepsilon^2)$, uniformly in $u \in B_r$ and $\varepsilon < \varepsilon_0$.

(ii) $\mathcal{G}$ is smooth in v, and $D_v\mathcal{G}(0; u, \varepsilon) : H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1}) \to H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1})$ is bounded invertible with uniform bounds on the inverse for $|\varepsilon| < \varepsilon_0$ and $u \in B_r$.

For (i), since L^{ε} is uniformly bounded in ε, there exist a constant C such that $\|L^{\varepsilon}_{hc}\|_{H^{\ell} \to H^{\ell}} + \|L^{\varepsilon}_{hh}\|_{H^{\ell} \to H^{\ell}} \le C$, we then have

$$\begin{aligned}\|\mathcal{G}(0, u; \varepsilon)\|_{H^{\ell}} \le\ & \varepsilon^2 C(\|\mathcal{B}_c(u, 0; \varepsilon)\|_{H^{\ell}} + \|\mathcal{B}_h(u, 0; \varepsilon)\|_{H^{\ell}}) \\ & + \varepsilon^{-4} C(\|\mathcal{R}_c(u, 0; \varepsilon)\|_{H^{\ell}} + \|\mathcal{R}_h(u, 0; \varepsilon)\|_{H^{\ell}}).\end{aligned}$$

Using estimates (2.7) and (2.8), we have $\|\mathcal{G}(0; u, \varepsilon)\|_{H^{\ell}} \le C(r)\varepsilon^2$ uniformly in $u \in B_r$ and ε small.

For (ii), we conclude the smoothness of $\mathcal{G}$ in v by the smoothness of the superposition operator and the fact that L^{ε}_j are bounded linear operators. We compute the Frechet derivative of $\mathcal{G}$ to obtain

$$D_v\mathcal{G}(v; u, \varepsilon)w = w + \sum_{j=c,h} L^{\varepsilon}_{hj}(\varepsilon^2 D_v\mathcal{B}_j(u, v; \varepsilon) + \varepsilon^{-2} D_v\mathcal{R}_j(u, v; \varepsilon))w$$

for $w \in H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1})$. Using estimate (2.8), we see that $D_v\mathcal{G}(0; u, \varepsilon)$ is an $\mathrm{O}(\varepsilon^2)$ perturbation of the identity as an operator on $H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1})$ uniformly for $u \in B_r$. Thus, if ε_0 is small enough, then for all ε with $|\varepsilon| < \varepsilon_0$, we have that $D_v\mathcal{G}(0; u, \varepsilon)$ is bounded invertible with uniform bounds in ε.

After establishing (i) and (ii), fix $\delta > 0$ and $u \in B_r$. Let N_{δ} denote the closed ball of radius δ around 0 in $H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1})$, we introduce a map $\mathcal{S}(\cdot; u, \varepsilon) : H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1}) \to H^{\ell}(\mathbb{R}^n, \mathbb{R}^{k-1})$ as follows

$$\mathcal{S}(v; u, \varepsilon) = v - D_v\mathcal{G}(0; u, \varepsilon)^{-1}[\mathcal{G}(v; u, \varepsilon)].$$

We then find

$$\|\mathcal{S}(0; u, \varepsilon)\|_{H^{\ell}} \le \|D_v\mathcal{G}(0; u, \varepsilon)^{-1}\|_{H^{\ell} \to H^{\ell}} \|\mathcal{G}(0; u, \varepsilon)\|_{H^{\ell}} = \mathrm{O}(\varepsilon^2).$$

Also, $D_v\mathcal{S}(0; u, \varepsilon) = 0$ by definition, and $\mathcal{S}$ is smooth in v by (ii). Therefore, if δ is small and $v \in N_{\delta}$, it then follows that $\|D_v S(v; u, \varepsilon)\|_{H^{\ell} \to H^{\ell}} \le C\delta$ for some constant C independent of δ.

Then we start our iteration with $v_0 = 0$, $v_{n+1} = \mathcal{S}(v_n; u, \varepsilon)$, $n \geq 0$. Suppose by induction $v_k \in N_\delta$ for $1 \leq k \leq n$, then

$$\|v_{n+1} - v_n\|_{H^\ell} \leq C\delta \|v_n - v_{n-1}\|_{H^\ell},$$

by the mean value theorem. Therefore

$$\|v_{n+1}\|_{H^\ell} \leq \frac{C}{1 - C\delta} \|v_1 - v_0\|_{H^\ell} = \frac{C}{1 - C\delta} \|\mathcal{S}(0; u, \varepsilon)\|_{H^\ell}.$$

This implies that for ε small and $u \in B_r$, we have $v_{n+1} \in N_\delta$, and that $\mathcal{S}$ is a contraction for δ sufficiently small. Then, as in the proof of Banach's fixed point theorem, we conclude that $v_n \to v = \psi(u, \varepsilon)$ as $n \to \infty$ and v is a fixed point of $\mathcal{S}$. Note that we automatically get $\|\psi(u, \varepsilon)\|_{H^\ell} = \mathrm{O}(\varepsilon^2)$ uniformly in $u \in B_r$.

To show the smooth dependence of $\psi(u, \varepsilon)$, we note that $\mathcal{G}(v; u, \varepsilon)$ is also smooth in u by Hypothesis (TC). By choosing ε small, the contraction constant for $\mathcal{S}$ can be chosen uniformly in $u \in B_r$. Hence by the uniform contraction principle (see for instance [10], Theorem 1.244), we conclude that ψ depends smoothly on u as well.

