Modulating pulse solutions to quadratic quasilinear wave equations over exponentially long length scales

Modulating pulse solutions to quadratic quasilinear wave equations over exponentially long length scales

Mark D. Groves
Department of Mathematical Sciences
Loughborough University
Loughborough LE11 3TU
UK

Guido Schneider
Mathematisches Institut I
Universität Karlsruhe
76128 Karlsruhe
Germany

Abstract

This paper presents an existence proof for modulating pulse solutions to a wide class of quadratic quasilinear Klein-Gordon equations of the form

$$\partial_t^2 u = \partial_x^2 u - u + f_1(u, \partial_x u, \partial_t u)\partial_x^2 u + f_2(u, \partial_x u, \partial_t u).$$

Modulating pulse solutions consist of a pulse-like envelope advancing in the laboratory frame and modulating an underlying wave-train; they are also referred to as 'moving breathers' since they are time-periodic in a moving frame of reference. The problem is formulated as an infinite-dimensional dynamical system with three stable, three unstable and infinitely many neutral directions. By transforming part of the equation into a normal form with an exponentially small remainder term and using a generalisation of local invariant-manifold theory to the quasilinear setting, we prove the existence of small-amplitude modulating pulses on domains in space whose length is exponentially large compared to the magnitude of the pulse.

MSC: 35L05, 37G05, 37K50

Keywords: Quasilinear wave equations, spatial dynamics, moving breathers

1 Introduction

1.1 Breathers and modulating pulses

A *breather solution* of the nonlinear wave equation

$$\partial_t^2 u = \partial_x^2 u - u + g(u), \qquad x, t \in \mathbb{R},$$

in which $g : \mathbb{R} \to \mathbb{R}$ is a smooth function, is a solution $u(x,t) \in \mathbb{R}$ which is periodic in t and decays to zero as $x \to \pm\infty$; in particular the completely integrable sine-Gordon equation

($g(u) = u - \sin u$) admits explicit breather solutions. It seems natural to ask whether wave equations obtained by perturbing the sine-Gordon nonlinearity also have breather solutions; a negative answer to this question was given by Denzler [2] and Birnir, McKean & Weinstein [1], who showed that the only perturbations of the sine-Gordon equation admitting breather solutions are those which can be transformed back into the sine-Gordon equation by a rescaling.

The 'non-existence of breathers' result is remarkable since the existence of small-amplitude breather solutions is predicted by multiple scaling analysis. Making the *Ansatz*

$$u(x,t) = \varepsilon A(\varepsilon(x - c_{\mathrm{g}}' t), \varepsilon^2 t)\mathrm{e}^{\mathrm{i}k_0 x - \mathrm{i}\omega_0 t} + \text{c.c.},$$

in which $0 < \varepsilon \ll 1$ is a small perturbation parameter, k_0 is the basic spatial wavenumber, $\omega_0 = \omega_0(k_0) = \sqrt{1 + k_0^2}$ is the basic frequency and $c_{\mathrm{p}}' = c_{\mathrm{p}}'(k_0) = \omega_0/k_0$, $c_{\mathrm{g}}' = c_{\mathrm{g}}'(k_0) = k_0/\omega_0$ are the linear phase and group velocities, one finds that at leading order $A(X,T) \in \mathbb{C}$ satisfies the nonlinear Schrödinger equation

$$\partial_T A = \mathrm{i}\nu_1 \partial_x^2 A + \mathrm{i}\nu_2 A|A|^2 \tag{1}$$

with coefficients $\nu_j = \nu_j(k_0) \in \mathbb{R}$. Equation (1) is said to be *focusing* if $\nu_1\nu_2 > 0$; in this case it possesses a family of breather solutions

$$A(X,T) = B(X)e^{\mathrm{i}\gamma T}, \qquad \gamma \in \mathbb{R},$$

where $B(X) \in \mathbb{R}$ satisfies $\lim_{X\to\pm\infty} B(X) = 0$, and it is well known that these solutions correctly approximate solutions of the nonlinear wave equation on length- and time-scales of $O(1/\varepsilon^2)$ (see e.g. Kalyakin [8] and Schneider [12]). The breathers considered by Denzler and Birnir, McKean & Weinstein have a basic wavenumber $k_0 = 0$; for $k_0 \neq 0$, so that $c_{\mathrm{g}}'(k_0) \neq 0$, the solutions are called *moving breathers* (because they are time-periodic in a moving frame of reference) or *modulating pulses* (because they consist of a pulse like envelope advancing in the laboratory frame and modulating an underlying periodic wave train). The 'non-existence of breathers' result shows that the above breathers are destroyed in the full equation by higher order terms which are neglected in the derivation of the nonlinear Schrödinger equation.

The 'non-persistence' phenomenon is easily understood in terms of the geometrical theory of differential equations. Let us examine the equation

$$\partial_t^2 u = \partial_x^2 u - u + uf(u^2), \tag{2}$$

where $f : \mathbb{R} \to \mathbb{R}$ is another smooth function and seek moving breather solutions $u(x,t) = v(x - c_{\mathrm{g}}t, k_0(x - c_{\mathrm{p}}t))$, where v is 2π-periodic and odd in its second variable and $c_{\mathrm{g}} = c_{\mathrm{g}}' + \mathcal{O}(\varepsilon^2)$, $c_{\mathrm{p}} = c_{\mathrm{p}}' + \mathcal{O}(\varepsilon^2)$. Writing the equation for v as an evolutionary equation in which the unbounded spatial variable $\xi = x - c_{\mathrm{g}}t$ plays the role of the time-like variable ('spatial dynamics'), one obtains a system of equations of the form

$$\begin{aligned} \partial_\xi z &= \Lambda z + F(z,w), && (3)\\ \partial_\xi w &= \Lambda w + G(z,w) + H(w) && (4) \end{aligned}$$

(see Groves & Schneider [4]), in which Λ is a linear operator whose spectrum consists of a pair of real eigenvalues of opposite signs and infinitely many purely imaginary eigenvalues, z

and w are the projections of v onto the hyperbolic and centre subspaces of Λ and F, G, H are nonlinearities with $G(z,0) = 0$. A moving breather corresponds to a *homoclinic solution* of (3), (4), that is a solution (z,w) such that $(z(\xi),w(\xi)) \to (0,0)$ as $\xi \to \infty$, and arises as the intersection of the stable and unstable manifolds in the phase space of 2π-periodic functions. Equations (3), (4) have the property that $\{w = 0\}$ is a two-dimensional invariant subspace when the higher-order terms $H(z)$ are neglected; the stable and unstable manifolds are contained in this two-dimensional subspace, and in fact intersect whenever $f'(0) > 0$, giving rise to homoclinic solutions of the two-dimensional dynamical system

$$\partial_\xi z = \Lambda z + F(z,0).$$

On the other hand $\{w = 0\}$ is no longer invariant when $H(z)$ is included, and the intersection of the one-dimensional stable and unstable manifolds in the infinite-dimensional phase space is a rare phenomenon; the homoclinic solution for $H(z) = 0$ does not persist for $H(z) \neq 0$.

Any further analysis of the situation clearly has to take the infinite-dimensional centre space and hence the variable w into account. Groves & Schneider [4] proved the existence of modulating pulse solutions to (2) which remain $\mathcal{O}(\varepsilon^n)$-close to the approximate solutions of amplitude $\mathcal{O}(\varepsilon)$ obtained by setting $H(z) = 0$ but do not decay to zero as $\xi \to \pm\infty$, so that their 'tails' are $\mathcal{O}(\varepsilon^n)$; here $n \in \mathbb{N}$ is arbitrary but fixed. The proof involves using a sequence of normal-form transformations which eliminate successive terms in the Taylor expansion of $H(z)$, so that it can be made $\mathcal{O}(\varepsilon^N)$, where N is arbitrary but fixed (and determined by the choice of n). A standard construction for semilinear evolutionary equations yields a family of solutions on $\xi \in [0,\infty)$ whose hyperbolic parts are $\mathcal{O}(\varepsilon^n)$ and whose centre part may experience secular growth; the initial-values of these solutions form the *centre-stable manifold*. The $\mathcal{O}(\varepsilon^n)$-boundedness of their centre parts follows using an auxiliary argument: a solution $v(\xi)$ with initial data on the centre-stable manifold converges to a solution on an appropriately defined *centre manifold*, a graph in phase space upon which all solutions remain so long as they are $O(\varepsilon^n)$; the existence of a Lyapunov function (the Hamiltonian function for the wave equation) shows that the centre manifold is actually globally invariant; and the rate of convergence of $v(\xi)$ to the centre manifold is shown to be faster than the rate of secular growth of its centre part. Finally, the *reversibility* of (2) is exploited to extend the above solutions to symmetric solutions on $\xi \in (-\infty,\infty)$. Our result identifies a sense in which the modulating pulses for $H(z) = 0$ persist for $H(z) \neq 0$ and we generalise our definition of 'modulating pulses' and 'moving breathers' accordingly. An alternative approach to persistence, which is based upon scattering theory, is given by McLaughlin & Shatah [10].

Further complications arise when studying *quasilinear* wave equations. Our technique in reference [4] relies heavily upon semilinearity, in particular that global existence theory is available for globally Lipschitz nonlinearities with small Lipschitz constant; this method is therefore not applicable to quasilinear problems. Progress was however made in our study of the prototype quasilinear equation

$$\partial_t^2 u = \partial_x^2 u - u + \partial_x^2(u^3)$$

(Groves & Schneider [5]). The theory is analogous to that for semilinear equations: a normal-form transformation eliminates terms up to $\mathcal{O}(\varepsilon^N)$ in the Taylor expansion of $H(z)$; an iteration scheme and energy estimates are used to construct solutions which exist on $\xi \in [0,\varepsilon^{-n}]$ and are $\mathcal{O}(\varepsilon^n)$ close to the approximate solutions of amplitude $\mathcal{O}(\varepsilon)$ obtained by setting $H(z) = 0$; and reversibility is used to extend these solutions to symmetric solutions on $\xi \in [-\varepsilon^{-n}, \varepsilon^{-n}]$. The definition of 'modulating pulses' and 'moving breathers' is thus generalised further, namely to include pulses which exist on large but finite spatial intervals in a frame of reference moving with the pulse.

In the present paper we present an existence theory for modulating pulses which greatly improves and generalises the result given in reference [5]. We show that a large class of quasilinear wave equations with analytic nonlinearities and whose quadratic terms do not necessarily vanish admit modulating pulse solutions which exist on $\xi \in [-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]$ and are $\mathcal{O}(\mathrm{e}^{-c^\star/2\sqrt{\varepsilon}})$-close to the $\mathcal{O}(\varepsilon)$ approximate modulating pulses. Our solutions are therefore *exponentially close* to the approximate pulses and exist on *exponentially long* length scales in a frame of reference moving with the pulse. The main tool is a normal-form transformation which makes $H(z)$ exponentially rather than algebraically small.

1.2 The result

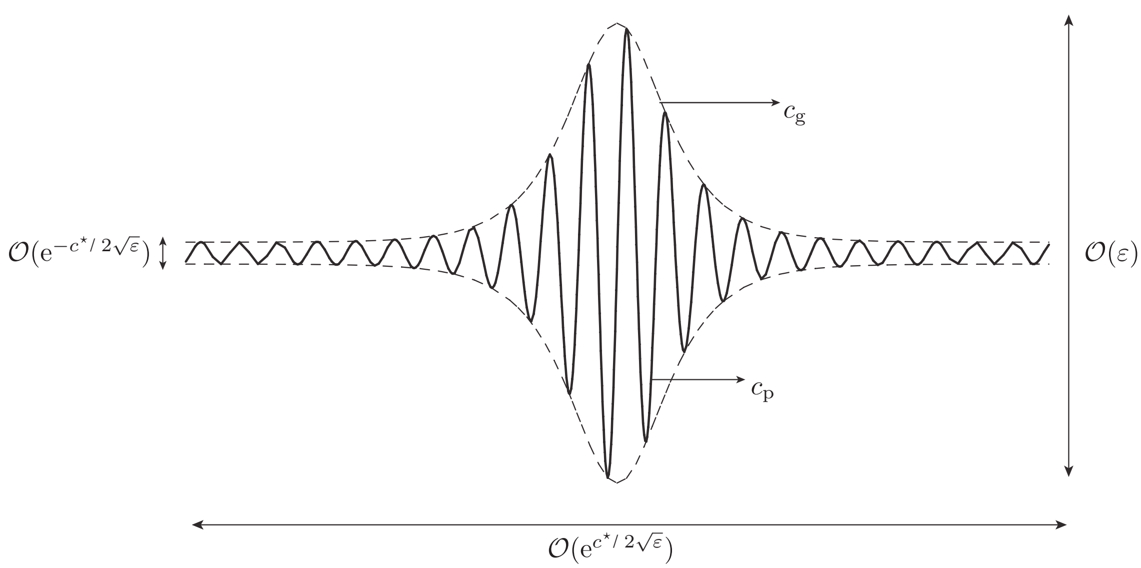

Figure 1: A modulating pulse solution guaranteed by Theorem 1.1.

We seek modulating pulse solutions to the quasilinear wave equation

$$\partial_t^2 u = \partial_x^2 u - u + f_1(u, \partial_x u, \partial_t u)\partial_x^2 u + f_2(u, \partial_x u, \partial_t u), \tag{5}$$

in which f_1, $f_2 : \mathbb{R}^3 \to \mathbb{R}$ are analytic functions which satisfy

$$f_i(a, -b, -c) = f_i(a, b, c), \qquad i = 1, 2; \tag{6}$$

this hypothesis ensures that the spatial dynamics formulation of (5) is reversible, and is an essential requirement for the construction of symmetric modulating pulses. Our result is stated in Theorem 1.1 and illustrated in Figure 1.

Theorem 1.1 *Fix a positive real number k_0. There exist positive constants ε_0 and c^* with the property that for each $\varepsilon \in (0, \varepsilon_0)$ equation (5) admits an infinite-dimensional, continuous family of modulating pulse solutions of the form*

$$u(x,t) = v(x - c_{\mathrm{g}}t, k_0(x - c_{\mathrm{p}}t)),$$

where v is 2π-periodic in its second argument and

$$c_{\mathrm{p}} = c_{\mathrm{p}}' + \gamma_1\varepsilon^2, \qquad c_{\mathrm{g}} = \frac{1}{c_{\mathrm{p}}}.$$

These solutions satisfy

$$v(\xi,\eta) = v(-\xi,-\eta), \qquad |v(\xi,\eta) - h^\varepsilon(\xi,\eta)| \le e^{-c^\star/2\sqrt{\varepsilon}}$$

for all $\eta \in \mathbb{R}$ and $\xi \in [-e^{c^\star/2\sqrt{\varepsilon}}, e^{c^\star/2\sqrt{\varepsilon}}]$, in which

$$h^\varepsilon(\xi,\eta) = \pm\varepsilon\left(\frac{2\check{C}_1}{\pi\check{C}_2}\right)^{1/2}\operatorname{sech}(\check{C}_1^{1/2}\varepsilon\xi)\cos\eta + \mathcal{O}(\varepsilon^{3/2}\mathrm{e}^{-\varepsilon\theta\xi}), \qquad 0 < \theta < \check{C}_1^{1/2}$$

(so that $\lim_{\xi\to\pm\infty} h^\varepsilon(\xi,\eta) = 0$ uniformly in $\eta \in \mathbb{R}$). Here $\check{C}_1 = -2k_0\gamma_1(1+k_0^2)^{3/2}$ is positive for $\gamma_1 < 0$ and $\check{C}_2$ is a normal-form coefficient which is defined in equation (29) and required to be positive.

Example 1.2 *The quasilinear wave equation*

$$\partial_t^2 u = \partial_x^2 u - u - \alpha u \partial_x^2 u - \beta(\partial_x u)^2\partial_x^2 u,$$

in which α, β are real parameters, satisfies the hypotheses of Theorem 1.1 with

$$\check{C}_2 = -\frac{\alpha^2 k_0^2}{4\pi}(2 + 7k_0^2 + 5k_0^4) + \frac{k_0^4(1+k_0^2)\beta}{4\pi};$$

this coefficient is positive for sufficiently large values of k_0 whenever $\beta > 5\alpha^2$.

Remark 1.3 *The coefficient $\check{C}_2$ is positive whenever the asssociated nonlinear Schrödinger equation (1) is focusing.*

The proof of Theorem 1.1 has five main steps.

I. Spatial dynamics formulation (Section 2) The equation for v is formulated as an evolutionary system in which the unbounded spatial variable ξ is the time-like variable. The linear operator in this evolutionary system has two geometrically double real eigenvalues of $\mathcal{O}(\varepsilon)$, two simple strongly hyperbolic eigenvalues of $\mathcal{O}(1)$ and an infinite number of geometrically double purely imaginary eigenvalues. In terms of the projections $z = P_{\mathrm{wh}}v$ and $q = P_{\mathrm{sh,c}}v$ of v onto the weakly hyperbolic and strongly hyperbolic/centre subspaces, we write the system as the coupled fourth-order dynamical system

$$\partial_\xi z = Kz + F^\varepsilon(z,q), \qquad K = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{7}$$

and quasilinear wave equation

$$\begin{aligned}\partial_\xi q_1 &= q_2, &(8)\\ \partial_\xi q_2 &= -c_3^\varepsilon k_0^2\partial_\eta^2 q_1 - c_4^\varepsilon q_1 + P_{\mathrm{sh,c}}(d_1(z)q_1) + P_{\mathrm{sh,c}}(d_2(z)\partial_\eta q_1) + P_{\mathrm{sh,c}}(d_3(z)q_2)\\ &\quad + P_{\mathrm{sh,c}}(g_3^\varepsilon(z,q)\partial_\eta^2 q_1) + g_4^\varepsilon(z,q) + P_{\mathrm{sh,c}}(g_5^\varepsilon(z,q)\partial_\eta q_2) + h^\varepsilon(z), &(9)\end{aligned}$$

where the notation for the nonlinearities has been designed to help with the careful book-keeping which is needed later. The nonlinearities F^ε, g_j^ε and h^ε are analytic functions of their arguments and the parameter ε, where $g_4^\varepsilon(z,q)$ contains no ε-independent terms which are linear in z. The functions d_j are linear and independent of ε, so that $d_1(z)q_1$, $d_2(z)\partial_\eta q_1$ and $d_3(z)q_2$ represent the terms 'missing' in the Taylor expansion of $g_4^\varepsilon(z,q)$. Hypothesis (6) implies that $g_j^\varepsilon(z_1,-z_2,q_1,-\partial_\eta q_1,-q_2) = g_j^\varepsilon(z_1,-z_2,q_1,-\partial_\eta q_1,-q_2)$, $h^\varepsilon(z_1,-z_2) = h^\varepsilon(z_1,z_2)$, and this restriction is an essential requirement, its purpose being to guarantee the *reversibility* of equations (7)–(9), that is their invariance under the transformation $\xi \mapsto -\xi$, $(v_1,v_2) \mapsto S(v_1,v_2)$, where $S(v_1(\eta),v_2(\eta)) = (v_1(-\eta),-v_2(-\eta))$.

Theorem 1.1 evidently requires an 'almost global-wellposedness' result for (8), (9). It is well known that the presence of quadratic terms in wave equations causes difficulties in constructing existence theories of this kind, and the usual approach is to construct a normal-form transformation which eliminates them (e.g. see Shatah [13]). Unfortunately elimination of the quasilinear quadratic terms in equation (9) in this fashion would cause a loss of regularity and complicate our analysis. In fact we do not require a complete theory for the initial-value problem for equations (7)–(9) since we are only interested in solutions of a certain type, and it is actually not necessary to eliminate the quadratic terms to solve the initial-value problem for such solutions (see step IV below).

II. Identification of approximate modulating pulses (Section 3) According to the discussion in Section 1.1, approximate modulating pulses exist as homoclinic solutions of the equation

$$\partial_\xi z = Kz + F^\varepsilon(z,0), \tag{10}$$

and the approximations increase in accuracy as a sequence of transformations is constructed to remove terms of order 2, 3, … from h^ε. The transformation eliminating the quadratic part of h^ε affects the cubic part of F^ε, which in turn controls homoclinic bifurcation in equation (10). It is therefore necessary to carry out this preliminary transformation separately, after which dynamical systems arguments show that (10) admits a pair of homoclinic solutions provided that a coefficient $\check{C}_2$ in the cubic part of F^ε is positive.

III. Normal-form theory (Section 4) We proceed by using a sequence of normal-form transformations to eliminate terms of order 3, 4, … in the Taylor expansion of h^ε. One cannot expect to eliminate the whole of h^ε in this fashion, because our equations would then admit homoclinic solutions whose existence would contradict the 'non-existence of breathers' result. By restricting attention to a neighbourhood of the origin (which is large enough to contain the approximate homoclinic solutions), one can however optimise over the order of the eliminated terms so that the remainder is exponentially small. The necessary transformation theory (Section 4) is a generalisation of a theory for finite-dimensional dynamical analytic vector fields given by Iooss & Lombardi [7], and here we adopt their notation and make frequent reference to their paper for needed results of a combinatorial nature. A central requirement of Iooss & Lombardi's result is that the linearised vector field should be diagonalisable (this condition ensures that certain estimates hold uniformly in the order of terms eliminated from the vector field). In the present context the corresponding requirement is that the matrix K should be diagonalisable, a condition which is clearly not met. This difficulty is overcome by writing $\varepsilon = \mu^2$ and introducing scaled parameters which convert the equation for z into

$$\partial_\xi z = F^\mu(z, q),$$

in which the linear part of the vector field is the (trivially diagonalisable) zero matrix. A similar device was used by Iooss & Lombardi [6] in an application of their normal-form theory to the $0^2\mathrm{i}\omega$ resonance.

The transformation theory in Section 4 amounts to a *partial normal form* since only certain higher-order terms (the q-independent terms in the equations for q) are eliminated. A complete normal form would involve eliminating *all* 'non-resonant' terms in the vector field, and this task is known to be impossible because of a small-divisor problem arising from asymptotic resonances among the frequencies, that is the magnitudes of the purely imaginary eigenvalues (e.g. see Pöschel [11]). By contrast, the frequencies interact in a helpful way in our partial normal form: they guarantee that the transformation itself is *smoothing of degree one* (see Proposition 4.1), and this property in turn ensures that the transformed equation for q is again a quasilinear wave equation.

IV. Existence theory (Section 5) The next step is to construct an existence theory for solutions of (7)–(9) which remain exponentially close to one of the approximate modulating pulses identified in step II over an exponentially long time scale. For this purpose we use an iteration scheme for quasilinear systems of the type suggested by Kato [9], and here the main task is to prove that the iterative sequence $\{w_{(m)}\}_{m\in\mathbb{N}_0}$ for the central part $w = P_\mathrm{c} q$ of q converges; in particular we show that

$$\|w_{(m)}(\xi)\| \leq \mathrm{e}^{-c^\star/2\mu}, \qquad \xi \in [0, \mathrm{e}^{c^\star/2\mu}], \tag{11}$$

$$\|w_{(m+1)}(\xi) - w_{(m)}(\xi)\| \leq \frac{1}{2}\|w_{(m)}(\xi) - w_{(m-1)}(\xi)\|, \qquad \xi \in [0, \mathrm{e}^{c^\star/2\mu}] \tag{12}$$

for each $m \in \mathbb{N}_0$, so that $w = \lim_{m\to\infty} w_{(m)}$ exists and satisfies $\|w(\xi)\| \leq \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$.

The analysis of the sequence $\{w_{(m)}\}_{m\in\mathbb{N}_0}$ is complicated by the presence of quadratic terms in our nonlinearities. In proving (11) one arrives at the differential inequality

$$\partial_\xi \|w_{(m)}\|^2 \leq c\mu(\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi})\|w_{(m)}\| + c\mu(\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-\mu^2\theta\xi})\|w_{(m)}\|^2 \tag{13}$$

for $w_{(m)}$, and it is necessary to deduce that $\|w_{(m)}(\xi)\| \leq \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$. A better inequality is obtained for equations with cubic nonlinearities, namely

$$\partial_\xi \|w_{(m)}\|^2 \leq c\mu(\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi})\|w_{(m)}\| + c\mu(\mathrm{e}^{-c^\star/2\mu} + \mu^2\mathrm{e}^{-\mu^2\theta\xi})\|w_{(m)}\|^2;$$

integrating and using the means inequality, one finds that

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|w_{(m)}(\xi)\|^2 \leq \|w_{(m)}(0)\|^2 + c\mu\mathrm{e}^{-c^\star/\mu} + c\mu \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|w_{(m)}(\xi)\|^2,$$

so that $\|w_{(m)}(\xi)\| \leq c\mu e^{-c^\star/2\mu}$ for $\xi \in [0, e^{c^\star/2\mu}]$ whenever $\|w_{(m)}(0)\| \leq \mu e^{-c^\star/2\mu}$. Notice however that this calculation does not yield the required result for inequality (13) (the $O(\mu)$ coefficient of the third term on the right-hand side of the deduced inequality is replaced by an $O(1)$ coefficient).

