Existence and stability of traveling pulse solutions of the FitzHugh-Nagumo equation

Gianni Arioli [1] and Hans Koch [2]

Abstract. The FitzHugh-Nagumo model is a reaction-diffusion equation describing the propagation of electrical signals in nerve axons and other biological tissues. One of the model parameters is the ratio ϵ of two time scales, which takes values between 0.001 and 0.1 in typical simulations of nerve axons. Based on the existence of a (singular) limit at $\epsilon = 0$, it has been shown that the FitzHugh-Nagumo equation admits a stable traveling pulse solution for sufficiently small $\epsilon > 0$. Here we prove the existence of such a solution for $\epsilon = 0.01$, both for circular axons and axons of infinite length. This is in many ways a completely different mathematical problem. In particular, it is non-perturbative and requires new types of estimates. Some of these estimates are verified with the aid of a computer. The methods developed in this paper should apply to many other problems involving homoclinic orbits, including the FitzHugh-Nagumo equation for a wide range of other parameter values.

1. Introduction and main results

The FitzHugh-Nagumo equation

$$\partial_t w_1 = \partial_x^2 w_1 + f(w_1) - w_2\,, \qquad \partial_t w_2 = \epsilon(w_1 - \gamma w_2)\,, \tag{1.1}$$

models the propagation of electric signals in nerve axons [1,2,3]. Here $w_1(x,t)$ is the voltage inside the axon at position $x \in \mathbb{R}$ and time t. The first equation in (1.1) is Kirchhoff's law, expressing that the change $\partial_x^2 w_1$ of the current $\partial_x w_1$ along the axon is compensated by the currents passing through the cell membrane: a capacitance based current $\partial_t w_1$ and a resistance based current $w_2 - f(w_1)$. The function w_2 describes a part of the transmembrane current that passes through slowly adapting ion channels. The remaining part of the resistor is modeled by a simple cubic

$$f(r) = r(r-a)(1-r)\,, \qquad 0 < a < \tfrac{1}{2}\,. \tag{1.2}$$

Simulations based on this model are abundant, both with analog and digital computers. The choice of parameters $\epsilon, \gamma, a > 0$ depends on the type of system being considered. For our investigation of FitzHugh-Nagumo equation we choose the values

$$\epsilon = \tfrac{1}{100}\,, \qquad \gamma = 5\,, \qquad a = \tfrac{1}{10}\,, \tag{1.3}$$

which produce traveling waves in numerical simulations that look very similar to those observed in biological experiments.

A traveling wave is a solution of the form $w_1(x,t) = \phi_1(x-ct)$ and $w_2(x,t) = \phi_2(x-ct)$. We are interested here only in traveling waves that can be identified as a single localized pulse. There appear to be exactly two values of the velocity c for which such solutions

[1] Department of Mathematics and MOX, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano.
[2] Department of Mathematics, University of Texas at Austin, Austin, TX 78712

exist, one corresponding to a stable "fast" pulse and the other to an unstable "slow" pulse. Both have (singular) limits as $\epsilon \to 0$ that can be computed explicitly. For $\epsilon > 0$ sufficiently small, it is possible to use perturbative methods to prove the existence of the fast pulse [8,9,10,11,17]. This pulse was shown to be stable in [14,15]. The existence and instability of the slow pulse, for sufficiently small $\epsilon > 0$, was proved in [8,9] and [16], respectively. Multiple pulse solutions, turning points, and bifurcations are described in [12,13,19,20]. Further information and references can be found in [18].

The equation (1.1) is usually considered for axons of infinite length. In addition to this case, we also consider a circular axon of length $\ell = 128$. Our first result is the following. Let $\mathbb{S}_\ell = \mathbb{R}/(\ell\mathbb{Z})$.

Theorem 1.1. *The equation (1.1) on $\mathbb{S}_{128} \times \mathbb{R}$ with parameter values (1.3) has a traveling pulse solution with velocity $c = 0.470336308\ldots$ Furthermore, this solution is real analytic and exponentially stable.*

By stability we mean nonlinear stability, as defined below. A proof of this Theorem will be given in Section 4, based in part on estimates that have been carried out by computer.

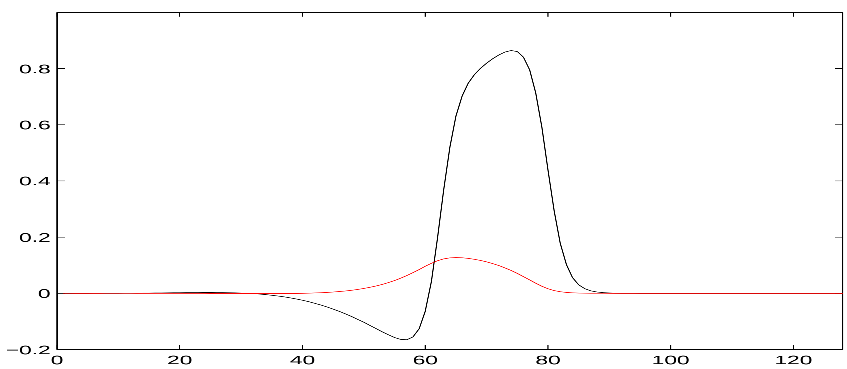

Figure 1. Components ϕ_1 (large) and ϕ_2 (small) of the periodic pulse.

The equation for a traveling pulse solution with velocity c is

$$\phi_1'' + c\phi_1' = -f(\phi_1) + \phi_2\,, \qquad c\phi_2' = -\epsilon(\phi_1 - \gamma\phi_2)\,. \tag{1.4}$$

In the case of a pulse on $\mathbb{R}$, one also imposes the conditions $\phi_j(\pm\infty) = 0$. As usual, a second order equation like (1.4) can be written as a system of first order equations: Using the auxiliary function $\phi_0 = \phi_1'$, we have

$$\phi' = X(\phi)\,, \qquad \phi = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \end{bmatrix}, \quad X(\phi) = \begin{bmatrix} -c\phi_0 - f(\phi_1) + \phi_2 \\ \phi_0 \\ -c^{-1}\epsilon(\phi_1 - \gamma\phi_2) \end{bmatrix}. \tag{1.5}$$

This defines a dynamical system in $\mathbb{R}^3$. For the value of c given in Theorem 1.1, this dynamical system has a periodic orbit of length 128.

Notice that the origin is an equilibrium point, since $X(0) = 0$. It is not hard to check that for every $c > 0$, the derivative $DX(0)$ has one real negative eigenvalue, and two complex conjugate eigenvalues with positive real parts. Thus, the origin has a stable manifold $\mathcal{W}_c^s$ of dimension 1 and an unstable manifold $\mathcal{W}_c^u$ of dimension 2. We will show that there exists a value of c for which these two manifolds intersect at some point other than the origin. (In fact $\mathcal{W}_c^u \subset \mathcal{W}_c^s$, since both manifolds are invariant under the flow.) To be more precise, we prove the following

Theorem 1.2. *The equation (1.1) on $\mathbb{R} \times \mathbb{R}$ with parameter values (1.3) has a traveling pulse solution with velocity $c = 0.470336270\ldots$ Furthermore, this solution is real analytic, decreases exponentially at infinity, and is exponentially stable.*

A proof of this Theorem will be given in Section 4, based in part on estimates that have been carried out by computer. The stability proof involves estimates on the Evans function [7]. For the existence proof we follow the approach used in [22]. The model considered in [22] is a modification of the FitzHugh-Nagumo model, proposed and studied first in [21], which takes into account electro-mechanical effects that are important in tissues such as the heart muscle. Our choice of parameter values (1.3) is the same that was used in [21] for numerical comparisons. A result similar to Theorem 1.2, but with local uniqueness in place of stability, was proved in [33] for a system of equations that descibes the propagation of a pulse in relaxing media.

For the value of c described in Theorem 1.2, the negative eigenvalue of $DX(0)$ is larger in absolute value than the real parts of the two complex conjugate eigenvalues. Under this condition, it was shown in [13] that the existence of a homoclinic pulse implies the existence of a wide variety of multiple pulse solutions for the same velocity[†].

We use the standard notion of stability, described e.g. in [4]. It takes into account that the equation (1.1) is invariant under translations, so that any translate of a solution is also a solution. Let $\underline{\phi} = [\phi_1\ \phi_2]^\top$ be a traveling pulse solution of (1.1) with velocity c. To be more precise, $\underline{\phi}$ satisfies the equation (1.4), and the asymptotic condition $\underline{\phi}(\pm\infty) = 0$ in the case of a pulse on $\mathbb{R}$. For any bounded solution $\underline{\mathrm{w}} = [w_1\ w_2]^\top$ of (1.1) define

$$\delta(\underline{\mathrm{w}}, t, s) = \sup_{j,x} \bigl| w_j(x,t) - \phi_j(x - s - ct) \bigr| . \tag{1.6}$$

The sup in this equation is taken over all $x \in \mathbb{R}$ (or $x \in \mathbb{S}_\ell$ in the periodic case) and $j \in \{1, 2\}$. The pulse $\underline{\phi}$ is said to be *exponentially stable* if there exist constants $\alpha, b, C > 0$ such that for $0 \le \beta \le b$, and for any solution w that satisfies $\delta(\underline{\mathrm{w}}, 0, 0) \le \beta$, there exists $s \in [-C\beta, C\beta]$ such that $\delta(\underline{\mathrm{w}}, t, s) \le C\beta e^{-\alpha t}$ for all $t > 0$. Here, and in what follows, a solution w of (1.1) is always assumed to be uniformly continuous and to have uniformly continuous bounded derivatives $\partial_x w_1$, $\partial_x^2 w_1$ and $\partial_x w_2$.

The remaining part of the paper is organized as follows.

In Section 2 we consider the existence of the pulse solutions. For the periodic pulse, the problem is reduced to solving an appropriate fixed point equation in a space of real analytic periodic functions in the variable $y = x - ct$. The homoclinic pulse is obtained

† We thank Stuart Hastings for bringing this result to our attention.

via Taylor expansions (after a suitable change of variables) at $y = \pm\infty$ whose terms can be computed order by order.

In Sections 3 and 4 we consider the stability of the pulse solutions. As one would expect, the stability problem can be reduced to proving that the generator L_ϕ of the linearized flow has no eigenvalues in a half-plane. But more importantly, we can exclude eigenvalues outside some small region R. In the periodic case, the problem is reduced via perturbation theory to estimates on matrices. In the homoclinic case, we determine the number of eigenvalues in R by estimating the Evans function along the boundary of R and then applying the argument principle.

What remains to be proved at this point are four technical lemmas. In Section 5 we show how this problem is reduced further, to a point where the remaining steps can be mechanized and carried out by a computer. This includes a description of how the mechanical part is organized. The complete details can be found in [37].

2. Existence of pulse solutions

2.1. Existence of a periodic pulse

Our goal here is to reduce the existence part of Theorem 1.1 to a suitable fixed point problem. Let $\eta = \ell/(2\pi)$. Substituting $\phi_1(y) = \varphi(y/\eta)$ and $\phi_2(y) = \psi(y/\eta)$ into the equation (1.4) we get

$$\eta^{-2}\varphi'' + c\eta^{-1}\varphi' = \psi - f(\varphi)\,, \qquad c\eta^{-1}\psi' = \epsilon(\gamma\psi - \varphi)\,. \tag{2.1}$$

We need to find two 2π-periodic functions φ and ψ that satisfy (2.1) for some positive value of c. Notice that $\psi - f(\varphi)$ and $\gamma\psi - \varphi$ must have average zero. Thus,

$$\langle \gamma f(\varphi) - \varphi \rangle = 0\,, \tag{2.2}$$

where $\langle . \rangle$ denotes averaging. As a first step we rewrite (2.1) as an equation for the function φ alone. Let D^{-1} be the antiderivative operator on the space of 2π-periodic continuous functions with average zero. Then $\psi_0 = \psi - \langle\psi\rangle$ can be computed from φ as follows:

$$\begin{aligned} c\psi_0 &= \epsilon\eta D^{-1}(\gamma\psi - \varphi) \\ &= \epsilon\eta D^{-1}\bigl(\gamma[\psi - f(\varphi)] + [\gamma f(\varphi) - \varphi]\bigr) \\ &= \epsilon\gamma\eta D^{-1}[\psi - f(\varphi)] + \epsilon\eta D^{-1}[\gamma f(\varphi) - \varphi] \\ &= \epsilon\gamma[\eta^{-1}\varphi' + c\varphi_0] + \epsilon\eta D^{-1}[\gamma f(\varphi) - \varphi]\,, \end{aligned} \tag{2.3}$$

where $\varphi_0 = \varphi - \langle\varphi\rangle$. Substituting this into the first identity in (2.1) we have

$$\begin{aligned} \eta^{-2}\varphi'' + c\eta^{-1}\varphi' &= \langle\psi\rangle + \psi_0 - f(\varphi) \\ &= \langle\psi\rangle + c^{-1}\epsilon\gamma\eta^{-1}\varphi' + N_c(\varphi_0)\,, \end{aligned} \tag{2.4}$$

where

$$N_c(\varphi_0) = -f(\varphi) + \epsilon\gamma\varphi_0 + \epsilon c^{-1}\eta D^{-1}[\gamma f(\varphi) - \varphi] \,. \tag{2.5}$$

In writing $N_c(\varphi_0)$ instead of $N_c(\varphi)$, we are assuming that the function φ is determined by its zero-average part φ_0 via the condition (2.2). This assumption will be verified later. The equation (2.4) can also be written as

$$\varphi'' + \kappa\varphi' = \eta^2 \mathrm{I}_0 N_c(\varphi_0) \,, \qquad \kappa = \eta\bigl(c - c^{-1}\epsilon\gamma\bigr) \,, \tag{2.6}$$

where $\mathrm{I}_0 h = h - \langle h \rangle$ denotes the zero-average part of a continuous 2π-periodic function. Finally, applying $(D^2 + \kappa D)^{-1}$ to both sides of this equation, we obtain the fixed point equation

$$\varphi_0 = \mathcal{N}_c(\varphi_0) \,, \qquad \mathcal{N}_c(\varphi_0) = \eta^2 \bigl(D^2 - \kappa^2 \mathrm{I}\bigr)^{-1} \bigl(\mathrm{I} - \kappa D^{-1}\bigr) \mathrm{I}_0 N_c(\varphi_0) \,. \tag{2.7}$$

Here we have used the identity $(D^2 + \kappa D)^{-1} = (D^2 - \kappa^2 \mathrm{I})^{-1}(\mathrm{I} - \kappa D^{-1})$. This factorization is convenient with respect to our representation (2.8), since the first factor is diagonal (preserves parity) and the second factor is close to the identity at high frequencies k.

We consider the fixed point equation for $\mathcal{N}_c$ in a space $\mathcal{F}_0 = \mathrm{I}_0\mathcal{F}$ of real analytic functions. To be more precise, given $\rho > 0$, let $\mathcal{S}_\rho$ be the strip in $\mathbb{C}$ defined by $|\mathrm{Im}\,(x)| < \rho$. Every 2π-periodic analytic function on $\mathcal{S}_\rho$ admits a Fourier series

$$h(x) = \sum_{k=0}^{\infty} h_k \cos(kx) + \sum_{k=1}^{\infty} h_{-k} \sin(kx) \,. \tag{2.8}$$

Denote by $\mathcal{F}$ the space of all real analytic functions (2.8) that have continuous extensions to the closure of $\mathcal{S}_\rho$ and a finite norm

$$\|h\| = \sum_{k=-\infty}^{\infty} |h_k| \cosh(\rho k) \,. \tag{2.9}$$

Our reason for considering a space with this norm, as opposed e.g. to the sup-norm on $\mathcal{S}_\rho$, is that the sum (2.9) is easy to estimate from our bounds (5.26) on periodic functions.