Finally, to get the estimate $\|D_u \psi\|_{H^\ell \to H^\ell} = \mathrm{O}(\varepsilon^2)$, we differentiate the equation $0 = \mathcal{G}(\psi(u, \varepsilon); u, \varepsilon)$ in u for $u \in B_r$ to see $D_u \psi$ satisfies the equation

$$D_v \mathcal{G}(\psi(u, \varepsilon); u, \varepsilon) D_u \psi(u, \varepsilon) + D_u \mathcal{G}(\psi(u, \varepsilon); u, \varepsilon) = 0.$$

Now, $D_u \mathcal{G}(v; u, \varepsilon)$ is of the form

$$D_u \mathcal{G}(v; u, \varepsilon) w = \sum_{j=c,h} L^\varepsilon_{hj} (\varepsilon^2 D_u \mathcal{B}_j(u, v; \varepsilon) + \varepsilon^{-2} D_u \mathcal{R}_j(u, v; \varepsilon)) w,$$

hence, for $u \in B_r$ and $v = \psi(u, \varepsilon) \in N_\delta$, we have $\|D_u \mathcal{G}(v; u, \varepsilon)\|_{H^\ell \to H^\ell} = \mathrm{O}(\varepsilon^2)$ again by estimate (2.7) and (2.8).

On the other hand, $D_v \mathcal{G}(v; u, \varepsilon)$ is uniformly invertible in ε for $v = \psi(u, \varepsilon) \in N_\delta$ and $u \in B_r$ as calculated previously. Therefore we can write $D_u \psi(u, \varepsilon) = -[D_v \mathcal{G}]^{-1} D_u \mathcal{G}$ and conclude that

$$\|D_u \psi(u, \varepsilon)\|_{H^\ell \to H^\ell} \leq C(r, \delta) \varepsilon^2.$$

This finishes the proof. ∎

Remark 3.2 *Because of the dependence of the convolution operator $L^\varepsilon_{hc}, L^\varepsilon_{hh}$ on ε is not smooth at $\varepsilon = 0$, we cannot use the usual implicit function theorem directly to solve the equation $\mathcal{G}(v; u, \varepsilon) = 0$. We follow the Newton iteration scheme in [13] to circumvent this problem.*

3.2 Preconditioning the bifurcation equation

Using lemma 3.1, we substitute $v_h = \psi(v_c, \varepsilon)$ into equation (2.5). We obtain the following scalar equation

$$0 = \varepsilon^{-2} M^\varepsilon v_c + \sum_{j=c,h} L^\varepsilon_{cj} \left[B_j(v_c, \psi(v_c, \varepsilon)) + \varepsilon^{-4} \mathcal{R}_j(v_c, \psi(v_c, \varepsilon); \varepsilon) \right]. \tag{3.1}$$

It is now crucial to understand the behavior of the operator M^ε as $\varepsilon \to 0$. Recall that by definition

$$\widehat{M^\varepsilon v}(\xi) = m(\varepsilon\xi) \widehat{v}(\xi) = \frac{|\varepsilon\xi|^2}{1 + |\varepsilon\xi|^2} \widehat{v}(\xi),$$

for any $v \in H^\ell(\mathbb{R}^n, \mathbb{R}^k)$. We then define a new operator $\mathcal{M}^\varepsilon$ through

$$\widehat{\mathcal{M}^\varepsilon v}(\xi) = \frac{m(\varepsilon\xi)}{|\varepsilon\xi|^2} \widehat{v}(\xi) = \frac{1}{1 + |\varepsilon\xi|^2} \widehat{v}(\xi).$$

Since $1/(1+|\varepsilon\xi|^2)$ is a bounded function on $\mathbb{R}^n$, $\mathcal{M}^\varepsilon$ maps $H^\ell(\mathbb{R}^n,\mathbb{R}^k)$ into itself.

For $v \in H^\ell(\mathbb{R}^n,\mathbb{R}^k)$, $(\mathcal{M}^\varepsilon)^{-1}$ is defined through

$$\widehat{(\mathcal{M}^\varepsilon)^{-1}}v(\xi) = \frac{|\varepsilon\xi|^2}{m(\varepsilon\xi)}\widehat{v}(\xi) = (1+|\varepsilon\xi|^2)\widehat{v}(\xi),$$

moreover we have:

$$\begin{aligned}\|((\mathcal{M}^\varepsilon)^{-1}-1)v\|_{H^{\ell-2}} &= \left\|\left(1+|\varepsilon\xi|^2-1\right)\widehat{v}(\xi)(1+|\xi|^2)^{\frac{\ell-2}{2}}\right\|_{L^2} \\ &\le \sup_\ell \left|\frac{|\varepsilon\xi|^2}{1+|\xi|^2}\right| \|\widehat{v}(\xi)(1+\xi^2)^{\frac{\ell}{2}}\|_{L^2} \\ &\le \varepsilon^2\|v\|_{H^\ell}.\end{aligned}$$

Therefore, considered as an operator from $H^\ell(\mathbb{R}^n,\mathbb{R}^k)$ to $H^{\ell-2}(\mathbb{R}^n,\mathbb{R}^k)$, $(\mathcal{M}^\varepsilon)^{-1}$ is well-defined, and we have $\|(\mathcal{M}^\varepsilon)^{-1}v - v\|_{H^{\ell-2}} \to 0$ as $\varepsilon \to 0$ for $v \in H^\ell(\mathbb{R}^n,\mathbb{R}^k)$. This simple observation is central to identifying the leading-order terms and we state it as a lemma.