In fact the required result does follow from inequality (13), but a more careful two-step estimation technique is required. In the first step we define $\xi^\star$ so that $e^{-\mu^2\theta\xi^\star} = \mu^\alpha$, where α is an appropriately chosen positive constant; a straightforward application of Gronwall's inequality shows that $\|w_{(m)}(\xi)\|^2 \leq c\mu|\log\mu|e^{-c^\star/\mu}$ for $\xi \in [0, \xi^\star]$ whenever $\|w_{(m)}(0)\| \leq \mu e^{-c^\star/2\mu}$. In the second step we integrate (13) over $[0, e^{c^\star/2\mu}]$ and split the range of integration into $[0, \xi^\star]$ and $[\xi^\star, e^{c^\star/2\mu}]$. Satisfactory estimates for the integrals over $[\xi^\star, e^{c^\star/2\mu}]$ are obtained by an optimal choice of α (and hence $\xi^\star$), while the integrals over $[0, \xi^\star]$ are handled using the result from the first step; the final result is that $\|w_{(m)}(\xi)\| \leq c\mu^{1/2}|\log\mu|e^{-c^\star/2\mu}$ for $\xi \in [0, e^{c^\star/2\mu}]$. Full details of this estimate are given in the final part of Lemma 5.3, and the corresponding calculation needed for (12) is discussed in Lemma 5.4.

V. Extension to symmetric modulating pulses (Section 6) The final step is is based upon the following observations: (i) solutions $v(\xi) = (z(\xi), q(\xi))$ of (7)–(9) on $\xi \in [0, \xi_0]$ with the property that $v(0)$ lies on the *symmetric section* $\Sigma = \operatorname{Fix} S$ can be extended to symmetric solutions on $\xi \in [-\xi_0, \xi_0]$; (ii) the initial values $v(0)$ of the solutions identified in step IV define a manifold in phase space (the *centre-stable manifold*) which is parameterised by the projections w^0 of $v(0)$ onto the infinite-dimensional centre subspace and (Z^0, r_1^0, r_2^0) of $v(0)$ onto an appropriately defined three-dimensional stable subspace. An intersection of the centre-stable manifold with the symmetric section therefore guarantees the existence of symmetric modulating pulse solutions on $\xi \in [-e^{c^\star/2\mu}, e^{c^\star/2\mu}]$. Denote the spectral projections of the phase space onto the centre and hyperbolic subspaces by respectively P_c and P_h. Because $P_\mathrm{c}v(0) = w^0$ we have that $v(0) \in \Sigma_\mathrm{c} := P_\mathrm{c}\Sigma$ whenever $w^0 \in \Sigma_\mathrm{c}$ and fixed-point arguments are used to prove that there is a unique value of (Z^0, r_1^0, r_2^0) such that $P_\mathrm{h}v(0) \in \Sigma_\mathrm{h} := P_\mathrm{h}\Sigma$ (and additional regularity requirements on w^0 beyond those used in step IV are necessary here). In this fashion we obtain the result announced in Theorem 1.1, namely the existence of an infinite-dimensional, continuous family of modulating pulse solutions parameterised by $w^0 \in \Sigma_\mathrm{c}$.

Acknowledgements The work of Guido Schneider is partially supported by the Deutsche Forschungsgemeinschaft DFG under the grant Schn 520/3-1/2.

2 Spatial dynamics formulation

We look for modulating pulse solutions of the nonlinear wave equation (5) of the form

$$u(x,t) = v_1(x - c_\mathrm{g}t, k_0(x - c_\mathrm{p}t)) = v_1(\xi, \eta),$$

where v_1 is periodic in η with period 2π and k_0 is a fixed positive number. Making this *Ansatz*, one arrives at the equation

$$\begin{aligned}(1-c_\mathrm{g}^2)\partial_\xi^2 v_1 &+ 2(1-c_\mathrm{g}c_\mathrm{p})k_0\partial_\xi\partial_\eta v_1 + (1-c_\mathrm{p}^2)k_0^2\partial_\eta^2 v_1 - v_1 \\ &+ f_3(v_1, \partial_\eta v_1, \partial_\xi v_1)(\partial_\xi^2 v_1 + 2k_0\partial_\eta\partial_\xi v_1 + k_0^2\partial_\eta^2 v_1) + f_4(v_1, \partial_\eta v_1, \partial_\xi v_1) = 0,\end{aligned}$$

where

$$
\begin{aligned}
f_3(v_1, \partial_\eta v_1, \partial_\xi v_1) &= f_1(v_1, \partial_\xi v_1 + k_0 \partial_\eta v_1, -c_{\mathrm{g}} \partial_\xi v_1 - k_0 c_{\mathrm{p}} \partial_\eta v_1), \\
f_4(v_1, \partial_\eta v_1, \partial_\xi v_1) &= f_2(v_1, \partial_\xi v_1 + k_0 \partial_\eta v_1, -c_{\mathrm{g}} \partial_\xi v_1 - k_0 c_{\mathrm{p}} \partial_\eta v_1);
\end{aligned}
$$

notice that f_3, f_4 are analytic real-valued functions of their arguments with the property that $f_i(a, -b, -c) = f_i(a, b, c)$, $i = 3, 4$. It is convenient to choose

$$
c_{\mathrm{p}} = c_{\mathrm{p}}' + \gamma_1 \varepsilon^2, \qquad c_{\mathrm{g}} = 1/c_{\mathrm{p}},
$$

so that c_{p} is a small perturbation of the phase velocity c_{p}' of the linearised problem and the equation simplifies to

$$
\begin{aligned}
&\partial_\xi^2 v_1 + \frac{1 - c_{\mathrm{p}}^2 + f_3(v_1, \partial_\eta v_1, \partial_\xi v_1)}{1 - c_{\mathrm{g}}^2 + f_3(v_1, \partial_\eta v_1, \partial_\xi v_1)} k_0^2 \partial_\eta^2 v_1 \\
&\qquad + \frac{2 f_3(v_1, \partial_\eta v_1, \partial_\xi v_1)}{1 - c_{\mathrm{g}}^2 + f_3(v_1, \partial_\eta v_1, \partial_\xi v_1)} k_0 \partial_\eta \partial_\xi v_1 + \frac{f_4(v_1, \partial_\eta v_1, \partial_\xi v_1) - v_1}{1 - c_{\mathrm{g}}^2 + f_3(v_1, \partial_\eta v_1, \partial_\xi v_1)} = 0.
\end{aligned}
$$

Introducing the new variable $v_2 = \partial_\xi v_1$, we can write the above equation as the evolutionary system

$$
\partial_\xi v_1 = v_2, \tag{14}
$$

$$
\begin{aligned}
\partial_\xi v_2 = &-c_3^\varepsilon k_0^2 \partial_\eta^2 v_1 - c_4^\varepsilon v_1 + (c_0 v_1 + g_0^\varepsilon(v)) \partial_\eta^2 v_1 + (c_2 v_1 + g_2^\varepsilon(v)) \partial_\eta v_2 \\
&+ c_{1,1} v_1^2 + c_{1,2} (\partial_\eta v_1)^2 + c_{1,3} v_2^2 + c_{1,4} v_2 \partial_\eta v_1 + g_1^\varepsilon(v),
\end{aligned} \tag{15}
$$

which we study in the phase space

$$
\mathcal{X}^s = \{ v = (v_1, v_2) \in H_{\mathrm{per}}^{s+1}(0, 2\pi) \times H_{\mathrm{per}}^s(0, 2\pi) \}, \qquad s > 0,
$$

the domain of the vector field on the right-hand side of (14), (15) being $\mathcal{D}^s = \mathcal{X}^{s+1}$. Here

$$
c_3^\varepsilon = \frac{1 - c_{\mathrm{p}}^2}{1 - c_{\mathrm{g}}^2}, \qquad c_4^\varepsilon = \frac{-1}{1 - c_{\mathrm{g}}^2}
$$

are negative constants and the analytic functions g_0^ε, g_1^ε, $g_2^\varepsilon : \mathcal{X}^{s+1} \to H_{\mathrm{per}}^{s+1}(\mathbb{R}^2)$ and constants $c_0, c_{1,1}, \ldots c_{1,4}, c_2$ are defined by

$$
\begin{aligned}
-c_3^\varepsilon + c_0 v_1 + g_0^\varepsilon(v) &= -k_0^2 \frac{1 - c_{\mathrm{p}}^2 + f_3(v_1, \partial_\eta v_1, v_2)}{1 - c_{\mathrm{g}}^2 + f_3(v_1, \partial_\eta v_1, v_2)}, \\
-c_4^\varepsilon v_1 + c_{1,1} v_1^2 + c_{1,2} (\partial_\eta v_1)^2 + c_{1,3} v_2^2 &+ c_{1,4} v_2 \partial_\eta v_1 + g_1^\varepsilon(v) \\
&= -\frac{f_4(v_1, \partial_\eta v_1, v_2) - v_1}{1 - c_{\mathrm{g}}^2 + f_3(v_1, \partial_\eta v_1, v_2)}, \\
c_2 v_1 + g_2^\varepsilon(v) &= -\frac{2 k_0 f_3(v_1, \partial_\eta v_1, v_2)}{1 - c_{\mathrm{g}}^2 + f_3(v_1, \partial_\eta v_1, v_2)},
\end{aligned}
$$

so that $g_j^\varepsilon(0) = 0$, $\mathrm{d}g_j^\varepsilon[0] = 0$, $j = 0,1,2$ and $\mathrm{d}^2 g_1^0[0] = 0$. (We denote parameter dependence, which is always analytic, of functions and constants by the superscript ε; functions and constants without the superscript do not depend upon ε.) The evolutionary system (14), (15) has a discrete symmetry which plays an important role in the following theory. It is *reversible,* that is invariant with respect to the transformation $\xi \mapsto -\xi$, $(v_1, v_2) \mapsto S(v_1, v_2)$, where the *reverser* S is defined by the formula

$$S(v_1(\eta), v_2(\eta)) = (v_1(-\eta), -v_2(-\eta)).$$

This symmetry has the consequence that $(v_1(-\xi, -\eta), -v_2(-\xi, -\eta))$ solves the equation whenever $(v_1(\xi, \eta), v_2(\xi, \eta))$ is a solution.

We may express an element of $H^s_{\mathrm{per}}(0, 2\pi)$ as a Fourier series

$$v_1(\eta) = \sqrt{\frac{1}{2\pi}} v_{1,0} + \sqrt{\frac{1}{\pi}} \sum_{m=1}^{\infty} \{v_{1,m,\mathrm{o}} \sin(m\eta) + v_{1,m,\mathrm{e}} \cos(m\eta)\}$$

and define projections P_0, $P_{m,\mathrm{o}}, P_{m,\mathrm{e}}, P_m : H^s_{\mathrm{per}}(0, 2\pi) \to H^s_{\mathrm{per}}(0, 2\pi)$ by the formulae

$$P_0 \left(\sqrt{\frac{1}{2\pi}} v_{1,0} + \sqrt{\frac{1}{\pi}} \sum_{j=1}^{\infty} (v_{1,j,\mathrm{o}} \sin(j\eta) + v_{1,j,\mathrm{e}} \cos(j\eta)) \right) = \sqrt{\frac{1}{2\pi}} v_{1,0}$$

and

$$\begin{aligned} P_{m,\mathrm{o}} \left(\sqrt{\frac{1}{2\pi}} v_{1,0} + \sqrt{\frac{1}{\pi}} \sum_{j=1}^{\infty} (v_{1,j,\mathrm{o}} \sin(j\eta) + v_{1,j,\mathrm{e}} \cos(j\eta)) \right) &= \sqrt{\frac{1}{\pi}} v_{1,m,\mathrm{o}} \sin(m\eta), \\ P_{m,\mathrm{e}} \left(\sqrt{\frac{1}{2\pi}} v_{1,0} + \sqrt{\frac{1}{\pi}} \sum_{j=1}^{\infty} (v_{1,j,\mathrm{o}} \sin(j\eta) + v_{1,j,\mathrm{e}} \cos(j\eta)) \right) &= \sqrt{\frac{1}{\pi}} v_{1,m,\mathrm{e}} \cos(m\eta) \end{aligned}$$

with $P_m = P_{m,\mathrm{o}} + P_{m,\mathrm{e}}$ for $m = 1, 2, \ldots$. By extending the Fourier series coordinatewise to vector-valued functions we find that $\mathcal{X}^s$ decomposes into a direct sum $\oplus_{m \in \mathbb{N}_0} E_m$ of subspaces, where

$$E_m = E_{m,\mathrm{o}} \oplus E_{m,\mathrm{e}}, \qquad E_{m,\mathrm{o}} = \{(v_{1,m,\mathrm{o}}, v_{2,m,\mathrm{o}})\}, \quad E_{m,\mathrm{e}} = \{(v_{1,m,\mathrm{e}}, v_{2,m,\mathrm{e}})\}.$$

We may therefore write

$$\mathcal{X}^s = \ell^{s+1} \times \ell^s,$$

in which

$$\ell^t = \{u \mid \|u\|_t^2 := |u_0|^2 + \sum_{m=1}^{\infty} m^{2t}(|u_{m,\mathrm{o}}|^2 + |u_{m,\mathrm{e}}|^2) < \infty\},$$

and $P_{m,\mathrm{o}}, P_{m,\mathrm{e}}, P_m$ also extend naturally to projections $\mathcal{X}^s \to \mathcal{X}^s$ which are denoted by the same symbols. Notice that P_m is infinitely smoothing due to its finite-dimensional range, so that

$$\|P_m v_1\|_{t_1} \leq C_{m,t_1,t_2} \|v_1\|_{t_2}, \qquad t_1 \geq t_2;$$

the same smoothing property is enjoyed by $P_{m,\mathrm{o}}$ and $P_{m,\mathrm{e}}$. The action of the reverser S in the new coordinate system is readily confirmed to be

$$S(v_{1,\mathrm{o}}, v_{1,\mathrm{e}}, v_{2,\mathrm{o}}, v_{2,\mathrm{e}}) = (-v_{1,\mathrm{o}}, v_{1,\mathrm{e}}, v_{2,\mathrm{o}}, -v_{2,\mathrm{e}}),$$

where $(v_{1,\mathrm{o}}, v_{2,\mathrm{o}}) = \{(v_{1,m,\mathrm{o}}, v_{2,m,\mathrm{o}})\}$, $(v_{1,\mathrm{e}}, v_{2,\mathrm{e}}) = \{(v_{1,m,\mathrm{e}}, v_{2,m,\mathrm{e}})\}$. Note also that the periodicity in y combines with the translation invariance in this variable to give an $O(2)$ symmetry represented in the new coordinates by

$$\{(v_{1,m,\mathrm{o}}, v_{1,m,\mathrm{e}}, v_{2,m,\mathrm{o}}, v_{2,m,\mathrm{e}})\} \mapsto \{(R_{ma}(v_{1,m,\mathrm{o}}, v_{1,m,\mathrm{e}}), R_{ma}(v_{2,m,\mathrm{o}}, v_{2,m,\mathrm{e}}))\}, \qquad a \in \mathbb{R},$$

where R_θ is the 2×2 matrix representing a rotation through the angle θ.

The spectrum of the linearised system

$$\partial_\xi v = L^\varepsilon v, \qquad L^\varepsilon \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_2 \\ -c_3^\varepsilon k_0^2 \partial_\eta^2 v_1 - c_4^\varepsilon v_1 \end{pmatrix}$$

associated with (14), (15) was calculated by Groves & Schneider [4]; we recall the complete result since extensive use is made of it in the following analysis. The mth Fourier component satisfies the ordinary differential equations

$$\begin{aligned} \partial_\xi v_{1,m} &= v_{2,m}, \\ \partial_\xi v_{2,m} &= \frac{m^2 k_0^2 (1 - c_\mathrm{p}^2) + 1}{(1 - c_\mathrm{g}^2)} v_{1,m}, \end{aligned}$$

and the associated eigenvalues $\lambda_{m,\varepsilon}$ of this system of equations are given by

$$\begin{aligned} \lambda_{m,\varepsilon}^2 &= \frac{m^2 k_0^2 (1 - c_\mathrm{p}^2) + 1}{(1 - c_\mathrm{g}^2)} \\ &= (k_0^2 + 1)(1 - m^2) - 2k_0 (1 + k_0^2)^{1/2} (k_0^2 + m^2) \gamma_1 \varepsilon^2 + \mathcal{O}(\varepsilon^4), \end{aligned}$$

in which the $\mathcal{O}(\varepsilon^4)$ estimate on the remainder term holds uniformly in m.

$\boldsymbol{m = 0}$: We have two simple, real eigenvalues $\pm\lambda_{0,\varepsilon} = \pm(1+k_0^2)^{1/2} + \mathcal{O}(\varepsilon^2)$. The corresponding eigenvectors are given by

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \pm\lambda_{0,\varepsilon} \end{pmatrix}.$$

$\boldsymbol{m = 1}$: For $\varepsilon = 0$ we have a geometrically simple and algebraically double zero eigenvalue in $E_{1,\mathrm{o}}$. The eigenvector and associated generalised eigenvector are given by

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \sin\eta, \qquad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \sin\eta.$$

For $\varepsilon > 0$ we have two simple eigenvalues $\pm\lambda_{1,\varepsilon}$ which satisfy the equation $(\lambda_{1,\varepsilon})^2 = -2k_0\gamma_1\varepsilon^2(1+k_0^2)^{3/2} + \mathcal{O}(\varepsilon^4)$; they are therefore real if $\gamma_1 < 0$. The eigenvectors are

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \pm\lambda_{1,\varepsilon} \end{pmatrix} \sin\eta.$$

The same result holds in $E_{1,\mathrm{e}}$ with $\sin\eta$ replaced by $\cos\eta$.

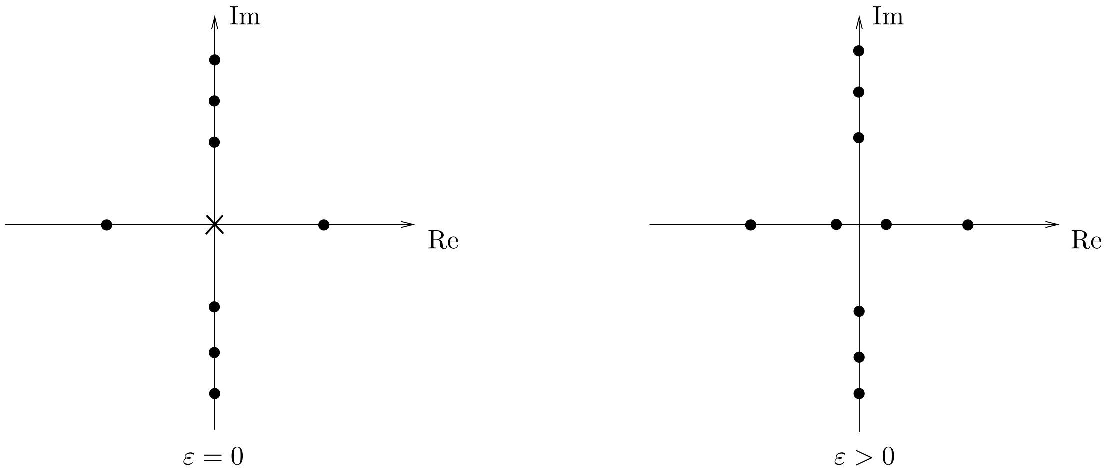

Figure 2: The spectrum of the linearised problem consists of infinitely many semisimple purely imaginary eigenvalues and two real semisimple real eigenvalues together with two Jordan blocks of length two at the origin for $\varepsilon = 0$ or two additional semisimple real eigenvalues for $\varepsilon > 0$; with the exception of the geometrically simple real eigenvalues of largest magnitude all eigenvalues have geometric multiplicity two.

$\boldsymbol{m > 1}$: We have two simple purely imaginary eigenvalues in $E_{m,\mathrm{o}}$ given by $\pm\mathrm{i}\omega_{m,\varepsilon}$, where $\omega_{m,\varepsilon} = \pm\mathrm{i}(m^2-1)^{1/2}(k_0^2+1)^{1/2} + \mathcal{O}(\varepsilon^2)$. The eigenvectors are

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \pm\lambda_{m,\varepsilon} \end{pmatrix} \sin(m\eta).$$

The same result holds in $E_{m,\mathrm{e}}$ with $\sin(m\eta)$ replaced by $\cos(m\eta)$.

The eigenvalue picture is summarised in Figure 2; for $\varepsilon > 0$ we have a two-dimensional strongly hyperbolic part $\mathcal{X}^s_{\mathrm{sh}} = E_0$, a four-dimensional weakly hyperbolic part $\mathcal{X}^s_{\mathrm{wh}} = E_1$ and an infinite-dimensional central part $\mathcal{X}^s_{\mathrm{c}} = \oplus_{m=2}^{\infty} E_m$ of phase space. Notice that P_0 is the projection onto the strongly hyperbolic subspace $\mathcal{X}^s_{\mathrm{sh}}$ along the central and weakly hyperbolic subspace $\mathcal{X}^s_{\mathrm{wh}} \cup \mathcal{X}^s_{\mathrm{c}}$, while P_1 is the projection onto the weakly hyperbolic subspace $\mathcal{X}^s_{\mathrm{wh}}$ along the central and strongly hyperbolic subspace $\mathcal{X}^s_{\mathrm{sh}} \cup \mathcal{X}^s_{\mathrm{c}}$. In the theory below we therefore write P_{sh} for P_0, P_{wh} for P_1, P_{c} for $I - P_{\mathrm{sh}} - P_{\mathrm{wh}}$ and also define $P_{\mathrm{sh,c}} = P_{\mathrm{sh}} + P_{\mathrm{c}}$; we use the notation $(Z_1, Z_2) = P_{\mathrm{sh}}(v_1, v_2)$, $(z_1, z_2) = P_{\mathrm{wh}}(v_1, v_2)$, $(w_1, w_2) = P_{\mathrm{c}}(v_1, v_2)$ and $(q_1, q_2) = P_{\mathrm{sh,c}}(v_1, v_2)$, so that $L^\varepsilon(Z_1, Z_2) = (Z_2, \lambda_{0,\varepsilon}^2 Z_1)$, $L^\varepsilon(z_1, z_2) = (z_2, \lambda_{1,\varepsilon}^2 z_1)$ and $L^\varepsilon\{(q_{1,m}, q_{2,m})\}_{m\geq 2} = \{(q_{2,m}, -\omega_{m,\varepsilon}^2 q_{1,m})\}_{m\geq 2}$.

One may formulate equations (14), (15) as the coupled four-dimensional dynamical system

$$\partial_\xi z = Kz + F^\varepsilon(z, q), \tag{16}$$

where

$$K = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad F^\varepsilon(z, q) = \begin{pmatrix} 0 \\ \lambda_{1,\varepsilon}^2 z + f^\varepsilon(z, q) \end{pmatrix},$$

$$\begin{aligned} f^\varepsilon(z, q) = P_{\mathrm{wh}}[&(c_0(z_1 + q_1) + g_0^\varepsilon(z + q))\partial_\eta^2(z_1 + q_1) \\ &+ c_{1,1}(z_1 + q_1)^2 + c_{1,2}(\partial_\eta(z_1 + q_1))^2 + c_{1,3}(z_2 + q_2)^2 + c_{1,4}(z_2 + q_2)\partial_\eta(z_1 + q_1) \\ &+ g_1^\varepsilon(z + q) + (c_2(z_1 + q_1) + g_2^\varepsilon(z + q))\partial_\eta(z_2 + q_2)], \end{aligned}$$

and quasilinear wave equation

$$\begin{aligned} \partial_\xi q_1 &= q_2, && (17)\\ \partial_\xi q_2 &= -c_3^\varepsilon k_0^2 \partial_\eta^2 q_1 - c_4^\varepsilon q_1 + P_{\mathrm{sh,c}}(d_1(z)q_1) + P_{\mathrm{sh,c}}(d_2(z)\partial_\eta q_1) + P_{\mathrm{sh,c}}(d_3(z)q_2) \\ &\qquad + P_{\mathrm{sh,c}}(g_3^\varepsilon(z,q)\partial_\eta^2 q_1) + g_4^\varepsilon(z,q) + P_{\mathrm{sh,c}}(g_5^\varepsilon(z,q)\partial_\eta q_2) + h^\varepsilon(z), && (18) \end{aligned}$$

where

$$\begin{aligned} d_1(z) &= c_0\partial_\eta^2 z_1 + 2c_{1,1}z_1 + c_2\partial_\eta z_2,\\ d_2(z) &= 2c_{1,2}\partial_\eta z_1 + c_{1,4}z_2,\\ d_3(z) &= 2c_{1,3}z_2 + c_{1,4}\partial_\eta z_1,\\ g_3^\varepsilon(z,q) &= c_0(q_1+z_1) + g_0^\varepsilon(z+q),\\ g_4^\varepsilon(z,q) &= P_{\mathrm{sh,c}}[(g_0^\varepsilon(z+q) - g_0^\varepsilon(z))\partial_\eta^2 z_1 + (g_2^\varepsilon(z+q) - g_2^\varepsilon(z))\partial_\eta z_2 \\ &\qquad + c_{1,1}q_1^2 + c_{1,2}(\partial_\eta q_1)^2 + c_{1,3}q_2^2 + c_{1,4}q_2\partial_\eta q_1 + g_1^\varepsilon(z+q) - g_1^\varepsilon(z)],\\ g_5^\varepsilon(z,q) &= c_2(q_1+z_1) + g_2^\varepsilon(z+q),\\ h^\varepsilon(z) &= P_{\mathrm{sh,c}}[(c_0 z_1 + g_0^\varepsilon(z))\partial_\eta^2 z_1 + (c_2 z_1 + g_2^\varepsilon(z))\partial_\eta z_2 \\ &\qquad + c_{1,1}z_1^2 + c_{1,2}(\partial_\eta z_1)^2 + c_{1,3}z_2^2 + c_{1,4}z_2\partial_\eta z_1 + g_1^\varepsilon(z)]; \end{aligned}$$

the linear functions d_j, $j=1,2,3$ and nonlinearities in equations (16)–(18) satisfy the estimates

$$\begin{aligned} \|d_j(z)\|_{s+1} &= \mathcal{O}(|z|), \qquad j=1,2,3,\\ |f^\varepsilon(z,q)| &= \mathcal{O}(\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}\|(z,q)\|_{\mathcal{X}^{s+1}} + \|(z,q)\|_{\mathcal{X}^{s+1}}^3),\\ |F^\varepsilon(z,q)| &= \mathcal{O}(\varepsilon^2|z| + \|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}\|(z,q)\|_{\mathcal{X}^{s+1}} + \|(z,q)\|_{\mathcal{X}^{s+1}}^3),\\ \|g_j^\varepsilon(z,q)\|_{s+1} &= \mathcal{O}(\|(z,q)\|_{\mathcal{X}^{s+1}}), \qquad j=3,5,\\ \|g_4^\varepsilon(z,q)\|_{s+1} &= \mathcal{O}(\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}^2 + \|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}|z||(z,\varepsilon)|),\\ \|h^\varepsilon(z)\|_{s+1} &= \mathcal{O}(|z|^2). \end{aligned}$$

We use this formulation in the remainder of the article.