Remark 1. As defined above, $\mathcal{F}$ is a Banach space over $\mathbb{R}$. When discussing eigenvectors of linear operators on $\mathcal{F}$, we will also need the corresponding space over $\mathbb{C}$. Since it should be clear from the context which number field is being used, we will denote both spaces by $\mathcal{F}$. Since we are only interested in in real solutions to (1.1), the default field is $\mathbb{R}$. The same applies to the spaces $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{V}$ considered later.

Our goal is to solve the fixed point equation $\mathcal{N}_c(\varphi_0) = \varphi_0$ via a Newton-type procedure. However, there is an obvious obstruction that needs to be taken care of first. Namely, $\mathcal{N}_c$ commutes with translations $\mathcal{T}_s : h \mapsto h(\,\boldsymbol{.} - s)$. Thus, if φ_0 is a fixed point of $\mathcal{N}_c$ then so is any translate $\mathcal{T}_s\varphi_0$. Differentiating the identity $\mathcal{N}_c(\mathcal{T}_s\varphi_0) = \mathcal{T}_s\varphi_0$ with respect to s, we see that $D\mathcal{N}_c(\varphi_0)$ must have an eigenvalue 1,

$$D\mathcal{N}_c(\varphi_0)\varphi_0' = \varphi_0' \,. \tag{2.10}$$

We intend to eliminate this (simple) eigenvalue 1 by projecting onto the orthogonal complement of a function close to φ_0'. The inner product used here is $Q(g,h)=\sum_k g_k h_k$. Notice that Q is invariant under translations, and thus $Q(g,g')=0$ for all $g\in\mathcal{F}$.

Assume that we have found numerically a Fourier polynomial $p_0\in\mathcal{F}_0$ that is an approximate fixed point of $\mathcal{N}_c$ for some value $c_{\rm num}$ of c. Let $p_1\in\mathcal{F}_0$ be a Fourier polynomial close to a constant multiple of p_0', normalized in such a way that $Q(p_1,p_1)=1$. Now define

$$\mathcal{N}_c'(g)=\mathbb{P}\mathcal{N}_c(g)\,,\qquad g\in\mathcal{F}_0\,,\tag{2.11}$$

where $\mathbb{P}$ is the projection operator on $\mathcal{F}$ defined by $\mathbb{P}g=g-Q(g,p_1)p_1$.

By assumption, p_1 is approximately Q-orthogonal to p_0. Thus p_0 is an approximate fixed point of $\mathcal{N}_c'$, for values of c close to $c_{\rm num}$. Furthermore, we can expect $D\mathcal{N}_c'(g)$ to have no eigenvalue 1, for g near p_0 and c near $c_{\rm num}$. In particular, $\mathcal{N}_c'$ can be expected to have a fixed point near p_0, for every value of c near $c_{\rm num}$. The value of c for which a fixed point g_c of $\mathcal{N}_c'$ is also a fixed point of $\mathcal{N}_c$ is determined by the condition

$$Q\bigl(\mathcal{N}_c'(g_c),p_1\bigr)=0\,.\tag{2.12}$$

In order to solve the fixed point problem for $\mathcal{N}_c'$ we consider a quasi-Newton map

$$\mathcal{M}_c(h)=h+\mathcal{N}_c'(p_0+Ah)-(p_0+Ah)\,,\qquad h\in\mathcal{F}_0\,,\tag{2.13}$$

with A an approximation to $[\mathrm{I}-D\mathcal{N}_c'(p_0)]^{-1}$. A solution of the fixed point equation $\mathcal{N}_c'(g)=g$ is then obtained by solving $\mathcal{M}_c(h)=h$ and setting $g=p_0+Ah$.

Given $r>0$ and $h\in\mathcal{F}_0$ we denote by $B_r(h)$ the open ball of radius r in $\mathcal{F}_0$, centered at h. Let $\rho=\log\bigl(1+2^{-10}\bigr)$ and $\delta=2^{-60}$.

Lemma 2.1. *There exist two Fourier polynomials p_0,p_1, an open interval J, and two constants $c_0=0.4703363082\ldots$ and $R>0$, such that the following holds. For every function $g\in B_R(p_0)$ there exists a unique real number $\bar g\in J$, such that $\varphi=g+\bar g$ satisfies the condition (2.2). The resulting map $(c,g)\mapsto\mathcal{N}_c(g)$ is differentiable on some open neighborhood of $I\times B_R(p_0)$, where $I=[c_0-\delta,c_0+\delta]$. The function $c\mapsto Q(\mathcal{N}_c'(g),p_1)$ changes sign on I, for every function $g\in B_r(p_0)$. Furthermore, there exists a continuous linear operator A on $\mathcal{F}_0$, and positive constants K,r,ε satisfying $\varepsilon+Kr<r$ and $\|A\|r\le R$, such that the map $\mathcal{M}_c$ defined by (2.13) satisfies*

$$\|\mathcal{M}_c(0)\|<\varepsilon\,,\qquad\|D\mathcal{M}_c(h)\|<K\,,\qquad h\in B_r(0)\,.\tag{2.14}$$

This lemma, together the contraction mapping theorem, implies the claims in Theorem 1.1 concerning the existence of the periodic pulse. A more detailed argument will be given in Subsection 4.1. Our proof of Lemma 2.1 involves estimates that have been carried out by computer. More details can be found in Section 5.

2.2. Existence of a homoclinic pulse

The equation (1.5) for a pulse solution on $\mathbb{R}$ can be written as

$$\partial_y\phi=DX(0)\phi+B(\phi_1)\,,\tag{2.15}$$

where

$$DX(0)=\begin{bmatrix}-c & a & 1\\ 1 & 0 & 0\\ 0 & -c^{-1}\epsilon & c^{-1}\epsilon\gamma\end{bmatrix}\,,\qquad B(\phi_1)=\begin{bmatrix}-f(\phi_1)-a\phi_1\\ 0\\ 0\end{bmatrix}\,. \tag{2.16}$$

Here we have used that $f'(0)=-a$. Notice that $DB(0)=0$. We restrict the analysis to real velocity parameters c in a ball $|c-c_0|\le\varrho$ of radius $\varrho=2^{-96}$. For these values of c, the matrix $DX(0)$ has a negative real eigenvalue μ_0 and two complex conjugate eigenvalues ν_0 and $\bar\nu_0$ with positive real part. The numerical values are

$$\begin{aligned} c_0 &= 0.4703362702\ldots\,,\\ \mu_0 &= -0.6628889605\ldots\,,\\ \nu_0 &= 0.1494298049\ldots - i*0.1605660907\ldots\,. \end{aligned} \tag{2.17}$$

Our first goal is to construct a solution $\phi=\phi^s$ that satisfies $\phi^s(+\infty)=0$. This solution parametrizes the one-dimensional local stable manifold $\mathcal{W}^s$ of the origin. A similar problem was solved in [22] via the ansatz

$$\phi^s(y)=\ell^s(r)+Z^s(r)\,,\qquad r=e^{\mu_0 y}\,, \tag{2.18}$$

with $\ell^s(r)=r\mathtt{U}_0$ and $Z^s(r)=\mathcal{O}(r^2)$. Here $\mathtt{U}_0$ is the eigenvector of $DX(0)$ associated with the eigenvalue μ_0. The equation (2.15) for $\phi=\phi^s$ can now be written as

$$Z^s=\bigl[\partial_y-DX(0)\bigr]^{-1}B(\ell_1^s+Z_1^s)\,,\qquad \partial_y=\mu_0 r\partial_r\,. \tag{2.19}$$

Notice that the monomials $r\mapsto r^k$ are eigenfunctions of ∂_y. Furthermore, the second component of $\mathtt{U}_0$ is nonzero and thus can be chosen to be 1, making $\ell_1^s(r)=r$. Thus, the equation (2.19) can be solved order by order in powers of r, yielding a unique formal power series solution $Z^s(r)=\sum_{k\ge 2}Z_k r^k$. We will show that this solution is actually analytic in a disk centered at the origin, which includes the point $e^{\mu_0 y_0}$ with $y_0=\frac{5}{2}$. This holds for all values of c in the interval mentioned above.

The stable manifold is now "prolonged" backwards in time, using a simple Taylor integrator: For $n=0,1,\ldots,26$ we determine the Taylor expansion of $\phi^s(y)$ in powers of $y-y_n$ and evaluate the result at y_{n+1}. As described in Section 5, this is done order by order up to a certain degree, and the reminder is estimated rigorously. The endpoint is $y_{27}=-43$, and $\phi^s(-43)$ is again close to the origin.

The next goal is to compute the two-dimensional local unstable manifold. Following again the procedure in [22], we make the ansatz

$$\phi^u(y)=\Phi^u\bigl(Re^{\nu_0 y},Re^{\bar\nu_0 y}\bigr)\,, \tag{2.20}$$

with $R>0$ fixed but arbitrary, and

$$\Phi^u(s)=\ell^u(s)+Z^u(s)\,,\qquad s=(s_1,s_2)\,, \tag{2.21}$$

where $\ell^u(s) = s_1\mathtt{V}_0 + s_2\bar{\mathtt{V}}_0$ and $Z^u(s) = \mathcal{O}(|s|^2)$. Here, $\mathtt{V}_0$ and $\bar{\mathtt{V}}_0$ are eigenvectors of $DX(0)$ for the eigenvalues ν_0 and $\bar{\nu}_0$, respectively. The equation for Z^u is analogous to the equation (2.19) for Z^s, namely

$$Z^u = \bigl[\partial_y - DX(0)\bigr]^{-1} B(\ell_1^u + Z_1^u)\,, \qquad \partial_y = \nu_0 s_1 \partial_{s_1} + \bar{\nu}_0 s_2 \partial_{s_2}\,. \tag{2.22}$$

Notice that the monomials $s \mapsto s_1^k s_2^m$ are eigenfunctions of ∂_y. Furthermore, the second component of $\mathtt{V}_0$ is nonzero and thus can be chosen to be 1, making $\ell_1^u(s) = s_1 + s_2$. Thus, the equation (2.22) can be solved order by order in powers of s_1 and s_2, yielding a unique formal power series solution for Z^u. That is, each component $h = Z_j^u$ of Z^u has a unique expansion

$$h(s) = \sum_{k,m} h_{k,m} s_1^k s_2^m\,, \tag{2.23}$$

with $h_{k,m} = 0$ for $k + m \le 1$. Notice that (2.22) reduces by projection to an equation for the component Z_1^u alone. We will prove that the solution Z^u is real analytic in a polydisk $\mathcal{D}_\rho^2$ characterized by $|s_1|, |s_2| < \rho$. The analysis is carried out in the space $\mathcal{A}_\rho^2$ of all real analytic function h on $\mathcal{D}_\rho^2$ that extend continuously to the closure of $\mathcal{D}_\rho^2$, and which have a finite norm

$$\|h\|_\rho = \sum_{k,m} \|h_{k,m}\| \rho^{k+m}\,. \tag{2.24}$$

To be more precise, our functions $h = Z_j^u$ also depends on the velocity parameter c. The function in $\mathcal{A}_\rho^2$ take values in a space $\mathcal{B}$ of function that are real analytic in the disk $|c - c_0| < \varrho$. The norm of a function $g : c \mapsto \sum_n g_n (c - c_0)^n$ in the space $\mathcal{B}$ is given by $\|g\| = \sum_n |g_n| \varrho^n$. This is the norm that appears in (2.24) for the coefficients $h_{k,m}$.

To emphasize the dependence on the velocity parameter c, we will now include c as an extra argument. So $r \mapsto \phi^s(c, r)$ parametrizes the local stable manifold $\mathcal{W}_c^s$ of the origin, and $s_1 \mapsto \Phi^u(c, s_1, \bar{s}_1)$ is a parametrization of the local stable manifold $\mathcal{W}_c^s$. The two manifolds intersect if the difference

$$\Upsilon(c, \sigma, \tau) = \Phi^u(c, \sigma + i\tau, \sigma - i\tau) - \phi^s(c, -43) \tag{2.25}$$

vanishes for some real values of σ and τ. Let now $\rho = 2^{-5}$, $\varrho = 2^{-96}$, and $R = 1$.

Lemma 2.2. *For $c_0 - \varrho < c < c_0 + \varrho$ the eigenvalues of $DX(0)$ satisfy the bounds (2.17). The functions Z_j^u that define the unstable manifold belong to $\mathcal{A}_\rho^2$. Similarly, $y \mapsto \phi^s(\,.\,, y)$ is a real analytic function on $(-45, +\infty)$ taking values in $\mathcal{B}$. So in particular, Υ is well defined and differentiable on the domain $|c - c_0| < \varrho$ and $|\sigma + i\tau| < \rho$. In this domain there exists a cube where Υ has a unique zero, and $|c - c_0| < 2^{-172}$ for all points in this cube.*

Our proof of this lemma involves estimates that have been carried out by computer. More details will be given in Section 5. Clearly Lemma 2.2 implies the claims in Theorem 1.2 concerning the existence of the periodic pulse.

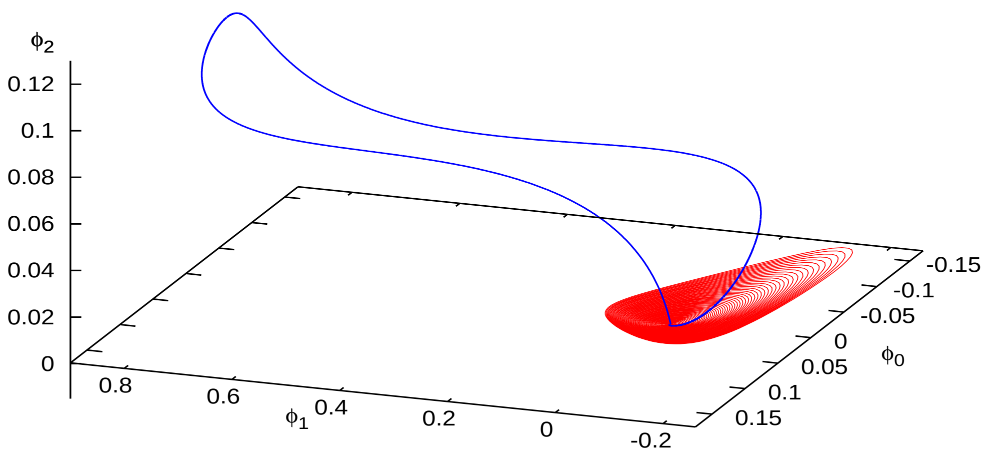

Figure 2. The $1d$ stable manifold ϕ^s and $2d$ local unstable manifold Φ^u.

3. Stability of the pulse solution

Following a common strategy, we reduce the stability problem for the nonlinear equation (1.1) to the problem of proving that the generator L_ϕ of the linearized flow has no spectrum in a half-plane $\mathrm{Re}\,(z) > -\alpha$, except for a simple eigenvalue 0. The spectrum of L_ϕ will be discussed in the next section.