Lemma 3.3 *The multiplier operator $(\mathcal{M}^\varepsilon)^{-1}$ with symbol $\dfrac{|\varepsilon\xi|^2}{m(\varepsilon\ell)} = 1+|\varepsilon\xi|^2$ is well defined, maps from $H^\ell(\mathbb{R}^n,\mathbb{R}^k)$ into $H^{\ell-2}(\mathbb{R}^n,\mathbb{R}^k)$, and satisfies the estimate*

$$\|(\mathcal{M}^\varepsilon)^{-1} - I\|_{H^\ell\to H^{\ell-2}} = \mathrm{O}(\varepsilon^2). \tag{3.2}$$

Since we are seeking solutions which inherit the symmetry of $\mathcal{K}$, we shall work with the subspace of H^ℓ which respect the symmetry in $\Gamma \subset \mathbf{O}(n)$. So we define

$$H^\ell_\Gamma(\mathbb{R}^n,\mathbb{R}^k) := \{u \in H^\ell(\mathbb{R}^n,\mathbb{R}^k): \quad u(\cdot) = u(\gamma\cdot) \text{ for any } \gamma \in \Gamma.\}$$

We remark that $(\mathcal{M}^\varepsilon)^{-1}$ takes H^ℓ_Γ into $H^{\ell-2}_\Gamma$ since its symbol is radial.

Let now v_* be the unique positive ground sate solution of the equation $\Delta v - v + v^2 = 0$. Using the nondegenracy results for the ground state solution of the equation $\Delta v - v + v^2 = 0$, we can prove the following invertibility lemma, which is crucial in setting up the fixed point argument.

If we now set $\mathcal{L} = -\Delta + 1 - 2v_*$, the linearized operator at the ground state v_*. We then have

Lemma 3.4 *The opearator $\mathcal{L}$ is nondegenerate in the following sense:* $\ker\mathcal{L} \cap L^2_\Gamma(\mathbb{R}^n) = \{0\}$*. Consequently, $\mathcal{L}$ is bounded invertible from $H^\ell_\Gamma(\mathbb{R}^n)$ to $H^{\ell-2}_\Gamma(\mathbb{R}^n)$ for $\ell > n/2$.*

Proof. By [27], any element $\eta(x) \in \ker\mathcal{L}$ is of the form $\langle a, \nabla v_*(x)\rangle$ for some vector $a \in \mathbb{R}^n$. Let $\eta \in \ker\mathcal{L}\cap L^2_\Gamma$ be of this form for some a, we have $\eta(x) = \eta(\gamma x)$ for all $\gamma \in \Gamma$. On the other hand, using radial symmetry of v_*, we have $v_*(\gamma x) = v(x)$ for all γ, differentiate the above equality in x, we have $\nabla v_*(x) = \gamma\nabla v_*(\gamma x)$. Hence we have

$$\langle a, \nabla v_*(x)\rangle = \eta(x) = \eta(\gamma x) = \langle a, \nabla v_*(\gamma x)\rangle = \langle \gamma a, \nabla v_*(x)\rangle.$$

So $\gamma a - a$ is orthogonal to the vector $\nabla v_*(x)$ for any $x \in \mathbb{R}^n$. However, using again the radial symmetry of v_*, we have $\nabla v_*(x) = \frac{V'(|x|)}{|x|}x$ for some scalar function $V = V(r)$ defined for $r \ge 0$. Since $x \in \mathbb{R}^n$ is arbitrary, we conclude that $\gamma a = a$, which implies $a \in \mathrm{Fix}\,\Gamma = \{0\}$.

Next we show that $\mathcal{L}$ is Fredholm with index zero from $H^\ell(\mathbb{R}^n)$ to $H^{\ell-2}(\mathbb{R}^n)$ by writing it as the compact perturbation of an invertible operator. For $h \in H^\ell(\mathbb{R}^n)$, we decompose $\mathcal{L}$ as

$$\mathcal{L}_1 h = -\Delta h + h, \qquad \mathcal{L}_2 h = (2v_*)h.$$

We first show $\mathcal{L}_2 : H^\ell(\mathbb{R}^n) \to H^\ell(\mathbb{R}^n)$ is compact. Fix an integer $k > 0$, let χ_k be a positive smooth cutoff function which equals 1 on the cube $[-k, k]^n$ and vanishes outside of $[-2k, 2k]^n$. Consider the sequence of operators $\mathcal{L}_2^{(k)} : H^\ell(\mathbb{R}^n) \to H^{\ell-2}(\mathbb{R}^n)$ defined by $\mathcal{L}_2^{(k)} h = \chi_k(2v_*)h$. We claim that $\mathcal{L}_2^{(k)}$ are compact operators for each k and converges to $\mathcal{L}_2$ in the operator norm.

To see $\mathcal{L}_2^{(k)}$ are compact for each k, let $\{h_i\}_{i=1}^\infty$ be a bounded sequence of functions in $H^\ell(\mathbb{R}^n)$ with $\|h_i\|_{H^\ell} \le 1$. We want to show that there exist a subsequence $h_{i'}$ so that $\mathcal{L}_2^{(k)} h_{i'} = \chi_k(2v_*)h_{i'}$ is convergent in $H^{\ell-2}(\mathbb{R}^n)$. From the Sobolev embedding $H^\ell(\mathbb{R}^n) \to \mathscr{C}_0^1(\mathbb{R}^n)$, the h_i and their derivatives are uniformly bounded on $\mathbb{R}^n$. Therefore, $\mathcal{L}_2^{(k)} h_i = \chi_k(2v_*)h_i$ is a sequence of functions that is uniformly bounded and equicontinuous on the compact set $[-2k, 2k]^n$, we conclude from the Arzela-Ascoli theorem that there exist a subsequence i' such that $\mathcal{L}_2^{(k)} h_{i'}$ converges uniformly on $[-2k, 2k]^n$. Since $\mathcal{L}_2^{(k)} h_{i'}$ are continuous and have a common compact support $[-2k, 2k]$, it is an convergent sequence in $H^{\ell-2}(\mathbb{R}^n)$ as well. Therefore $\mathcal{L}_2^{(k)}$ is compact for each k.