3 Approximate modulating pulses

In this section we construct a pair of approximate modulating pulse solutions to equations (16)–(18) from which we later obtain genuine modulating pulses by perturbation arguments. Consider the approximate problem obtained by removing the term $h^\varepsilon(z)$ from equation (18). This approximate problem has the property that $E_1 = \{q=0\}$ is an invariant subspace, the flow in which is given by the equation

$$\partial_\xi z = Kz + F^\varepsilon(z,0),$$

where $|F^\varepsilon(z,0)| = O(|z|^3)$. Under a sign condition on the cubic part of the nonlinearity, the above equation admits a pair of small-amplitude homoclinic orbits $p^{\varepsilon\pm}$ for small, positive values of ε (see below), and each of these orbits serves as an approximate modulating pulse.

Our strategy is therefore to use a sequence of changes of variable which systematically removes the term $h_j^\varepsilon(z)$ that is homogeneous of degree j in (z,ε) from h^ε while preserving the overall structure of the equations. This procedure is carried out in Section 4 below, where it is shown that the remaining terms in h^ε can be made exponentially small in comparison to ε, so that $p^{\varepsilon\pm}$ become very good approximations to genuine modulating pulses. The transformation which eliminates the term $h_2^0(z)$ (the term that is homogeneous of degree two in z and does not depend upon ε) affects the coefficient in $F^\varepsilon(z,0)$ whose sign determines whether homoclinic bifurcation takes place; we therefore consider this transformation as a separate preliminary step.

Lemma 3.1 *There is a near-identity, finite-dimensional change of coordinates which transforms the coupled systems (16), and (17), (18) into*

$$\partial_\xi z = Kz + \tilde{F}^\varepsilon(z,q), \tag{19}$$

$$\partial_\xi q_1 = q_2 + g_6^\varepsilon(q,z) + \tilde{h}_1^\varepsilon(z), \tag{20}$$

$$\begin{aligned}\partial_\xi q_2 &= -c_3^\varepsilon k_0^2\partial_\eta^2 q_1 - c_4^\varepsilon q_1 + P_{\mathrm{sh,c}}(d_1(z)q_1) + P_{\mathrm{sh,c}}(d_2(z)\partial_\eta q_1) + P_{\mathrm{sh,c}}(d_3(z)q_2)\\ &\quad + P_{\mathrm{sh,c}}(\tilde{g}_3^\varepsilon(z,q)\partial_\eta^2 q_1) + \tilde{g}_4^\varepsilon(z,q) + P_{\mathrm{sh,c}}(\tilde{g}_5^\varepsilon(z,q)\partial_\eta q_2) + \tilde{h}_2^\varepsilon(z)\end{aligned} \tag{21}$$

and preserves the reversibility. The nonlinearities $\tilde{F}^\varepsilon$, $\tilde{g}_3^\varepsilon$, $\tilde{g}_4^\varepsilon$, $\tilde{g}_5^\varepsilon$ satisfy the same estimates as respectively F^ε, g_3^ε, g_4^ε, g_5^ε, while

$$\|g_6^\varepsilon(z,\tilde{q})\|_{s+2} = \mathcal{O}(|z|\|\tilde{q}\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}\|(z,\tilde{q})\|_{\mathcal{X}^{s+1}}), \qquad \|\tilde{h}^\varepsilon(z)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} = \mathcal{O}(|z|^2|(z,\varepsilon)|).$$

Proof. Write $h^\varepsilon(z) = h_2^0(z) + \hat{h}^\varepsilon(z)$, where $h_2^0(z)$ is the part of $h^\varepsilon(z)$ that is homogeneous of degree j in (z,ε), so that $\|\hat{h}^\varepsilon(z)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} = \mathcal{O}(|z|^2|(z,\varepsilon)|)$. Observe that h_2^0 is a mapping from E_1 to $E_0 \oplus E_2$, and this fact suggests using a finite-dimensional change of coordinates of the form

$$\tilde{q} = q + \Gamma(z),$$

where $P_m\Gamma(z) = 0$ for $m \neq 0,2$. Substituting $q = \tilde{q} - \Gamma(z)$ into (16) and (17), (18), one finds that they are transformed into respectively (19) and (20), (21) with

$$\begin{aligned}
\tilde{F}^\varepsilon(z,q) &= \begin{pmatrix} 0 \\ \lambda_{1,\varepsilon}^2 z + \tilde{f}^\varepsilon(z,q)\end{pmatrix},\\
\tilde{f}^\varepsilon(z,\tilde{q}) &= f^\varepsilon(z,\tilde{q}-\Gamma(z)),\\
\tilde{g}_j^\varepsilon(z,\tilde{q}) &= g_j^\varepsilon(z,\tilde{q}-\Gamma(z)), \qquad j=3,5,\\
\tilde{g}_4^\varepsilon(z,\tilde{q}) &= g_4^\varepsilon(z,\tilde{q}-\Gamma(z)) - g_4^\varepsilon(z,-\Gamma(z))\\
&\qquad - P_{\mathrm{sh,c}}(g_3^\varepsilon(z,\tilde{q}-\Gamma(z)) - g_3^\varepsilon(z,-\Gamma(z)))\partial_\eta^2\Gamma_1(z)\\
&\qquad - P_{\mathrm{sh,c}}(g_5^\varepsilon(z,\tilde{q}-\Gamma(z)) - g_5^\varepsilon(z,-\Gamma(z)))\partial_\eta\Gamma_2(z)\\
&\qquad + \mathrm{d}\Gamma_2[z](F^\varepsilon(z,\tilde{q}-\Gamma(z)) - F^\varepsilon(z,-\Gamma(z))),\\
g_6^\varepsilon(z,\tilde{q}) &= \mathrm{d}\Gamma_1[z](F^\varepsilon(z,\tilde{q}-\Gamma(z)) - F^\varepsilon(z,-\Gamma(z))),
\end{aligned}$$

and

$$
\begin{aligned}
\tilde{h}^\varepsilon(z) = {} & -L^\varepsilon\Gamma(z) + \mathrm{d}\Gamma[z](Kz) + \mathrm{d}\Gamma[z](F^\varepsilon(z,-\Gamma(z))) \\
& + \begin{pmatrix} 0 \\ -P_{\mathrm{sh,c}}(d_1(z)\Gamma_1(z)) - P_{\mathrm{sh,c}}(d_2(z)\partial_\eta\Gamma_1(z)) - P_{\mathrm{sh,c}}(d_3(z)\Gamma_2(z)) \end{pmatrix} \\
& + \begin{pmatrix} 0 \\ g_4^\varepsilon(z,-\Gamma(z)) - P_{\mathrm{sh,c}}[g_3^\varepsilon(z,-\Gamma(z))\partial_\eta^2\Gamma_1(z) + g_5^\varepsilon(z,-\Gamma(z))\partial_\eta\Gamma_2(z)] + \hat{h}^\varepsilon(z) \end{pmatrix};
\end{aligned}
$$

our objective is therefore achieved by choosing $\Gamma(z)$ to be a polynomial which is homogeneous of degree two and satisfies

$$
L^0\Gamma(z) - \mathrm{d}\Gamma[z](Kz) = \begin{pmatrix} 0 \\ h_2^0(z) \end{pmatrix}. \tag{22}
$$

Notice that (22) decomposes into component equations for $P_0\Gamma(z)$, $P_{2,\mathrm{o}}\Gamma(z)$ and $P_{2,\mathrm{e}}\Gamma(z)$ in respectively E_0, $E_{2,\mathrm{o}}$ and $E_{2,\mathrm{e}}$. Let $\mathcal{R}^2$ denote the space of $\mathbb{R}^2$-valued polynomials of degree two in the variables $z_{1,\mathrm{o}}$, $z_{1,\mathrm{e}}$, $z_{2,\mathrm{o}}$, $z_{2,\mathrm{e}}$, equip $\mathcal{R}^2$ with the basis

$$
\mathcal{B} = \{(1,0)^{\mathrm{T}} z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^\ell, (0,1)^{\mathrm{T}} z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^\ell\}_{i+j+k+\ell=2}
$$

and consider the linear operator $\mathcal{L} : \mathcal{R}^2 \to \mathcal{R}^2$ defined by $(\mathcal{L}\Gamma)(z) = L_0\Gamma(z) - \mathrm{d}\Gamma[z](Kz)$. Using the calculations

$$
\begin{aligned}
& \mathcal{L}((1,0)^{\mathrm{T}} z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^\ell) \\
& \qquad = (0,1)^{\mathrm{T}} z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^\ell - (1,0)^{\mathrm{T}}(i z_{1,\mathrm{o}}^{i-1} z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^{k+1} z_{2,\mathrm{e}}^\ell + j z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^{j-1} z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^{\ell+1}), \\
& \mathcal{L}((0,1)^{\mathrm{T}} z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^\ell) \\
& \qquad = \lambda_{1,\varepsilon}^2 (1,0)^{\mathrm{T}} z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^\ell - (0,1)^{\mathrm{T}}(i z_{1,\mathrm{o}}^{i-1} z_{1,\mathrm{e}}^j z_{2,\mathrm{o}}^{k+1} z_{2,\mathrm{e}}^\ell + j z_{1,\mathrm{o}}^i z_{1,\mathrm{e}}^{j-1} z_{2,\mathrm{o}}^k z_{2,\mathrm{e}}^{\ell+1}),
\end{aligned}
$$

to compute the matrix of $\mathcal{L}$ with respect to $\mathcal{B}$, we find that this matrix is invertible, so that the component equation of (22) for $P_0\Gamma(z)$ has a unique solution. A similar argument yields $P_{2,\mathrm{o}}\Gamma(z)$ and $P_{2,\mathrm{e}}\Gamma(z)$, and one concludes that (22) admits a unique solution for $\Gamma(z)$. □

Let us now examine the system of ordinary differential equations

$$
\partial_\xi z_{1,\mathrm{o}} = z_{2,\mathrm{o}}, \qquad \partial_\xi z_{1,\mathrm{e}} = z_{2,\mathrm{e}}, \tag{23}
$$

$$
\partial_\xi z_{2,\mathrm{o}} = \lambda_{1,\varepsilon}^2 z_{1,\mathrm{o}} + \tilde{f}_{\mathrm{o}}^\varepsilon(z_{\mathrm{o}}, z_{\mathrm{e}}), \qquad \partial_\xi z_{2,\mathrm{e}} = \lambda_{1,\varepsilon}^2 z_{1,\mathrm{e}} + \tilde{f}_{\mathrm{e}}^\varepsilon(z_{\mathrm{o}}, z_{\mathrm{e}}), \tag{24}
$$

where $\tilde{f}_{\mathrm{o}}^\varepsilon = P_{1,\mathrm{o}}\tilde{f}^\varepsilon|_{q=0}$ and $\tilde{f}_{\mathrm{e}}^\varepsilon = P_{1,\mathrm{e}}\tilde{f}^\varepsilon|_{q=0}$ are $\mathcal{O}(|(z_{\mathrm{o}}, z_{\mathrm{e}})|^2|(z_{\mathrm{o}}, z_{\mathrm{e}}, \varepsilon^2)|)$, which is obtained by neglecting $\tilde{h}^\varepsilon$ and setting $q = 0$ in equation (19). This system inherits the reversibility and $O(2)$ symmetry of equations (16)–(18): it is invariant under the transformation $\xi \mapsto -\xi$, $(z_{1,\mathrm{o}}, z_{1,\mathrm{e}}, z_{2,\mathrm{o}}, z_{2,\mathrm{e}}) \mapsto S_{\mathrm{h}}(z_{1,\mathrm{o}}, z_{1,\mathrm{e}}, z_{2,\mathrm{o}}, z_{2,\mathrm{e}})$, where the reverser S_{h} is defined by

$$
S_{\mathrm{h}}(z_{1,\mathrm{o}}, z_{1,\mathrm{e}}, z_{2,\mathrm{o}}, z_{2,\mathrm{e}}) = (-z_{1,\mathrm{o}}, z_{1,\mathrm{e}}, z_{2,\mathrm{o}}, -z_{2,\mathrm{e}}),
$$

and under the transformation

$$
\begin{pmatrix} z_{1,\mathrm{o}} \\ z_{1,\mathrm{e}} \end{pmatrix} \mapsto R_a \begin{pmatrix} z_{1,\mathrm{o}} \\ z_{1,\mathrm{e}} \end{pmatrix}, \quad \begin{pmatrix} z_{2,\mathrm{o}} \\ z_{2,\mathrm{e}} \end{pmatrix} \mapsto R_a \begin{pmatrix} z_{2,\mathrm{o}} \\ z_{2,\mathrm{e}} \end{pmatrix}
$$

for each $a \in [0, 2\pi)$. Introducing the scaled variables

$$\check{\xi} = \varepsilon\xi, \quad z_1(\xi) = \varepsilon \check{z}_1(\check{\xi}), \quad z_2(\xi) = \varepsilon^2 \check{z}_2(\check{\xi}),$$

one finds from (23)–(24) that

$$\begin{aligned} \partial_{\check{\xi}} \check{z}_{1,\mathrm{o}} &= \check{z}_{2,\mathrm{o}}, & (25)\\ \partial_{\check{\xi}} \check{z}_{2,\mathrm{o}} &= \check{C}_1 \check{z}_{1,\mathrm{o}} - \check{C}_2 \check{z}_{1,\mathrm{o}}(\check{z}_{1,\mathrm{o}}^2 + \check{z}_{1,\mathrm{e}}^2) + \mathcal{R}_{\mathrm{o}}^{\varepsilon}(\check{z}_{1,\mathrm{o}}, \check{z}_{1,\mathrm{e}}, \check{z}_{2,\mathrm{o}}, \check{z}_{2,\mathrm{e}}), & (26) \end{aligned}$$

$$\begin{aligned} \partial_{\check{\xi}} \check{z}_{1,\mathrm{e}} &= \check{z}_{2,\mathrm{e}}, & (27)\\ \partial_{\check{\xi}} \check{z}_{2,\mathrm{e}} &= \check{C}_1 \check{z}_{1,\mathrm{e}} - \check{C}_2 \check{z}_{1,\mathrm{e}}(\check{z}_{1,\mathrm{o}}^2 + \check{z}_{1,\mathrm{e}}^2) + \mathcal{R}_{\mathrm{e}}^{\varepsilon}(\check{z}_{1,\mathrm{o}}, \check{z}_{1,\mathrm{e}}, \check{z}_{2,\mathrm{o}}, \check{z}_{2,\mathrm{e}}), & (28) \end{aligned}$$

in which

$$\check{C}_1 = -2k_0\gamma_1(1 + k_0^2)^{3/2} > 0$$

and the remainder terms $\mathcal{R}_{\mathrm{o}}^{\varepsilon}$ and $\mathcal{R}_{\mathrm{e}}^{\varepsilon}$ are both $\mathcal{O}(\varepsilon^2)$ and respectively odd and even in $(\check{z}_{1,\mathrm{o}}, \check{z}_{2,\mathrm{e}})$. The remaining coefficient $\check{C}_2$ is is given by the formulae

$$\begin{aligned} \check{C}_2 &= \frac{1}{\pi^{3/2}} f^{0;3,0}[\{\cos\eta\}^{(3)}, 0] + \frac{1}{\pi^{3/2}} f^{0;1,1}[\cos\eta, -\Gamma_{0200}] \\ &= \frac{1}{\pi^{3/2}} f^{0;3,0}[\{\sin\eta\}^{(3)}, 0] + \frac{1}{\pi^{3/2}} f^{0;1,1}[\sin\eta, -\Gamma_{2000}], \end{aligned} \qquad (29)$$

and is required to be positive. Here we have adopted the notation

$$f^{0;n_1,n_2} = \frac{1}{(n_1+n_2)!} \mathrm{d}_{1,2}^{n_1,n_2} f^0[0,0], \qquad \Gamma(z) = \sum_{i+j+k+\ell=2} \Gamma_{ijk\ell} z_{1\mathrm{o}}^i z_{1\mathrm{e}}^j z_{2\mathrm{o}}^k z_{2\mathrm{e}}^\ell$$

and $\{u\}^{(n)}$ is an abbreviation for the n-tuple $(u, \ldots, u)$. In the limit $\varepsilon \to 0$ the system (25)–(26) has the property that the $(\check{z}_{1,\mathrm{e}}, \check{z}_{2,\mathrm{e}})$ coordinate plane is invariant; its phase portrait is shown in Figure 3. In fact each orbit in the four-dimensional phase space of the limiting equations is obtained from an orbit in the $(\check{z}_{1,\mathrm{e}}, \check{z}_{2,\mathrm{e}})$ coordinate plane by a rotation R_a for some $a \in (0, 2\pi)$ (so that each subspace $(R_a(0, \check{z}_{1,\mathrm{e}}), R_a(0, \check{z}_{2,\mathrm{e}}))$, $a \in (0, 2\pi)$ is invariant). Notice in particular that the $(\check{z}_{1,\mathrm{e}}, \check{z}_{2,\mathrm{e}})$ coordinate plane contains two homoclinic orbits $\check{p}^{\pm}$ given by the explicit formulae

$$\check{p}_{1,\mathrm{e}}^{\pm}(\check{\xi}) = \pm\left(\frac{2\check{C}_1}{\check{C}_2}\right)^{1/2} \mathrm{sech}(\check{C}_1^{1/2}\check{\xi}), \qquad \check{p}_{2,\mathrm{e}}^{\pm}(\check{\xi}) = \mp\left(\frac{2\check{C}_1^2}{\check{C}_2}\right)^{1/2} \mathrm{sech}(\check{C}_1^{1/2}\check{\xi})\tanh(\check{C}_1^{1/2}\check{\xi}).$$

These orbits are *reversible*, that is they satisfy $S_{\mathrm{h}}\check{p}^{\pm}(-\xi) = \check{p}^{\pm}(\xi)$, and this feature can be exploited to prove their persistence for small values of ε. The necessary argument is given by Groves & Schneider [5, §4] and yields Lemma 3.2 below; it is based upon the fact that the stable manifold to the zero equilibrium of the limiting equations (which is the two-dimensional manifold $\{(R_a(0, \check{p}_{1,\mathrm{e}}^{\pm}(\xi)), R_a(0, \check{p}_{2,\mathrm{e}}^{\pm}(\xi)) \mid a \in [0, 2\pi), \xi \in \mathbb{R}\}$) intersects the symmetric section Fix S_{h} transversally at the points $\check{p}^{\pm}(0)$.

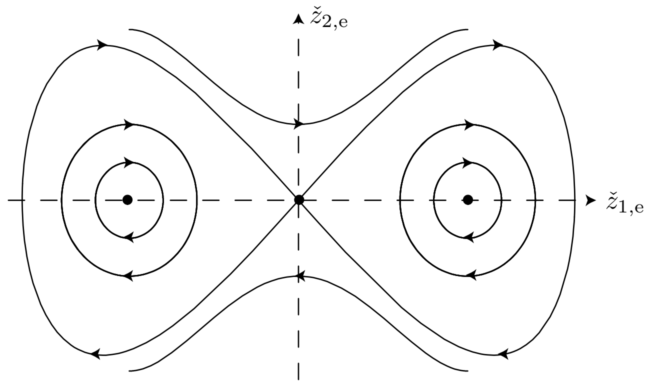

Figure 3: Dynamics in the $(\check{z}_{1,\mathrm{e}}, \check{z}_{2,\mathrm{e}})$ coordinate plane.

Lemma 3.2 *Equations (23), (24) admit a pair $p^{\varepsilon\pm}$ of reversible homoclinic orbits of the form*

$$\begin{pmatrix} p_1^{\varepsilon\pm}(\xi) \\ p_2^{\varepsilon\pm}(\xi) \end{pmatrix} = \begin{pmatrix} \varepsilon\tilde{p}_1^{\varepsilon\pm}(\varepsilon\xi) \\ \varepsilon^2\tilde{p}_2^{\varepsilon\pm}(\varepsilon\xi) \end{pmatrix},$$

where $\tilde{p}_1^{\varepsilon\pm}$, $\tilde{p}_2^{\varepsilon\pm}$ are smooth functions with bounded derivatives. These homoclinic orbits satisfy

$$|p_1^{\varepsilon\pm}(\xi)| \leq c\varepsilon \mathrm{e}^{-\theta\varepsilon|\xi|}, \quad |p_2^{\varepsilon\pm}(\xi)| \leq c\varepsilon^2 \mathrm{e}^{-\theta\varepsilon|\xi|}, \qquad \xi \in \mathbb{R}$$

for any $\theta \in (0, (-2k_0\gamma_1)^{1/2}(1+k_0^2)^{3/4})$.

The next step is a partial normal-form theory which eliminates sufficiently many terms in the Taylor expansion of $h^\varepsilon(z)$ that the remaining terms are exponentially small in comparison with ε. A central requirement of the relevant transformation theory, which is based upon a theory for finite-dimensional dynamical systems given by Iooss & Lombardi [7], is that the linearised vector field in the dynamical system for z should be diagonalisable; this condition ensures that certain estimates hold uniformly in the order of the terms eliminated from the vector field (see Lemma 4.1). The matrix K clearly does not meet this criterion, and this difficulty is overcome by introducing the following scaled variables. Writing $\varepsilon = \mu^2$ and defining

$$z_1' = \mu^{-1}z_1, \quad z_2' = \check{C}_1^{-1/2}\mu^{-3}z_2, \quad (q_1', q_2') = \mu^{-2}(q_1, q_2),$$

one finds that (19)–(21) are transformed into

$$\partial_\xi z = F^\mu(z, q), \tag{30}$$

$$\begin{aligned} \partial_\xi q_1 &= q_2 + g_6^\mu(z,q) + h_1^\mu(z), && (31)\\ \partial_\xi q_2 &= -c_3^\mu k_0^2 \partial_\eta^2 q_1 - c_4^\mu q_1 + \mu P_{\mathrm{sh,c}}(d_1(z)q_1) + \mu P_{\mathrm{sh,c}}(d_2(z)\partial_\eta q_1) + \mu P_{\mathrm{sh,c}}(d_3(z)q_2) \\ &\quad + P_{\mathrm{sh,c}}(g_3^\mu(z,q)\partial_\eta^2 q_1) + g_4^\mu(z,q) + P_{\mathrm{sh,c}}(g_5^\mu(z,q)\partial_\eta q_2) + h_2^\mu(z), && (32) \end{aligned}$$

in which

$$F^\mu(z,q) = \begin{pmatrix} \check{C}_1^{1/2}\mu^2 z_2 \\ \check{C}_1^{1/2}\mu^2 z_1 + (\lambda_{1,\mu}^2 - \check{C}_1\mu^4)z_1 + \check{C}_1^{-1/2}\mu^{-3}\tilde{f}^\varepsilon(\mu z_1, \check{C}_1^{1/2}\mu^3 z_2, \mu^2 q) \end{pmatrix},$$

$$
\begin{aligned}
g_j^\mu(z,q) &= \tilde{g}_j^\varepsilon(\mu z_1, \mu^3 \check{C}_1^{1/2} z_2, \mu^2 q), \qquad j=3,5,\\
g_j^\mu(z,q) &= \mu^{-2}\tilde{g}_j^\varepsilon(\mu z_1, \check{C}_1^{1/2}\mu^3 z_2, \mu^2 q), \quad j=4,6,\\
h^\mu(z) &= \mu^{-2}\tilde{h}^\varepsilon(\mu z_1, \check{C}_1^{1/2}\mu^3 z_2)
\end{aligned}
$$

and, with a slight abuse of notation, we have abbreviated $\lambda_{m,\varepsilon}|_{\varepsilon=\mu^2}$, $\omega_{m,\varepsilon}|_{\varepsilon=\mu^2}$, $c_3^\varepsilon|_{\varepsilon=\mu^2}$ and $c_4^\varepsilon|_{\varepsilon=\mu^2}$ to respectively $\lambda_{m,\mu}$, $\omega_{m,\mu}$, c_3^μ and c_4^μ (the primes have been dropped for notational simplicity). The linear part of the vector field on the right-hand side of (30) at $\mu=0$ is the zero matrix, which is trivially diagonalisable, while the nonlinearities in (30)–(32) satisfy the estimates

$$
\begin{aligned}
|F^\mu(z,q)| &= O(\mu^2|z| + \|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}\|(z,q)\|_{\mathcal{X}^{s+1}} + \|(z,q)\|_{\mathcal{X}^{s+1}}^3),\\
\|g_j^\mu(z,q)\|_{s+1} &= \mathcal{O}(\mu\|(z,q)\|_{\mathcal{X}^{s+1}}), \qquad j=3,5,\\
\|g_4^\mu(z,q)\|_{s+1} &= O(\mu^2\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}^2 + \mu^2\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}|z||(z,\mu)|),\\
\|g_6^\mu(z,q)\|_{s+2} &= O(\mu^2|z|\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}\|(z,q)\|_{\mathcal{X}^{s+1}}),\\
\|h^\mu(z)\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}} &= \mathcal{O}(\mu|z|^2|(z,\mu)|).
\end{aligned}
$$

The homoclinic orbits detected in Lemma 3.2 are denoted in the new variables by $p^{\mu\pm}$ and satisfy the estimate

$$
|p^{\mu\pm}(\xi)| \le c_{\mathrm{h}}\mu \mathrm{e}^{-\mu^2\theta|\xi|}, \qquad \xi \in \mathbb{R};
$$

for notational simplicity we henceforth use the symbol p^μ to denote either of the functions $p^{\mu+}$, $p^{\mu-}$.