3.1. Reduction to linear stability

It is convenient to rewrite the FitzHugh-Nagumo equation in moving coordinates $y = x - tc$. Substituting $w_j(x,t) = u_j(x - ct, t)$ into (1.1) yields the equation

$$\partial_t \underline{u} = \begin{bmatrix} \partial_y^2 u_1 + c\partial_y u_1 + f(u_1) - u_2 \\ c\partial_y u_2 + \epsilon[u_1 - \gamma u_2] \end{bmatrix}, \qquad \underline{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{3.1}$$

A stationary solution $u_j(y,t) = \phi_j(y)$ of this equation corresponds to a moving pulse solution for the original equation (1.1). The definition of (exponential) stability for (3.1) is analogous to the definition for the original equation, and the two are manifestly equivalent.

The linearization of (3.1) about such a stationary solution is

$$\partial_t \underline{u} = L_\phi \underline{u}\,, \tag{3.2}$$

where

$$L_\phi \underline{u} = \begin{bmatrix} \partial_y^2 + c\partial_y + f'(\phi_1) & -1 \\ \epsilon & c\partial_y - \epsilon\gamma \end{bmatrix} \underline{u}\,, \qquad \underline{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{3.3}$$

The time-t map $\underline{\mathrm{u}}(0) \mapsto \underline{\mathrm{u}}(t)$ will be denoted by e^{tL_ϕ}. Since the equation (3.1) is invariant under translations, $\underline{\phi}'$ is a stationary solution of the equation (3.2). Notice that $\underline{\phi}'$ has the same regularity property as $\underline{\phi}$, as can be seen e.g. from equation (1.5). Similarly for the decay at infinity in the homoclinic case.

For a proper discussion of the equation (3.3) we need to introduce some functions spaces. Let $\mathcal{C}_n$ be the space of all bounded and uniformly continuous functions $h : \mathbb{S} \to \mathbb{R}^n$, equipped with the sup-norm $\|\underline{\mathrm{v}}\| = \sup_{j,y} |v_j(y)|$. Notice that orbits $s \mapsto \underline{\mathrm{v}}(\,\boldsymbol{.} - s)$ for the translation group are continuous in this space. This implies in particular that the C^∞ functions with bounded derivatives are dense in $\mathcal{C}_n$.

In what follows, the space $\mathcal{C}_2$ will also be denoted by $\mathcal{C}$. See Remark 1 concerning the complexification of $\mathcal{C}$. Let $\mathcal{C}'$ to be the the set of all functions $\underline{\mathrm{v}} \in \mathcal{C}$ with the property that the derivatives v_1', v_1'', and v_2' exist and belong to $\mathcal{C}_1$. We say that the stationary solution $\underline{\phi}'$ of the linearized system (3.3) is *exponentially stable* if there exists a continuous linear functional $p : \mathcal{C} \to \mathbb{R}$, and two constants $C, \omega > 0$, such that $\left\|e^{tL_\phi}\underline{\mathrm{v}} - p(\underline{\mathrm{v}})\underline{\phi}'\right\| \le Ce^{-t\omega}\|\underline{\mathrm{v}}\|$ for all $\underline{\mathrm{v}} \in \mathcal{C}'$ and all $t \ge 0$.

Assume now exponential stability of $\underline{\phi}'$ for the linearized system. This stability condition implies in particular that e^{tL_ϕ} extend to a bounded linear operator on $\mathcal{C}$. Assuming also that $t \mapsto e^{tL_\phi}$ is a semigroup, one easily finds that $p(\underline{\phi}') = 1$, and that $p\bigl(e^{tL_\phi}\underline{\mathrm{v}}\bigr) = p(\underline{\mathrm{v}})$ for all $v \in \mathcal{C}$ and all $t \ge 0$. Thus $P_\parallel \underline{\mathrm{v}} = p(\underline{\mathrm{v}})\underline{\phi}'$ defines a projection onto $\mathrm{span}(\underline{\phi}')$ that is invariant under the linearized flow.

The following lemma was proved in [4] for pulse solutions on the real line. Instead of complementing this with a proof for the circle only, we will give a proof that covers both cases.

Lemma 3.1. *Let $\underline{\phi} \in \mathcal{C}'$ be a traveling pulse. That is, $\underline{\phi}$ satisfies the equation (1.4) and vanishes at $\pm\infty$. If $\underline{\phi}'$ is exponentially stable for the linear system (3.3) then the traveling pulse $\underline{\phi}$ is exponentially stable as well.*

Proof. A curve $\underline{\mathrm{u}} = \underline{\mathrm{u}}(t)$ in $\mathcal{C}'$ is a solution of (3.1) if and only if $\underline{\mathrm{v}} = \underline{\mathrm{u}} - \underline{\phi}$ satisfies the equation

$$\partial_t \underline{\mathrm{v}} = L_\phi \underline{\mathrm{v}} + Q(v_1)\,, \qquad Q(v_1) = \begin{bmatrix} f(\phi_1 + v_1) - f(\phi_1) - f'(\phi_1)v_1 \\ 0 \end{bmatrix}. \tag{3.4}$$

The corresponding integral equation with fixed initial condition at time $\tau \ge 0$ is

$$\underline{\mathrm{v}}(\tau + t) = e^{tL_\phi}\underline{\mathrm{v}}(\tau) + \int_0^t e^{(t-s)L_\phi} Q(v_1(\tau + s))\,ds\,, \qquad t \ge 0. \tag{3.5}$$

Here we have used that the orbits $t \mapsto e^{tL_\phi}\underline{\mathrm{c}}$ are continuous for every $\underline{\mathrm{c}} \in \mathcal{C}$. This follows from standard results on semigroups; see also the next subsection. Define

$$\underline{\mathrm{w}} = \underline{\mathrm{v}} + \sigma\underline{\phi}' = \underline{\mathrm{u}} - \underline{\phi} + \sigma\underline{\phi}'\,, \qquad \sigma = -p(\underline{\mathrm{v}}(\tau))\,. \tag{3.6}$$

Here $\underline{\phi}$ and $\underline{\phi}'$ stand for the constant curves $t \mapsto \underline{\phi}$ and $t \mapsto \underline{\phi}'$, respectively. Then the equation (3.5) can be written as

$$\underline{\mathrm{w}}(\tau + t) = e^{tL_\phi} P_\perp \underline{\mathrm{w}}(\tau) + \int_0^t e^{(t-s)L_\phi} Q\bigl(w_1(\tau + s) - \sigma\phi_1'\bigr)\, ds\,, \tag{3.7}$$

where $P_\perp = \mathrm{I} - P_\parallel$. This equation can be regarded as a fixed point problem $\underline{\mathrm{w}} = \mathcal{N}_c(\underline{\mathrm{w}})$ on the space of continuous $\mathcal{C}$-valued functions on a given interval $[\tau, \tau + T]$, equipped with the norm $\|\underline{\mathrm{w}}\| = \sup_{0 \le t \le T} \|\underline{\mathrm{w}}(\tau + t)\|$. To be more precise, $\mathcal{N}$ acts on the affine subspace where $\underline{\mathrm{w}}(\tau)$ is fixed. Let $\underline{\mathrm{w}}_0(\tau + t) = e^{tL_\phi} P_\perp \underline{\mathrm{w}}(\tau)$. Notice that Q is a polynomial with a zero of order 2 at the origin. Thus, it is clear that there exists $C_0 > 0$ such that if $\|\underline{\mathrm{w}}(\tau)\| < \delta$ and $|\sigma| < \delta$, with $\delta > 0$ sufficiently small, then $\mathcal{N}$ is a contraction on the ball $\|\underline{\mathrm{w}} - \underline{\mathrm{w}}_0\| \le C_0\delta$. Assume now that the solution $\underline{\phi}'$ of the linearized system (3.3) is exponentially stable with exponent $\omega > 0$. Then the integral in (3.7) grows at most linearly with t. Thus, the above constant C_0 can be chosen to be independent of T, provided that $T \le \delta^{-1/2}$.

Let $0 < \omega' < \omega$. Assume that

$$\|\underline{\mathrm{u}}(\tau) - \underline{\phi}\| \le \varepsilon\,. \tag{3.8}$$

Then $|\sigma| \le \|p\|\varepsilon$ and $\|\underline{\mathrm{w}}(\tau)\| \le \|P_\perp\|\varepsilon$. Thus, if $\varepsilon > 0$ is sufficiently small, then the solution $\underline{\mathrm{w}}$ of the equation (3.7) satisfies a bound

$$\|\underline{\mathrm{w}}(\tau + t)\| \le C_1\varepsilon\,, \qquad 0 \le t \le C_2\varepsilon^{-1/2}\,. \tag{3.9}$$

Here, and in what follows, $C_1, C_2, \ldots$ are universal positive constants that depend only on the model parameters (1.3), the solution $\underline{\phi}$, the constant in the linear stability condition, and on ω'. Using the bound (3.9) to estimate the right hand side of (3.7), we find that

$$\|\underline{\mathrm{w}}(\tau + t)\| \le C_3 e^{-t\omega}\varepsilon + C_4 t\varepsilon^2 \le C_5 e^{-t\omega'}\bigl[e^{-t(\omega-\omega')} + t\varepsilon e^{t\omega'}\bigr]\varepsilon \le e^{-t\omega'}\varepsilon\,, \tag{3.10}$$

for $0 < \varepsilon \le C_6$ and $C_7 \le t \le C_8|\log\varepsilon|$. This can now be used to estimate the difference

$$\underline{\mathrm{u}} - \mathcal{T}_\sigma\phi = \underline{\mathrm{w}} - \bigl[\,\mathcal{T}_\sigma\underline{\phi} - \underline{\phi} + \sigma\underline{\phi}'\,\bigr]\,, \tag{3.11}$$

where $\mathcal{T}_\sigma$ denotes translation by σ. Since the term in square brackets is bounded in norm by $C_9|\sigma|^2$, we obtain from (3.10) a bound

$$\bigl\|\underline{\mathrm{u}}(\tau + t) - \mathcal{T}_\sigma\underline{\phi}\bigr\| \le e^{-t\omega'}\varepsilon\,, \tag{3.12}$$

for $\varepsilon \le C_{10}$ and $C_{11} \le t \le C_{12}|\log\varepsilon|$.

This procedure can now be iterated: Choose $k > C_{11}$. At step 0 consider the solution $\underline{\mathrm{u}}$ with initial condition $\underline{\mathrm{u}}(0)$ satisfying $\|\underline{\mathrm{u}}(0) - \underline{\phi}\| < \exp(-2k/C_{12})$. At steps $n = 0, 1, 2, \ldots$ we use $\tau = nk$, denote $\underline{\mathrm{u}}$ and ε by $\underline{\mathrm{u}}_n$ and ε_n, respectively, and define $\sigma_{n+1} = p(\underline{\phi} - u_n(nk))$. The initial condition at step $n \ge 1$ is given by $\underline{\mathrm{u}}_n(nk) = \mathcal{T}_{\sigma_n}\underline{\mathrm{u}}_{n-1}(nk)$. By (3.12), and by the translation invariance of the norm in $\mathcal{C}$, we have

$$\varepsilon_n = \|\underline{\mathrm{u}}_n(nk) - \underline{\phi}\| = \|\underline{\mathrm{u}}_{n-1}(nk) - \mathcal{T}_{\sigma_n}\underline{\phi}\| \le \varepsilon_{n-1}e^{-k\omega'}\,. \tag{3.13}$$

Thus the procedure can be iterated ad infinitum. The values ε_n and σ_n converge to 0 at least geometrically with ratio $e^{-k\omega'}$. In particular, if $s=\sum_{j\ge 1}\sigma_j$ then

$$|s|\le C_{13}\varepsilon_0\,,\qquad |s_n-s|\le C_{14}\varepsilon_0 e^{-nk\omega'}\,,\qquad s_n=\sum_{j=1}^{n}\sigma_j\,. \tag{3.14}$$

By the translation invariance of the equation (3.1) we have $\mathcal{T}_{-\sigma_n}\underline{\mathrm{u}}_n(nk+t)=\underline{\mathrm{u}}_{n-1}(nk+t)$ for all $n\ge 1$ and for $0\le t\le k$. Thus, each of the functions $\mathcal{T}_{-s_n}\underline{\mathrm{u}}_n$ extends the original solution $\underline{\mathrm{u}}=\underline{\mathrm{u}}_0$, and

$$\|\underline{\mathrm{u}}(t)-\mathcal{T}_{s_n}\underline{\phi}\|\le\varepsilon_0 e^{-t\omega'}\,,\qquad nk\le t\le (n+1)k\,, \tag{3.15}$$

for $n\ge 1$. For $n=0$ the same holds by (3.9) with an extra factor $C_{15}\ge 1$. Combining this inequality with (3.14) yields the bound

$$\|\underline{\mathrm{u}}(t)-\mathcal{T}_{s}\underline{\phi}\|\le C_{15}\|\underline{\mathrm{u}}(0)-\underline{\phi}\|e^{-t\omega'}\,,\qquad t\ge 0\,.$$

This completes the proof of Lemma 3.1. **QED**

3.2. Reduction to a problem about eigenvalues

The problem considered here is a common one. The sequence of steps in this subsection follows roughly those in [5,6]. But we need to cover both the periodic and the homoclinic case, and our function spaces are not exactly the same as those used in [5,6].

Consider the linear operator $L_\phi:\mathcal{C}'\to\mathcal{C}$ defined by equation (3.3). It is convenient to split

$$L_\phi=L_0+F\,, \tag{3.16}$$

where

$$L_0=\begin{bmatrix} D^2+cD-\theta & -1\\ \epsilon & cD-\epsilon\gamma\end{bmatrix}\,,\qquad F=\begin{bmatrix} f'(\phi_1)+\theta & 0\\ 0 & 0\end{bmatrix}\,, \tag{3.17}$$

and where $\theta=-f'(0)$ unless specified otherwise. Here we have replaced ∂_y by D, since we are now considering functions of one variable only. And $f'(\phi_1)+\theta$ stands for the multiplication operator $v_1\mapsto[f'(\phi_1)+\theta]v_1$. Similarly for the other scalar entries in L_0. Clearly $L_0:\mathcal{C}'\to\mathcal{C}$ is closed, so L_ϕ is closed as well, since F is bounded.