To see $\mathcal{L}_2^{(k)} \to \mathcal{L}_2$ in the operator norm, fix $h \in H^\ell(\mathbb{R}^n; \mathbb{R})$ with $\|h\|_{H^\ell} = 1$, we have

$$\|(\mathcal{L}_2^{(k)} - \mathcal{L}_2)h\|_{L^2} = \|(\chi_k - 1)(2v_*)h\|_{L^2} \le \sup_{x \in \mathbb{R}^n} (\chi_k(x) - 1)(2v_*(x))\|h\|_{L^2}.$$

Since $\chi_k = 1$ on $[-k, k]^n$ and $v_*(x) \to 0$ as $|x| \to \infty$, it follows that $\sup\limits_{x \in \mathbb{R}^n} (\chi_k(x) - 1)(2v_*(x)) \to 0$ as $k \to \infty$. Thus $\|\mathcal{L}_2^{(k)} - \mathcal{L}_2\|_{H^\ell \to H^{\ell-2}} \to 0$ and we conclude that $\mathcal{L}_2 : H^\ell(\mathbb{R}^n) \to H^{\ell-2}(\mathbb{R}^n)$ is compact.

The fact that $\mathcal{L}_1 = -\Delta + 1$ is invertible from $H^\ell(\mathbb{R}^n)$ to $H^{\ell-2}(\mathbb{R}^n)$ follows directly by looking at its Fourier symbol. This shows $\mathcal{L}$ is the compact perturbation of an invertible operator. It follows that $\mathcal{L}$ is Fredholm with index 0 from $H^\ell(\mathbb{R}^n)$ to $H^{\ell-2}(\mathbb{R}^n)$.

Therefore, when restricted to the subspace $H^\ell_\Gamma(\mathbb{R}^n)$, which has trivial intersection with Ker $\mathcal{L}$ as shown earlier, $\mathcal{L}$ becomes invertible as an operator from $H^\ell_\Gamma(\mathbb{R}^n)$ to $H^{\ell-2}_\Gamma(\mathbb{R}^n)$, this concludes the proof. ■

We now set up and solve the fixed point problem by preconditioning the system with $(\mathcal{M}^\varepsilon)^{-1}$. More precisely, we prove

Proposition 3.5 *Assume $n < 6$ and $\ell > n/2$. If $\varepsilon_1 > 0$ is sufficiently small, then for $0 < \varepsilon < \varepsilon_1$, there exist a family of solutions to* (3.1) *of the form $v_c(\cdot; \varepsilon) = v_*(\cdot) + w(\cdot; \varepsilon)$. Here $w = w(\cdot, \varepsilon) \in H^\ell_\Gamma(\mathbb{R}^n, \mathbb{R}^k)$ is a family of correctors parametrised by ε such that $\|w(\cdot, \varepsilon)\|_{H^\ell} \to 0$ as $\varepsilon \to 0$.*

Proof. We substitute the ansatz $v_c = v_* + w$ into (3.1), where v_* is as stated in the lemma and $w \in H^\ell_\Gamma$. We will determine an equation for w and ε and show that it can be solved using Newton iteration scheme near $(w, \varepsilon) = (0, 0)$. First, for the term $\varepsilon^{-2} M^\varepsilon v_c$ with $v_c \in H^\ell$, we apply Fourier transform to obtain

$$\varepsilon^{-2} m(\varepsilon\xi)\widehat{v}_c(\xi) = -\frac{m(\varepsilon\xi)}{|\varepsilon\xi|^2}(-|\xi|^2)\widehat{v}_c(\xi) = -\widehat{\mathcal{M}^\varepsilon \Delta v_c},$$

and equation (3.1) becomes

$$0 = -\mathcal{M}^\varepsilon \Delta v_c + \left(L^\varepsilon_{cc} + L^\varepsilon_{ch} a^h_{101}\right) v_c + \left(L^\varepsilon_{cc} a^c_{110} + L^\varepsilon_{ch} a^h_{110}\right) v_c^2 + \mathcal{R}(v_c, \psi; \varepsilon),$$

where $\mathcal{R}(v_c, \psi; \varepsilon)$ contains all the terms of order ε^2 and higher,

$$\mathcal{R}(v_c, \psi; \varepsilon) = \sum_{j=c,h} L^\varepsilon_{cj} \left[\frac{a^j_{011}}{\alpha}\psi + a^j_{110} v_c \psi + a^j_{020}(-\beta)[\psi, \psi] + \varepsilon^{-4}\mathcal{R}_j(v_c, \psi; \varepsilon) \right].$$

Indeed, for w with $v_* + w \in B_r$, we claim that $\mathcal{R}$ satisfies the estimate $\|\mathcal{R}\|_{H^\ell} = \mathrm{O}(\varepsilon^2)$. To see this, we first apply Lemma 3.1 with $r = 2\|v_*\|_{H^\ell}$ to obtain $\psi = \psi(v_* + w, \varepsilon)$ which satisfies $\|\psi(v_* + w, \varepsilon)\|_{H^\ell} = \mathrm{O}(\varepsilon^2)$.

The linear operators $L^\varepsilon_{cc}, L^\varepsilon_{ch}$ are uniformly bounded in ε, so that we have

$$\left\| \sum_{j=c,h} L^\varepsilon_{cj} \left(\frac{a^j_{011}}{\alpha}\psi + a^j_{110} v_c \psi + a^j_{020}(-\beta)[\psi, \psi] \right) \right\|_{H^\ell} \le C(\|\psi\|_{H^\ell} + \|\psi\|^2_{H^\ell}) = \mathrm{O}(\varepsilon^2),$$

for some constant C from Lemma 3.1.