4 Normal-form theory

4.1 Construction of the normal-form transformation

Our normal-form theory consists of a sequence of changes of variable which systematically removes the terms $h_j^\mu(z)$, $j=3,\ldots,p$ that are homogeneous of degree j in (z,μ) from $h^\mu(z)$ while preserving the overall structure of equations (30)–(32). It is possible to make an optimal choice of p so that the remaining terms are exponentially small in comparison to μ; the functions $p^{\mu\pm}$ found in Section 3 therefore become very good approximations to genuine modulating pulse solutions and can be used as the starting point for a perturbation argument to find genuine modulating pulses. Our analysis is based upon a theory for finite-dimensional dynamical systems given by Iooss & Lombardi [7], and we use their notation and refer to several of their combinatorial results here.

The dependence of our equations upon μ is accommodated by introducing the new variable $y=(z,\mu)$ and attaching the additional equation

$$
\partial_\xi \mu = 0
$$

to equation (30); in this notation equations (30)–(32) are written as

$$
\partial_\xi y = F(y,q), \tag{33}
$$

$$\begin{aligned}\partial_\xi q_1 &= q_2 + g_6(y,q) + h_1(y), &(34)\\ \partial_\xi q_2 &= -c_3^\mu k_0^2 \partial_\eta^2 q_1 - c_4^\mu q_1 + P_{\rm sh,c}(d_4(y)q_1) + P_{\rm sh,c}(d_5(y)\partial_\eta q_1) + P_{\rm sh,c}(d_6(y)q_2)\\ &\qquad + P_{\rm sh,c}(g_3(y,q)\partial_\eta^2 q_1) + g_4(y,q) + P_{\rm sh,c}(g_5(y,q)\partial_\eta q_2) + h_2(y), &(35)\end{aligned}$$

where

$$F(y,q) = \begin{pmatrix} F^\mu(z,q) \\ 0 \end{pmatrix},$$

$$g_j(y,q) = g_j^\mu(z,q), \quad j=3,\ldots,6, \qquad h_j(y) = h_j^\mu(z), \quad j=1,2,$$

$$d_{j+3}(y) = \mu d_j(z), \quad j=1,2,3.$$

We use a change of variable of the form

$$\tilde{q} = q + \Phi(y),$$

where $\Phi : \mathbb{R}^5 \to \mathcal{X}_{\rm sh,c}^{s+2}$, which transforms equations (33)–(35) into

$$\partial_\xi y = \tilde{F}(y,\tilde{q}), \tag{36}$$

$$\begin{aligned}\partial_\xi \tilde{q}_1 &= \tilde{q}_2 + \tilde{g}_6(y,\tilde{q}) + \tilde{h}_1(y), &(37)\\ \partial_\xi \tilde{q}_2 &= -c_3^\mu k_0^2 \partial_\eta^2 \tilde{q}_1 - c_4^\mu \tilde{q}_1 + P_{\rm sh,c}(d_4(y)\tilde{q}_1) + P_{\rm sh,c}(d_5(y)\partial_\eta \tilde{q}_1) + P_{\rm sh,c}(d_6(y)\tilde{q}_2)\\ &\qquad + P_{\rm sh,c}(\tilde{g}_3(y,\tilde{q})\partial_\eta^2 \tilde{q}_1) + \tilde{g}_4(y,\tilde{q}) + P_{\rm sh,c}(\tilde{g}_5(y,\tilde{q})\partial_\eta \tilde{q}_2) + \tilde{h}_2(y). &(38)\end{aligned}$$

Here

$$\begin{aligned}\tilde{F}(y,\tilde{q}) &= F(y,\tilde{q}-\Phi(y)), &(39)\\ \tilde{g}_j(y,\tilde{q}) &= g_j(y,\tilde{q}-\Phi(y)), \qquad j=3,5, &(40)\\ \tilde{g}_4(y,\tilde{q}) &= g_4(y,\tilde{q}-\Phi(y)) - g_4(y,-\Phi(y)) + \mathrm{d}\Phi_2[y](F(y,\tilde{q}-\Phi(y)) - F(y,-\Phi(y)))\\ &\qquad - P_{\rm sh,c}(g_3(y,\tilde{q}-\Phi(y)) - g_3(y,-\Phi(y)))\partial_\eta^2\Phi_1(y)\\ &\qquad - P_{\rm sh,c}(g_5(y,\tilde{q}-\Phi(y)) - g_5(y,-\Phi(y)))\partial_\eta\Phi_2(y), &(41)\\ \tilde{g}_6(y,\tilde{q}) &= g_6(y,\tilde{q}-\Phi(y)) - g_6(y,-\Phi(y)) + \mathrm{d}\Phi_1[y](F(y,\tilde{q}-\Phi(y)) - F(y,-\Phi(y))), &(42)\end{aligned}$$

and

$$\tilde{h}(y) = -L^0\Phi(y) + N(y), \tag{43}$$

in which $N : \mathbb{R}^5 \to \mathcal{X}_{\rm sh,c}^{s+1}$ is defined by the formula

$$\begin{aligned}N(y) = &-(L^\mu - L^0)\Phi(y) + \mathrm{d}\Phi[y](F(y,-\Phi(y))) + h(y)\\ &+ \begin{pmatrix} g_6(y,-\Phi(y)) \\ -P_{\rm sh,c}(d_4(y)\Phi_1(y)) - P_{\rm sh,c}(d_5(y)\partial_\eta\Phi_1(y)) - P_{\rm sh,c}(d_6(y)\Phi_2(y)) \end{pmatrix}\\ &+ \begin{pmatrix} 0 \\ g_4(y,-\Phi(y)) - P_{\rm sh,c}[g_3(y,-\Phi(y))\partial_\eta^2\Phi_1(y) + g_5(y,-\Phi(y))\partial_\eta\Phi_2(y)] \end{pmatrix}\end{aligned}$$

and L^μ is an abbreviation for $L^\varepsilon|_{\varepsilon=\mu^2}$.

Let us write

$$\Phi(y) = \sum_{k=2}^{p} \Phi^k(y),$$

where $\Phi^k(y)$ is a polynomial which is homogeneous of degree k in y and takes values in $\mathcal{X}^{s+2}_{\mathrm{sh,c}}$. We denote the space of such polynomials by $\mathcal{P}^k_{s+2}$ and equip it with the inner product and norm

$$\langle P, Q\rangle_{\mathcal{P}^k_{s+2}} = P(\partial_y).Q(y)|_{y=0}, \qquad |P|_2^{s+2} = \langle P, P\rangle^{1/2}_{\mathcal{P}^k_{s+2}},$$

in which the period denotes the $\mathcal{X}^{s+2}_{\mathrm{sh,c}}$ inner product, and for later use we also introduce the symbol $\mathcal{Q}_k$ for the space of polynomials $\mathbb{R}^5 \to \mathbb{R}^5$ that are homogenous of degree k in y. We proceed by choosing Φ^k so that

$$\mathcal{L}\Phi^k = N^k, \qquad k = 2, \ldots, p, \tag{44}$$

where

$$(\mathcal{L}\Phi^k)(y) = L^0\Phi^k(y) \tag{45}$$

and $N^k(y)$ is the part of $N(y)$ which is homogeneous of degree k in y; this choice of Φ ensures that the Taylor expansion of $\tilde{h}(y)$ does not contain any terms of order less than p (see equation (43)). Notice that $N^2(y) = h^2(y) = 0$, so that the term Φ^2 is not actually needed; certain combinatorial aspects of the following theory are however simplified by allowing this zero term to remain in the expansion for Φ^2. It is also important to note that Φ^k affects N^{k+1}, …, N^p, so that Φ^k must be chosen systematically for $k = 2, \ldots, p$.)

The following result shows that equation (44) admits a unique solution for any value for k and yields an estimate for Φ^k in terms of N^k which is independent of k. This estimate, which plays a crucial role in the following analysis, follows from the simple formula for the operator $\mathcal{L}$, which is in turn a consequence of the fact that the linearisation of the vector field on the right-hand of equation (36) is the zero matrix.

Proposition 4.1 *The operator $\mathcal{L} : \mathcal{P}^k_{s+2} \to \mathcal{P}^k_{s+1}$ defined by (45) is invertible and its operator norm*

$$|||\mathcal{L}^{-1}||| = \sup_{|\Phi^k|_2^{s+1}=1} |\mathcal{L}^{-1}\Phi^k|_2^{s+2}$$

is less than unity (and in particular is independent of k).

Proof. Let $\{e_{0,1}, e_{0,2}\}$, $\{e_{m,1,\mathrm{o}}, e_{m,2,\mathrm{o}}\}$ and $\{e_{m,1,\mathrm{e}}, e_{m,2,\mathrm{e}}\}$ be the usual bases for respectively E_0, $E_{m,\mathrm{o}}$, $m \geq 2$ and $E_{m,\mathrm{e}}$, $m \geq 2$ and consider the orthonormal basis

$$\mathcal{B}_k = \{P^\alpha e_{0,1}, P^\alpha e_{0,2}, P^\alpha e_{m,1,\mathrm{o}}, P^\alpha e_{m,2,\mathrm{o}}, P^\alpha e_{m,1,\mathrm{e}}, P^\alpha e_{m,2,\mathrm{e}} : |\alpha| = k,\ m \geq 2\}$$

for $\mathcal{P}^k_s$, where $\alpha \in \mathbb{N}_0^5$ is a multi-index and

$$P^\alpha = \frac{1}{\alpha_1!\alpha_2!\alpha_3!\alpha_4!\alpha_5!}\mu^{\alpha_1} z_{1,\mathrm{o}}^{\alpha_2} z_{1,\mathrm{e}}^{\alpha_3} z_{2,\mathrm{o}}^{\alpha_4} z_{2,\mathrm{e}}^{\alpha_5}.$$

Observe that

$$\mathcal{P}^k_s = \bigoplus_{|\alpha|=k} \mathcal{P}^{\alpha,k}_0 \bigoplus_{\substack{|\alpha|=k\\ m\geq 2}} \mathcal{P}^{\alpha,k}_{m,\mathrm{o}} \bigoplus_{\substack{|\alpha|=k\\ m\geq 2}} \mathcal{P}^{\alpha,k}_{m,\mathrm{e}}, \tag{46}$$

in which the subspaces $\mathcal{P}_0^{\alpha,k} = \operatorname{span}\{P^\alpha e_{0,1}, P^\alpha e_{0,2}\}$, $\mathcal{P}_{m,\mathrm{o}}^{\alpha,k} = \operatorname{span}\{P^\alpha e_{m,1,\mathrm{o}}, P^\alpha e_{m,2,\mathrm{o}}\}$ and $\mathcal{P}_{m,\mathrm{o}}^{\alpha,k} = \operatorname{span}\{P^\alpha e_{m,1,\mathrm{e}}, P^\alpha e_{m,1,\mathrm{e}}\}$, are invariant under $\mathcal{L}$; furthermore $\mathcal{L}_0^{\alpha,k} = \mathcal{L}|_{\mathcal{P}_{k,0}}^{\alpha}$ admits an inverse whose matrix with respect to the basis $\{P^\alpha e_{0,1}, P^\alpha e_{0,2}\}$ for $\mathcal{P}_0^{\alpha,k}$ is

$$(\mathcal{L}_0^{\alpha,k})^{-1} = \begin{pmatrix} 0 & 1/\lambda_{0,0}^2 \\ 1 & 0 \end{pmatrix}$$

and $\mathcal{L}_{m,\mathrm{o}}^{\alpha,k} = \mathcal{L}|_{\mathcal{P}_{m,\mathrm{o}}^{\alpha,k}}$, $\mathcal{L}_{m,\mathrm{e}}^{\alpha,k} = \mathcal{L}|_{\mathcal{P}_{m,\mathrm{e}}^{\alpha,k}}$ admit inverses whose matrices with respect to the bases $\{P^\alpha e_{m,1,\mathrm{o}}, P^\alpha e_{m,2,\mathrm{o}}\}$ for $\mathcal{P}_{m,\mathrm{o}}^{\alpha,k}$ and $\{P^\alpha e_{m,1,\mathrm{e}}, P^\alpha e_{m,2,\mathrm{e}}\}$ for $\mathcal{P}_{m,\mathrm{e}}^{\alpha,k}$ are

$$(\mathcal{L}_{m,\mathrm{o}}^{\alpha,k})^{-1} = \begin{pmatrix} 0 & -1/\omega_{m,0}^2 \\ 1 & 0 \end{pmatrix}, \qquad (\mathcal{L}_{m,\mathrm{e}}^{\alpha,k})^{-1} = \begin{pmatrix} 0 & -1/\omega_{m,0}^2 \\ 1 & 0 \end{pmatrix}.$$

Let us write

$$\Phi^k = \sum_{|\alpha|=k} \Phi_0^{\alpha,k} + \sum_{\substack{|\alpha|=k \\ m\geq 2}} \Phi_{m,\mathrm{o}}^{\alpha,k} + \sum_{\substack{|\alpha|=k \\ m\geq 2}} \Phi_{m,\mathrm{e}}^{\alpha,k}$$

in accordance with the orthogonal decomposition (46), so that

$$(|\Phi^k|_2^s)^2 = \sum_{|\alpha|=k} |\Phi_0^{\alpha,k}|^2 + \sum_{\substack{|\alpha|=k \\ m\geq 2}} m^{2(s+1)}|(\Phi_{m,\mathrm{o}}^{\alpha,k})_1|^2 + |(\Phi_{m,\mathrm{e}}^{\alpha,k})_1|^2) + \sum_{\substack{|\alpha|=k \\ m\geq 2}} m^{2s}|(\Phi_{m,\mathrm{o}}^{\alpha,k})_2|^2 + |(\Phi_{m,\mathrm{e}}^{\alpha,k})_2|^2),$$

where the symbols $(v)_1$ and $(v)_2$ denote the first and second components of a vector $v \in \mathcal{X}^s$. Since

$$\mathcal{L}^{-1}\Phi^k = \sum_{|\alpha|=k} (\mathcal{L}_0^{\alpha,k})^{-1}\Phi_0^{\alpha,k} + \sum_{\substack{|\alpha|=k \\ m\geq 2}} (\mathcal{L}_{m,\mathrm{o}}^{\alpha,k})^{-1}\Phi_{m,\mathrm{o}}^{\alpha,k} + \sum_{\substack{|\alpha|=k \\ m\geq 2}} (\mathcal{L}_{m,\mathrm{e}}^{\alpha,k})^{-1}\Phi_{m,\mathrm{e}}^{\alpha,k}$$

it follows that

$$\begin{aligned}
&(|\mathcal{L}^{-1}\Phi^k|_2^{s+2})^2 \\
&\quad = \sum_{|\alpha|=k} |(\mathcal{L}_0^{\alpha,k})^{-1}\Phi_0^{\alpha,k}|^2 \\
&\qquad\quad + \sum_{\substack{|\alpha|=k \\ m\geq 2}} m^{2(s+3)}\{|((\mathcal{L}_{m,\mathrm{o}}^{\alpha,k})^{-1}\Phi_{m,\mathrm{o}}^{\alpha,k})_1|^2 + |((\mathcal{L}_{m,\mathrm{e}}^{\alpha,k})^{-1}\Phi_{m,\mathrm{e}}^{\alpha,k})_1|^2\} \\
&\qquad\quad + \sum_{\substack{|\alpha|=k \\ m\geq 2}} m^{2(s+2)}\{|((\mathcal{L}_{m,\mathrm{o}}^{\alpha,k})^{-1}\Phi_{m,\mathrm{o}}^{\alpha,k})_2|^2 + |((\mathcal{L}_{m,\mathrm{e}}^{\alpha,k})^{-1}\Phi_{m,\mathrm{e}}^{\alpha,k})_2|^2\} \\
&\quad \leq \sum_{|\alpha|=k} |\Phi_0^{\alpha,k}|^2 + \sum_{\substack{|\alpha|=k \\ m\geq 2}} \frac{m^{2(s+3)}}{\omega_{m,0}^4}|(\Phi_{m,\mathrm{o}}^{\alpha,k})_2|^2 + |(\Phi_{m,\mathrm{e}}^{\alpha,k})_2|^2) \\
&\qquad\quad + \sum_{\substack{|\alpha|=k \\ m\geq 2}} m^{2(s+2)}|(\Phi_{m,\mathrm{o}}^{\alpha,k})_1|^2 + |(\Phi_{m,\mathrm{e}}^{\alpha,k})_1|^2)
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{|\alpha|=k} |\Phi_0^{\alpha,k}|^2 + \sum_{\substack{|\alpha|=k\\ m\geq 2}} m^{2(s+1)}|(\Phi_{m,\mathrm{o}}^{\alpha,k})_2|^2 + |(\Phi_{m,\mathrm{e}}^{\alpha,k})_2|^2) \\
&\qquad + \sum_{\substack{|\alpha|=k\\ m\geq 2}} m^{2(s+2)}|(\Phi_{m,\mathrm{o}}^{\alpha,k})_1|^2 + |(\Phi_{m,\mathrm{e}}^{\alpha,k})_1|^2) \\
&= (|\Phi^k|_2^{s+1})^2,
\end{aligned}
$$

where we have used the fact that $\omega_{m,0} \sim m$ as $m \to \infty$. $\square$

4.2 Estimates for the transformation

The next task is to estimate the size of Φ, and for this purpose we use the norms $|\Phi^m|_2^{s+2}$, $\|\Phi^m(y)\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+2}}$ and

$$
|\Phi^m|_0^{s+2} = \sup_{y\in\mathbb{R}^5} \frac{\|\Phi^m(y)\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+2}}}{|y|^m}
$$

for $m = 2, \ldots, p$. Our estimates for these quantities are used in Section 4.3 below to show that $\tilde{h}(y)$ is exponentially small with respect to y.

We begin by estimating $|N^m|_2^{s+1}$, from which an estimate for $|\Phi^m|_2^{s+2}$ is obtained using Proposition 4.1. A straightforward calculation shows that

$$
\begin{aligned}
&N_1^m(y) \\
&\quad = h_1^m[\{y\}^{(m)}] + \sum_{2\leq q\leq m} \sum_{i=0}^{q} \sum_{i+p_1+\ldots p_{q-i}=m} g_6^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}] \\
&\qquad + \sum_{2\leq k\leq m-1} \mathrm{d}\Phi_1^k[y]\Bigg(\sum_{2\leq q\leq m-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots p_{q-i}\\ =m-k+1}} F^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}]\Bigg), \qquad (47)
\end{aligned}
$$

$$
\begin{aligned}
&N_2^m(y) \\
&\quad = h_2^m[\{y\}^{(m)}] \\
&\qquad + \sum_{2\leq k\leq m-1} \mathrm{d}\Phi_2^k[y]\Bigg(\sum_{2\leq q\leq m-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots p_{q-i}\\ =m-k+1}} F^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}]\Bigg) \\
&\qquad - \sum_{1\leq q\leq m-1} \sum_{i=0}^{q} \sum_{\substack{p_1+\ldots p_{q-i}\\ +i+r=m\\ r>0}} P_{\mathrm{sh,c}}(g_3^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}]\partial_\eta^2 \Phi_1^r) \\
&\qquad + \sum_{2\leq q\leq m} \sum_{i=0}^{q} \sum_{i+p_1+\ldots p_{q-i}=m} g_4^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}] \\
&\qquad - \sum_{1\leq q\leq m-1} \sum_{i=0}^{q} \sum_{\substack{p_1+\ldots p_{q-i}\\ +i+r=m\\ r>0}} P_{\mathrm{sh,c}}(g_5^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}]\partial_\eta \Phi_2^r)
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{2\le k\le m-1} c_3^{m-k}\mu^{m-k}k_0^2\partial_\eta^2\Phi_1^k+\sum_{2\le k\le m-1} c_4^{m-k}\mu^{m-k}\Phi_1^k\\
&-P_{\mathrm{sh,c}}(d_4(y)\Phi_1^{m-2})-P_{\mathrm{sh,c}}(d_5(y)\partial_\eta\Phi_1^{m-2})-P_{\mathrm{sh,c}}(d_6(y)\Phi_2^{m-2}),
\end{aligned}
\tag{48}
$$

in which we use the notation

$$
\begin{aligned}
f^n[y_1,\dots,y_n] &= \frac{1}{n!}\mathrm{d}^n f[0](y_1,\dots,y_n),\\
f^{n_1,n_2}[y_1,\dots,y_{n_1},w_1,\dots,w_{n_2}] &= \frac{1}{(n_1+n_2)!}\mathrm{d}_{1,2}^{n_1,n_2} f[0](y_1,\dots,y_{n_1},w_1,\dots,w_{n_2})
\end{aligned}
$$

for derivatives of functions of one and two variables and write

$$
c_3^\mu=\sum_{i=0}^\infty c_3^i\mu^i,\qquad c_4^\mu=\sum_{i=0}^\infty c_4^i\mu^i.
$$

An estimate for $|N^m|_2^{s+1}$ can be obtained from formulae (47), (48) using the following lemma, whose proof is readily deduced from that of Lemmata 2.10 and 2.11 of Iooss & Lombardi [7].

Lemma 4.2 *Define*

$$
|P|_{2,m}^s=\frac{1}{\sqrt{m!}}|P|_2^s,\qquad P\in\mathcal{P}_s^m
$$

and

$$
|P|_{2,m}=\frac{1}{\sqrt{m!}}|P|_2,\qquad P\in\mathcal{Q}^m.
$$

(i) The estimates

$$
|\Phi^k|_0^s\ \le\ |\Phi^k|_{2,k}^s\ \le\ \sqrt{5}k^2|\Phi^k|_0^s
$$

hold for each $\Phi^k\in P_k^s$.

(ii) Suppose that $q\in\mathbb{N}$, $i\in\{0,\dots,q\}$, $\{p_\ell\}_{1\le\ell\le q-i}\subset\mathbb{N}$ and that R_q is a bounded, q-linear operator $(\mathbb{R}^5\times\mathcal{X}_{\mathrm{sh,c}}^s)^q\to\mathcal{X}_{\mathrm{sh,c}}^s$ with operator norm $\|\!|R_q|\!\|$. For each choice of $\Phi_{p_\ell}\in\mathcal{P}_s^\ell$, $\ell=1,\dots,q-i$ the polynomial $R_q[\{y\}^i,\Phi^{p_1},\dots,\Phi^{p_{q-i}}]$ lies in $\mathcal{P}_s^n$ with $n=p_1+\dots+p_{q-i}+i$ and satisfies the estimate

$$
|R_q[\{y\}^i,\Phi^{p_1},\dots,\Phi^{p_{q-i}}]|_{2,n}^s\le\|\!|R_q|\!\|\sqrt{5}^i|\Phi^{p_1}|_{2,p_1}^s\dots|\Phi^{p_{q-i}}|_{2,p_{q-i}}^s.
$$

The analogous result holds when $\mathcal{P}_s^m$ is replaced by $\mathcal{Q}^m$.