Proposition 3.2. *The operator L_0 generates a* C^0 *(strongly continuous) semigroup on $\mathcal{C}$ that satisfies $\|e^{tL_0}\|\le Ce^{-t\epsilon\gamma}$ for some $C>0$ and all $t\ge 0$.*

Proof. Consider the decomposition $L_0=A_2+A_1+A_0+B$, where

$$A_2=\begin{bmatrix} D^2 & 0\\ 0 & 0\end{bmatrix},\;\; A_1=\begin{bmatrix} cD & 0\\ 0 & cD\end{bmatrix},\;\; A_0=\begin{bmatrix} -\theta & 0\\ 0 & -\epsilon\gamma\end{bmatrix},\;\; B=\begin{bmatrix} 0 & -1\\ \epsilon & 0\end{bmatrix}. \tag{3.18}$$

The operators A_1 and A_2 (with obvious domains) generate the translation group and the heat flow (in the first component), respectively. Both are easily seen to be C^0 contraction semigroups on $\mathcal{C}$. Given that these two semigroups commute, it follows that their product defines a C^0 contraction semigroup whose generator is $H = A_2 + A_1$. The same holds if we change the norm on $\mathcal{C}$ to the equivalent norm $\|\underline{\mathrm{v}}\|' = \sup_y \bigl(\epsilon^{1/2}|v_1(y)|^2 + \epsilon^{-1/2}|v_2(y)|^2\bigr)^{1/2}$. For this new norm, B generates a uniformly continuous group of rotations that are isometries. Since B is bounded, $H + B$ generates a C^0 semigroup as well [26]. In fact, this is again a contraction semigroup, as one can see e.g. from the Trotter product formula $e^{t(H+B)} = \lim_n \bigl(e^{tH/n}e^{tB/n}\bigr)^n$, which holds in the strong sense [27,28]. Using that $\theta > \epsilon\gamma$ for our choice of parameters (1.3), we have $\|e^{tA_0}\underline{\mathrm{v}}\|' \leq e^{-t\epsilon\gamma}\|\underline{\mathrm{v}}\|'$ for all $\underline{\mathrm{v}} \in \mathcal{C}$ and all $t \geq 0$. The assertion now follows by applying Trotter's formula to the sum $L_0 = (H + B) + A_0$. **QED**

Proposition 3.3. *F is compact relative to L_0.*

Proof. In the periodic case where $\mathbb{S}$ is compact, the assertion follows e.g. from the fact that $(D^2 + cD - 2c^2)^{-1}$ is a compact convolution operator on $\mathrm{C}^0(\mathbb{S})$.

Consider now the case $\mathbb{S} = \mathbb{R}$. Given $w \in \mathcal{C}$ define $W = \mathrm{diag}(w, 0)$. Notice that $W : \mathcal{C} \to \mathcal{C}$ is bounded in norm by $\sup_y |w(y)|$. Thus, since $f'(\phi_1(y)) + \theta \to 0$ as $|y| \to 0$, we can find $w \in \mathcal{C}'$ with compact support, such that $\|W - F\| < \varepsilon$, for any given $\varepsilon > 0$. Given that the compact operators on $\mathcal{C}$ constitute a closed subspace of $\mathfrak{B}(\mathcal{C})$, it suffices to prove that any such W is compact relative to L_0. But this follows from the fact that the composition of $(D^2 + cD - 2c^2)^{-1}$ with $h \mapsto wh$ is compact as an operator on $\mathrm{C}^0(\mathbb{S})$. **QED**

Proposition 3.4. *Let $t \mapsto e^{tA}$ be a C^0 semigroup on a Banach space X. Let B and F be bounded linear operators on X. Then $A + B$ generates a C^0 semigroup. If Fe^{tA} is compact for all $t > 0$, then so are $Fe^{t(A+B)}$ and $e^{t(A+F)} - e^{tA}$.*

Proof. The fact that the perturbation of A by a bounded operator B generates a C^0 semigroup is standard [26]. Similarly for the identity

$$Me^{t(A+B)} = Me^{tA} + \int_0^t Me^{(t-s)A}Be^{s(A+B)}\,ds\,, \qquad t > 0\,, \tag{3.19}$$

where M can be any bounded linear operator on X. The integral in this identity is to be understood in the sense of equation (3.20) below. The compactness properties claimed above are a consequence of the following result [29, Theorem 1.3]: Let (Ω, Σ, μ) be a complete positive finite measure space. Let U be a bounded function from Ω to $\mathcal{B}(X)$, such that $s \mapsto U(s)x$ is measurable for every $x \in X$. If $U(s)$ is compact for every $s \in \Omega$ then so is $\int_\Omega U(s)\,d\mu(s)$, where by definition,

$$\biggl[\int_S U(s)\,d\mu(s)\biggr]x = \int_S \bigl[U(s)x\bigr]\,d\mu(s)\,. \tag{3.20}$$

Applying this theorem to the integral in (3.19), with $M = F$, yields the compactness of $Fe^{t(A+B)}$ for $t > 0$. Applying it with $M = \mathrm{I}$, and with B replaced by F, yields the compactness of $e^{t(A+F)} - e^{tA}$ for $t \geq 0$. Here μ is Lebesgue measure on $\Omega = (0,t)$, and we have used the fact that the semigroups involved are strongly continuous. **QED**

Given $\alpha > 0$, denote by H_α the half-plane in $\mathbb{C}$ defined by $\mathrm{Re}\,(z) > -\alpha$.

Proposition 3.5. *Let $0 < \alpha \leq \epsilon\gamma$. Assume that L_ϕ has no spectrum in H_α except for a simple eigenvalue 0. Denote by $P_\perp$ the spectral projection associated with the spectrum of L_ϕ in $\mathbb{C} \setminus \{0\}$. Then for every $\omega < \alpha$ there exists a constant $C_\omega > 0$ such that $\|e^{tL_\phi}P_\perp\| \leq C_\omega e^{-t\omega}$ for all $t \geq 0$.*

Proof. Consider the decomposition $L_0 = A + B$ and $A = A_2 + A_1 + A_0$ as defined by (3.18). Approximating F by localized operators as in the proof of Proposition 3.3, we see that Fe^{tA} is compact for all $t > 0$. Thus Fe^{tL_0} is compact for all $t > 0$ by Proposition 3.4. Applying Proposition 3.4 again, with L_0 in place of A, shows that $e^{tL_\phi} - e^{tL_0}$ is compact for all $t > 0$. By Proposition 3.2 the operator e^{tL_0} has no spectrum outside the disk $|z| \leq e^{-\epsilon\gamma}$. Thus, the spectrum of e^{tL_ϕ} outside this disk consists of isolated eigenvalues only [24].

Let e^λ be an eigenvalue of e^{tL_ϕ} with $\mathrm{Re}\,(\lambda) > -\alpha$. The goal is to show that $\lambda = 0$. Let $L = L_\phi - \lambda\mathrm{I}$. Then e^L has an eigenvector $\underline{\mathrm{v}}$ with eigenvalue 1. Since $e^{tL}\underline{\mathrm{v}} = \underline{\mathrm{v}}$ whenever t is a positive integer, the orbit of $\underline{\mathrm{v}}$ under the flow for L is either a point or an invariant circle. In either case, L has an eigenvalue on the imaginary axis. By our assumption on the spectrum of L_ϕ, this implies that $\mathrm{Re}\,(\lambda) = 0$. Thus the orbit of $\underline{\mathrm{v}}$ under the flow for L_ϕ is either a point or an invariant circle. The circle is excluded, since 0 is the only eigenvalue of L_ϕ on the imaginary axis. Thus $\lambda = 0$.

Let P be the spectral projection for e^{L_ϕ} associated with the eigenvalue 1. Given that the eigenvalue 0 of L_ϕ is simple, the Laplace transform of $t \mapsto e^{tL_\phi}P$ has a pole of order 1 at the origin. Thus P has rank 1, implying that $P = \mathrm{I} - P_\perp$.

Let $0 \leq \omega < \alpha$. Since the spectral radius of $e^{L_\phi}P_\perp$ is bounded above by $e^{-\alpha}$, we have $\|e^{nL_\phi}P_\perp\|^{1/n} \leq e^{-\omega}$ for $n \in \mathbb{N}$ sufficiently large, and thus $\|e^{tL_\phi}P_\perp\| \leq C_\omega e^{-t\omega}$ for some $C_\omega > 0$ and all $t \geq 0$. **QED**

4. Spectral properties

4.1. Bounds on the eigenvalues

In this subsection we consider eigenvalues of L_0 and L_ϕ with eigenfunctions that are square-integrable. Denote by $\mathcal{H}$ the Hilbert space of all $\mathbb{C}^2$-valued functions on $\mathbb{S}$ whose components are square-integrable, equipped with the inner product and norm

$$\langle u, v\rangle = \sum_{j=1}^{2} \int u_j(y)\overline{v_j(y)}\,dy\,, \qquad \|u\| = \langle u, u\rangle^{1/2}\,. \tag{4.1}$$

Let $\mathcal{H}'$ be the subspace of all functions $u \in \mathcal{H}$ that have square-integrable derivatives u_1', u_1'', and u_2'. Clearly $\mathcal{H}'$ is dense in $\mathcal{H}$. Moreover, $L_0 : \mathcal{H}' \to \mathcal{H}$ is closed and $F = L_\phi - L_0$ is bounded. Thus $L_\phi : \mathcal{H}' \to \mathcal{H}$ is closed as well.

Proposition 4.1. *If λ is an eigenvalue of $L_\phi : \mathcal{H}' \to \mathcal{H}$ then*

$$\mathrm{Re}\,(\lambda) \le \Lambda\,, \qquad \Lambda = \sup_r f'(r) = \tfrac{91}{300}\,. \tag{4.2}$$

Proof. To simplify notation we consider a scaling $S : [u_1, u_2]^\top \mapsto [\epsilon^{-1/4}u_1, \epsilon^{1/4}u_2]^\top$. Then $L = S^{-1}L_\phi S$ can be written as

$$L = L^s + L^a, \quad L^s = \begin{bmatrix} D^2 + f'(\phi_1) & 0 \\ 0 & -\epsilon\gamma \end{bmatrix}, \quad L^a = \begin{bmatrix} cD & -\sqrt{\epsilon} \\ \sqrt{\epsilon} & cD \end{bmatrix}. \tag{4.3}$$

Of course the spectra of L_ϕ and L are the same. Notice that L^s is symmetric, and $L^s \le \Lambda$ in the sense of quadratic forms. And the operator L^a is antisymmetric. Thus, if $Lu = \lambda u$ with $\|u\|_2 = 1$, then

$$\mathrm{Re}\,(\lambda) = \tfrac{1}{2}\langle Lu, u\rangle + \tfrac{1}{2}\langle u, Lu\rangle = \langle L^s u, u\rangle \le \langle \Lambda u, u\rangle = \Lambda\,, \tag{4.4}$$

as claimed. **QED**

Proposition 4.2. *For every $\delta > 0$ there exists $\omega > 0$ such that the following holds. If λ is any eigenvalue of $L_\phi : \mathcal{H}' \to \mathcal{H}$, then either* $\mathrm{Re}\,(\lambda) < -\omega$ *or else*

$$\big|\mathrm{Im}\,(\lambda)\big| \le \sqrt{c^2 + \gamma^{-1}}\,\Lambda^{1/2} + \delta\,. \tag{4.5}$$

Proof. Consider the operator L defined in (4.4), which has the same eigenvalues as L_ϕ. Let u be an eigenvector of L with eigenvalue λ. Writing just the first component of the identity $(L - \lambda)u = 0$, we have

$$\big[D^2 + f'(\phi_1) + cD - \lambda\big]u_1 - \sqrt{\epsilon}u_2 = 0\,. \tag{4.6}$$

We may assume that u_1 is nonzero. Taking the inner product of each side with u_1 in this equation, we obtain

$$\big\langle \big[D^2 + f'(\phi_1) + cD - \lambda\mathrm{I}\big]u_1, u_1\big\rangle = \sqrt{\epsilon}\langle u_2, u_1\rangle\,. \tag{4.7}$$

Taking the imaginary part on both sides yields

$$c\,\mathrm{Im}\,\langle Du_1, u_1\rangle - \mathrm{Im}\,(\lambda)\|u_1\|^2 = \sqrt{\epsilon}\,\mathrm{Im}\,\langle u_2, u_1\rangle\,. \tag{4.8}$$

Here we have used that $\langle D^2u_1, u_1\rangle = -\langle Du_1, Du_1\rangle$ is real. Thus,

$$\big|\mathrm{Im}\,(\lambda)\big|\|u_1\|^2 \le c\|u_1\|\|Du_1\| + \sqrt{\epsilon}\|u_1\|\|u_2\|\,. \tag{4.9}$$

In addition we have

$$\begin{aligned} \mathrm{Re}\,(\lambda)\|u\|^2 = \langle L^s u, u\rangle &= \int_{\mathbb{S}} \big[\bar{u}_1 D^2 u_1 + \bar{u}_1 f'(\phi_1)u_1 - \epsilon\gamma\bar{u}_2 u_2\big] \\ &\le -\|Du_1\|^2 + \Lambda\|u_1\|^2 - \epsilon\gamma\|u_2\|^2\,. \end{aligned} \tag{4.10}$$

Assume now that $\mathrm{Re}\,(\lambda) \ge -\omega$ with $0 \le \omega < \epsilon\gamma$. Then from (4.10) we obtain

$$\|Du_1\|^2 + \epsilon\gamma\|u_2\|^2 \le \Lambda\|u_1\|^2 + \omega\|u\|^2\,. \tag{4.11}$$

To simplify notation, let us normalize $\|u_1\| = 1$. Then (4.11) becomes

$$\|Du_1\|^2 + \bigl(\sqrt{\epsilon\gamma - \omega}\,\|u_2\|\bigr)^2 \le \Lambda + \omega\,. \tag{4.12}$$

And the inequality (4.9) can be written as

$$\bigl|\mathrm{Im}\,(\lambda)\bigr| \le c\|Du_1\| + \sqrt{\epsilon}\|u_2\| = \begin{bmatrix} c \\ (\gamma - \omega/\epsilon)^{-1/2} \end{bmatrix}^{\top} \begin{bmatrix} \|Du_1\| \\ \sqrt{\epsilon\gamma - \omega}\,\|u_2\| \end{bmatrix}. \tag{4.13}$$

Applying the Cauchy-Schwarz inequality in $\mathbb{R}^2$ to the product on the right hand side, and using (4.12), yields the bound

$$\bigl|\mathrm{Im}\,(\lambda)\bigr| \le \sqrt{c^2 + (\gamma - \omega/\epsilon)^{-1}}\,(\Lambda + \omega)^{1/2}\,. \tag{4.14}$$

Given that the right hand side depends continuously on ω for $0 \le \omega < \epsilon\gamma$, the assertion follows. **QED**

Remark 2. The operator $Q : u_1 \mapsto f'(\phi_1)u_1$ that appears in the definition of (3.3) of L_ϕ is symmetric on L^2 and bounded above by Λ. This is the only property of Q that was used in the proof of Proposition 4.2. Thus, the bound (4.5) remains valid if we replace Q by $sQ + (1-s)PQP$, where P is any orthogonal projection on L^2 and $0 \le s \le 1$. This fact will be used later.