On the other hand, the remainders $\mathcal{R}_c$ and $\mathcal{R}_h$ satisfy $\|\mathcal{R}_c\|_{H^\ell} = \mathrm{O}(\varepsilon^6), \|\mathcal{R}_h\|_{H^\ell} = \mathrm{O}(\varepsilon^6)$ uniformly for v_* and w such that $v_* + w \in B_r$ as $\varepsilon \to 0$ by estimates (2.7) and (2.8). Therefore we conclude that $\|\mathcal{R}(v_c, \psi; \varepsilon)\|_{H^\ell} = \mathrm{O}(\varepsilon^2)$ for $v_c = v_* + w \in B_r$.

Next, add the equation $\Delta v_* - v_* + v_*^2 = 0$ to the right hand side of (3.1) and precondition with the operator $(\mathcal{M}^\varepsilon)^{-1}$. Set $\alpha^\varepsilon = L^\varepsilon_{cc} + \frac{a^h_{101}}{\alpha} L^\varepsilon_{ch}, \beta^\varepsilon = -L^\varepsilon_{cc} + \frac{a^h_{200}}{-\beta} L^\varepsilon_{ch}$ and we find

$$\begin{aligned} 0 &= (\mathcal{M}^\varepsilon)^{-1}\left[(1-\mathcal{M}^\varepsilon)\Delta v_* - \mathcal{M}^\varepsilon \Delta w + \alpha^\varepsilon(v_*+w) - v_* + \beta^\varepsilon(v_*+w)^2 + v_*^2 + \mathcal{R}\right] \\ &= [(\mathcal{M}^\varepsilon)^{-1} - 1]\mathcal{M}^\varepsilon \Delta v_* + (\mathcal{M}^\varepsilon)^{-1}\left[(\alpha^\varepsilon - 1)v_* + (\beta^\varepsilon+1)v_*^2 + \mathcal{R}\right] + \\ &\qquad -\Delta w + (\mathcal{M}^\varepsilon)^{-1}\left[\alpha^\varepsilon w + \beta^\varepsilon(2v_* w + w^2)\right], \\ &\qquad := F_1(w;\varepsilon) + F_2(w;\varepsilon) := F(w;\varepsilon). \end{aligned} \tag{3.3}$$

By Lemma 3.3, we have that F maps $H^\ell_\Gamma(\mathbb{R}^n)$ to $H^{\ell-2}_\Gamma(\mathbb{R}^n)$. Our goal is to set up a Newton iteration scheme to solve $F(w,\varepsilon) = 0$ for w in terms of ε as a fixed point problem.

Following the strategy of Lemma 3.1, we shall show

(i) $\|F(0,\varepsilon)\|_{H^{\ell-2}} \to 0$ as $\varepsilon \to 0$.

(ii) $F(w,\varepsilon)$ is continuously differentiable in w and $D_w F(0,\varepsilon) : H^\ell_\Gamma(\mathbb{R}^n) \to H^{\ell-2}_\Gamma(\mathbb{R}^n)$ is uniformly invertible in ε.

For (i), we note that

$$F(0,\varepsilon) = F_2(0;\varepsilon) = [(\mathcal{M}^\varepsilon)^{-1} - 1]\mathcal{M}^\varepsilon d\Delta v_* + (\mathcal{M}^\varepsilon)^{-1}[(\alpha^\varepsilon - 1)v_* + (\beta^\varepsilon + 1)v_*^2 + \mathcal{R}(v_*, \psi; \varepsilon)].$$

By [reference on ground state], $\Delta v_* \in H^\ell(\mathbb{R}^n)$ for all ℓ, since $\mathcal{M}^\varepsilon$ take $H^\ell(\mathbb{R}^n)$ into itself and is uniformly bounded in ε, we conclude from Lemma 3.3 that $\|[(\mathcal{M}^\varepsilon)^{-1} - 1]\mathcal{M}^\varepsilon \Delta v_*\|_{H^\ell} \to 0$ as $\varepsilon \to 0$.

Moreover, by (2.9), it holds that $\|\alpha^\varepsilon v - v\|_{H^\ell} \to 0$ and $\|\beta^\varepsilon v + v\|_{H^\ell} \to 0$ as $\varepsilon \to 0$ for any $v \in H^\ell(\mathbb{R}^n)$, and the remainder $\mathcal{R}(v_*, \psi; \varepsilon)$ satisfies $\|\mathcal{R}\|_{H^\ell} = \mathrm{O}(\varepsilon^2)$ as proved earlier. Hence, we conclude that

$$\|F(0;\varepsilon)\|_{H^{\ell-2}} = \|F_2(0;\varepsilon)\|_{H^{\ell-2}} \to 0$$

as $\varepsilon \to 0$, which proves (i).

For (ii), we first verify that F is continuously differentiable in w from $H^\ell(\mathbb{R}^n)$ to $H^{\ell-2}(\mathbb{R}^n)$. Indeed, take $h, w_0 \in H^\ell(\mathbb{R}^n)$ with w_0 fixed. We observe that $D_w F(w_0;\varepsilon)h : H^\ell(\mathbb{R}^n) \to H^{\ell-2}(\mathbb{R}^n)$ is given by

$$D_w F(w_0;\varepsilon)h = -\Delta h + (\mathcal{M}^\varepsilon)^{-1}\left[(a^\varepsilon h) + 2v_*\beta^\varepsilon h + 2w_0 h) + D_w \mathcal{R} h\right],$$

which depends continuously in w_0 since the superposition operator induced by $\mathcal{N}$ is of class $\mathscr{C}^1$.