(iii) Suppose that $p\in\mathbb{N}$, $k\in\mathbb{N}_0$ and $\Phi^k\in\mathcal{P}_s^k$, $N^p\in\mathcal{Q}^p$. The polynomial $\mathrm{d}\Phi^k[y](N^p)$ lies in $\mathcal{P}_s^n$ with $n=k-p+1$ and satisfies the estimate

$$
|\mathrm{d}\Phi^k[y](N^p)|_{2,n}^s\le\sqrt{k^2+4k}|\Phi^k|_{2,k}^s|N^p|_{2,p}.
$$

Because $g_6 : \mathbb{R}^5 \times \mathcal{X}^{s+1}_{\mathrm{sh,c}} \to H^{s+2}_{\mathrm{per}}(\mathbb{R}^2)$, $j = 3, 4, 5$ is analytic, it satisfies the inequality

$$\|g_6^{n_1,n_2}[y_1, \ldots, y_{n_1}, q_1, \ldots, q_{n_2}]\|_{s+2} \leq \frac{a}{\rho^{n_1+n_2}} |y_1| \ldots |y_{n_1}| \|q_1\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} \ldots \|q_{n_2}\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}$$

for each $n_1, n_2 \in \mathbb{N}_0$, where $a > 1$ and $\rho < 1$ are universal constants, and the analogous estimates apply to $F : \mathbb{R}^5 \times \mathcal{X}^{s+1}_{\mathrm{sh,c}} \to \mathbb{R}^5$ and $h : \mathbb{R}^5 \to \mathcal{X}^{s+1}_{\mathrm{sh,c}}$; for notational simplicity later we estimate

$$\begin{aligned}
&\|g_4^{n_1,n_2}[y_1, \ldots, y_{n_1}, q_1, \ldots, q_{n_2}]\|_{s+1} \\
&\qquad \leq \frac{a}{3\rho^{n_1+n_2}} |y_1| \ldots |y_{n_1}| \|q_1\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} \ldots \|q_{n_2}\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}, \qquad j = 3, 4, 5
\end{aligned}$$

and

$$c_j^n \leq \frac{a}{2\rho^n}, \quad j = 3, 4, \qquad \|d_j(y)\|_{s+1} \leq \frac{a}{3\rho^2} |y|^2, \quad j = 4, 5, 6.$$

Using these estimates together with Proposition 4.1 and Lemma 4.2, we find from equations (47), (48) that

$$\begin{aligned}
\phi_m \;\leq\; & \frac{a}{\rho^m} \sqrt{5}^m + \sum_{2 \leq k \leq m-1} (k^2 + 4k)^{1/2} \phi_k \sum_{2 \leq q \leq m-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots+p_{q-i} \\ = m-k+1}} \frac{a}{\rho^q} \sqrt{5}^i \phi_{p_1} \ldots \phi_{p_{q-i}} \\
& + \frac{2}{3} \sum_{2 \leq q \leq m-1} \sum_{i=0}^{q} \sum_{\substack{p_1+\ldots p_{q-i} \\ +i+r=m \\ r>0}} \frac{a}{\rho^q} \sqrt{5}^i \phi_{p_1} \ldots \phi_{p_{q-i}} \phi_r \\
& + \frac{1}{3} \sum_{2 \leq q \leq m} \sum_{i=0}^{q} \sum_{i+p_1+\ldots p_{q-i}=m} \frac{a}{\rho^q} \sqrt{5}^i \phi_{p_1} \ldots \phi_{p_{q-i}} \\
& + \frac{a}{\rho^2} \sqrt{5}^2 \phi^{m-2} + \sum_{2 \leq k \leq m-1} \frac{a}{\rho^{m-k}} \sqrt{5}^{m-k} \phi^k \\
\leq\; & \frac{a}{\rho^m} \sqrt{5}^m + \sum_{2 \leq k \leq m-1} (k^2 + 4k)^{1/2} \phi_k \sum_{2 \leq q \leq m-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots+p_{q-i} \\ = m-k+1}} \frac{a}{\rho^q} \sqrt{5}^i \phi_{p_1} \ldots \phi_{p_{q-i}} \\
& + 3 \sum_{2 \leq q \leq m} \sum_{i=0}^{q} \sum_{i+p_1+\ldots p_{q-i}=m} \frac{a}{\rho^q} \sqrt{5}^i \phi_{p_1} \ldots \phi_{p_{q-i}},
\end{aligned} \tag{49}$$

where

$$\phi_m = |\Phi^m|^{s+2}_{2,m}.$$

The above inequalities are converted into an estimate for ϕ_m in Propositions 4.3 and 4.4 below, the first of which is proved by straightforward mathematical induction; we note that

$$\beta_2 = 4, \qquad \phi_2 \leq 4\sqrt{5} \left(\frac{\sqrt{5}a}{\rho^2} \right) = \sqrt{5} \left(\frac{\sqrt{5}a}{\rho^2} \right) \beta_2,$$

so that the result holds for $m = 2$, and proceed inductively using (49).

Proposition 4.3 *Consider the sequence $\{\beta_n\}$ defined recursively by the formulae*

$$
\begin{aligned}
\beta_1 &= 1, \\
\beta_m &= \left(\frac{\rho}{a}\right)^{m-2} + 3 \sum_{2\le q\le m} \sum_{i=0}^{q} \sum_{i+p_1+\ldots p_{q-i}=m} \left(\frac{\rho}{a}\right)^{q-2} \beta_{p_1}\ldots\beta_{p_{q-i}} \\
&\quad + 5 \sum_{2\le k\le m-1} k\beta_k \sum_{2\le q\le m-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots+p_{q-i}\\ =m-k+1}} \left(\frac{\rho}{a}\right)^{q-2} \beta_{p_1}\ldots\beta_{p_{q-i}}, \qquad m\ge 2.
\end{aligned}
$$

The quantity ϕ_m satisfies the estimate

$$
\phi_m \le \sqrt{5}\left(\frac{\sqrt{5}a}{\rho^2}\right)^{m-1} \beta_m, \qquad m\ge 2.
$$

Proposition 4.4 *Consider the sequence $\{\alpha_n\}$ defined by the formulae*

$$
\begin{aligned}
\alpha_1 &= 1, \\
\alpha_m &= \Theta^{m-2}(m-2)!, \qquad m\ge 2,
\end{aligned}
$$

where $\Theta \ge 1$ is a constant. The estimate

$$
\beta_m \le 2^m \alpha_m, \qquad m\ge 1 \tag{50}
$$

holds for $\Theta \ge 25 + 13\rho/a$.

Proof. This result is also established using mathematical induction. We note that

$$
\beta_1 = 1 < 2 = 2\alpha_1, \qquad \beta_2 = 4 = 2^2\alpha_2,
$$

so that (50) holds for $m = 1$ and $m = 2$, and proceed inductively by choosing $m \ge 3$ and supposing that $\beta_k \le 2^k\alpha_k$ for $1 \le k < m$.

Observe that

$$
\sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots+p_{q-i}=m\\ p_j\ge 2}} \beta_{p_1}\ldots\beta_{p_{q-i}} = \sum_{\substack{p_1+\ldots+p_q=m\\ p_j\ge 1}} \beta_{p_1}\ldots\beta_{p_q}
$$

because $\beta_1 = 1$; the recursion relation for β_m can therefore be rewritten as

$$
\begin{aligned}
\beta_m &= \left(\frac{\rho}{a}\right)^{m-2} + 3 \sum_{2\le q\le m} \sum_{i=0}^{q} \sum_{\substack{p_1+\ldots+p_q=m\\ p_j\ge 1}} \left(\frac{\rho}{a}\right)^{q-2} \beta_{p_1}\ldots\beta_{p_q} \\
&\quad + 5 \sum_{2\le k\le m-1} k\beta_k \sum_{2\le q\le m-k+1} \sum_{i=0}^{q} \sum_{\substack{p_1+\ldots+p_q\\ =m-k+1\\ p_j\ge 1}} \left(\frac{\rho}{a}\right)^{q-2} \beta_{p_1}\ldots\beta_{p_q}, \qquad m\ge 2,
\end{aligned}
$$

and it follows that

$$\beta_m \leq 2^m[\Delta_m^1 + \Delta_m^2 + \Delta_m^3 + \Delta_m^4] + \left(\frac{\rho}{a}\right)^{m-2},$$

where

$$\begin{aligned}
\Delta_m^1 &= 3 \sum_{3\leq q\leq m} \sum_{\substack{p_1+\ldots p_q=m \\ p_j\geq 1}} \left(\frac{\rho}{a}\right)^{q-2} \alpha_{p_1}\ldots\alpha_{p_q}, \\
\Delta_m^2 &= 5 \sum_{2\leq k\leq m-1} k\alpha_k \sum_{3\leq q\leq m-k+1} \sum_{\substack{p_1+\ldots+p_q \\ =m-k+1 \\ p_j\geq 1}} \left(\frac{\rho}{a}\right)^{q-2} \alpha_{p_1}\ldots\alpha_{p_q}, \\
\Delta_m^3 &= 3 \sum_{1\leq k\leq m-1} \alpha_k\alpha_{m-k}, \\
\Delta_m^4 &= 5 \sum_{2\leq k\leq m-1} k\alpha_k \left(\sum_{1\leq j\leq m-k} \alpha_j\alpha_{m-k+1-j}\right).
\end{aligned}$$

To estimate the quantities $\Delta_m^1, \ldots, \Delta_m^4$ we use the inequalities

$$\sum_{\substack{p_1+\ldots p_q=m \\ p_j\geq 1}} \alpha_{p_1}\ldots\alpha_{p_q} \leq \frac{2}{\Theta^{q-2}}\alpha_m, \qquad 3\leq q\leq m, \tag{51}$$

$$\sum_{2\leq k\leq m-1} k\alpha_k\alpha_{m-k+1} \leq \frac{5}{2\Theta}\alpha_m, \qquad m\geq 3, \tag{52}$$

$$\sum_{1\leq k\leq m-1} \alpha_k\alpha_{m-k} \leq \frac{2}{\Theta}\alpha_m, \qquad m\geq 3, \tag{53}$$

which were established by Iooss & Lombardi [7, Lemma 2.13]. We find from inequality (51) that

$$\Delta_m^1 \leq \sum_{3\leq q\leq m} 6\left(\frac{\rho}{a\Theta}\right)^{q-2}\alpha_m \leq 6\left(\frac{\frac{\rho}{a\Theta}}{1-\frac{\rho}{a\Theta}}\right)\alpha_m \leq \frac{1}{2}\alpha_m$$

whenever $\rho/(a\Theta) \leq 1/13$. Inequalities (51), (52) similarly yield

$$\Delta_m^2 \leq \frac{5}{2\Theta}\sum_{2\leq k\leq m-1} k\alpha_k\alpha_{m-k+1} \leq \frac{25}{4\Theta^2}\alpha_m,$$

inequality (53) shows that

$$\Delta_m^3 \leq \frac{6}{\Theta}\alpha_m$$

and it follows from (52), (53) that

$$\Delta_m^4 \leq \frac{10}{\Theta}\sum_{2\leq k\leq m-1} k\alpha_k\alpha_{m-k+1} \leq \frac{25}{\Theta^2}\alpha_m.$$

Finally, note that

$$\left(\frac{\rho}{a}\right)^{m-2} = \left(\frac{\rho}{a\Theta}\right)^{m-2}\Theta^{m-2} \;\le\; \frac{1}{4}2^m\alpha_m,$$

and choosing $\Theta \ge 25$, one concludes that

$$\beta_m \le \left(\frac{3}{4}+\frac{6}{\Theta}+\frac{125}{4\Theta^2}\right)2^m\alpha_m \le 2^m\alpha_m. \qquad \square$$

In keeping with Proposition 4.4 we choose

$$\Theta = 25 + \frac{13\rho}{a} \tag{54}$$

and fix this value of Θ for the remainder of the article. The proposition implies that

$$\phi_m \le \frac{20a}{\rho^2}\left(\frac{2\sqrt{5}a\Theta}{\rho^2}\right)^{m-2}(m-2)!, \qquad m = 2,3,\ldots, \tag{55}$$

and by imposing a mutual constraint upon the order p of the normal form and the maximum size δ of $|y|$, we can use this fact to obtain another estimate for Φ.

Proposition 4.5 *Suppose that $\delta > 0$ and $p \ge 2$ satisfy*

$$\delta p \le \frac{\rho^2}{4\sqrt{5}a\Theta}. \tag{56}$$

The estimates

$$\left\|\sum_{2\le k\le p}\Phi^k(y)\right\|_{\mathcal{X}^{s+2}_{\mathrm{sh,c}}} \le \frac{\sqrt{5}\delta}{\Theta}, \qquad \left|\!\left|\!\left|\sum_{2\le k\le p}\mathrm{d}\Phi^k[y]\right|\!\right|\!\right| \le \frac{10}{\Theta}$$

hold for every $y \in \mathbb{R}^5$ such that $|y| \le \delta$. Here $|||\sum_{2\le k\le p}\mathrm{d}\Phi^k[y]|||$ denotes the operator norm of $\sum_{2\le k\le p}\mathrm{d}\Phi^k[y] : \mathbb{R}^5 \to \mathcal{X}^{s+2}_{\mathrm{sh,c}}$.

Proof. Observe that

$$\begin{aligned}
\left\|\sum_{2\le k\le p}\Phi^k(y)\right\|_{\mathcal{X}^{s+2}_{\mathrm{sh,c}}} &\le \sum_{2\le k\le p}|\Phi^k|^{s+2}_{0,k}|y|^k \\
&\le \sum_{2\le k\le p}|\Phi^k|^{s+2}_{2,k}|y|^k \\
&\le \sum_{2\le k\le p}\phi_k\delta^k \\
&\le \sum_{2\le k\le p}\frac{20a\delta^2}{\rho^2}\left(\frac{2\sqrt{5}a\Theta\delta}{\rho^2}\right)^{k-2}(k-2)! \\
&\le \frac{2\sqrt{5}\delta}{\Theta}\sum_{2\le k\le p}\left(\frac{1}{2p}\right)^{k-1}(k-2)! \\
&\le \frac{2\sqrt{5}\delta}{\Theta p}\sum_{2\le k\le p}\left(\frac{1}{2}\right)^{k-1} \\
&\le \frac{\sqrt{5}\delta}{\Theta},
\end{aligned}$$

where the inequalities $|\Phi^k|_0^{s+2} \leq |\Phi^k|_{2,k}^{s+2}$ (see Lemma 4.2(i)) and $(k-2)!/p^{k-1} \leq 1/p$ have been used. Furthermore, it follows from the estimate

$$\begin{aligned}\frac{\|\mathrm{d}\Phi^k[y](\tilde{y})\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+2}}}{|y|^{k-1}} &\leq |\mathrm{d}\Phi^k[y](\tilde{y})|_{0,k}^{s+2} \\ &\leq |\mathrm{d}\Phi^k[y](\tilde{y})|_{2,k}^{s+2} \\ &\leq \sqrt{k^2+4k}|\Phi^k|_{2,k}^{s+2}|\tilde{y}|_{2,0} \\ &\leq \sqrt{5}k\phi_k|\tilde{y}| \end{aligned} \tag{57}$$

that

$$\Bigg\|\Bigg| \sum_{2\leq k\leq p} \mathrm{d}\Phi^k[y] \Bigg|\Bigg\| \leq \sum_{2\leq k\leq p} \sqrt{5}k\phi_k\delta^{k-1},$$

and one obtains the estimate for $\||\sum_{2\leq k\leq p} \mathrm{d}\Phi^k[y]|\|$ from this inequality by applying the arguments used in the previous calculation. $\square$

4.3 Estimates for the transformed nonlinearities

In this section we use the above estimates for $\Phi = \sum_{k=2}^{p} \Phi^k$ with an optimal choice of p to derive estimates for the terms $\tilde{F}$, $\tilde{g}_j$, $j = 3, \ldots, 6$ and $\tilde{h}$ appearing in equations (37), (38); in particular we show that $\|\tilde{h}(y)\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}$ is exponentially small with respect to y.

It follows from that fact that $\tilde{h}^m$ is identically zero for $m \leq p$ that

$$\begin{aligned}
&\tilde{h}_1(y) \\
&= \sum_{p+1\leq q} h_1^q[\{y\}^{(q)}] \\
&\quad + \sum_{2\leq q\leq p} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots p_{q-i}\\ \geq p+1}} g_6^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}] \\
&\quad + \sum_{p+1\leq q} \sum_{i=0}^{q} g_6^{i,q-i}\left[\{y\}^{(i)}, -\Bigg\{\sum_{2\leq k\leq p} \Phi^k\Bigg\}^{(q-i)}\right] \\
&\quad + \sum_{\substack{2\leq k\leq p\\ p+1\leq n\leq p+k-1}} \mathrm{d}\Phi_1^k[y]\Bigg(\sum_{2\leq q\leq n-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots p_{q-i}\\ =n-k+1}} F^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}]\Bigg) \\
&\quad + \Bigg(\sum_{k=2}^{p} \mathrm{d}\Phi_1^k[y]\Bigg)\Bigg(\sum_{2\leq q\leq p} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots p_{q-i}\\ \geq p+1}} F^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}] \\
&\qquad\qquad + \sum_{p+1\leq q} \sum_{i=0}^{q} F^{i,q-i}\left[\{y\}^{(i)}, -\Bigg\{\sum_{2\leq k\leq p} \Phi^k\Bigg\}^{(q-i)}\right]\Bigg),
\end{aligned} \tag{58}$$

$$
\begin{aligned}
&\tilde{h}_2(y)\\
&= \sum_{p+1\le q} h_2^q[\{y\}^{(q)}] + \sum_{p\le q} c_4^q \mu^q \sum_{2\le k\le p} \Phi_1^k + \sum_{\substack{1\le q\le p-1\\ r\ge p+1-q}} c_4^q\mu^q\Phi_1^r\\
&\quad + \sum_{p\le q} c_3^q\mu^q k_0^2\partial_\eta^2\Big(\sum_{2\le k\le p}\Phi_1^k\Big) + \sum_{\substack{1\le q\le p-1\\ r\ge p+1-q}} c_3^q\mu^q k_0^2\partial_\eta^2\Phi_1^r\\
&\quad - \sum_{1\le q\le p-1}\sum_{i=0}^{q}\sum_{\substack{i+p_1+\ldots+p_{q-i}\\ +r\ge p+1\\ r>0}} g_3^{i,q-i}[\{y\}^{(i)},-\Phi^{p_1},\ldots,-\Phi^{p_{q-i}}]\partial_\eta^2\Phi_1^r\\
&\quad - \sum_{1\le q\le p-1}\sum_{i=0}^{q}\sum_{\substack{i+p_1+\ldots+p_{q-i}\\ +r\ge p+1\\ r>0}} g_5^{i,q-i}[\{y\}^{(i)},-\Phi^{p_1},\ldots,-\Phi^{p_{q-i}}]\partial_\eta\Phi_2^r\\
&\quad + \sum_{2\le q\le p}\sum_{i=0}^{q}\sum_{\substack{i+p_1+\ldots p_{q-i}\\ \ge p+1}} g_4^{i,q-i}[\{y\}^{(i)},-\Phi^{p_1},\ldots,-\Phi^{p_{q-i}}]\\
&\quad - \sum_{p\le q}\sum_{i=0}^{q} g_3^{i,q-i}\Bigg[\{y\}^{(i)},-\Big\{\sum_{2\le k\le p}\Phi^k\Big\}^{(q-i)}\Bigg]\partial_\eta^2\Big(\sum_{2\le k\le p}\Phi_1^k\Big)\\
&\quad - \sum_{p\le q}\sum_{i=0}^{q} g_5^{i,q-i}\Bigg[\{y\}^{(i)},-\Big\{\sum_{2\le k\le p}\Phi^k\Big\}^{(q-i)}\Bigg]\partial_\eta\Big(\sum_{2\le k\le p}\Phi_2^k\Big)\\
&\quad + \sum_{p+1\le q}\sum_{i=0}^{q} g_4^{i,q-i}\Bigg[\{y\}^{(i)},-\Big\{\sum_{2\le k\le p}\Phi^k\Big\}^{(q-i)}\Bigg]\\
&\quad + \sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}} \mathrm{d}\Phi_2^k[y]\Bigg(\sum_{2\le q\le n-k+1}\sum_{i=0}^{q}\sum_{\substack{i+p_1+\ldots p_{q-i}\\ =n-k+1}} F^{i,q-i}[\{y\}^{(i)},-\Phi^{p_1},\ldots,-\Phi^{p_{q-i}}]\Bigg)\\
&\quad + \Big(\sum_{k=2}^{p}\mathrm{d}\Phi_2^k[y]\Big)\Bigg(\sum_{2\le q\le p}\sum_{i=0}^{q}\sum_{\substack{i+p_1+\ldots p_{q-i}\\ \ge p+1}} F^{i,q-i}[\{y\}^{(i)},-\Phi^{p_1},\ldots,-\Phi^{p_{q-i}}]\\
&\qquad\qquad + \sum_{p+1\le q}\sum_{i=0}^{q} F^{i,q-i}\Bigg[\{y\}^{(i)},-\Big\{\sum_{2\le k\le p}\Phi^k\Big\}^{(q-i)}\Bigg]\Bigg)\\
&\quad + \sum_{p-1\le q\le p} P_{\mathrm{sh,c}}(d_4(y)\Phi_1^q) + \sum_{p-1\le q\le p} P_{\mathrm{sh,c}}(d_5(y)\partial_\eta\Phi_1^q) + \sum_{p-1\le q\le p} P_{\mathrm{sh,c}}(d_6(y)\Phi_2^q). \qquad (59)
\end{aligned}
$$

Suppose that $|y|\le\delta$ and p, δ satisfy the constraint (56). Using Proposition 4.5 and the rule

$$
\|\Phi^k(y)\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+2}} \;\le\; |\Phi^k|_0^{s+2}|y|^k \;\le\; \phi_k|y|^k,
$$

one finds from the above formulae that

$$
\begin{aligned}
\|\tilde{h}_1(y)\|_{s+2} &\le \sum_{p+1\le q} \frac{a}{\rho^q}(\sqrt{5}\delta)^q + \left(1+\frac{10}{\Theta}\right) \sum_{2\le q\le p} \sum_{i=0}^{q} \sum_{\substack{n=i+p_1+\ldots+p_{q-i}\\ \ge p+1}} \frac{a\delta^n}{\rho^q}\sqrt{5}^i \phi_{p_1}\ldots\phi_{p_{q-i}} \\
&\quad + \left(1+\frac{10}{\Theta}\right) \sum_{p+1\le q} \sum_{i=0}^{q} \frac{a}{\rho^q}(\sqrt{5}\delta)^i \left(\frac{\sqrt{5}\delta}{\Theta}\right)^{q-i} \\
&\quad + \sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}} \sqrt{5}k\phi_k \sum_{2\le q\le n-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots+p_{q-i}\\ =n-k+1}} \frac{a\delta^n}{\rho^q}\sqrt{5}^i \phi_{p_1}\ldots\phi_{p_{q-i}} \\
&\le \left(2+\frac{10}{\Theta}\right)\Delta_p^1 + \left(1+\frac{10}{\Theta}\right)\Delta_p^2 + \Delta_p^3, \qquad (60)
\end{aligned}
$$

$$
\begin{aligned}
\|\tilde{h}_2(y)\|_{s+1} &\le \sum_{p+1\le q} \frac{a}{\rho^q}(\sqrt{5}\delta)^q + \sum_{p\le q} \frac{a}{\rho^q}(\sqrt{5}\delta)^q \left(\frac{\sqrt{5}\delta}{\Theta}\right) + \sum_{\substack{1\le q\le p-1\\ r\ge p+1-q}} \frac{a}{\rho^q}\sqrt{5}^q \phi_r \delta^{q+r} \\
&\quad + \left(\frac{1}{3}+\frac{10}{\Theta}\right) \sum_{2\le q\le p} \sum_{i=0}^{q} \sum_{\substack{n=i+p_1+\ldots+p_{q-i}\\ \ge p+1}} \frac{a\delta^n}{\rho^q}\sqrt{5}^i \phi_{p_1}\ldots\phi_{p_{q-i}} \\
&\quad + \frac{2}{3} \sum_{1\le q\le p-1} \sum_{i=0}^{q} \sum_{\substack{n=i+p_1+\ldots+p_{q-i}\\ +r\ge p+1\\ r>0}} \frac{a\delta^n}{\rho^q}\sqrt{5}^i \phi_{p_1}\ldots\phi_{p_{q-i}}\phi_r \\
&\quad + \left(\frac{1}{3}+\frac{10}{\Theta}\right) \sum_{p+1\le q} \sum_{i=0}^{q} \frac{a}{\rho^q}(\sqrt{5}\delta)^i \left(\frac{\sqrt{5}\delta}{\Theta}\right)^{q-i} \\
&\quad + \frac{2}{3} \sum_{p\le q} \sum_{i=0}^{q} \frac{a}{\rho^q}(\sqrt{5}\delta)^i \left(\frac{\sqrt{5}\delta}{\Theta}\right)^{q-i} \left(\frac{\sqrt{5}\delta}{\Theta}\right) \\
&\quad + \sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}} \sqrt{5}k\phi_k \sum_{2\le q\le n-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots+p_{q-i}\\ =n-k+1}} \frac{a\delta^n}{\rho^q}\sqrt{5}^i \phi_{p_1}\ldots\phi_{p_{q-i}} \\
&\quad + \sum_{p-1\le q\le p} \frac{a}{\rho^2}\sqrt{5}^2 \phi_1^q \\
&\le \left(3+\frac{10}{\Theta}\right)\Delta_p^1 + \left(3+\frac{10}{\Theta}\right)\Delta_p^2 + \Delta_p^3, \qquad (61)
\end{aligned}
$$

where

$$
\Delta_1^p = \sum_{p+1\le q} \frac{a}{\rho^q}(q+1)\sqrt{5}^q \delta^q
$$

$$\Delta_2^p = \sum_{2\le q\le p} \sum_{\substack{n=p_1+\ldots+p_q\\ \ge p+1\\ p_j\ge 1}} \frac{a\delta^n}{\rho^q}\phi_{p_1}\ldots\phi_{p_q}$$

$$\Delta_3^p = \sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}} \sqrt{5}k\phi_k \sum_{2\le q\le n-k+1} \sum_{\substack{i+p_1+\ldots+p_q\\ =n-k+1\\ p_j\ge 1}} \frac{a\delta^n}{\rho^q}\sqrt{5}^i\phi_{p_1}\ldots\phi_{p_q}$$

and we have defined $\phi_1 = \sqrt{5}$.