Remark 3. Using the bounds on c given in Lemma 2.1 and Lemma 2.2, it is easy to check that

$$\sqrt{c^2 + \gamma^{-1}}\,\Lambda^{1/2} < \Theta \stackrel{\text{def}}{=} 0.35745\,. \tag{4.15}$$

Next consider the operator L_0 defined in (3.17), with $\theta \ge 0$ to be specified later. Given that the operator L_0 commutes with translations, it can be diagonalized using the Fourier transform $\mathcal{F} : \mathcal{H} \to \mathcal{H}$. For $\underline{\mathrm{v}} \in \mathcal{H}'$ we have $(\mathcal{F}L_0\underline{\mathrm{v}})(p) = \widetilde{L}_0(p)(\mathcal{F}\underline{\mathrm{v}})(p)$, with

$$\widetilde{L}_0(p) = \begin{bmatrix} -p^2 - \theta & -1 \\ \epsilon & -\epsilon\gamma \end{bmatrix} + icp\,\mathrm{I}\,, \tag{4.16}$$

where $p \in \mathbb{R}$ in the homoclinic case $\mathbb{S} = \mathbb{R}$, and $p \in \frac{2\pi}{\ell}\mathbb{Z}$ in the periodic case $\mathbb{S} = \mathbb{S}_\ell$. The eigenvalues of $\widetilde{L}_0(p)$ are

$$\begin{aligned} \lambda^-(p) &= -p^2 + icp - \theta - R(p^2)\,, \\ \lambda^+(p) &= icp - \epsilon\gamma + R(p^2)\,, \end{aligned} \tag{4.17}$$

where

$$R(p^2) = \frac{1}{2}\Bigl[\sqrt{(p^2 + \theta - \epsilon\gamma)^2 - 4\epsilon} - (p^2 + \theta - \epsilon\gamma)\Bigr]\,. \tag{4.18}$$

Define

$$J(p^2) = -p^2 - \theta - \epsilon\gamma - R(p^2)\,. \tag{4.19}$$

Assumption. Here we assume that $\theta \geq 0$ is chosen sufficiently large, such that the argument of the square root in (4.18) is nonnegative, and such that $J(p^2)^2 > \epsilon$.

Then the following two matrices are well defined and are inverses of each other:

$$\widetilde{M}_0(p)^{\pm 1} = \Bigl(1 - \epsilon J(p^2)^{-2}\Bigr)^{-1/2} \begin{bmatrix} 1 & \pm J(p^2)^{-1} \\ \pm\epsilon J(p^2)^{-1} & 1 \end{bmatrix}. \tag{4.20}$$

In fact, the column vectors of $\widetilde{M}_0(p)$ are the eigenvectors of $\widetilde{L}_0(p)$. Thus

$$\widetilde{M}_0(p)^{-1}\widetilde{L}_0(p)\widetilde{M}_0(p) = \begin{bmatrix} \lambda^-(p) & 0 \\ 0 & \lambda^+(p) \end{bmatrix}. \tag{4.21}$$

The corresponding diagonalization of L_0 is

$$M_0^{-1}L_0M_0 = \begin{bmatrix} D^2 + cD - \theta - R(-D^2) & 0 \\ 0 & cD - \epsilon\gamma + R(-D^2) \end{bmatrix}. \tag{4.22}$$

Notice that $R(-D^2)$ is bounded, since $R(p^2) = \mathcal{O}(p^{-2})$ for large $|p|$. In the periodic case $\mathbb{S} = \mathbb{S}_\ell$, the above shows that $L_0 : \mathcal{H}' \to \mathcal{H}$ has only discrete spectrum, and that the resolvents of L_0 are compact.

4.2. Eigenvalues for the periodic pulse

First we show that every eigenvalue of L_ϕ has an analytic eigenvector. For convenience we perform the same scaling as in Subsection 2.1 to change the periodicity from ℓ to 2π, but the various functions and operators will not be renamed to indicate the scaling. In particular, $\mathcal{C}$ is now a space of 2π-periodic functions. And the rescaled operator $L_\phi = L_0 + F$ is given by

$$L_0 = \begin{bmatrix} c_2D^2 + c_1D - \theta & -1 \\ \epsilon & c_1D - \epsilon\gamma \end{bmatrix}, \qquad F = \begin{bmatrix} f'(\phi_1) + \theta & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.23}$$

with $c_1 = c\eta^{-1}$ and $c_2 = \eta^{-2}$, where $\eta = \ell/(2\pi)$. Of course the spectrum of L_0 and of L_ϕ are not affected by this change of variables. Denote by $\mathcal{V}$ the space of all 2π-periodic functions $\underline{v} : \mathcal{S}_\rho \to \mathbb{R}^2$ whose components v_1 and v_2 belong to the space $\mathcal{F}$ defined after (2.8). We equip $\mathcal{V}$ with the norm $\|\underline{v}\| = \max\{\|v_1\|, \|v_2\|\}$. See Remark 1 concerning the complexification of $\mathcal{V}$.

Let $\mathcal{V}'$ be the set of functions $\underline{v} \in \mathcal{V}$ with the property that v_1', v_1'', and v_2' belong to $\mathcal{F}$. Clearly the operator $L_0 : \mathcal{V}' \to \mathcal{V}$ is closed. Furthermore, F is a bounded linear operator on $\mathcal{V}$, due to the fact that $\mathcal{F}$ is a Banach algebra containing ϕ_1. Thus, $L_\phi : \mathcal{V}' \to \mathcal{V}$ is closed as well.

Proposition 4.3. *$L_\phi : \mathcal{C}' \to \mathcal{C}$ and $L_\phi : \mathcal{V}' \to \mathcal{V}$ have the same eigenvectors.*

Proof. The following applies to both $X = \mathcal{H}$ or $X = \mathcal{V}$. As the computations in the preceding subsection show, $L_0 : X' \to X$ has a set of analytic eigenfunctions whose span is dense in X. The eigenvalues of L_0 have finite multiplicities and accumulate only at infinity. The same is true for the eigenvalues of $L_\phi : X' \to X$, since F is a bounded linear operator on X.

Denote the operators $L_\phi : \mathcal{H}' \to \mathcal{H}$ and $L_\phi : \mathcal{V}' \to \mathcal{V}$ by $L_\phi^{\mathcal{H}}$ and $L_\phi^{\mathcal{V}}$, respectively. Then the resolvents $z \mapsto (z\mathrm{I}-L_\phi^{\mathcal{H}})^{-1}$ and $z \mapsto (z\mathrm{I}-L_\phi^{\mathcal{V}})^{-1}$ are both analytic outside some discrete set Z. Furthermore, the restriction of $(z\mathrm{I}-L_\phi^{\mathcal{H}})^{-1}$ to $\mathcal{V}$ agrees with $(z\mathrm{I}-L_\phi^{\mathcal{V}})^{-1}$, since $z\mathrm{I}-L_\phi^{\mathcal{H}}$ is one-to-one on $\mathcal{H}' \supset \mathcal{V}'$ for $z \in Z$. This carries over to the spectral projections for points in Z by contour integration. Thus, since $\mathcal{V}$ is dense in $\mathcal{H}$, the corresponding spectral subspaces agree. The claim now follows from the fact that $\mathcal{V} \subset \mathcal{C} \subset \mathcal{H}$. **QED**

Our goal is to prove that $L_\phi : \mathcal{V}' \to \mathcal{V}$ has no nonzero eigenvalues in a half-plane $\mathrm{Re}\,(\lambda) > -\alpha$, for some $\alpha > 0$. By Proposition 4.2 and Remark 3, it suffices to prove that every nonzero eigenvalue λ of $L_\phi : \mathcal{V}' \to \mathcal{V}$ with $|\mathrm{Im}\,(\lambda)| \le \Theta$ has a negative real part. We approach this problem using perturbation theory.

The basic idea is to split F into a low-mode approximation and a high-mode remainder. This is best described in terms of the Fourier series

$$\underline{\mathrm{v}}(x) = \sum_{k=0}^{\infty} \underline{\mathrm{v}}_k \cos(kx) + \sum_{k=1}^{\infty} \underline{\mathrm{v}}_{-k} \sin(kx)\,, \quad x \in \mathcal{S}_\rho\,, \tag{4.24}$$

associated with every function $\underline{\mathrm{v}} \in \mathcal{V}$. An operator M on $\mathcal{V}$ that commutes with translations will be called a multiplier operator; it is defined by its symbol: a sequence of 2×2 matrices $\widetilde{M}(k)$, such that $(M\underline{\mathrm{v}})_k = \widetilde{M}(k)\underline{\mathrm{v}}_k$ for all $\underline{\mathrm{v}} \in \mathcal{V}$ and $k \in \mathbb{Z}$.

Let κ be some positive integer, to be specified later. Then the projection onto the "low modes" is the multiplier operator P with symbol $\widetilde{P}(k) = \chi(|k| \le \kappa)\mathrm{I}$, where $\chi(\mathrm{true}) = 1$ and $\chi(\mathrm{false}) = 0$. If κ is chosen sufficiently large, then the operator $L_0 + PFP$ should be a good approximation of L_ϕ. At high modes it agrees with L_0, and the low-mode part is "just" a matrix. With the help of a computer, it is possible to diagonalize this matrix approximately. For practical purposes, it is useful to diagonalize the high-mode part as well. This is achieved by the multiplier operator $M : \mathcal{V} \to \mathcal{V}$ with symbol

$$\widetilde{M}(k) = \chi(|k| \le \kappa)\,\mathrm{I} + \chi(|k| > \kappa)\,\widetilde{M}_0(k/\eta)\,, \tag{4.25}$$

where $\widetilde{M}_0(p)$ is the 2×2 matrix given in (4.20). To be more precise: Here, and in what follows, we choose $\theta = 0$ in the definition (4.23) of the operator L_0. Since the matrices $\widetilde{M}_0(p)$ are being used for $p > \kappa/\eta$ only, the Assumption made after (4.19) is satisfied. Clearly, the operator M is bounded and has a bounded inverse. Thus, $M^{-1}L_\phi M$ has the same spectrum as L_ϕ. As indicated above, we now split $M^{-1}L_\phi M$ into a sum $\mathcal{L}_0 + \mathcal{K}$, where

$$\begin{aligned} \mathcal{L}_0 &= M^{-1}L_0 M + PFP\,, \\ \mathcal{K} &= (\mathrm{I}-P)M^{-1}FM + PM^{-1}FM(\mathrm{I}-P)\,. \end{aligned} \tag{4.26}$$

We also define

$$\mathcal{L}_1 = M^{-1}L_\phi M\,, \qquad \mathcal{L}_s = \mathcal{L}_0 + s\mathcal{K}\,, \qquad 0 \le s \le 1\,. \tag{4.27}$$

The goal is to verify the hypothesis of Proposition 3.5. Let $\vartheta = 2^{-6}$. Let Γ be the shortest path in $\mathbb{C}$ passing through the points $i\Theta$, $i\vartheta$, $-\vartheta$, $-i\vartheta$, and $-i\Theta$, in this order. Define $\Omega \subset \mathbb{C}$ to be the open neighborhood of 0, bounded by Γ and by the set of points z on the imaginary axis with $|z| > \Theta$. As we will describe later, $\mathcal{L}_0$ has exactly one eigenvalue in Ω, and this eigenvalue is simple. By Proposition 4.2 and Remark 3, it suffices to show that $\mathcal{L}_1$ has the same property. By Remark 2, this could fail only if one of the operators $\mathcal{L}_s$ had an eigenvalue on Γ. So the task left is to exclude this possibility as well.

For $z \in \Gamma$ we have

$$z\mathrm{I} - \mathcal{L}_s = \bigl[\mathrm{I} - s\mathcal{K}(z\mathrm{I} - \mathcal{L}_0)^{-1}\bigr](z\mathrm{I} - \mathcal{L}_0)\,. \tag{4.28}$$

Thus, in order to show that z is not an eigenvalue of $\mathcal{L}_s$, it suffices to show that the operator in square brackets has a bounded inverse.

This task is more delicate than it may seem, for the following reason. Notice that $\mathcal{K}$ couples high modes to low modes and vice versa. If $z\mathrm{I} - \mathcal{L}_0$ were uniformly large at high modes, this would be no problem. However, $\mathcal{L}$ has an infinite number of eigenvalues $\lambda^+(p) = icp - \epsilon\gamma + \mathcal{O}(p^{-1})$, and many of them are not far from Γ. Thus, it is crucial to control the coupling terms of $\mathcal{K}$ efficiently.

To this end, we perform yet another conjugacy, by the multiplier operator $\mathcal{U}$ with symbol

$$\widetilde{\mathcal{U}}(k) = \chi\bigl(|k|<\kappa'\bigr)\,\mathrm{I} + \chi\bigl(\kappa'\le|k|<\kappa\bigr)\,U\bigl(\varepsilon e^{r|k|-r\kappa}\bigr) + \chi\bigl(\kappa\le|k|\bigr)\,U(\varepsilon)\,, \tag{4.29}$$

where $U(t) = \left[\begin{smallmatrix} t & 0 \\ 0 & 1 \end{smallmatrix}\right]$. For the constants r and ε in this definition we use the values 0.035 and 0.015, respectively. The constant κ' is determined by the equation $\varepsilon e^{r\kappa'-r\kappa} = 1$.

In view of (4.28), we would like to show that the operator $\mathcal{K}(z\mathrm{I} - \mathcal{L}_0)^{-1}$ has a spectral radius less than 1, for all $z \in \Gamma$. It suffices to prove this for the operator

$$\mathcal{U}\bigl[\mathcal{K}(z\mathrm{I} - \mathcal{L}_0)^{-1}\bigr]\mathcal{U}^{-1} = \bigl(\mathcal{U}\mathcal{K}\mathcal{U}\bigr)\bigl[\mathcal{U}^{-1}(z\mathrm{I} - \mathcal{L}_0)^{-1}\mathcal{U}^{-1}\bigr]\,. \tag{4.30}$$

Lemma 4.4. *The operator $\mathcal{L}_0 : \mathcal{V}' \to \mathcal{V}$ has no spectrum in Ω except for a simple eigenvalue. Furthermore, there exists $\delta > 0$ such that*

$$\|\mathcal{U}\mathcal{K}\mathcal{U}\| < \delta\,, \qquad \bigl\|\mathcal{U}^{-1}(z\mathrm{I} - \mathcal{L}_0)^{-1}\mathcal{U}^{-1}\bigr\| < \delta^{-1}\,, \qquad z \in \Gamma\,. \tag{4.31}$$

Our proof of this lemma is computer-assisted; see Section 5 for more details. The value of δ is approximately $\frac{1}{1800}$.

These estimates show that no eigenvalue of $\mathcal{L}_s$ crosses Γ as s is increased from 0 to 1. The same "homotopy principle" was used also in [30,31], but the operators considered in these papers had only finitely many eigenvalues within any fixed distance of the imaginary axis. This made it possible in [30] to estimate the relevant resolvents along the entire imaginary axis. The same methods would not work here, but what saves the situation are the eigenvalue bounds in Proposition 4.1 and Proposition 4.2.

Using Lemma 4.4 and Lemma 2.1, we can now give a

Proof of Theorem 1.1. Existence of the pulse: The estimates (2.14) ensure that $\mathcal{M}_c$ is a contraction on $B_r(0)$. Thus, $\mathcal{M}_c$ has a unique fixed point $h_c \in B_r(0)$. This holds for every $c \in I$. The differentiability of $(c, g) \mapsto \mathcal{N}_c(g)$ ensures that the map $c \mapsto h_c$ is continuous on I. Define $g_c = p_0 + Ah_c$. Given that the function $c \mapsto Q(\mathcal{N}_c'(g_c), p_1)$ changes sign on the interval I, it follows that (2.12) holds for some $c \in I$. For this value of c, the function g_c is a fixed point of $\mathcal{N}_c$. The function ψ is reconstructed from $\varphi = g_c + \bar{g}_c$ as described in Subsection 2.1. Thus, the existence part of Theorem 1.1 is proved.