Now, at $w_0 = 0$, we see that, $D_w F(0;\varepsilon)h \to -\Delta h + h - 2v_* h = \mathcal{L}h$ in $H^{\ell-2}(\mathbb{R}^n)$ as $\varepsilon \to 0$ for $h \in H^\ell$ because $\|D_w \mathcal{R} h\|_{H^\ell} = \mathrm{O}(\varepsilon^2)$ as remarked earlier. By Lemma 3.4, the operator $\mathcal{L} : H^\ell_\Gamma(\mathbb{R}^n) \to H^{\ell-2}_\Gamma(\mathbb{R}^n)$ is bounded invertible. We notice that $D_w F(0;\varepsilon)$ respects the symmetry and is a small perturbation of $\mathcal{L}$, therefore invertible with uniform bounds on the inverse for ε small enough. This shows (ii).

We now set up the Newton iteration scheme, define $\tilde{\mathcal{S}}$ through

$$\tilde{\mathcal{S}}(w;\varepsilon) = w - D_w F(0;\varepsilon)^{-1}[F(w;\varepsilon)].$$

Note that $\tilde{\mathcal{S}}$ respects the symmetry as well: $\tilde{\mathcal{S}} : H^\ell_\Gamma(\mathbb{R}^n) \to H^{\ell-2}_\Gamma(\mathbb{R}^n)$. Therefore we can proceed as in Lemma 3.1 to obtain $w = w(\varepsilon)$ which solves $F(w(\varepsilon);\varepsilon) = 0$ for ε small enough and satisfies $\|w(\varepsilon)\|_{H^\ell} \to 0$ as $\varepsilon \to 0$. ■

Finally, we prove Theorem 1.

Proof. [Of theorem 1] We now write the tildes for the rescaled variables. From proposition 3.5, we know that (3.1) has a solution of the form $\tilde{v}_c(\cdot) = v_*(\cdot) + w(\cdot;\varepsilon)$. Together with $\tilde{v}_h = \psi(\tilde{v}_c, \varepsilon)$, reverting the rescaling, we obtain $v_c(\cdot) = -\frac{\alpha}{\beta}\mu\tilde{v}_c(\sqrt{\alpha\mu}\cdot)$ and $v_h(\cdot) = \alpha\mu\tilde{v}_h(\sqrt{\alpha\mu}\cdot)$ as solutions to (2.2) and (2.3).

Now, recall that $V = (v_c, v_h)^T$ and the original variable U are related by $U = QV$ where Q is defined in the proof of Lemma 2.1. We conclude that $U(\cdot) = v_c(\cdot)\mathcal{E}_0 + v_\perp(\cdot)$, where $v_\perp$ takes values in the complement of $\mathcal{E}_0$. The behavior of $v_c, v_\perp$ as $\mu \to 0$ is a direct consequence of Lemma 3.1 and Proposition 3.5. Lastly, we restore to the original variable $x = T_0 y$, thus getting the desired form of the bifurcating solution.

■

4 Applications

We outline how our main result applies rather immediately to a variety of specific model equations.

Neural fields. In the simplest setup, neural field equations involve nonlocal coupling through a sigmoidal response function S, with dynamics

$$u_t = -u + \mathcal{K} * S(u;\mu),$$

where $\mathcal{K}$ is a not necessarily positive convolution kernel sampling input from firing neighboring neurons, and S is the firing rate depending on the state u of neurons, typically a sigmoidal, strictly montone function [7]. The sign of $\mathcal{K}$ may change depending on an excitatory or inhibitory coupling. Vector-valued generalizations of the equations have been proposed to model functionally different populations of neurons.

Looking for stationary spike-like solutions of this equation, we set $u_t = 0$ and substitute $U = S(u;\mu)$, with inverse $u = \Psi(U;\mu)$, obtaining

$$-U + \mathcal{K} * U + (U - \Psi(U;\mu)) = 0,$$

which is of the form (1.1). Assuming $\int \mathcal{K} = 1$, saddle-node bifurcations occur when the nonlinearity $U - \Psi(U;\mu)$ has a double zero, which is equivalent to a double zero in $-u + S(u;\mu)$. Hypotheses (L) and (TC) directily translate into assumptions on $\mathcal{K}$ and S. A generic saddle-node bifurcation can be easily transformed into a transcritical bifurcation as outlined in the discussion following Hypothesis (TC).

The spikes constructed in this fashion would be expected to be unstable, of Morse index 1, their stable manifold separating spatially uniformly quiescent and spatially uniformly excited populations of neurons.

Material science. Phase separation in multi-component alloys has been modeled by free energy functionals for concentrations of species $W(u)$, $u \in \mathbb{R}^k$, together with a local or nonlocal interaction term. For nonlocal interactions, together with an H^{-1}-gradient flow, one obtains nonlocal Cahn-Morral systems

$$u_t = -\Delta(-u + J * u - W'(u));$$

see [15]. Equilibria satisfy

$$-u + J * u - W'(u) = \mu,$$

with chemical potential $\mu \in \mathbb{R}^k$. Saddle-node bifurcations in $W'(u) + \mu = 0$ now lead to bifurcation of spikes as constructed here. We also note that anisotropic versions, respecting discrete crystallographic symmetries Γ have also been proposed, at least in a context of local coupling [26].

Dispersive solitary waves. In a slightly different direction, nonlocal coupling can encode dispersion, such as in modles for shallow wavter waves generalizing KdV

$$u_t = (Mu - u^2)_x,$$

where Mu is a nonlocal pseudo-differential operator generalizing ∂_{xx} in the KdV equation. Traveling waves satisfy

$$Mu + cu - u^2 = 0,$$

which in the case of $M = I_k + \mathcal{K}*$ reduces to the problem studied here [8]. Other examples include systems of nonlinear Schrödinger equations with nonlocal dispersion,

$$\mathrm{i}v_t = -v + \mathcal{K} * v + N(v) \in \mathbb{R}^k,$$

with possible nonlocal nonlinear dispersion $N(v)$. Here, the simplest case $v = \mathrm{e}^{\mathrm{i}\mu t}u,\ u \in \mathbb{R},\ N(u) = u|u|^2$, leads to

$$-u + \mathcal{K} * u - \mu u + u^3 = 0,$$

which is amenable to the analsysis presented here, slightly changing scalings to account for the cubic nonlinearity.