Proposition 4.6 *Suppose that $\delta > 0$ and $p \ge 2$ satisfy the stronger mutual constraint*

$$\delta p \le \frac{\rho^2}{4\sqrt{5}\mathrm{e}a\Theta}.$$

The estimate

$$\|\tilde{h}(y)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} \le c\left((C\delta)^{p+1}p! + \frac{1}{\mathrm{e}^{p+1}p^2}\right), \qquad C = \frac{4\sqrt{5}a\Theta}{\rho^2}$$

holds for every $y \in \mathbb{R}^5$ such that $|y| \le \delta$.

Proof. We proceed by estimating the quantities Δ_1^p, Δ_2^p, Δ_3^p, making use of the inequalities

$$\sum_{2\le q\le p} r^q \sum_{n\ge p+1} \sum_{\substack{p_1+\ldots+p_q\le n\\ 1\le p_j\le p}} (C\delta)^n(p_1-2)!\ldots(p_q-2)! \;\le\; \frac{4ar^2}{\mathrm{e}^{p+1}}.\frac{1}{p^2}.\frac{1}{1-r}, \tag{62}$$

where $r = \rho/(a\Theta)$, and

$$\sum_{2\le k\le p} k(k-2)!(C\delta)^{p+1} \sum_{p+1\le n\le p+k-1} (C\delta)^{n-p-1}(n-k-1)! \;\le\; 2(C\delta)^{p+1}p! \tag{63}$$

which were established by Iooss & Lombardi [7, pp. 30–32].

Observe that $q+1 \le 2^q$ and

$$\frac{2\sqrt{5}\delta}{\rho} \;\le\; \frac{r}{2p} \;\le\; \frac{r}{4}$$

for $p \ge 2$, so that

$$\Delta_p^1 \;\le\; a\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1} \sum_{q\ge 0}\left(\frac{r}{4}\right)^q \;=\; \frac{4a}{4-r}\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1} \le\; \frac{8a}{7}\left(\frac{\sqrt{5}\delta}{\rho}\right)^{p+1}$$

for $r \le 1/2$ (in accordance with (54)). The quantity Δ_p^2 is estimated using the calculation

$$\Delta_p^2 \;\le\; \sum_{2\le q\le p}\sum_{n\ge p+1}\sum_{\substack{p_1+\ldots+p_q\le n\\ 1\le p_j\le p}} \frac{a\delta^n}{\rho^q}\left(\frac{20a}{\rho^2}\right)^q\left(\frac{2\sqrt{5}a}{\rho^2}\right)^{n-2q} \alpha_{p_1}\ldots\alpha_{p_q}$$

$$
\begin{aligned}
&\le \sum_{2\le q\le p}\sum_{n\ge p+1}\frac{a\delta^n}{\rho^q}(2\sqrt{5})^q\left(\frac{2\sqrt{5}a\Theta}{\rho^2}\right)^{n-q}(p_1-2)!\dots(p_q-2)!\\
&\le a\sum_{2\le q\le p} r^q\sum_{n\ge p+1}\sum_{\substack{p_1+\ldots+p_q\le n\\ 1\le p_j\le p}}(C\delta)^n(p_1-2)!\dots(p_q-2)!\\
&\le \frac{4ar^2}{\mathrm{e}^{p+1}}.\frac{1}{p^2}.\frac{1}{1-r}\\
&\le \frac{2a}{p^2\mathrm{e}^{p+1}}
\end{aligned}
$$

for $r\le 1/2$, where the fourth line follows from the third by (62). Finally, we find from (51) that

$$
\begin{aligned}
&\sum_{2\le q\le n-k+1}\ \sum_{\substack{i+p_1+\ldots+p_q\\ =n-k+1\\ p_j\ge 1}}\left(\frac{\rho}{a}\right)^{q-2}\alpha_1\dots\alpha_{p_q}\\
&\qquad\le \sum_{2\le q\le n-k+1}2\left(\frac{\rho}{a\Theta}\right)^{q-2}\alpha_{n-k+1}\ \le\ \frac{2}{1-r}\alpha_{n-k+1}\ \le\ 4\alpha_{n-k+1}
\end{aligned}
$$

for $r\le 1/2$, and combining this estimate with (63) yields

$$
\begin{aligned}
\Delta_3^p\ &\le \sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}}\sqrt{5}k.\frac{20a}{\rho^2}\left(\frac{2\sqrt{5}a}{\rho^2}\right)^{k-2}\alpha_k\\
&\qquad\times\sum_{2\le q\le n-k+1}\ \sum_{\substack{i+p_1+\ldots+p_q\\ =n-k+1\\ p_j\ge 1}}\frac{a\delta^n}{\rho^q}\left(\frac{20a}{\rho^2}\right)^q\left(\frac{2\sqrt{5}a}{\rho^2}\right)^{n-k+1-2q}\alpha_1\dots\alpha_{p_q}\\
&= \sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}}20\sqrt{5}\delta^n\left(\frac{2\sqrt{5}a}{\rho^2}\right)^{k-2}k\alpha_k\sum_{2\le q\le n-k+1}\ \sum_{\substack{i+p_1+\ldots+p_q\\ =n-k+1\\ p_j\ge q}}\left(\frac{\rho}{a}\right)^{q-2}\alpha_1\dots\alpha_{p_q}\\
&\le 80\sqrt{5}\sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}}\delta^n\left(\frac{2\sqrt{5}a}{\rho^2}\right)^{n-1}k\alpha_k\alpha_{n-k+1}\\
&\le \frac{40\rho^2}{a\Theta^3}\sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}}(C\delta)^n k(k-2)!(n-k-1)!\\
&= \frac{40}{a\Theta^3}\sum_{2\le k\le p}k(k-2)!(C\delta)^{p+1}\sum_{p+1\le n\le p+k-1}(C\delta)^{n-p-1}(n-k-1)!\\
&\le \frac{80\rho^2}{a\Theta^3}(C\delta)^{p+1}p!.
\end{aligned}
$$

The result follows from inequalities (60), (61) and the above estimates for Δ_p^1, Δ_p^2, Δ_p^3. $\square$

Remark 4.7 *Inspecting the proof of the above proposition, one finds that it can proved under the weaker hypothesis*

$$\delta p \leq \frac{\rho^2}{2\sqrt{5}\mathrm{e}a\Theta}$$

and with C replaced by $2\sqrt{5}a\Theta/\rho^2$. The stronger constraint and larger value of C are however required later (see Proposition 4.12) and we introduce them from the outset for notational simplicity.

The final step in the derivation of our estimate for $\|\tilde{h}\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}$ is given by the following proposition, which was proved by Iooss & Lombardi [7, Lemma 2.18].

Proposition 4.8 *Choose $\vartheta > 0$. The function $f_\vartheta : \mathbb{N} \to \mathbb{R}$ defined by*

$$f_\vartheta(p) = \vartheta^{p+1} p!$$

satisfies

$$f_\vartheta\left(\left[\frac{1}{\vartheta \mathrm{e}}\right]\right) = m\sqrt{\frac{\vartheta}{\mathrm{e}}}\mathrm{e}^{-2/\vartheta \mathrm{e}}, \qquad m = \sup_{p \in \mathbb{N}} \frac{\mathrm{e}^2 p!}{p^{p+1/2}\mathrm{e}^{-p}},$$

where $[\cdot]$ denotes the integer part of a real number.

Define

$$p_{\mathrm{opt}} = \left[\frac{1}{\mathrm{e}C\delta}\right],$$

and note that p_{opt} satisfies

$$2 \leq p_{\mathrm{opt}} \leq \frac{1}{\mathrm{e}C\delta};$$

applying Proposition 4.6, we therefore find that

$$\begin{aligned}
\|\tilde{h}(y)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} &\leq c\left((C\delta)^{p_{\mathrm{opt}}+1} p_{\mathrm{opt}}! + \frac{1}{\mathrm{e}^{p_{\mathrm{opt}}+1} p_{\mathrm{opt}}^2}\right) \\
&\leq c\left(m\sqrt{\frac{C\delta}{\mathrm{e}}}\mathrm{e}^{-2/\mathrm{e}C\delta} + (2\mathrm{e}C\delta)^2 \mathrm{e}^{-1/\mathrm{e}C\delta}\right) \\
&\leq cC^2\left(m\sqrt{\frac{27}{8\mathrm{e}}} + 4\mathrm{e}^2\right)\delta^2 \mathrm{e}^{-1/\mathrm{e}C\delta}, \qquad (64)
\end{aligned}$$

in which the second line follows from the first by Proposition 4.8 with $\vartheta = C\delta$ and the inequalities

$$\frac{1}{\mathrm{e}C\delta} \leq p_{\mathrm{opt}} + 1, \qquad \frac{1}{p_{\mathrm{opt}}} \leq 2\mathrm{e}C\delta.$$

Let us now return to the original notation by writing equations (36)–(38) as

$$\partial_\xi z = \tilde{F}^\mu(z, \tilde{q}), \qquad (65)$$

$$\partial_\xi \tilde{q}_1 = \tilde{q}_2 + \tilde{g}_6^\mu(z, \tilde{q}) + \tilde{h}_1^\mu(z), \qquad (66)$$

$$\begin{aligned}
\partial_\xi \tilde{q}_2 = &-c_3^\mu k_0^2 \partial_\eta^2 \tilde{q}_1 - c_4^\mu \tilde{q}_1 + \mu P_{\mathrm{sh,c}}(d_1(z)\tilde{q}_1) + \mu P_{\mathrm{sh,c}}(d_2(z)\partial_\eta \tilde{q}_1) + \mu P_{\mathrm{sh,c}}(d_3(z)\tilde{q}_2) \\
&+ P_{\mathrm{sh,c}}(\tilde{g}_3^\mu(z,\tilde{q})\partial_\eta^2 \tilde{q}_1) + \tilde{g}_4^\mu(z,\tilde{q}) + P_{\mathrm{sh,c}}(\tilde{g}_5^\mu(z,\tilde{q})\partial_\eta \tilde{q}_2) + \tilde{h}_2^\mu(z), \qquad (67)
\end{aligned}$$

where $y = (z, \mu)$; these equations are valid for $|(z, \mu)| \leq \delta$. Recall that the approximate modulating pulses p^μ found in Section 3 satisfies

$$|p^\mu(\xi)| \leq c_{\rm h}\mu \mathrm{e}^{-\mu^2\theta\xi}, \qquad \xi \in [0, \infty),$$

and it is therefore necessary to choose δ so that $|(p^\mu(\xi), \mu)| \leq \delta$ for $\xi \in [0, \infty)$. This task is accomplished by defining $\delta = (3c_{\rm h} + 1)\mu_0$ and restricting (z, μ) to $\{|z| \leq 2c_{\rm h}\mu_0,\ 0 \leq \mu \leq \mu_0\}$; without loss of generality we henceforth suppose that $\mu = \mu_0$. It follows from inequality (64) that

$$\|\tilde{h}^\mu(z)\|_{\mathcal{X}^{s+1}_{\rm sh,c}} \leq c\mu^2 \mathrm{e}^{-c^\star/\mu}, \tag{68}$$

where $c^\star = (\mathrm{e}C(3c_{\rm h} + 1))^{-1}$, and inequality (68) is the requisite estimate showing that $\tilde{h}^\mu$ is exponentially small with respect to μ for $p = p_{\rm opt}$.

We conclude this section by stating estimates for the remaining nonlinearities $\tilde{F}^\mu$ and $\tilde{g}^\mu_j$, $j = 3, \ldots, 6$ appearing in the transformed equations (36)–(38). These estimates are obtained as a corollary of the following proposition, which refines the results of Proposition 4.5.

Proposition 4.9 *Suppose that $\delta > 0$ and $p \geq 2$ satisfy (56). The estimates*

$$\Bigg\| \sum_{2 \leq k \leq p} \Phi^k(y) \Bigg\|_{\mathcal{X}^{s+2}_{\rm sh,c}} \leq \frac{40a}{\rho^2}\sqrt{10}|z|^2, \qquad \Bigg|\Bigg|\Bigg| \sum_{2 \leq k \leq p} \mathrm{d}\Phi^k[y] \Bigg|\Bigg|\Bigg| \leq 800\frac{a^2\Theta}{\rho^4}|y|^2$$

hold for every $y \in \mathbb{R}^5$ such that $|y| \leq \delta$.

Proof. The key step in the first estimate is the inequality

$$\|\Phi^k(y)\|_{\mathcal{X}^{s+2}_{\rm sh,c}} \leq \sqrt{10k(k-1)}|\Phi^k|^{s+2}_{2,k}|y|^{k-2}|z|^2, \qquad 2 \leq k \leq p,$$

which is obtained by a straightforward calculation using the fact that $\Phi^k(0, \mu) = 0$, $\mathrm{d}_1\Phi^k[0, \mu] = 0$ (see Iooss & Lombardi [7, Lemma A.3] for a similar calculation). It follows that

$$\Bigg\| \sum_{2 \leq k \leq p} \Phi^k(y) \Bigg\|_{\mathcal{X}^{s+2}_{\rm sh,c}} \leq \sum_{2 \leq k \leq p} \sqrt{10k(k-1)}\phi_k \delta^{k-2}|z|^2,$$

while the inequality

$$\Bigg|\Bigg|\Bigg| \sum_{2 \leq k \leq p} \mathrm{d}\Phi^k[y] \Bigg|\Bigg|\Bigg| \leq \sum_{3 \leq k \leq p} \sqrt{5}k\phi_k\delta^{k-3}|y|^2$$

is a direct consequence of (57) and the fact that $\Phi^2 = 0$. The stated estimates are derived by applying the arguments used in Proposition 4.5 to the above inequalities. $\square$

Corollary 4.10 *The transformed nonlinearities $\tilde{F}^\mu$ and g^μ_j, $j = 3, \ldots, 6$ satisfy the inequalities*

$$\begin{aligned}
|\tilde{F}^\mu(z, \tilde{q})| &= O(\mu^2|z| + \|\tilde{q}\|_{\mathcal{X}^{s+1}_{\rm sh,c}}\|(z, \tilde{q})\|_{\mathcal{X}^{s+1}} + \|(z, \tilde{q})\|^3_{\mathcal{X}^{s+1}}),\\
\|\tilde{g}^\mu_j(z, \tilde{q})\|_{s+1} &\leq c\mu\|(z, \tilde{q})\|_{\mathcal{X}^{s+1}_{\rm sh,c}}, \qquad j = 3, 5,\\
\|\tilde{g}^\mu_4(z, \tilde{q})\|_{s+1} &\leq c\mu^2\|\tilde{q}\|_{\mathcal{X}^{s+1}_{\rm sh,c}}\|(z, \tilde{q})\|_{\mathcal{X}^{s+1}},\\
\|\tilde{g}^\mu_6(z, \tilde{q})\|_{s+2} &\leq c\mu^2\|\tilde{q}\|_{\mathcal{X}^{s+1}_{\rm sh,c}}\|(z, \tilde{q})\|_{\mathcal{X}^{s+1}}.
\end{aligned}$$

Proof. Recall that $\tilde{F}^\mu$ and $\tilde{g}_j^\mu$, $j = 3, \ldots, 6$ are defined in terms of F^μ, g^μ, $j = 3, \ldots, 6$ and $\Phi^\mu(z)$ by formulae (39)–(42). The stated estimates for the transformed nonlinearities follow directly from these formulae, the estimates

$$\begin{aligned}
|F^\mu(z,q)| &= O(\mu^2|z| + \|q\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}\|(z,q)\|_{\mathcal{X}^{s+1}} + \|(z,q)\|^3_{\mathcal{X}^{s+1}}), \\
\|g_j^\mu(z,q)\|_{s+1} &= \mathcal{O}(\mu\|(z,q)\|_{\mathcal{X}^{s+1}}), \qquad j = 3, 5, \\
\|g_4^\mu(z,q)\|_{s+1} &= O(\mu^2\|q\|^2_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} + \mu^2\|q\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}|z||(z,\mu)|), \\
\|g_6^\mu(z,q)\|_{s+2} &= O(\mu^2|z|\|q\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}\|(z,q)\|_{\mathcal{X}^{s+1}})
\end{aligned}$$

for the original nonlinearities and the estimates for $\Phi(y)$ given in Proposition 4.9. $\square$

4.4 Estimates for the derivatives of the transformed nonlinearities

The existence theory for modulating pulses presented in Sections 5 and 6 below is based upon perturbation arguments around an approximate modulating pulse p^μ. In order to use perturbation theory of this kind we require additional estimates upon the derivatives of the nonlinearities in the transformed equations (36)–(38), and the appropriate estimates are derived in this section.

We begin by estimating the derivative $\partial\Phi(y)$ of $\Phi(y)$ with respect to $y = (z, \mu)$. It follows from equation (44) that

$$\mathcal{L}(\partial\Phi^m) = \partial N^m, \qquad m = 2, \ldots, p,$$

and differentiating (47), (48), one finds that

$$\begin{aligned}
&\partial N_1^m(y) \\
&= m h_1^m[\{y\}^{(m-1)}, 1] \\
&\quad + \sum_{2\le q\le m} \sum_{i=0}^{q} \sum_{i+p_1+\ldots p_{q-i}=m} \Bigg(i g_6^{i,q-i}[\{y\}^{(i-1)}, 1, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}] \\
&\qquad\qquad + \sum_{j=1}^{q-i} g_6^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\partial\Phi^{p_j}, \ldots - \Phi^{p_{q-i}}]\Bigg) \\
&\quad + \sum_{2\le k\le m-1} \mathrm{d}\Phi_1^k[y]\Bigg(\sum_{2\le q\le m-k+1} \sum_{i=0}^{q} \sum_{\substack{i+p_1+\ldots p_{q-i}\\ =m-k+1}} \Bigg(i F^{i,q-i}[\{y\}^{(i-1)}, 1, -\Phi^{p_1}, \ldots, -\Phi^{p_{q-i}}] \\
&\qquad\qquad + \sum_{j=1}^{q-i} F^{i,q-i}[\{y\}^{(i)}, -\Phi^{p_1}, \ldots, -\partial\Phi^{p_j}, \ldots - \Phi^{p_{q-i}}]\Bigg)\Bigg),
\end{aligned}$$

together with a similar expression for $\partial N_2^m(y)$. The methods used to establish Proposition 4.3 show that $\psi_{m-1} = |\partial\Phi_m|^{s+2}_{2,m-1}$ satisfies

$$\psi_{m-1} \le \sqrt{5}m\left(\frac{\sqrt{5}a}{\rho^2}\right)^{m-1}\beta_m,$$

so that

$$\psi_{m-1} \;\le\; \frac{20a}{\rho^2}\left(\frac{2\sqrt{5}a\Theta}{\rho^2}\right)^{m-2} m(m-2)!, \qquad m = 2,3,\ldots, \tag{69}$$

and the arguments in the proof of Proposition 4.5 yield another estimate for $\partial\Phi$.

Proposition 4.11 *Suppose that $\delta > 0$ and $p \ge 2$ satisfy*

$$\delta p \le \frac{\rho^2}{4\sqrt{5}a\Theta}.$$

The estimates

$$\Big\| \sum_{2\le k\le p} \partial\Phi^k(y) \Big\|_{\mathcal{X}^{s+2}_{\rm sh,c}} \le \frac{\sqrt{5}}{\Theta}, \qquad \Big|\Big|\Big| \sum_{2\le k\le p} \mathrm{d}(\partial\Phi^k)[y] \Big|\Big|\Big| \le \frac{10}{\Theta\delta}$$

hold for every $y \in \mathbb{R}^5$ such that $|y| \le \delta$.

The next step is to derive an estimate for $\|\partial\tilde{h}(y)\|_{\mathcal{X}^{s+1}_{\rm sh,c}}$. Differentiating equations (58), (59) and proceeding as in Section 4.3, we find that

$$\begin{aligned}
\|\partial\tilde{h}_1(y)\|_{s+2} \;&\le\; \left(2+\frac{20}{\Theta}\right)\hat{\Delta}^1_p + \left(1+\frac{20}{\Theta}\right)\hat{\Delta}^2_p + \hat{\Delta}^3_p, \\
\|\partial\tilde{h}_2(y)\|_{s+1} \;&\le\; \left(3+\frac{20}{\Theta}\right)\hat{\Delta}^1_p + \left(3+\frac{20}{\Theta}\right)\hat{\Delta}^2_p + \hat{\Delta}^3_p,
\end{aligned}$$

where

$$\begin{aligned}
\hat{\Delta}^p_1 \;&=\; \sum_{p+1\le q} \frac{a}{\rho^q} q(q+1)\sqrt{5}^q\delta^{q-1} \\
\hat{\Delta}^p_2 \;&=\; \sum_{2\le q\le p}\; \sum_{\substack{n=p_1+\ldots+p_q\\ \ge p+1\\ p_j\ge 1}} \frac{an\delta^{n-1}}{\rho^q}\phi_{p_1}\ldots\phi_{p_q} \\
\hat{\Delta}^p_3 \;&=\; \sum_{\substack{2\le k\le p\\ p+1\le n\le p+k-1}} \sqrt{5}k\phi_k \sum_{2\le q\le n-k+1}\; \sum_{\substack{i+p_1+\ldots+p_q\\ =n-k+1\\ p_j\ge 1}} \frac{an\delta^{n-1}}{\rho^q}\sqrt{5}^i\phi_{p_1}\ldots\phi_{p_q}.
\end{aligned}$$

(In deriving these estimates we have replaced ψ_{m-1} by $m\phi_m$; this procedure is permissible in view of inequalities (55) and (69), which are used to estimate ψ_{m-1} and ϕ_m in the subsequent analysis.) An estimate for $\|\partial\tilde{h}(y)\|_{\mathcal{X}^{s+1}_{\rm sh,c}}$ is obtained from the above calculation by estimating $q(q+1)\sqrt{5}^q \le (2\sqrt{5})^q$ and repeating the proof of Proposition 4.6.

Proposition 4.12 *Suppose that $\delta > 0$ and $p \ge 2$ satisfy the stronger mutual constraint*

$$\delta p \le \frac{\rho^2}{4\sqrt{5}\mathrm{e}a\Theta}.$$

The estimate

$$\|\partial\tilde{h}(y)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} \leq \frac{c}{\delta}\left((C\delta)^{p+1}p! + \frac{1}{\mathrm{e}^{p+1}p^2}\right), \qquad C = \frac{4\sqrt{5}a\Theta}{\rho} \tag{70}$$

holds for every $y \in \mathbb{R}^5$ such that $|y| \leq \delta$.

Proposition 4.12 and the calculation above inequality (64) yield the inequality

$$\|\partial\tilde{h}(y)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} \leq cC^2\left(m\sqrt{\frac{27}{8\mathrm{e}}} + 4\mathrm{e}^2\right)\delta\mathrm{e}^{-1/\mathrm{e}C\delta}$$

for $p = p_{\mathrm{opt}}$, from which it follows that

$$\|\partial\tilde{h}^\mu(z)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} \leq c\mu\mathrm{e}^{-c^\star/\mu}. \tag{71}$$

It remains to derive the corresponding estimates for the derivatives of g_j^μ, $j = 3, \ldots, 6$, and this task is accomplished by repeating the arguments used in Proposition 4.9 and Corollary 4.10.

Proposition 4.13 *The transformed nonlinearities $\tilde{g}_j^\mu$, $j = 3, \ldots, 6$ satisfy the inequalities*

$$\begin{aligned}
\|\partial_i\tilde{g}_j^\mu(z,\tilde{q})\|_{s+1} &\leq c\mu, \qquad j = 3, 5, \\
\|\partial_i\tilde{g}_4^\mu(z,\tilde{q})\|_{s+1} &\leq c\mu^2\|(z,\tilde{q})\|_{\mathcal{X}^{s+1}}, \\
\|\partial_i\tilde{g}_6^\mu(z,\tilde{q})\|_{s+2} &\leq c\mu^2\|(z,\tilde{q})\|_{\mathcal{X}^{s+1}}
\end{aligned}$$

for $i = 1, 2$.