Stability: The bound (4.31) implies that the operator $\mathcal{K}(z\mathrm{I} - \mathcal{L}_0)$ has a spectral radius less than 1. Thus, by (4.28), the operator $\mathcal{L}_s$ for $0 \le s \le 1$ has no eigenvalue on Γ. As explained after (4.27), this shows that $\mathcal{L}_\phi : \mathcal{V}' \to \mathcal{V}$ has no spectrum in a half-plane $\mathrm{Re}\,(z) > -\alpha$ with $\alpha > 0$, except for a simple eigenvalue 0. The same holds for $\mathcal{L}_\phi : \mathcal{C}' \to \mathcal{C}$ by Proposition 4.3. By Proposition 3.5 this implies that the solution ϕ' is exponentially stable for the linear system (3.2). The exponential stability of the periodic pulse ϕ now follows from Lemma 3.1. **QED**

4.3. Eigenvalues for the homoclinic pulse

Let $0 < \omega < \epsilon\gamma$, so that L_0 has no spectrum in the half-plane H_ω. The goal is to show that L_ϕ has no nonzero eigenvalues in H_α for some positive $\alpha \le \omega$.

In [7] Evans introduces a function that is analytic on H_ω (in our case) and whose zeros are precisely the eigenvalues of L_ϕ in H_ω. After describing the method and introducing some notation, we will prove exponential stability of the homoclinic pulse based on estimates on the Evans function.

The eigenvalue equation $(L_\phi - \lambda\mathrm{I})\underline{u} = 0$, written in terms of the components u_1 and u_1, is

$$\begin{bmatrix} cD + D^2 + f'(\phi_1) - \lambda & -1 \\ \epsilon & cD - \epsilon\gamma - \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} . \tag{4.32}$$

Setting $u_0 = Du_1$, this equation can also be written as

$$\begin{bmatrix} -1 & D & 0 \\ c + D & f'(\phi_1) - \lambda & -1 \\ 0 & \epsilon & cD - \epsilon\gamma - \lambda \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \tag{4.33}$$

Or equivalently,

$$u' = A_{\phi_1}(\lambda)u \,, \tag{4.34}$$

where

$$A_\varphi(z) = \begin{bmatrix} -c & -f'(\varphi) + z & 1 \\ 1 & 0 & 0 \\ 0 & -c^{-1}\epsilon & c^{-1}(\epsilon\gamma + z) \end{bmatrix} . \tag{4.35}$$

Notice that $A_{\phi_1}(0)$ is the linearized vector field $DX(\phi)$.

The differential equation $u' = A_{\phi_1}(z)u$ can be solved even if z is not an eigenvalue of L_ϕ. But for $z = \lambda$ to be an eigenvalue of L_ϕ, the solution of (4.34) has to be bounded. To see what this entails, consider first the simpler equation $u' = A_0(z)u$, which is the analogue of (4.34) for the rest solution $\phi = 0$ of (1.5). If z belongs to the half-plane H_ω, then we know from Proposition 3.2 that this equation has no bounded solutions. Thus, the matrix $A_0(z)$ is hyperbolic (has no purely imaginary eigenvalues) for all $z \in H_\omega$. The eigenvalues for $z = 0$ are given by (2.17). Consequently, $A_0(z)$ has one eigenvalue μ_z with negative real part, and two eigenvalues $\nu_z^\pm$ with positive real part, for each $z \in H_\alpha$. Denote by $\mathtt{U}_z$ and $\mathtt{W}_z^\pm$ the corresponding eigenvectors (or generalized eigenvectors if $\nu_z^+ = \nu_z^-$). Then the dynamical system $u' = A_0(z)u$ has a linear stable manifold $\hat{\mathtt{U}}_z = \operatorname{span}(\mathtt{U}_z)$ and a linear unstable manifold $\hat{\mathtt{W}}_z = \operatorname{span}(\mathtt{W}_z^\pm)$.

Given that the equation $u' = A_{\phi_1}(z)u$ is linear, it is convenient to consider the associated equation for a function $\hat{u}$ that takes values in the projective plane $\mathbb{C}\mathrm{P}^2$. Notice that the coefficient function $y \mapsto A_{\phi_1(y)}(z)$ is bounded. Thus, a solution of (4.34) cannot vanish at any point without being identically zero. In order to simplify the discussion, let us identify $\hat{u}(y)$ with $\operatorname{span}(u(y))$. Then our dynamical system becomes

$$\phi' = X(\phi)\,, \qquad \hat{u}' = A_{\phi_1}(z)\hat{u}\,. \tag{4.36}$$

We have included here the evolution equation (1.5) for the pulse ϕ, in order to make the system autonomous. This system has a fixed point $(0, \hat{\mathtt{U}}_z)$ and an invariant projective line $\{0\} \times \hat{\mathtt{W}}_z$. The fixed point is hyperbolic, with a 1-dimensional local stable manifold that is tangent at $(0, \hat{\mathtt{U}}_z)$ to the line $\hat{\mathtt{U}}_0 \times \hat{\mathtt{U}}_z$. Thus, the only way for a solution u of (4.34) to stay bounded as $y \to +\infty$ is for $\hat{u}(y)$ to be on the local stable manifold, in which case $\hat{u}(y) \to \hat{\mathtt{U}}_z$ and $u(y) \to 0$.

The invariant line $\{0\} \times \hat{\mathtt{W}}_z$ is hyperbolic as well, with 3-dimensional local stable and unstable manifolds that are tangent (at this line) to $\hat{\mathtt{W}}_0 \times \hat{\mathtt{W}}_z$ and $\hat{\mathtt{U}}_0 \times \mathbb{C}\mathrm{P}^2$, respectively. So a solution u of $u' = A_{\phi_1}(z)u$ stays bounded as $y \to -\infty$ only if $u(y)$ lies on the local unstable manifold, in which case $\hat{u}(y)$ approaches the projective line $\hat{\mathtt{W}}_z$ as $y \to -\infty$, and $u(y) \to 0$. And by hyperbolicity, the only alternative is $\hat{u}(y) \to \hat{\mathtt{U}}_z$, with $u(y)$ growing exponentially.

What will be needed from this discussion is summarized in the proposition below. Notice that $\hat{\mathtt{W}}_z$ is the set of all vectors $w \in \mathbb{C}^3$ with the property that $\mathtt{V}_z^\top w = 0$, where $\mathtt{V}_z$ is the eigenvector of the transposed matrix $A_0(z)^\top$ for the eigenvalue μ_z.

Proposition 4.5. *If $z \in H_\omega$ then the equation $u' = A_{\phi_1}(z)u$ has a solution $u = u_z$ satisfying*

$$\lim_{y \to +\infty} u_z(y) e^{-y\mu_z} = \mathtt{U}_z\,. \tag{4.37}$$

$\lambda \in H_\omega$ is an eigenvalue of L_ϕ if and only if $\mathtt{V}_\lambda^\top u_\lambda(y) \to 0$ as $y \to -\infty$.

This was proved already in [7] for a more general class of equations, under the assumption that the rest solution $\underline{\phi} = 0$ is exponentially stable. Notice that (ϕ, u_z) is a parametrization of the local stable manifold for the system $\phi' = X(\phi)$ and $u' = A_{\phi_1}(z)u$.

Proposition 4.5 by itself is of limited practical value, due to the asymptotic nature of the condition on u_λ. The idea is to "propagate" this condition from $y = -\infty$ to finite y. More specifically, consider the "adjoint" equation

$$v' = -A_{\phi_1}(z)^\top v\,. \tag{4.38}$$

The limiting behavior of its solutions, as $y \to \pm\infty$, is governed by the matrix $-A_0(z)^\top$. This matrix has an eigenvalue $-\mu_z$ with positive real part, and two eigenvalues $-\nu_z^\pm$ with negative real part. The discussion preceding Proposition 4.5 shows that we can find a solution $v = v_z$ that satisfies

$$\lim_{y\to-\infty} v_z(y)e^{y\mu_z} = \mathtt{V}_z\,. \tag{4.39}$$

The main reason for solving (4.38) is that

$$[v_z^\top u_z]' = \bigl[v_z'\bigr]^\top u_z + v_z^\top u_z' = [-A_{\phi_1}(z)^\top v_z]^\top + v_z^\top A_{\phi_1}(z)u_z' = 0\,. \tag{4.40}$$

So the product $v_z^\top u_z$ is a constant function. In particular, if $v_z(y)^\top u_z(y) = 0$ for some y then $v_z^\top u_z = 0$. This property remains the same if we replace $v_z(y)$ by $v_z(y)e^{y\mu_z}$, which tends to $\mathtt{V}_z$ as $y \to \infty$. This shows that $z \in H_\omega$ is an eigenvalue of L_ϕ if and only if $v_z^\top u_z = 0$. A stronger and more general result was proved in [7]. The following theorem only describes the parts that are relevant to our problem.

Theorem 4.6. [7] *Given any $y \in \mathbb{R}$, define the function $\Delta : H_\omega \to \mathbb{C}$ by the equation $\Delta(z) = v_z(y)^\top u_z(y)$. Then Δ is analytic and independent of the choice of y. Furthermore, $\lambda \in H_\omega$ is an eigenvalue of L_ϕ with algebraic multiplicity m, if and only if λ is a zero of Δ of order m.*

Let $r = \frac{2485}{8192}$ and $\vartheta = \frac{5857}{16384}$. Denote by R the closed rectangle in $\mathcal{C}$ with corners at $\pm i\vartheta$ and $r \pm i\vartheta$. Let D be the closed disk in $\mathbb{C}$, centered at the origin, with radius $\frac{1}{32}$.

Lemma 4.7. *The function Δ has a simple zero at 0 and no other zeros in D. Furthermore, the restriction of Δ to the boundary of $R \setminus D$ takes no real values in the interval $[0,\infty)$.*

Our proof of this lemma is computer-assisted; see Section 5 for more details.

Using Lemma 4.7 and Lemma 2.2, we can now give a

Proof of Theorem 1.2. The existence and real analyticity of the homoclinic pulse follows from Lemma 2.2.

Stability: Lemma 4.7 together with the argument principle imply that function Δ has no zeros in $R \cup D$ besides the simple zero at 0. So by Theorem 4.6, the operator L_ϕ has a simple eigenvalue 0 and no other eigenvalues in $R \cup D$. Given that $r < \Lambda$ and $\vartheta < \Theta$, we conclude by Proposition 4.1 and Proposition 4.2 that L_ϕ has no nonzero eigenvalue in H_α for some $\alpha > 0$. By Proposition 3.5 this implies that the solution ϕ' is exponentially stable for the linear system (3.2). The exponential stability of the homoclinic pulse ϕ now follows from Lemma 3.1. **QED**

5. Further reduction of the problem

5.1. The periodic pulse

What remains to be proved are Lemma 2.1 and Lemma 4.4.

Notice that f is a polynomial, so the map $\mathcal{N}_c$ defined in (2.7) involves little more than products of functions in $\mathcal{F}$ and multiplier operators. These are rather easy to compute. By "computing" an element x in some space X we mean finding a set $x^\flat \subset X$, that contains x. The types of enclosures $x^\flat$ that we use, and the procedures for finding them, will be described in Subsection 5.3.

What can be surprisingly difficult in computer-assisted proofs are estimates on linear operators. It is often necessary to "condition" an operator via conjugacies as in (4.27) or (4.30). But once things boil down to computing operator norms, the norm (2.9) used in the spaces $\mathcal{F}$ and $\mathcal{V}$ makes the task easy: Let $\{e_0, e_1, \ldots\}$ be an enumeration of the Fourier modes $c_k \cos(k\boldsymbol{.})$ and $s_k \sin(k\boldsymbol{.})$, with c_k and s_k chosen in such a way that $\|e_j\| = 1$ for all j. Then the operator norm of a bounded linear operator $U : \mathcal{F} \to \mathcal{F}$ can be bounded by using that

$$\|U\| = \sup_j \|Ue_j\| \leq \max \left\{ \|Ue_0\|, \|Ue_1\|, \ldots, \|Ue_{n-1}\|, \|UE_n\| \right\}, \tag{5.1}$$

where $E_n = \{e_n, e_{n+1}, \ldots\}$. As we will describe later, computing a set UE_n is not more involved than computing a function Ue_j. So it suffices to choose n sufficiently large, such that the bound on $\|UE_n\|$ is smaller than the bound on $\|Ue_j\|$ for some $j < n$.

This procedure is used to obtain the estimate (2.14) on the operator $D\mathcal{M}_c(h)$. One reason why this is possible is that $D\mathcal{M}_c(h)$ is compact, due to a factor $(D^2 - \kappa^2 \mathrm{I})^{-1}$. But this alone is not sufficient. What contributes is that the pulse $\underline{\phi}$ has effectively a much larger domain of analyticity than the strip $|\mathrm{Im}\,(x)| < \log(1 + 2^{-10})$ considered here. Thus, multiplication operators like $h \mapsto \varphi h$, when viewed as a "matrix" for the Fourier coefficients, are very small far away from the diagonal.

The same procedure is used to obtain the estimate (4.31) on the operator $\mathcal{U}\mathcal{K}\mathcal{U}$. Here we do not have compactness, but the two factors $\mathcal{U}$ compensate for this.

This leaves the operator $\mathcal{L}_0$ and its resolvents. $\mathcal{L}_0$ is a direct sum of a high-mode part and a low-mode part. The high-mode part is trivial. Its eigenvalues are given by (4.17) with $p > \kappa/\theta$, and it is easy to check that all of them have a negative real part. But we use the computer to do this, which also verifies that the high-mode part of $\mathcal{U}^{-1}(z\mathrm{I} - \mathcal{L}_0)^{-1}\mathcal{U}^{-1}$ satisfies the bound (4.31) for all $z \in \Gamma$.

The low-mode part $L = P(L_0 + F)P$ of $\mathcal{L}_0$ is in effect just a matrix. After performing an approximate diagonalization $D = A^{-1}LA$, we use the Gerschgorin disks to check that exactly one (simple) eigenvalue of D lies in Ω and the others are a positive distance away from Ω.

Denote the low-mode part of $\mathcal{U}$ by U. The task left is to show that the matrix $U^{-1}(z\mathrm{I} - L)^{-1}U^{-1}$ satisfies the bound (4.31) for all $z \in \Gamma$. This is done by covering Γ with a finite collection of disks $|z - z_j| < \delta_j$ with centers $z_j \in \Gamma$. The resolvent matrices $R_j = (z_j\mathrm{I} - L)^{-1}$ are computed explicitly (with error estimates, of course) and shown to satisfy $\delta_j \|R_j\| < 1$. This bound implies that the matrix

$$z\mathrm{I} - L = \left[\mathrm{I} + (z - z_j)R_j\right](z_j\mathrm{I} - L) \tag{5.2}$$

is invertible whenever $|z - z_j| < \delta_j$. Estimating its inverse is straightforward.

5.2. The homoclinic pulse

In Subsection 2.2 we described our construction of the traveling pulse ϕ solution of the equation (1.5). It involves an expansion near $y = +\infty$ for the parametrized local stable manifold ϕ^s, an expansion near $y = -\infty$ for the parametrized local unstable manifold Φ^u, and a sequence of integration steps to prolong ϕ^s. All three problems were reduced to equations that can be solved order by order. At the end of this subsection we give a result that applies to such order by order computations. (As mentioned earlier, by "computation" we mean a process that leads to a rigorous enclosure.)