5 Discussion

We presented a direct and simple approach to the bifurcations of localized spikes in nonlocally coupled systems. While somewhat simpler and more general than approaches based on spatial dynamics, it does not offer insight into uniqueness questions, and, arguably, relies on an a priori understanding of the resulting phenomena. Our assumptions appear to be sharp in terms of localization of the convolution kernel. Interesting questions arise when studying kernels with less smoothness; see for instance [2] for a numerical exploration of kernel regularity on phenomena.

Large μ. Of course, the bifurcation theoretic approach here is limited to small values of the bifurcation parameter μ. Quite different phenomena are to be expected for large values of μ as the simple scalar example

$$-u + \mathcal{K} * u + \mu u - u^2 = 0, \tag{5.1}$$

shows. For $\mu = 1/\varepsilon$, we can scale $\varepsilon u = v$ and obtain

$$v - v^2 + \varepsilon(-v + \mathcal{K} * v) = 0.$$

At $\varepsilon = 0$, we have solutions $v(x) = 1$ for $x \in \Omega$, $v(x) = 0$ otherwise, for any measurable Ω. The linearization at such solutions in $L^\infty(\mathbb{R}^n)$ is invertible as a multiplication operator with values ± 1, and the solutions therefore can be continued in ε. This plethora of solution does of course not exist in the case of diffusive, local coupling; see [5]. The transition can in some cases be understood as a depinning transition of interfaces as analyzed in [2].

Tail expansions. At leading order, the solutions we find here are exponentially localized, with exponential rate of order $\sqrt{\mu}$. As is clear from the simple example (5.1), the actual solution will typically not be exponentially localized. To see this, it suffices to observe that $-u + \mu u - u^2$ is exponentially localized when u is, but $\mathcal{K} * u$ is not, ofr instance when both u and $\mathcal{K}$ are positive, and $\mathcal{K}$ does not decay exponentially.

Somewhat more precisely, our analysis shows that the leading order correction to the solution of

$$-u + \mathcal{K} * u - \mu u + u^2 = 0,$$

can be found by substituting $u(x) = \mu u_*(\sqrt{\mu}x) + \mu w(\sqrt{\mu}x)$, $u_*'' - u_* + u_*^2 = 0$, finding at leading order

$$w'' - w + 2u_*w = u_*'' - \frac{1}{\mu}(-u_* + \mathcal{K}_{\sqrt{\mu}} * u_*).$$

The right-hand side is algebraically localized with decay behavior dominated by $\mathcal{K}(x)\int u_*$. Solving for w and comparing decay rates shows the same behavior for the corrector $-w$. In principle, one can in this way obtain higher-order algebraic expansions for the decay of u, assuming and expansion for $\mathcal{K}$ in terms of x^{-k}.

Periodic spike patterns. The analysis of the bifurcation towards spatially periodic patterns is much simpler, due to the fact that the convolution acts as a compact perturbation of the identity. Therefore, classic Lyapunov-Schmidt bifurcation analysis will yield bifurcation of spatially periodic patterns, after restricting to appropriate symmetry planforms, such as hexgonal or square lattices in $\mathbb{R}^2$. More interesting and relevant for our technical questions here is the case of large spatial period, say $L = L_0/\sqrt{\mu}$. Imposing such boundary conditions allows for an analysis completely analogous to our analysis here, with linear Fourier multipliers restricted to periodic boundary conditions, thus allowing for the same bounds and convergence estimates. The solutions in the resaled system would be a periodic solution to

$$\Delta u - u + u^2 = 0,$$

which one can obtain, for periods $L_0 > 2\pi$, using a variety of methods, for instance bifurcation and global continuation [20].

We can now repeat the calculation for the next-order corrector w from above and find

TODO

- somehow the corrector should be significant here, in its L-dependence, letting say L_0 to infinity, not sure we can figure this out. periodically concatenating profiles one finds a "mismatch at the gluing points as large as the tails, say $L^{-\beta}$ if that's the decay of $\mathcal{K}$, which then leads to a correction of the amplitude $L^{-2\beta}$ I'd guess. That's what usually happens with exponential tails.

Traveling waves. In some cases, relevant coherent structures may not be stationary in time, such that we need to make assumptions on temporal dynamics. Considering for instance the neural field model

$$u_t = -u + \mathcal{K} * S(u),$$

we find the traveling-wave equation

$$cu_x - u + \mathcal{K} * S(u) = 0.$$

Applying $(1 - c\partial_x)^{-1}$ gives

$$-u + \tilde{K}_c * S(u) = 0, \qquad \tilde{K}_c = (1 - c\partial_x)^{-1} * \mathcal{K}.$$

Assuming that $\mathcal{K}$ is of class $W^{2,1}$, say, $\tilde{\mathcal{K}}_c$ is differentiable in c as a function in $W^{1,1}$, with expansion $\tilde{\mathcal{K}}_c = \mathcal{K} + c\mathcal{K}'$. In the long-wavelength scaling, $\mathcal{K}'$ converges to ∂_x, such that we obtain at leading order, after scaling the model equation

$$u'' + \tilde{c}u' + u - u^2 = 0.$$

This strategy has been carried out in the case of exponentially localized kernels in [14, §4.2], and we expect that the methods here allow for an adaptation to kernels with second moments, only.