The existence theory presented in Sections 5 and 6 below makes frequent use of the fact that the nonlinearities $\tilde{g}_j^\mu$, $j = 3, \ldots, 6$ and $\tilde{h}^\mu$ are Lipschitz functions of their arguments whose Lipschitz constants are estimated by inequality (71) and Proposition 4.13. The Lipschitz continuity of their derivatives is also required, but here the size of the Lipschitz constants is not important. In these circumstances we use estimates of the form

$$\|\partial_i\tilde{g}_6^\mu(z_1,\tilde{q}_1) - \partial_i\tilde{g}_6^\mu(z_2,\tilde{q}_2)\|_{s+2} \leq c_\mu\|(z_1 - z_2, \tilde{q}_1 - \tilde{q}_2)\|_{\mathcal{X}^{s+1}}, \qquad j = 1, 2,$$

in which the Lipschitz constant c_μ depends upon μ; these estimates follow from the analyticity of the nonlinearities and the restriction (70).

5 The local centre-stable manifold

In this section we construct solutions of equations (36)–(38) whose pointwise distance from an approximate pulse p^μ identified in Section 3 does not exceed $\mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$ (see Figure 4). The *local centre-stable manifold* is the set of initial data for such solutions, and we use it in the next section to extend these solutions to symmetric modulating pulses which exist for $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$. The centre-stable manifold is a generalisation of a concept familiar in dynamical-systems theory, although the methods used to find the solutions whose initial data defines the manifold is rather different here due to the quasilinear nature of our problem.

We begin by writing $z = p^\mu + r$, so that r is a perturbation around an approximate modulating pulse, and decomposing equations (37), (38) into equations for the strongly hyperbolic part $Z = P_{\mathrm{sh}} q$ and central part $w = P_{\mathrm{c}} q$ of q. Recall that the mth Fourier components of $d_1(z)$ vanish for $m \neq 1$, so that

$$P_{\mathrm{sh}}(d_1(z)q_1) = 0, \qquad P_{\mathrm{c}}(d_1(z)q_1) = P_{\mathrm{c}}(d_1(z)w_1).$$

Using this calculation and the corresponding results for d_2 and d_3, one finds that

$$\begin{aligned} \partial_\xi Z &= L_0^\mu Z + F_0^\mu(Z, p^\mu + r, w) + F_1^\mu(p^\mu + r), && (72)\\ \partial_\xi r &= \mathcal{L}_1^\mu r + N^\mu(Z, r, w), && (73) \end{aligned}$$

where

$$\begin{aligned} L_0^\mu &= L^\mu|_{\mathcal{X}^s_{\mathrm{sh}}},\\ F_0^\mu(Z,z,w) &= \begin{pmatrix} P_{\mathrm{sh}}(\tilde{g}_6^\mu(z, Z+w)) \\ P_{\mathrm{sh}}[\tilde{g}_3^\mu(z,Z+w)\partial_\eta^2(Z_1+w_1) + \tilde{g}_4^\mu(z,Z+w) + \tilde{g}_5^\mu(z,Z+w)\partial_\eta(Z_2+w_2)] \end{pmatrix},\\ F_1^\mu(z) &= \begin{pmatrix} P_{\mathrm{sh}}(h_1^\mu(z)) \\ P_{\mathrm{sh}}(h_2^\mu(z)) \end{pmatrix},\\ \mathcal{L}_1^\mu r &= \mathrm{d}_1 F^\mu[p^\mu, 0](r),\\ N^\mu(Z,r,w) &= F^\mu(p^\mu + r, Z+w) - F^\mu(p^\mu, Z+w) - \mathrm{d}_1 F^\mu[p^\mu,0](r), \end{aligned}$$

and

$$\begin{aligned} \partial_\xi w_1 &= w_2 + \hat{g}_6^\mu(Z, p^\mu + r, w) + \hat{h}_1^\mu(p^\mu + r), && (74)\\ \partial_\xi w_2 &= -c_3^\mu k_0^2 \partial_\eta^2 w_1 - c_4^\mu w_1 + \mu P_{\mathrm{c}}(d_1(z)w_1) + \mu P_{\mathrm{c}}(d_2(z)\partial_\eta w_1) + \mu P_{\mathrm{c}}(d_3(z)w_2)\\ &\quad + P_{\mathrm{c}}(\hat{g}_3^\mu(Z, p^\mu + r, w)\partial_\eta^2 w_1) + \hat{g}_4^\mu(Z, p^\mu + r, w)\\ &\quad + P_{\mathrm{c}}(\hat{g}_5^\mu(Z, p^\mu + r, w)\partial_\eta w_2) + \hat{h}_2^\mu(p^\mu + r), && (75) \end{aligned}$$

where

$$\begin{aligned} \hat{g}_j^\mu(Z,z,w) &= \tilde{g}^\mu(z, Z+w), \qquad j = 3,5,\\ \hat{g}_4^\mu(Z,z,w) &= P_{\mathrm{c}}[\tilde{g}_4^\mu(z,Z+w) + \tilde{g}_3^\mu(z,Z+w)\partial_\eta^2 Z_1 + \tilde{g}_5^\mu(z,Z+w)\partial_\eta Z_2],\\ \hat{g}_6^\mu(Z,z,w) &= P_{\mathrm{c}}\tilde{g}_6^\mu(z,Z+w), \qquad \hat{h}_j^\mu(z) = P_{\mathrm{c}}\tilde{h}_j^\mu(z), \quad j = 1,2. \end{aligned}$$

Estimates for the nonlinearities in equations (72) and (74), (75) are obtained from Corollary 4.10. Notice that

$$\begin{aligned} |F_0^\mu(Z,z,w)| &\leq c(\mu\|(Z,w)\|^2_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} + \mu^2|z|\|(Z,w)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}), && (76)\\ |\partial_i F_0^\mu(Z,z,w)| &\leq c(\mu\|(Z,w)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} + \mu^2|z|), \qquad j = 1,2,3, && (77) \end{aligned}$$

and

$$\begin{aligned} \|\hat{g}_4^\mu(Z,z,w)\|_{s+1} &\leq c(\mu\|(Z,w)\|^2_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} + \mu^2|z|\|(Z,w)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}}), && (78)\\ \|\partial_i \hat{g}_4^\mu(Z,z,w)\|_{s+1} &\leq c(\mu\|(Z,w)\|_{\mathcal{X}^{s+1}_{\mathrm{sh,c}}} + \mu^2|z|), \qquad j = 1,2,3 && (79) \end{aligned}$$

because $P_{\rm sh}\tilde g_3(z,Z+w)\partial_\eta^2(Z_1+w_1)$, $P_{\rm sh}\tilde g_5(z,Z+w)\partial_\eta(Z_2+w_2)$, $P_{\rm c}\tilde g_3(z,Z+w)\partial_\eta^2 Z_1$ and $P_{\rm c}\tilde g_5(z,Z+w)\partial_\eta Z_2$ do not contain any terms which are linear in z. Clearly

$$|F_1^\mu(z)| \;\le\; c\mu^2 \mathrm{e}^{-c^\star/\mu}, \qquad |\partial F_1^\mu(z)| \;\le\; c\mu \mathrm{e}^{-c^\star/\mu}, \tag{80}$$

while $\hat g_3^\mu(Z,z,w)$, $\hat g_5^\mu(Z,z,w)$ and $\hat g_6^\mu(Z,z,w)$ satisfy the same estimates as respectively $\tilde g_3^\mu(z,q)$, $\tilde g_5^\mu(z,q)$ and $\tilde g_6^\mu(z,q)$ with $q=Z+w$. The following result gives estimates for N^μ.

Proposition 5.1 *The nonlinearity N^μ appearing in equation (73) satisfies the estimates*

$$\begin{aligned}
|N^\mu(Z,r,w)| &\le c\left(\frac{|r|^2}{\mu^2}+\|(Z,w)\|^2_{\mathcal{X}^{s+1}_{\rm sh,c}}\right),\\
|\partial_1 N^\mu(Z,r,w)| &\le c\left(\frac{|r|}{\mu}+\|(Z,w)\|_{\mathcal{X}^{s+1}_{\rm sh,c}}\right),\\
|\partial_2 N^\mu(Z,r,w)| &\le \frac{c}{\mu}\left(\frac{|r|}{\mu}+\|(Z,w)\|_{\mathcal{X}^{s+1}_{\rm sh,c}}\right),\\
|\partial_3 N^\mu(Z,r,w)| &\le c\left(\frac{|r|}{\mu}+\|(Z,w)\|_{\mathcal{X}^{s+1}_{\rm sh,c}}\right).
\end{aligned}$$

Proof. Define $(\hat z,\hat\mu)=\delta^{-1}(z,\mu)$ and suppose that $|\hat y|=|(\hat z,\hat\mu)|\le 1$. The estimate

$$\|\Phi^k(\delta\hat y)\|_{\mathcal{X}^{s+1}_{\rm sh,c}} \le \frac{\sqrt{5}\delta}{\Theta}|\hat y|^k$$

(see Proposition 4.5) implies that Φ is a polynomial in $\hat y$ whose coefficients are bounded independently of p. It follows that $\hat F^{\hat\mu}(\hat z,Z+w)=F^\mu(z,Z+w)$ is an analytic function of Z, $\hat y$ and w whose Taylor coefficients are bounded independently of p, and the same is true of the quadratic function

$$\hat N^{\hat\mu}(Z,\hat r,w)=\hat F^{\hat\mu}(\mu^{-1}p^\mu+\hat r,Z+w)-\hat F^{\hat\mu}(\hat r,Z+w)-\mathrm{d}_1\hat F^{\hat\mu}[\mu^{-1}p^\mu,0](\hat r),$$

which therefore satisfies

$$\begin{aligned}
|\hat N^{\hat\mu}(Z,\hat r,w)| &\le c(|\hat r|^2+\|(Z,w)\|^2_{\mathcal{X}^{s+1}_{\rm sh,c}}),\\
|\partial_j\hat N^{\hat\mu}(Z,\hat r,w)| &\le c(|\hat r|+\|(Z,w)\|_{\mathcal{X}^{s+1}_{\rm sh,c}}), \qquad j=1,2,3.
\end{aligned}$$

The stated estimates for $N^\mu(Z,r,w)=\hat N^{\hat\mu}(Z,\hat r,w)$ are obtained by returning to the unscaled variable $r=\delta\hat r=(3c_{\rm h}+1)\mu\hat r$ in the above inequalities. □

Our task is to find solutions (Z,r,w) of (72)–(75) for which $|Z(\xi)|$, $|r(\xi)|$ and $\|w(\xi)\|_{\mathcal{X}^{s+1}_{\rm c}}$ do not exceed $\mathrm{e}^{-c^\star/2\mu}$ for $\xi\in[0,\mathrm{e}^{c^\star/2\mu}]$, and for this purpose we require some information concerning the spectrum of the linear part of the vector fields on the right-hand sides of (72) and (73). Recall that L_0^μ has a pair $\pm\lambda_{0,\mu}=\pm(1+k_0^2)^{1/2}+\mathcal{O}(\mu^4)$ of simple eigenvalues with corresponding eigenvectors $u_0=(1,\lambda_{0,\mu})$ and $s_0=(1,-\lambda_{0,\mu})$ which define the stable

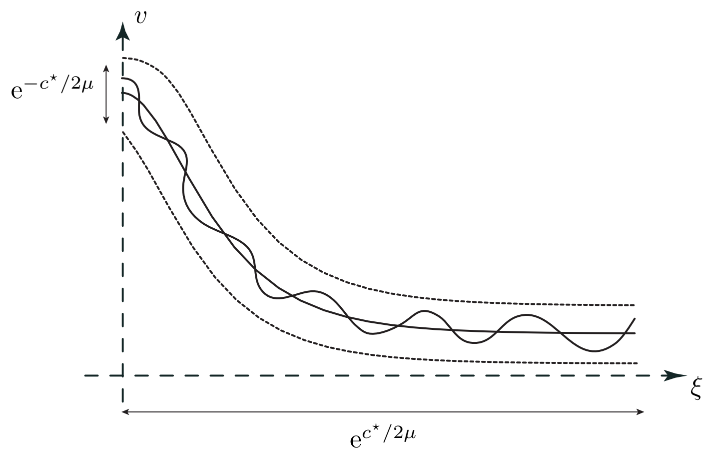

Figure 4: Solutions with initial data on the local centre-stable manifold $W^{\mathrm{cs}}_{\mathrm{loc}}$ remain within a distance of $\mathrm{e}^{-c^\star/2\mu}$ *of* p^μ *on the timescale* $[0, \mathrm{e}^{c^\star/2\mu}]$.

and unstable directions associated with this matrix. The projections onto these directions are constructed in the usual fashion using the dual basis $\{s_0^*, u_0^*\}$ to $\{s_0, u_0\}$ in $\mathcal{X}^s_{\mathrm{wh}}$, where

$$s_0^* = \frac{1}{2\lambda_{0,\mu}}(\lambda_{0,\mu}, 1), \qquad u_0^* = \frac{1}{2\lambda_{0,\mu}}(\lambda_{0,\mu}, -1).$$

The stable and unstable directions associated with the time-dependent linear operator $\mathcal{L}^\mu$ are described by the following result; it is proved by noting that

$$\|\mathcal{L}^\mu - L^\mu|_{\mathcal{X}^s_{\mathrm{wh}}}\|_{\mathcal{X}^s_{\mathrm{wh}} \to \mathcal{X}^s_{\mathrm{wh}}} \leq c\mu \mathrm{e}^{-\theta\mu^2|\xi|}, \qquad \xi \in \mathbb{R}$$

and using the method explained by Groves & Mielke [3, §4.3].

Proposition 5.2 *The equation*

$$\partial_\xi r = \mathcal{L}_1^\mu r$$

has solutions $s_{1,1}(\xi)$, $s_{1,2}(\xi)$, $u_{1,1}(\xi)$, $u_{1,2}(\xi)$ *on* $[0,\infty)$ *such that*

$$|s_{1,j}(\xi)| \leq c\mathrm{e}^{-\lambda_{1,\mu}\xi}, \quad |u_{1,j}(\xi)| \leq c\mathrm{e}^{\lambda_{1,\mu}\xi}, \qquad j = 1, 2, \quad \xi \in [0, \infty).$$

The dual basis $\{s_{1,1}^*(\xi), s_{1,2}^*(\xi), u_{1,1}^*(\xi), u_{1,2}^*(\xi)\}$ *to* $\{s_{1,1}(\xi), s_{1,2}(\xi), u_{1,1}(\xi), u_{1,2}(\xi)\}$ *in* $\mathcal{X}^s_{\mathrm{wh}}$ *satisfies*

$$|s_{1,j}^*(\xi)| \leq \frac{c}{\lambda_{1,\mu}}\mathrm{e}^{\lambda_{1,\mu}\xi}, \quad |u_{1,j}^*(\xi)| \leq \frac{c}{\lambda_{1,\mu}}\mathrm{e}^{-\lambda_{1,\mu}\xi}, \qquad j = 1, 2, \quad \xi \in [0, \infty).$$

The requisite solutions of (72)–(75) are constructed using the following iteration scheme. Choose real numbers Z^0, r_1^0, r_2^0 whose magnitude is at most $\mu\mathrm{e}^{-c^\star/2\mu}$ and $w^0 \in \mathcal{X}_{\mathrm{c}}^{s+1}$ such that $\|w^0\|_{\mathcal{X}_{\mathrm{c}}^{s+2}} \leq \mu\mathrm{e}^{-c^\star/2\mu}$. Set $Z_{(0)} = 0$, $r_{(0)} = 0$, $w_{(0)} = 0$ and for $m = 0, 1, 2, \ldots$ define $Z_{(m+1)} \in C([0, \mathrm{e}^{c^\star/2\mu}], \mathbb{R}^2)$, $r_{(m+1)} \in C([0, \mathrm{e}^{c^\star/2\mu}], \mathbb{R}^4)$ by the formulae

$$Z_{(m+1)}(\xi) = Z^0\, s_0 + \int_0^\xi \langle (F^\mu_{(0m)} + F^\mu_{(1m)})(\tau), s_0^* \mathrm{e}^{\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, s_0 \mathrm{e}^{-\lambda_{0,\mu}\xi}$$
$$- \int_\xi^{\mathrm{e}^{c^\star/2\mu}} \langle (F^\mu_{(0m)} + F^\mu_{(1m)})(\tau), u_0^* \mathrm{e}^{-\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, u_0 \mathrm{e}^{\lambda_{0,\mu}\xi}, \tag{81}$$

$$r_{(m+1)}(\xi) = r_1^0\, s_{1,1}(\xi) + r_2^0\, s_{1,2}(\xi)$$
$$+ \sum_{j=1}^2 \int_0^\xi \langle N^\mu_{(m)}(\tau), s^*_{1,j}(\tau)\rangle \,\mathrm{d}\tau\, s_{1,j}(\xi) - \sum_{j=1}^2 \int_\xi^{\mathrm{e}^{c^\star/2\mu}} \langle N^\mu_{(m)}(\tau), u^*_{1,j}(\tau)\rangle \,\mathrm{d}\tau\, u_{1,j}(\xi) \tag{82}$$

and let $w_{(m+1)} \in C([0, \mathrm{e}^{c^\star/2\mu}], \mathcal{X}_\mathrm{c}^{s+1})$ be the solution of the equations

$$\partial_\xi w_{1(m+1)} = w_{2(m+1)} + \hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)}, \tag{83}$$
$$\begin{aligned}\partial_\xi w_{2(m+1)} = {} & -c_3^\mu k_0^2 \partial_\eta^2 w_{1(m+1)} - c_4^\mu w_{1(m+1)} \\ & + P_\mathrm{c}(d_{1(m)} w_{1(m+1)}) + P_\mathrm{c}(d_{2(m)} \partial_\eta w_{1(m+1)}) + P_\mathrm{c}(d_{3(m)} w_{2(m+1)}) \\ & + P_\mathrm{c}(\hat{g}^\mu_{3(m)} \partial_\eta^2 w_{1(m+1)}) + \hat{g}^\mu_{4(m)} + P_\mathrm{c}(\hat{g}^\mu_{3(m)} \partial_\eta w_{2(m+1)}) + \hat{h}^\mu_{2(m)},\end{aligned} \tag{84}$$

with initial data $w_{(m+1)}|_{\xi=0} = w^0$; here $F_{j(m)}$, $N^\mu_{(m)}$, $d_{j(m)}$, $\hat{g}^\mu_{j(m)}$ and $\hat{h}^\mu_{j(m)}$ are abbreviations for respectively $F^\mu_j(Z_{(m)}, r_{(m)}, w_{(m)})$, $N^\mu(Z_{(m)}, r_{(m)}, w_{(m)})$, $d_j(z_{(m)})$, $\hat{g}^\mu_j(Z_{(m)}, r_{(m)}, w_{(m)})$ and $\hat{h}^\mu_j(Z_{(m)}, r_{(m)}, w_{(m)})$.

Lemma 5.3 *Suppose that $\|w^0\|_{\mathcal{X}_\mathrm{c}^{s+1}} \leq \mu \mathrm{e}^{-c^\star/2\mu}$. The estimates*

$$\sup_{m\in\mathbb{N}} \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |Z_{(m)}(\xi)| \leq c\mathrm{e}^{-c^\star/2\mu}, \qquad \sup_{m\in\mathbb{N}} \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |r_{(m)}(\xi)| \leq c\mathrm{e}^{-c^\star/2\mu}$$

and

$$\sup_{m\in\mathbb{N}} \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|w_{(m)}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \leq c\mathrm{e}^{-c^\star/2\mu}$$

hold for all sufficiently small values of μ.

Proof. We demonstrate that

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |Z_{(m+1)}(\xi)| \leq c\mu\mathrm{e}^{-c^\star/2\mu}, \qquad \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |r_{(m+1)}(\xi)| \leq c\mu\mathrm{e}^{-c^\star/2\mu}$$

and

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|w_{(m+1)}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \leq c\mu^{1/2}|\log\mu|\mathrm{e}^{-c^\star/2\mu}$$

whenever

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \Big(|Z_{(j)}(\xi)| + |r_{(j)}(\xi)| + \|w_{(j)}(\xi)\|_{\mathcal{X}_\mathrm{c}^s}\Big) \leq \mathrm{e}^{-c^\star/2\mu}$$

for $j = 0, \ldots, m$; the lemma follows inductively from this result.

Observe that

$$
\begin{aligned}
&|Z_{(m+1)}(\xi)| \\
&\quad\leq\; c\left(|Z^0| + \mu^3 \mathrm{e}^{-c^\star/2\mu}\int_0^\xi \mathrm{e}^{\lambda_{0,\mu}\tau}\,\mathrm{d}\tau\,\mathrm{e}^{-\lambda_{0,\mu}\xi} + \mu^3\mathrm{e}^{-c^\star/2\mu}\int_\xi^\infty \mathrm{e}^{-\lambda_{0,\mu}\tau}\,\mathrm{d}\tau\,\mathrm{e}^{\lambda_{0,\mu}\xi}\right) \\
&\quad\leq\; c(|Z^0| + \mu^3\mathrm{e}^{-c^\star/2\mu}) \\
&\quad\leq\; c\mu\mathrm{e}^{-c^\star/2\mu}
\end{aligned} \tag{85}
$$

and

$$
\begin{aligned}
&|r_{(m+1)}(\xi)| \\
&\quad\leq\; c\left(|r_1^0| + |r_2^0| + \frac{\mathrm{e}^{-c^\star/\mu}}{\mu^2\lambda_{1,\mu}}\int_0^\xi \mathrm{e}^{\lambda_{1,\mu}\tau}\,\mathrm{d}\tau\,\mathrm{e}^{-\lambda_{1,\mu}\xi} + \frac{\mathrm{e}^{-c^\star/\mu}}{\mu^2\lambda_{1,\mu}}\int_\xi^\infty \mathrm{e}^{-\lambda_{1,\mu}\tau}\,\mathrm{d}\tau\,\mathrm{e}^{\lambda_{1,\mu}\xi}\right) \\
&\quad\leq\; c\left(|r_1^0| + |r_2^0| + \frac{\mathrm{e}^{-c^\star/\mu}}{\mu^2\lambda_{1,\mu}^2}\right) \\
&\quad\leq\; c\mu\mathrm{e}^{-c^\star/2\mu}
\end{aligned} \tag{86}
$$

for $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$; here we have used the estimates

$$
|F^\mu_{0(m)}| \;\leq\; c\mu^3\mathrm{e}^{-c^\star/2\mu}, \qquad |F^\mu_{1(m)}| \;\leq\; c\mu^2\mathrm{e}^{-c^\star/\mu}, \qquad |N^\mu_{0(m)}| \;\leq\; \frac{c}{\mu^2}\mathrm{e}^{-c^\star/\mu},
$$

which are obtained from (76), (80) and Proposition 5.1. The corresponding result for $w_{(m+1)}$ is obtained by applying energy estimates to equations (83), (84).

Define the *energy* $\mathcal{E}_{s+1}$ by

$$
\begin{aligned}
\mathcal{E}_{s+1}(w) &= \int \{(\partial_\eta^{s+1}w_2)^2 - c_3^\mu k_0^2(\partial_\eta^{s+2}w_1)^2 + c_4^\mu(\partial_\eta^{s+1}w_1)^2\}\,\mathrm{d}\eta \\
&= \int (\partial_\eta^{s+1}w_2)^2\,\mathrm{d}\eta + \sum_{j=1}^\infty(-c_3^\mu k_0^2 j^2 + c_4^\mu)j^{2s+2}|w_{j,1}|^2
\end{aligned}
$$

and note that $\mathcal{E}_{s+1}$ is equivalent to the usual norm on $\mathcal{X}_\mathrm{c}^{s+1}$ because w has zero mean and $c_3^\mu < 0$. Applying the operator $\partial_\eta^{s+1}w_{2(m+1)}\partial_\eta^{s+1}$ to (84), integrating with respect to η over one period and using the calculation

$$
\begin{aligned}
&\int\{\partial_\eta^{s+1}w_{2(m+1)}\partial_\xi\partial_\eta^{s+1}w_{2(m+1)} + c_3^\mu k_0^2\partial_\eta^{s+1}w_{2(m+1)}\partial_\eta^{s+3}w_{1(m+1)} + c_4^\mu\partial_\eta^{s+1}w_{2(m+1)}\partial_\eta^{s+1}w_{1(m+1)}\}\,\mathrm{d}\eta \\
&\quad=\; \frac{1}{2}\partial_\xi\mathcal{E}_{s+1}(w_{(m+1)}) - c_3^\mu k_0^2\int \partial_\eta^{s+3}w_{1(m+1)}\partial_\eta^{s+1}(\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)})\,\mathrm{d}\eta \\
&\qquad\qquad - c_4^\mu\int \partial_\eta^{s+1}w_{1(m+1)}\partial_\eta^{s+1}(\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)})\,\mathrm{d}\eta,
\end{aligned}
$$

which is obtained by integrating by parts and using equation (83), one finds that

$$
\begin{aligned}
&\frac{1}{2}\partial_\xi \mathcal{E}_{s+1}(w_{(m+1)}) \\
&= c_3^\mu k_0^2 \int \partial_\eta^{s+3} w_{1(m+1)} \partial_\eta^{s+1}(\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)})\,\mathrm{d}\eta + c_4^\mu \int \partial_\eta^{s+1} w_{1(m+1)} \partial_\eta^{s+1}(\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)})\,\mathrm{d}\eta \\
&\quad + \mu \int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(d_{1(m)} w_{1(m+1)})\,\mathrm{d}\eta + \mu \int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(d_{2(m)} \partial_\eta w_{1(m+1)})\,\mathrm{d}\eta \\
&\quad + \mu \int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(d_{3(m)} w_{2(m+1)})\,\mathrm{d}\eta + \int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1}(\hat{g}^\mu_{4(m)} + \hat{h}^\mu_{2(m)})\,\mathrm{d}\eta \\
&\quad + \int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(\hat{g}^\mu_{3(m)} \partial_\eta^2 w_{1(m+1)})\,\mathrm{d}\eta + \int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(\hat{g}^\mu_{5(m)} \partial_\eta w_{2(m+1)})\,\mathrm{d}\eta.
\end{aligned}
$$

An estimate for $\mathcal{E}_{s+1}(w_{(m+1)})$ can be derived from this identity with the help of the estimates

$$
\begin{aligned}
\|d_{j(m)}\|_{s+1} &\leq c(\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\mu^2\theta\xi}), \qquad j = 1, 2, 3, \\
\|\hat{g}_{j(m)}\|_{s+1} &\leq c\mu(\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\mu^2\theta\xi}), \qquad j = 3, 5, \\
\|\hat{g}_{4(m)}\|_{s+1} &\leq c\mu(\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi}), \\
\|\hat{g}_{6(m)}\|_{s+2} &\leq c\mu(\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi}), \\
\|\hat{h}_{(m)}\|_{\mathcal{X}_\mathrm{c}^{s+1}} &\leq c\mu^2 \mathrm{e}^{-c^\star/\mu},
\end{aligned}
$$

which are obtained from (78) and Corollary 4.10.