The main ingredients needed to prove Theorem 4.6 are estimates on the functions u_z and v_z that enter the definition $\Delta(z) = v_z(y)^\top u_z(y)$ of the Evans function Δ. We will show that, after suitable reformulation, the equations for u_z and v_z can again be solved order by order. The eigenvector $\mathtt{U}_z$ of the matrix $A_0(z)$, used in (4.37), and the eigenvector $\mathtt{V}_z$ of the transposed matrix $A_0(z)^\top$, used in (4.39), are given by

$$\mathtt{U}_z = \begin{bmatrix} \mu_z \\ 1 \\ \epsilon(\epsilon\gamma + z - c\mu_z)^{-1} \end{bmatrix}, \qquad \mathtt{V}_z = \begin{bmatrix} 1 \\ c + \mu_z \\ -c(\epsilon\gamma + z - c\mu_z)^{-1} \end{bmatrix}, \tag{5.3}$$

where μ_z denotes the eigenvalue of $A_0(z)$ with negative real part. In what follows, we always assume that $z \in H_\omega$ with $0 < \omega < \epsilon\gamma$.

Given that ϕ is an analytic function of $r = e^{\mu_0 y}$ near $r = 0$, and that $e^{-\mu_z y} u_z(y) \to \mathtt{U}_z$ as $y \to +\infty$, we consider a re-normalized version of u_z,

$$U_z(y) = e^{(\mu_0 - \mu_z)y} u_z(y)\,. \tag{5.4}$$

Since $u_z' = A_{\phi_1}(z)u_z$, the function U_z has to satisfy the equation

$$U_z' = M^s_{\phi_1}(z)U_z\,, \qquad M^s_{\phi_1}(z) = A_{\phi_1}(z) + (\mu_0 - \mu_z)\mathrm{I}\,. \tag{5.5}$$

Recall that $A_0(z)\mathtt{U}_z = \mu_z\mathtt{U}_z$, and thus $M^s_0(z)\mathtt{U}_z = \mu_0\mathtt{U}_z$. Having in mind an expansion in powers of $r = e^{\mu_0 y}$, we decompose

$$U_z(y) = \ell^s(r) + \mathfrak{Z}^s(r)\,, \qquad \ell^s(r) = r\mathtt{U}_z\,, \tag{5.6}$$

with $\mathfrak{Z}^s(r) = \mathcal{O}(r^2)$. To simplify notation, we have suppressed here the dependence on z. Using that $\mu_0 r\partial_r\ell^s = \mu_0\ell^s = M^s_0(z)\ell^s$, the equation (5.5) can be rewritten as

$$\mu_0 s\partial_r\mathfrak{Z}^s = P_0\ell^s + M^s_{\phi_1}(z)\mathfrak{Z}^s\,, \tag{5.7}$$

where

$$P_0 = A_{\phi_1}(z) - A_0(z) = \begin{bmatrix} 0 & -f'(\phi_1) - a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.8}$$

Finally, we convert (5.7) to the integral equation

$$\mathfrak{Z}^s = \left[\partial_y - M^s_0(z)\right]^{-1} P_0(\ell^s + \mathfrak{Z}^s)\,, \qquad \partial_y = \mu_0 r\partial_r\,. \tag{5.9}$$

This equation for $\mathfrak{Z}^s = U_z - \ell^s$ is the analogue of the equation (2.19) for $Z^s = \phi^s - \ell^s$. Notice that, due to the special form of P_0, the equation (5.9) reduces by projection to an equation for the component $\mathfrak{Z}^s_1$ of $\mathfrak{Z}^s$ alone. Using that $\mathrm{Re}\,(\mu_z) < 0$, the operator $\partial_y - M^s_0(z)$ is easily seen to be invertible on the space of analytic functions $\mathcal{A}^1_\rho$ defined below. Recall also that $f'(0) = -a$. Thus $P_0(s) = \mathcal{O}(s)$. This shows that (5.9) can be solved order by order in powers of r.

We will prove that the solution $\mathfrak{Z}^s$ is analytic in a disk $\mathcal{D}_\rho$ characterized by $|r| < \rho$. The analysis is carried out in the space $\mathcal{A}^1_\rho$ of all analytic functions h on $\mathcal{D}_\rho$ that extend continuously to the boundary of $\mathcal{D}_\rho$ and have finite norm

$$\|h\|_\rho = \sum_{k=0}^\infty \|h_k\|\rho^k\,, \qquad h(r) = \sum_{k=0}^\infty h_k r^k\,. \tag{5.10}$$

To be more precise, our functions $h = \mathfrak{Z}^s_j$ also depends on the spectral parameter z. The function in $\mathcal{A}^1_\rho$ take values in a space $\mathcal{B}$ of function that are real analytic in the disk $|z - z_0| < \varrho$. The norm of a function $g : z \mapsto \sum_n g_n(z - z_0)^n$ in the space $\mathcal{B}$ is given by $\|g\| = \sum_n |g_n|\varrho^n$. This is the norm that appears in (5.10) for the coefficients h_k.

The procedure for computing the solution $v = v_z$ of the adjoint equation (4.38) satisfying (4.39) is similar. To match the asymptotic behavior of ϕ near $y = -\infty$ we consider

$$V_z(y) = e^{(\nu_0+\bar\nu_0+\mu_z)y} v_z(y)\,, \tag{5.11}$$

where ν_0 and $\bar\nu_0$ are the eigenvalues of $A_0(0)$ with positive real part. The function V_z has to satisfy the equation

$$V_z' = -M^u_{\phi_1}(z)^\top V_z\,, \qquad M^u_{\phi_1}(z) = A_{\phi_1}(z) - (\nu_0 + \bar\nu_0 + \mu_z)\mathrm{I}\,. \tag{5.12}$$

As in (2.20) we make the ansatz

$$V_z(y) = \boldsymbol{V}_z\big(Re^{\nu_0 y}, Re^{\bar\nu_0 y}\big)\,, \tag{5.13}$$

and

$$\boldsymbol{V}_z(s) = \ell^u(s) + \mathfrak{Z}^u(s)\,, \qquad s = (s_1, s_2)\,. \tag{5.14}$$

But here $\ell^u(s) = s_1 s_2 \mathtt{V}_z$ and $\mathfrak{Z}^u(s) = \mathcal{O}(|s|^3)$. Using that $A_0(z)^\top \mathtt{V}_z = \mu_z \mathtt{V}_z$, the equation (5.12) can be rewritten as

$$(\nu_0 s_1 \partial_{s_1} + \bar\nu_0 s_2 \partial_{s_2})\mathfrak{Z}^u = -P_0^\top \ell^u - M^u_{\phi_1}(z)^\top \mathfrak{Z}^u\,, \tag{5.15}$$

or

$$\mathfrak{Z}^u = -\big[\partial_y + M^u_0(z)^\top\big]^{-1} P_0^\top(\ell^u + \mathfrak{Z}^u)\,, \qquad \partial_y = \nu_0 s_1 \partial_{s_1} + \bar\nu_0 s_2 \partial_{s_2}\,. \tag{5.16}$$

Again, due to the special form of $P_0^\top$, the equation (5.16) reduces by projection to an equation for the component $\mathfrak{Z}^u_0$ of $\mathfrak{Z}^u$ alone. Using that $\mathrm{Re}\,(\nu_z) > 0$, the operator $\partial_y - M^s_0(z)$ is easily seen to be invertible. Recall also that $P_0(s) = \mathcal{O}(s)$. So (5.16) can be solved order by order in powers of s_1 and s_2. The function space considered here is the same space $\mathcal{A}_\rho^2$ that is used to solve the equation (2.22) for the unstable manifold. Again, each coefficient $h_{k,m}$ in the expansion (2.23) is a function in the space $\mathcal{B}$. But now the variable is the spectral parameter z, and not the velocity c. At this point, the velocity c is fixed to the value (interval) described in Lemma 2.2.

Recall that the local stable manifold ϕ^s obtained by solving (2.19) has to be "prolonged" in order to get it to meet the local unstable manifold Φ^u. Similarly, we have to prolong the curve V_z in order to be able to compute the Evans function $\Delta(z)$ via the product $V_z(y)^\top U_z(y)$. This is done for $y = -43$, which is the same value of y that is used to verify the homoclinic intersection.

As mentioned in Subsection 2.2, the prolongation of ϕ^s is carried out via repeated application of a Taylor integrator. The same is done for the prolongation of the curve V_z. The details are different in the two cases, but the principle is the same: We integrate a vector field $\mathcal{V}$,

$$x'(t) = \mathcal{V}(x(t)), \qquad x(0) = x_0, \tag{5.17}$$

where all functions involved are analytic near the origin. Decomposing

$$\mathcal{V}(x) = \ell + Q(x - x_0), \qquad \ell = \mathcal{V}(x_0), \quad Q(0) = 0, \tag{5.18}$$

and substituting

$$x(t) = x_0 + t\ell + Z(t), \qquad Z(t) = \sum_{k=2}^{\infty} Z_k t^k, \tag{5.19}$$

into the equation (5.17) yields the integral equation

$$Z(t) = \int_0^t Q\bigl(s\ell + Z(s)\bigr)\, ds. \tag{5.20}$$

Clearly this equation can be solved order by order, meaning that the coefficients Z_k can be determined inductively one after the other.

Unlike in the periodic case, we are not using the contraction mapping principle to solve fixed point equations. In the homoclinic case, all of our equations can be solved order by order. The details differ from one equation to the next, but the general principle is the same and can be described as follows.

Let $(X_k, \|.\|_k)$ be Banach spaces for $k = 0, 1, 2, \ldots$ and let $(X, \|.\|)$ be the Banach space of all functions $x : \mathbb{N} \to \bigcup_k X_k$ with $x(k) \in X_k$ for all k and $\|x\| = \sum_k \|x(k)\|_k$ finite. Denote by P_n the projection on X defined by setting $(P_n x)(k) = x(k)$ for $k \le n$ and $(P_n x)(k) = 0$ for $k > n$.

Lemma 5.1. *Let Y_0 be a closed bounded subset of X such that $P_n Y_0 \subset Y_0$ for all n, and $P_0 Y_0 = \{y_0\}$ for some $y_0 \in X$. Let $F : Y_0 \to Y_0$ be continuous, having the property* *that $P_{n+1}F = P_{n+1}FP_n$ for all n. Then F has a unique fixed point $y \in Y_0$, and $P_n y = P_n F^m(y_0)$ whenever $n \le m$.*

Notice that the set Y_0 in this lemma need not be compact or convex. The norm on X is of course very special. But this framework fits perfectly with the function spaces used here and in other computer-assisted proofs.

Proof. Let $y \in Y_0$. Then $P_0F^m(y) = y_0$ for all m. If $0 < n \le m$ then

$$\begin{aligned} P_nF^m(y) &= P_nF(P_{n-1}F^{m-1}(y)) = \ldots \\ &= P_nF(P_{n-1}F(\ldots P_1(F(P_0F^{m-n}(y)))\ldots)) \\ &= P_nF(P_{n-1}F(\ldots P_1(F(y_0))\ldots)). \end{aligned} \tag{5.21}$$

So $P_nF^m(y)$ is independent of y. For a fixed point y of F this yields $P_ny = P_nF^m(y_0)$ whenever $n \le m$. In particular, there can be at most one fixed point.

In order to prove that a fixed point exists, define $y_n = P_nF^n(y_0)$ for each n. By (5.21) we have $y_n = P_nF^m(y_0)$ and thus $y_n = P_ny_m$ for all $m \ge n$. This shows that

$$\|y_m\| = \|y_0\| + \sum_{k=1}^{m} \|y_k - y_{k-1}\|. \tag{5.22}$$

The sum in this equation is bounded uniformly in m, since y_m belongs to the bounded set Y_0. Thus, $n \mapsto y_n$ is Cauchy sequence. It converges since X is complete, and $y = \lim y_n$ belongs to Y_0 since Y_0 is closed. Using again that the right hand side of (5.21) is independent of $m \ge n$, we have

$$\begin{aligned} P_nF(y_m) &= P_nF(P_mF^m(y_0)) = P_nF(P_{n-1}F^m(y_0)) \\ &= P_nF(P_{n-1}F^{n-1}(y_0)) = P_nF^n(y_0) = y_n. \end{aligned} \tag{5.23}$$

Taking $m \to \infty$ and using the continuity of F, we get $P_nF(y) = y_n$ for all n, and thus $F(y) = y$. **QED**

5.3. Estimates done by computer

In Sections 2-4 we have reduced the problem of proving the existence and stability of pulse solution for the FitzHugh-Nagumo equation (1.1) and (1.3) to a proof of four lemmas: Lemma 2.1, Lemma 4.4 in the periodic case, and Lemma 2.2, Lemma 4.7 in the homoclinic case. In Subsections 5.1 and 5.2 we described how the claims made in these lemmas are reduced to concrete computations. Our aim here is to explain how these computations are carried out with the necessary error estimates. In essence, the problem is being reduced further, to a point where each step is as trivial as verifying an inequality $a * b < c$. These steps are then carried out by a computer. The full description, with all the necessary definitions (declarations) and propositions (procedures and functions), are written in the programming language Ada [32] and can be found in [37]. Here we only give a sketch of this reduction process and its organization.

At the highest level, our programs are essentially translations of expressions that appear in our equations like (2.5) and (5.9). At the lowest level, numbers are being added, multiplied, compared, etc., using interval arithmetic. The intermediate levels mostly reduce the problem. They are kept as independent from each other as possible.

Each level builds upon a generic type `Scalar` for which various operations are defined. These scalars are the "representable" subset of some algebra $\mathcal{S}$. By algebra we mean a commutative Banach algebra with unit. The collection of all scalars will be denoted by $R(\mathcal{S})$. Among the scalar operations is a function $\texttt{Sum} : R(\mathcal{S}) \times R(\mathcal{S}) \to R(\mathcal{S})$ with the property that $S_1 + S_2 \in \texttt{Sum}(S_1, S_2)$ whenever $S_1, S_2 \in R(\mathcal{S})$. Here $S_1 + S_2$ denotes the collection of all sums $s_1 + s_2$ with $s_1 \in S_1$ and $s_2 \in S_2$. If there exists no scalar in $R(\mathcal{S})$ that includes $S_1 + S_2$, then $\texttt{Sum}(S_1, S_2)$ returns a special element `Undefined`. (In practice, the program simply stops with an error message.) The same scalar `Undefined` is returned by other functions whenever some operation is undefined, either by necessity or choice.

It is not necessary to know the innards of a `Scalar`, except that it includes error estimates in the form of balls $\{s \in \mathcal{S}_i : \|s\| \le r\}$, where $\mathcal{S}_i$ can be any subspace of $\mathcal{S}$. Let $\mathcal{F}$ be an algebra of $\mathcal{S}$-valued functions on some domain D, and assume that $\mathcal{F}$ contains all constant functions. Then to every scalar $S \subset \mathcal{S}$ we associate an extended scalar $F \subset \mathcal{F}$ by replacing a ball $\{s \in \mathcal{S}_i : \|s\| \le r\}$ with the ball $\{f \in \mathcal{F}_i : \|f\| \le r\}$, where $\mathcal{F}_i$ denotes the subspace of all function in $\mathcal{F}$ that take values in $\mathcal{S}_i$. The collection of all such extended scalars is denoted by $R(\mathcal{S}, \mathcal{F})$. We assume that the above-mentioned function `Sum` extends to a function $\texttt{Sum} : R(\mathcal{S}, \mathcal{F}) \times R(\mathcal{S}, \mathcal{F}) \to R(\mathcal{S}, \mathcal{F})$ with the property that $F_1 + F_2 \in \texttt{Sum}(F_1, F_2)$ whenever $F_1\,, F_2 \in R(\mathcal{S}, \mathcal{F})$. This is assumed to hold for any choice of the algebra $\mathcal{F}$.