Stability. Focusing on the existence problem of stationary solution, we did not dicsuss dynamics in most of the exposition. From that perspective, a first relevant question would be the stability of bifurcating solutions. It is worth noting that an important tool for this analysis, the Evans function [12, 1, 24] is not available for the nonlocal equations considered here. On the other hands, an analysis as in [18, 17], exploiting Evans function analysis for the leading order expansion combined with a perturbation argument as presented here should yield stability information for the type of solutions constructed here for nonlocal equations. More ambitiously, it would be interesting to construct weak interaction manifolds [24, 6, 28], for spikes with algebraic tail decay constructed here.

References

[1] J. Alexander, R. Gardner, and C. Jones. A topological invariant arising in the stability analysis of travelling waves. *J. Reine Angew. Math.*, 410:167–212, 1990.

[2] T. Anderson, G. Faye, A. Scheel, and D. Stauffer. Pinning and unpinning in nonlocal systems. *J. Dynam. Differential Equations*, 28(3-4):897–923, 2016.

[3] F. Andreu-Vaillo, J. M. Mazón, J. D. Rossi, and J. J. Toledo-Melero. *Nonlocal diffusion problems*, volume 165 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI; Real Sociedad Matemática Española, Madrid, 2010.

[4] P. W. Bates. On some nonlocal evolution equations arising in materials science. In *Nonlinear dynamics and evolution equations*, volume 48 of *Fields Inst. Commun.*, pages 13–52. Amer. Math. Soc., Providence, RI, 2006.

[5] P. W. Bates, X. Chen, and A. Chmaj. Equilibria and traveling waves for bistable equations with non-local and discrete dissipation. *Sūrikaisekikenkyūsho Kōkyūroku*, (1178):48–71, 2000. Nonlinear diffusive systems—dynamics and asymptotic analysis (Japanese) (Kyoto, 2000).

[6] P. W. Bates, K. Lu, and C. Zeng. Approximately invariant manifolds and global dynamics of spike states. *Invent. Math.*, 174(2):355–433, 2008.

[7] P. C. Bressloff. Spatiotemporal dynamics of continuum neural fields. *J. Phys. A*, 45(3):033001, 109, 2012.

[8] J. C. Bronski, V. M. Hur, and M. A. Johnson. Modulational instability in equations of KdV type. In *New approaches to nonlinear waves*, volume 908 of *Lecture Notes in Phys.*, pages 83–133. Springer, Cham, 2016.

[9] R. S. Cantrell, C. Cosner, Y. Lou, and D. Ryan. Evolutionary stability of ideal free dispersal strategies: a nonlocal dispersal model. *Can. Appl. Math. Q.*, 20(1):15–38, 2012.

[10] C. Chicone. *Ordinary differential equations with applications*, volume 34 of *Texts in Applied Mathematics*. Springer, New York, second edition, 2006.

[11] S. N. Chow and J. K. Hale. *Methods of bifurcation theory*, volume 251 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York-Berlin, 1982.

[12] J. W. Evans. Nerve axon equations. I. Linear approximations. *Indiana Univ. Math. J.*, 21:877–885, 1971/72.

[13] G. Faye and A. Scheel. Existence of pulses in excitable media with nonlocal coupling. *Adv. Math.*, 270:400–456, 2015.

[14] G. Faye and A. Scheel. Center manifolds without a phase space. *arXiv preprint arXiv:1611.07487*, 2016.

[15] C. P. Grant. Slow motion in one-dimensional Cahn-Morral systems. *SIAM J. Math. Anal.*, 26(1):21–34, 1995.

[16] M. Haragus and G. Iooss. *Local bifurcations, center manifolds, and normal forms in infinite-dimensional dynamical systems*. Universitext. Springer-Verlag London, Ltd., London; EDP Sciences, Les Ulis, 2011.

[17] M. Haragus and A. Scheel. Finite-wavelength stability of capillary-gravity solitary waves. *Comm. Math. Phys.*, 225(3):487–521, 2002.

[18] M. Haragus and A. Scheel. Linear stability and instability of ion-acoustic plasma solitary waves. *Phys. D*, 170(1):13–30, 2002.

[19] B. Jourdain, S. Méléard, and W. A. Woyczynski. Nonlinear SDEs driven by Lévy processes and related PDEs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:1–29, 2008.

[20] H. Kielhöfer. *Bifurcation theory*, volume 156 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2012. An introduction with applications to partial differential equations.

[21] A. Mogilner and L. Edelstein-Keshet. A non-local model for a swarm. *J. Math. Biol.*, 38(6):534–570, 1999.

[22] L. Pitaevskii and S. Stringari. *Bose-Einstein condensation*, volume 116 of *International Series of Monographs on Physics*. The Clarendon Press, Oxford University Press, Oxford, 2003.

[23] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *De Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996.

[24] B. Sandstede. Stability of travelling waves. In *Handbook of dynamical systems, Vol. 2*, pages 983–1055. North-Holland, Amsterdam, 2002.

[25] A. Scheel. Radially symmetric patterns of reaction-diffusion systems. *Mem. Amer. Math. Soc.*, 165(786):viii+86, 2003.

[26] J. E. Taylor and J. W. Cahn. Diffuse interfaces with sharp corners and facets: phase field models with strongly anisotropic surfaces. *Phys. D*, 112(3-4):381–411, 1998. With an appendix by Jason Yunger.

[27] J. Wei and M. Winter. *Mathematical aspects of pattern formation in biological systems*, volume 189 of *Applied Mathematical Sciences*. Springer, London, 2014.

[28] S. Zelik and A. Mielke. Multi-pulse evolution and space-time chaos in dissipative systems. *Mem. Amer. Math. Soc.*, 198(925):vi+97, 2009.