Our scalars in $R(\mathbb{R})$ are all of the form $c + \{s \in \mathbb{R} : |s| \le r\}$ where c and r are real numbers that are representable in the given floating point environment. It is well-known how to implement a function `Sum` with the above-mentioned properties. Similarly for other elementary functions such as `SetZero`, `Neg`, `Diff`, `Prod`, `Quot`, `Exp`, ... The scalars in $R(\mathbb{C})$ are all sets of the form $X + iY$ with $X, Y \in R(\mathbb{R})$.

Our type `Taylor1` is now defined as follows. Consider a disk $D_r = \{z \in \mathbb{C} : |z| < r\}$, with $r > 0$ representable. Let $\mathcal{T}_r$ be the space of all analytic functions $f : D_r \to \mathcal{S}$ that extend continuously to the boundary of D_r and have a finite norm

$$\|f\| = \sum_{n=0}^{\infty} \|c_n\| r^n\,, \qquad f(z) = \sum_{n=0}^{\infty} c_n z^n\,, \qquad c_0, c_1, \ldots \in \mathcal{S}\,. \tag{5.24}$$

Clearly $\mathcal{T}_r$ is a Banach algebra under pointwise multiplication. Given a positive integer d, a `Taylor1` is any set of the form

$$F : z \mapsto \sum_{n=0}^{d} C(n) z^n\,, \qquad C(0\,..\,k-1) \in R(\mathcal{S})\,, \quad C(k\,..\,d) \in R(\mathcal{S}, \mathcal{T}_r)\,, \tag{5.25}$$

for some nonnegative integer $k \le d$. The collection of all these sets is denoted by $R(\mathcal{T}_r)$. Now a `Taylor1` is simply another type of `Scalar`, based on $\mathcal{T}_r$ instead of $\mathcal{S}$. An extended `Taylor1` is defined by replacing the algebra $\mathcal{T}_r$ in (5.25) by a function algebra $\mathcal{F}$, as described earlier. In our Ada implementation, `Taylor1` is a triple `F=(F.R,F.K,F.C)`, where `F.R` is the radius r of the domain D_r, `F.K` is the integer k in (5.25), and `F.C` is an array `F.C`$(0\,..\,d)$ of (extended) scalars. The component `T.R` is used for inclusions $\mathcal{T}_r \to \mathcal{T}_{r'}$ when $r > r' > 0$.

Implementing `Sum` : $R(\mathcal{T}_r) \times R(\mathcal{T}_r) \to R(\mathcal{T}_r)$ is trivial: Given `F1=(F1.R,F1.K,F1.C)` and `F2=(F2.R,F2.K,F2.C)` of type `Taylor1`, we set `F3.R := Min(F1.R,F2.R)`, `F3.K := Min(F1.K,F2.K)`, and `F3.C(N) := Sum(F1.C(N),F2.C(N))` for $N = 0, 1, \ldots, d$, and then return `F3`. The function `Prod` is equally trivial: just multiply `F1` and `F2` as if they were polynomials, except that terms of degree $n > d$ are added to `F3.C`(d), after being enclosed in zero-centered balls and multiplied by `F3.R`$^{n-d}$. For functions like `Quot` or `Exp` we compute the first N terms of the power series, using `Sum` and `Prod`. For the remainder of the Taylor series we use simple norm estimates, and the corresponding balls are added to the component `F.C`(d) of the result. Functions like `ArcCos` are implemented by first computing an approximation and then estimating the errors from the corresponding Newton map.

Our list of "standard" scalar operations also includes an upper bound on the norm, comparisons (which may be `Undefined` for some scalars), inclusion relations, certain decompositions, input/output, etc. After implementing each of them for the type `Taylor1`, this type can now be used as `Scalar` to construct new types. A type `Taylor1` with coefficients of type `Taylor1` is used e.g. to represent the functions ϕ^s whose Taylor coefficients depend on the velocity parameter $c \in \mathbb{R}$, or the functions u_z that depend on the spectral parameter $z \in \mathbb{C}$.

To give a simple example of how the type (5.25) can be used, consider the fixed point problem for the map $M : \mathcal{T}_1 \to \mathcal{T}_1$ defined by $(Mf)(z) = [z + f(z/2)]^2$ and $f(z) = \mathcal{O}(z^2)$. Here $\mathcal{S} = \mathbb{R}$. Starting with `F=(1,D,0)` and iterating `M` $d-2$ times, the coefficients `F.C(0..D-1)` stay the same under further iterations of `M`. Now enlarge `F.C(D)` to some closed ball `B` and apply `M` again. If `M(F).C(D)` is included in `B` then `M(F)` contains the fixed point of M, according to Lemma 5.1.

In addition to `Taylors1`, we also need a type `Taylor2` to represent sets of analytic functions on $D_r \times D_r$. For reasons of efficiency, we have implemented `Taylor2` directly, as opposed to deriving it from `Taylor1`. It is represented again by a triple `F=(F.R,F.K,F.C)`, but `F.C` is now a two-dimensional array of scalar coefficients `F.C`(m, n) with $m + n \le d$. The type `Taylor2` could also be used as scalar, but this is not needed for the problem at hand. A type `Taylor2` with coefficients of type `Taylor1` is used e.g. to represent the functions Φ_j^u and v_z whose Taylor coefficients depend on the spectral parameter $z \in \mathbb{C}$.

Besides the standard scalar operations, there are also some `Taylor1`-specific and/or `Taylor2`-specific operations such as as evaluation, derivatives (with result in $\mathcal{T}_{r'}$ for some $r' < r$), antiderivatives, and integrals. A quick look at the Ada packages `Taylors1` and `Taylors2` in [37] should make clear that all of this is rather straightforward. But as with other proofs, some essence is in the details as well.

Next, consider the space $\mathcal{F}$ of 2π-periodic analytic functions defined after (2.8). The representable sets in $\mathcal{F}$ are called `Fourier1` and are of the form

$$F = \sum_{k=0}^{d} C(k)\cos(k.) + \sum_{k=1}^{d} C(-k)\sin(k.) + \sum_{n=-2d}^{2d} B_n(r_n)\,, \tag{5.26}$$

with $C(-d\,..\,d) \in R(\mathcal{S})$, and with $B_n(r_n)$ defined as follows. Given $n \geq 0$, let $\mathcal{F}_n$ be the subspace of all even functions (cosine series) in $\mathcal{F}$ with frequencies $k \geq n$. For every representable real number $r_n \geq 0$, we define $B_n(r_n)$ to be the closed ball of radius r_n in $\mathcal{F}_n$. Similarly, the balls $B_n(r_n)$ for $n < 0$ contain all odd functions $f \in \mathcal{F}$ with frequencies $k \geq |n|$ and norm $\|f\| \leq r_n$. Notice that $B_n(1) \cup B_{-n}(1)$ is an enclosure for the set E_n described after equation (5.1).

Each `Fourier1` is represented by a triple `F=(F.R,F.C,F.E)`, where `F.R` $= e^\rho$, and where `F.C` and `F.E` are arrays containing the coefficients $C(k)$ and the radii r_n, respectively, that appear in the definition (5.26). We have implemented all standard scalar operations for the type `Fourier1`, plus a number other useful `Fourier1`-specific operations. For details we refer to the package `Fouriers1` in [37].

Basic linear algebra for matrices and vectors with scalar entries is covered in the packages `Matrices`, `Vectors`, and `ScalVectors`. The procedure `ScalVectors.FindZero3` is used e.g. to locate a zero of the function Υ defined in (2.25). Procedures that are specific to vectors with Taylor components are defined in `MultiTaylors1` and `MultiTaylors2`.

Some of the problem-specific procedures can be found in the child packages `Fouriers1.Periodic`, `MultiTaylors1.Homoclinic`, and `MultiTaylors2.Homoclinic`. Most high-level procedures are defined in the two packages `FHN_Periodic` and `FHN_Homoclinic`. They contain implementations of the procedures that have been discussed in earlier sections. (Our main programs merely call these procedures.) It is impossible to describe them all in detail here. To find out e.g. how $[\partial_y - M_0^s(z)]^{-1}$ is being estimated, there is no better way than to read `MultiTaylors1.Homoclinic.InvDyMinusM`.

For the centers of the balls in $R(\mathbb{R})$ we use a MPFR floating point type with up to 256 mantissa bits, depending on the program. MPFR is an open source multiple-precision floating-point library that supports controlled rounding [35]. Numeric types that do not require high accuracy, such as the radii of the balls in $R(\mathbb{R})$, we use a standard 80 bit extended precision format [36] provided by the Gnat compiler [34]. On a current personal computer with a 3.4GHz processor, our proof for the existence and stability of the periodic (homoclinic) pulse takes approximately 33 minutes and 140 hours (13 hours and 425 hours), respectively. With a multiprocessor machine, the effective computation time can be much shorter. Further details, including instructions on how to compile and run our source code can be found in [37].

References

[1] R. FitzHugh, *Mathematical models of threshold phenomena in the nerve membrane*, Bull. Math. Biophysics **17**, 257269 (1955).

[2] R. FitzHugh, *Impulses and physiological states in theoretical models of nerve membrane*, Biophys. J. **1** 445-466 (1961).

[3] J. Nagumo, S. Arimoto, S. Yoshizawa, *An active pulse transmission line simulating nerve axon*, Proc. IRE **50**, 2061-2070 (1962).

[4] J.W. Evans, *Nerve axon equations: I Linear approximations*, Indiana Univ. Math. J. **21**, 877-885 (1972).

[5] J.W. Evans, *Nerve axon equations: II Stability at rest*, Indiana Univ. Math. J. **22**, 75-90 (1973).

[6] J.W. Evans, *Nerve axon equations: III Stability of the nerve impulse*, Indiana Univ. Math. J. **22**, 577-594 (1973).

[7] J.W. Evans, *Nerve axon equations: IV The Stable and the Unstable Impulse*, Indiana Univ. Math. J. **24**, 1169-1190 (1975).

[8] S.P. Hastings, *On the existence of homoclinic and periodic orbits in the FitzHugh-Nagumo equations*, Quart. J. Math. Oxford **2**, 123-134 (1976).

[9] S.P. Hastings, *Single and Multiple Pulse Waves for the Fitzhugh-Nagumo Equations*, SIAM J. Appl. Math. **42** 247-260 (1982).

[10] G. Carpenter, *A geometric approach to singular perturbation problems with applications to nerve impulse equations*, J. Differ. Equations **23**, 152-173 (1977).

[11] R. Langer, *Existence and uniqueness of pulse solutions to the FitzHugh-Nagumo equations*, Ph.D. Thesis, Northeastern University (1980).

[12] J.W. Evans, N. Fenichel, J.A. Feroe, *Double impulse solutions in nerve axon equations*, SIAM J. Appl. Math. **42** 219-234 (1982).

[13] J.A. Feroe, *Existence and stability of multiple impulse solutions a nerve axon equation*, SIAM J. Appl. Math. **42** 225-246 (1982).

[14] C.K.R.T. Jones, *Stability of the traveling wave solution of the FitzHugh-Nagumo system*, Trans. Amer. Math. Soc. **286**, 431-469 (1984).

[15] E. Yanagida, *Stability of fast traveling pulse solutions of the FitzHugh-Nagumo equation*, J. Math. Biol. **22**, 81-104 (1985).

[16] G. Flores, *Stability analysis for the slow traveling pulse of the FitzHugh-Nagumo system*, SIAM J. Math. Anal. **22** 392-399 (1991).

[17] C. Jones, N. Kopell, R. Langer, *Construction of the FitzHugh-Nagumo pulse using differential forms.* In "Pattern and dynamics in reactive media", IMA Volumes in Mathematics and its Applications, Vol. 37, H. Swinney, G. Aris, D. Aronson (eds), Springer-Verlag, New York, 101-116 (1991)

[18 C. Rocşoreanu, A. Georgescu, N. Giurgiţeanu, *The FitzHugh-Nagumo model: bifurcation and dynamics*, Kluwer Academic Publishers, Boston, 2000.

[19] W. Liu, E. Van Vleck, *Turning points and traveling waves in FitzHughNagumo type equations*, J. Differ. Equations **225**, 381-410 (2006).

[20] J. Guckenheimer, C. Kuehn, *Homoclinic orbits of the FitzHugh-Nagumo equation: the singular limit*, Discrete Contin. Dynam. Systems S **2**, 851-872 (2009).

[21] D. Ambrosi, G. Arioli, F. Nobile, A. Quarteroni, *Electromechanical coupling in cardiac dynamics: the active strain approach*, SIAM J. Appl. Math. **71**, 605-621 (2011).

[22] D. Ambrosi, G. Arioli, H. Koch, *A homoclinic solution for excitation waves on a contractile substratum*, SIAM J. Appl. Dyn. Syst. **11**, 1533-1542 (2012).

[23] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag Berlin Heidelberg (1995).

[24] This follows e.g. from Theorems 4.5.26 and 4.5.31 in [23].

[25] E. Hille, R.S. Phillips, Functional analysis and semi-groups, American Mathematical Society (1957).

[26] See e.g. Theorem 13.2.1 in [25] or Theorem 9.2.1 in [23].

[27] H.F. Trotter, *On the product of semigroups of operators*, Proc. Am. Math. Soc. **10**, 545-551 (1959).

[28] P.R. Chernoff, *Note on product formulas for operator semigroups*, J. Funct. Anal. **2**, 238-242 (1968).

[29] J. Voigt, *On the convex compactness property for the strong operator topology*, Note di Matematica Vol. XII, 259-269 (1992).

[30] G. Arioli, H. Koch, *Computer-assisted methods for the study of stationary solutions in dissipative systems, applied to the Kuramoto-Sivashinski equation*, Arch. Ration. Mech. An. **197**, 1033-1051 (2010).

[31] G. Arioli, H. Koch, *Non-symmetric low-index solutions for a symmetric boundary value problem*, J. Differ. Equations **252**, 448-458 (2012).

[33] R. Szczelina, P. Zgliczyński, *A homoclinic orbit in a planar singular ODE — a computer assisted proof*, SIAM J. Applied Dynamical Systems **12**, 1541-1565 (2013).

[32] Ada Reference Manual, ISO/IEC 8652:2012(E), available e.g. at `http://www.ada-auth.org/arm.html`.

[34] A free-software compiler for the Ada programming language, which is part of the GNU Compiler Collection; see `http://gcc.gnu.org/software/gnat/`.

[35] The MPFR library for multiple-precision floating-point computations with correct rounding; see `http://www.mpfr.org/`.

[36] The Institute of Electrical and Electronics Engineers, Inc., *IEEE Standard for Binary Floating–Point Arithmetic*, ANSI/IEEE Std 754–2008.

[37] The computer programs are included with the preprint `mp_arc 13-91`. They are also available at `ftp://ftp.ma.utexas.edu/pub/papers/koch/fhn/index.html`.