Notice that

$$
\begin{aligned}
&\left|\int \partial_\eta^{s+3} w_{1(m+1)} \partial_\eta^{s+1}(\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)})\,\mathrm{d}\eta\right| \\
&\qquad = \left|\int \partial_\eta^{s+2} w_{1(m+1)} \partial_\eta^{s+2}(\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)})\,\mathrm{d}\eta\right| \\
&\qquad \leq \|w_{1(m+1)}\|_{s+2}(\|\hat{g}^\mu_{6(m)}\|_{s+2} + \|\hat{h}^\mu_{1(m)}\|_{s+2}) \\
&\qquad \leq c\mu \mathcal{E}_{s+1}(w_{(m+1)})^{1/2}(\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi}),
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\left|\int \partial_\eta^{s+1} w_{1(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(d_{1(m)} w_{1(m+1)})\,\mathrm{d}\eta\right| &\leq c(\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}_{s+1}(w_{(m+1)}), \\
\left|\int \partial_\eta^{s+1} w_{1(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(d_{2(m)} \partial_\eta w_{1(m+1)})\,\mathrm{d}\eta\right| &\leq c(\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}_{s+1}(w_{(m+1)}), \\
\left|\int \partial_\eta^{s+1} w_{1(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(d_{3(m)} w_{2(m+1)})\,\mathrm{d}\eta\right| &\leq c(\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}_{s+1}(w_{(m+1)}), \\
\left|\int \partial_\eta^{s+1} w_{1(m+1)} \partial_\eta^{s+1}(\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)})\,\mathrm{d}\eta\right| &\leq c\mu \mathcal{E}_{s+1}(w_{(m+1)})^{1/2}(\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi}), \\
\left|\int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1}(\hat{g}^\mu_{4(m)} + \hat{h}^\mu_{2(m)})\,\mathrm{d}\eta\right| &\leq c\mu \mathcal{E}_{s+1}(w_{(m+1)})^{1/2}(\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi}).
\end{aligned}
$$

Straightforward calculations show that

$$
\begin{aligned}
&\int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_c(\hat{g}^\mu_{3(m)} \partial_\eta^2 w_{1(m+1)})\, \mathrm{d}\eta \\
&\quad= \int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} (\hat{g}^\mu_{3(m)} \partial_\eta^2 w_{1(m+1)})\, \mathrm{d}\eta + s_1 \\
&\quad= \int \partial_\eta^{s+1} w_{2(m+1)} \hat{g}^\mu_{3(m)} \partial_\eta^{s+3} w_{1(m+1)}\, \mathrm{d}\eta + s_1 + s_2 \\
&\quad= -\int \partial_\eta^{s+2} w_{2(m+1)} \partial_\eta^{s+2} w_{1(m+1)} \hat{g}^\mu_{3(m)}\, \mathrm{d}\eta + s_1 + s_2 + s_3 \\
&\quad= -\int \partial_\eta^{s+2} (\partial_\xi w_{1(m+1)} - \hat{g}^\mu_{6(m)} - \hat{h}^\mu_{1(m)}) \partial_\eta^{s+2} w_{1(m+1)} \hat{g}^\mu_{3(m)}\, \mathrm{d}\eta + s_1 + s_2 + s_3 \\
&\quad= -\frac{1}{2} \partial_\xi \int (\partial_\eta^{s+2} w_{1(m+1)})^2 \hat{g}^\mu_{3(m)}\, \mathrm{d}\eta + s_1 + s_2 + s_3 + s_4 + s_5,
\end{aligned}
$$

where

$$
\begin{aligned}
s_1 &= -\int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} (P_\mathrm{h}(\hat{g}^\mu_{3(m)} \partial_\eta^2 w_{1(m+1)}))\, \mathrm{d}\eta, \\
s_2 &= \int \partial_\eta^{s+1} w_{2(m+1)} \left(\sum_{j=0}^{s} \begin{pmatrix} s+1 \\ j \end{pmatrix} \partial_\eta^{s+1-j} \hat{g}^\mu_{3(m)} \partial_\eta^{j+2} w_{1(m+1)} \right) \mathrm{d}\eta, \\
s_3 &= -\int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta \hat{g}^\mu_{3(m)} \partial_\eta^{s+1} w_{2(m+1)}\, \mathrm{d}\eta, \\
s_4 &= \int \partial_\eta^{s+2} (\hat{g}^\mu_{6(m)} + \hat{h}^\mu_{1(m)}) \partial_\eta^{s+2} w_{1(m+1)} \hat{g}^\mu_{3(m)}\, \mathrm{d}\eta \\
s_5 &= \frac{1}{2} \int (\partial_\eta^{s+2} w_{1(m+1)})^2 \partial_\xi \hat{g}^\mu_{3(m)}\, \mathrm{d}\eta,
\end{aligned}
$$

and

$$
\begin{aligned}
&\int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_\mathrm{c}(\hat{g}^\mu_{5(m)} \partial_\eta w_{2(m+1)})\, \mathrm{d}\eta \\
&\quad= \int \partial_\eta^{s+2} w_{2(m+1)} \partial_\eta^{s+1} (\hat{g}^\mu_{5(m)} \partial_\eta w_{2(m+1)})\, \mathrm{d}\eta + s_6 \\
&\quad= \int \partial_\eta^{s+1} w_{2(m+1)} \hat{g}^\mu_{5(m)} \partial_\eta^{s+2} \partial_\xi w_{2(m+1)}\, \mathrm{d}\eta + s_6 + s_7 \\
&\quad= \frac{1}{2} \int \partial_\eta ((\partial_\eta^{s+1} w_{2(m+1)})^2) \hat{g}^\mu_{5(m)}\, \mathrm{d}\eta + s_6 + s_7 \\
&\quad= s_6 + s_7 + s_8,
\end{aligned}
$$

where

$$
\begin{aligned}
s_6 &= -\int \partial_\eta^{s+1} w_{2(m+1)} \partial_\eta^{s+1} P_\mathrm{h}(\hat{g}^\mu_{5(m)} \partial_\eta w_{2(m+1)})\, \mathrm{d}\eta, \\
s_7 &= \int \partial_\eta^{s+1} w_{2(m+1)} \left(\sum_{j=0}^{s} \begin{pmatrix} s+1 \\ j \end{pmatrix} \partial_\eta^{s+1-j} \hat{g}^\mu_{5(m)} \partial_\eta^{j+1} w_{2(m+1)}) \right) \mathrm{d}\eta, \\
s_8 &= \frac{1}{2} \int \partial_\eta ((\partial_\eta^{s+1} w_{2(m+1)})^2) \hat{g}^\mu_{5(m)}\, \mathrm{d}\eta;
\end{aligned}
$$

these quantities satisfy the estimates

$$\begin{array}{rcll} |s_j| & \leq & c\mu^3\mathcal{E}_{s+1}(w_{(m+1)})^{1/2}(\mathrm{e}^{-c^\star/\mu}+\mu^2\mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi}), & j=1,2,3, \\ |s_j| & \leq & \mu\mathcal{E}_{s+1}(w_{(m+1)})(\mathrm{e}^{-c^\star/2\mu}+\mu\mathrm{e}^{-\mu^2\theta\xi}), & j=4,6,7,8 \end{array}$$

and

$$\begin{array}{rcl} |s_5| & \leq & c\mathcal{E}_{s+1}(w_{(m+1)})\|\partial_\xi\hat{g}^\mu_{3(m)}\|_s \\ & \leq & c\mathcal{E}_{s+1}(w_{(m+1)})\|\partial\hat{g}^\mu_{3(m)}\|_s(|\partial_\xi Z_{(m)}|+|\partial_\xi r_{(m)}|+|\partial_\xi p^\mu|+\|\partial_\xi w_{(m)}\|_{\mathcal{X}^s_c}) \\ & \leq & c\mu\mathcal{E}_{s+1}(w_{(m+1)})(\mathrm{e}^{-c^\star/2\mu}+\mu\mathrm{e}^{-\mu^2\theta\xi}) \end{array}$$

(the inequalities

$$|\partial_\xi Z_{(m)}|\leq c\mathrm{e}^{-c^\star/\mu},\qquad |\partial_\xi r_{(m)}|\leq c\mathrm{e}^{-c^\star/\mu},\qquad \|\partial_\xi w_{(m)}\|_{\mathcal{X}^s_c}\leq c\mathrm{e}^{-c^\star/\mu}$$

follow directly from the inductive hypothesis by means of equations (81)–(84) with m replaced by $m-1$).

Finally, let us define

$$\mathcal{E}^{\mathrm{e}}_{s+1}(w)=\mathcal{E}_{s+1}(w)-\frac{1}{2}\int(\partial^{s+2}_\eta w_1)^2\hat{g}^\mu_{3(m)}\,\mathrm{d}y$$

and note that

$$\frac{1}{c}\mathcal{E}_{s+1}(w)\leq\mathcal{E}^{\mathrm{e}}_{s+1}(w)\leq c\mathcal{E}_{s+1}(w) \tag{87}$$

since

$$\left|\int(\partial^{s+2}_\eta w_1)^2\hat{g}^\varepsilon_{3(m)}\,\mathrm{d}y\right| \;\leq\; \mathcal{E}_{s+1}(w)\|\hat{g}^\varepsilon_{3(m)}\|_{s+1} \;\leq\; \mu^2\mathcal{E}_{s+1}(w).$$

Altogether, we have that

$$\begin{aligned} &\partial_\xi\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)}) \\ &\quad\leq c\mu(\mathrm{e}^{-c^\star/\mu}+\mu^2\mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}_{s+1}(w_{(m+1)})^{1/2}+c\mu(\mathrm{e}^{-c^\star/2\mu}+\mu\mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}_{s+1}(w_{(m+1)}). \end{aligned} \tag{88}$$

We proceed by establishing an estimate for $\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})$ on the short interval $\xi\in[0,\xi^\star]$, where

$$\xi^\star=\frac{\alpha|\log\mu|}{\theta\mu^2},$$

so that $\mathrm{e}^{-\mu^2\theta\xi^\star}=\mu^\alpha$, and α is an appropriately chosen positive constant. It follows from inequality (88) that

$$\partial_\xi\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})\leq c_1\mu^2\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})+c_2\mu^4\mathrm{e}^{-c^\star/\mu},$$

and an application of Gronwall's inequality yields

$$\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\xi)\leq(\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(0)+c_2\mu^4\mathrm{e}^{-c^\star/\mu}\xi)\mathrm{e}^{c_1\mu^2\xi}.$$

Choosing $\alpha=\theta/c_1$, one finds that

$$e^{c_1\mu^2\xi^\star}=\frac{1}{\mu},$$

whereby

$$\begin{aligned}
\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\xi) &\leq \frac{1}{\mu}(\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(0) + c_2\mu^4 \mathrm{e}^{-c^\star/\mu}\xi^\star) \\
&\leq \frac{1}{\mu}(\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(0) + c\mu^2 \mathrm{e}^{-c^\star/\mu}|\log\mu|) \\
&\leq c\mu \mathrm{e}^{-c^\star/\mu}|\log\mu| \qquad (89)
\end{aligned}$$

for $\xi \in [0,\xi^\star]$. This intermediate step may now be used to deduce the desired estimate for $\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})$ on the long interval $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$.

Integrating inequality (88), we find that

$$\begin{aligned}
&\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\xi) \\
&\leq \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(0) + c\mu \sup_{\tau\in[0,\mathrm{e}^{c^\star/2\mu}]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau) \\
&\quad + \mu \mathrm{e}^{-c^\star/2\mu} \sup_{\tau\in[0,\mathrm{e}^{c^\star/2\mu}]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})^{1/2}(\tau) \\
&\quad + \mu^3 \mathrm{e}^{-c^\star/2\mu} \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\theta\mu^2\tau}\,\mathrm{d}\tau \sup_{\tau\in[0,\mathrm{e}^{c^\star/2\mu}]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})^{1/2}(\tau) \\
&\quad + c\mu^2 \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\theta\mu^2\tau} \mathcal{E}_{s+1}(w_{(m+1)})(\tau)\,\mathrm{d}\tau, \qquad \xi \in [0,\mathrm{e}^{c^\star/2\mu}]. \qquad (90)
\end{aligned}$$

Observe that

$$\begin{aligned}
&\mu^2 \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\mu^2\theta\tau} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau)\,\mathrm{d}\tau \\
&= \mu^2 \int_0^{\xi^\star} \mathrm{e}^{-\mu^2\theta\tau} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau)\,\mathrm{d}\tau + \mu^2 \int_{\xi^\star}^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\mu^2\theta\tau} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau)\,\mathrm{d}\tau \\
&\leq \mu^2\xi^\star \sup_{\tau\in[0,\xi^\star]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau) + \mu^2 \int_{\xi^\star}^{\infty} \mathrm{e}^{-\mu^2\theta\tau}\,\mathrm{d}\tau \sup_{\tau\in[0,\mathrm{e}^{c^\star/2\mu}]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau) \\
&= \mu^2\xi^\star \sup_{\tau\in[0,\xi^\star]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau) + \frac{\mathrm{e}^{-\theta\mu^2\xi^\star}}{\theta} \sup_{\tau\in[0,\mathrm{e}^{c^\star/2\mu}]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau) \\
&\leq c\mu \mathrm{e}^{-c^\star/\mu}|\log\mu|^2 + c\mu^\alpha \sup_{\tau\in[0,\mathrm{e}^{c^\star/2\mu}]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\tau),
\end{aligned}$$

in which the last step follows by (89) and the definition of $\xi^\star$. Inserting this estimate and the inequality

$$\int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\theta\mu^2\tau}\,\mathrm{d}\tau \;\leq\; \int_0^{\infty} \mathrm{e}^{-\theta\mu^2\tau}\,\mathrm{d}\tau \;=\; \frac{1}{\theta\mu^2}$$

into (90), one concludes that

$$\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\xi) \leq c\mu|\log\mu|^2 \mathrm{e}^{-c^\star/\mu} + c(\mu+\mu^\alpha) \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\xi), \qquad \xi \in [0,\mathrm{e}^{c^\star/2\mu}],$$

so that

$$\mathcal{E}^{\mathrm{e}}_{s+1}(w_{(m+1)})(\xi) \leq c\mu|\log\mu|^2 \mathrm{e}^{-c^\star/\mu}, \qquad \xi \in [0,\mathrm{e}^{c^\star/2\mu}]. \qquad \square$$

Lemma 5.4 *Suppose that* $\|w^0\|_{\mathcal{X}_c^{s+2}} \leq \mu e^{-c^\star/2\mu}$. *The iterates* $Z_{(m)}$, $r_{(m)}$ *and* $w_{(m)}$ *satisfy*

$$\sup_{\xi\in[0,e^{c^\star/2\mu}]}\left(|\tilde{Z}_{(m+1)}(\xi)|+|\tilde{r}_{(m+1)}(\xi)|+\|\tilde{w}_{(m+1)}(\xi)\|_{\mathcal{X}_c^{s+1}}\right)$$
$$\leq \frac{1}{2}\sup_{\xi\in[0,e^{c^\star/2\mu}]}\left(|\tilde{Z}_{(m)}(\xi)|+|\tilde{r}_{(m)}(\xi)|+\|\tilde{w}_{(m)}(\xi)\|_{\mathcal{X}_{\mathrm{sh,c}}^{s+1}}\right)$$

for each $m\in\mathbb{N}_0$, *where*

$$\tilde{Z}_{(m+1)} = Z_{(m+1)} - Z_{(m)}, \quad \tilde{r}_{(m+1)} = r_{(m+1)} - r_{(m)}, \quad \tilde{w}_{(m+1)} = w_{(m+1)} - w_{(m)}.$$

Proof. Examining the equations

$$\begin{aligned}\tilde{Z}_{(m+1)}(\xi) = {} & \int_0^\xi \langle (F^\mu_{0(m)} + F^\mu_{1(m)} - F^\mu_{0(m-1)} - F^\mu_{1(m-1)})(\tau), s_0^* e^{\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, s_0 e^{-\lambda_{0,\mu}\xi} \\ & - \int_\xi^{e^{c^\star/2\mu}} \langle (F^\mu_{0(m)} + F^\mu_{1(m)} - F^\mu_{0(m-1)} - F^\mu_{1(m-1)})(\tau), u_0^* e^{-\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, u_0 e^{\lambda_{0,\mu}\xi},\end{aligned}$$

$$\begin{aligned}\tilde{r}_{(m+1)}(\xi) = {} & \sum_{j=1}^2 \int_0^\xi \langle (N^\mu_{(m)} - N^\mu_{(m-1)})(\tau), s_{1,j}^*(\tau)\rangle \,\mathrm{d}\tau\, s_{1,j}(\xi) \\ & - \sum_{j=1}^2 \int_\xi^{e^{c^\star/2\mu}} \langle (N^\mu_{(m)} - N^\mu_{(m-1)})(\tau), u_{1,j}^*(\tau)\rangle \,\mathrm{d}\tau\, u_{1,j}(\xi),\end{aligned}$$

one finds that

$$\begin{aligned}&|\tilde{Z}_{(m+1)}(\xi)| \\ &\quad\leq c\left(\mu^3\int_0^\xi (|\tilde{Z}_{(m)}(\xi)|+|\tilde{r}_{(m)}|+\|\tilde{w}_{(m)}(\xi)\|_{\mathcal{X}_c^{s+1}}) e^{\lambda_{0,\mu}\tau}\,\mathrm{d}\tau\, e^{-\lambda_{0,\mu}\xi}\right. \\ &\qquad\qquad \left. + \mu^3\int_\xi^{e^{c^\star/2\mu}} (|\tilde{Z}_{(m)}(\xi)|+|\tilde{r}_{(m)}|+\|\tilde{w}_{(m)}(\xi)\|_{\mathcal{X}_c^{s+1}}) e^{-\lambda_{0,\mu}\tau}\,\mathrm{d}\tau\, e^{\lambda_{0,\mu}\xi}\right) \\ &\quad\leq c\mu^3 \sup_{\xi\in[0,e^{c^\star/2\mu}]}\left(|\tilde{Z}_{(m)}(\xi)|+|\tilde{r}_{(m)}|+\|\tilde{w}_{(m)}(\xi)\|_{\mathcal{X}_c^{s+1}}\right),\end{aligned}$$

$$\begin{aligned}&|\tilde{r}_{(m+1)}(\xi)| \\ &\quad\leq c\left(\frac{e^{-c^\star/2\mu}}{\mu^2\lambda_{1,\mu}}\int_0^\xi (|\tilde{Z}_{(m)}(\xi)|+|\tilde{r}_{(m)}|+\|\tilde{w}_{(m)}(\xi)\|_{\mathcal{X}_c^{s+1}}) e^{\lambda_{1,\mu}\tau}\,\mathrm{d}\tau\, e^{-\lambda_{1,\mu}\xi}\right. \\ &\qquad\qquad \left. + \frac{e^{-c^\star/2\mu}}{\mu^2\lambda_{1,\mu}}\int_\xi^{e^{c^\star/2\mu}} (|\tilde{Z}_{(m)}(\xi)|+|\tilde{r}_{(m)}|+\|\tilde{w}_{(m)}(\xi)\|_{\mathcal{X}_c^{s+1}}) e^{-\lambda_{1,\mu}\tau}\,\mathrm{d}\tau\, e^{\lambda_{1,\mu}\xi}\right) \\ &\quad\leq c\frac{e^{-c^\star/2\mu}}{\mu^2\lambda_{1,\mu}^2} \sup_{\xi\in[0,e^{c^\star/2\mu}]}\left(|\tilde{Z}_{(m)}(\xi)|+|\tilde{r}_{(m)}|+\|\tilde{w}_{(m)}(\xi)\|_{\mathcal{X}_c^{s+1}}\right)\end{aligned}$$

for $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$, in which the estimates

$$
\begin{aligned}
|F^\mu_{0(m)} - F^\mu_{0(m-1)}| &\leq c\mu^3(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}|), \\
|F^\mu_{1(m)} - F^\mu_{1(m-1)}| &\leq c\mu \mathrm{e}^{c^\star/\mu}(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}|), \\
|N^\mu_{(m)} - N^\mu_{(m-1)}| &\leq \frac{c\mathrm{e}^{c^\star/2\mu}}{\mu^2}(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}|)
\end{aligned}
$$

have been used (see equations (77), (80) and Proposition 5.1).

Similarly, the estimate for $\tilde{w}_{(m+1)}$ is obtained by studying the equations

$$
\partial_\xi \tilde{w}_{1(m+1)} = \tilde{w}_{2(m+1)} + \hat{g}^\mu_{6(m)} - \hat{g}^\mu_{6(m-1)} + \hat{h}^\mu_{1(m)} - \hat{h}^\mu_{1(m-1)}, \tag{91}
$$

$$
\begin{aligned}
\partial_\xi \tilde{w}_{2(m+1)} = {} & -c^\mu_3 k_0^2 \partial_\eta^2 \tilde{w}_{1(m+1)} - c^\mu_4 \tilde{w}_{1(m+1)} + \mu P_\mathrm{c}(d_{1(m)}\tilde{w}_{1(m+1)}) + \mu P_\mathrm{c}(d_{2(m)}\partial_\eta \tilde{w}_{1(m+1)}) \\
& + \mu P_\mathrm{c}(d_{3(m)}\tilde{w}_{2(m+1)}) + P_\mathrm{c}(\hat{g}^\mu_{3(m)}\partial_\eta^2 \tilde{w}_{1(m+1)}) + P_\mathrm{c}(\hat{g}^\mu_{5(m)}\partial_\eta \tilde{w}_{2(m+1)}) \\
& + \mu P_\mathrm{c}((d_{1(m)} - d_{1(m-1)})w_{1(m)} + \mu P_\mathrm{c}((d_{2(m)} - d_{2(m-1)})\partial_\eta w_{1(m)} \\
& + \mu P_\mathrm{c}((d_{3(m)} - d_{3(m-1)})w_{2(m)} + P_\mathrm{c}((\hat{g}^\mu_{3(m)} - \hat{g}^\mu_{3(m-1)})\partial_\eta^2 w_{1(m)}) \\
& + P_\mathrm{c}((\hat{g}^\mu_{5(m)} - \hat{g}^\mu_{5(m-1)})\partial_\eta w_{2(m)}) + \hat{g}^\mu_{4(m)} - \hat{g}^\mu_{4(m-1)} + \hat{h}^\mu_{2(m)} - \hat{h}^\mu_{2(m-1)}
\end{aligned} \tag{92}
$$

and using the additional estimates

$$
\begin{aligned}
\|d_{j(m)} - d_{j(m-1)}\|_{s+1} &\leq c|\tilde{r}_{(m)}|, \qquad j = 1, 2, 3, \\
\|\hat{g}^\mu_{4(m)} - \hat{g}^\mu_{4(m-1)}\|_{s+1} &\leq c\mu(\mathrm{e}^{-c^\star/2\mu} + \mu^2 \mathrm{e}^{-\mu^2\theta\xi})(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}|), \\
\|\hat{g}^\mu_{6(m)} - \hat{g}^\mu_{6(m-1)}\|_{s+2} &\leq c\mu(\mathrm{e}^{-c^\star/2\mu} + \mu^2 \mathrm{e}^{-\mu^2\theta\xi})(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}|), \\
\|\hat{g}^\mu_{j(m)} - \hat{g}^\mu_{j(m-1)}\|_{s+1} &\leq c\mu(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}|), \qquad j = 3, 5, \\
\|\hat{h}^\mu_{(m)} - \hat{h}^\mu_{(m-1)}\|_{\mathcal{X}_\mathrm{c}^{s+1}} &\leq c\mu \mathrm{e}^{-c^\star/\mu}|\tilde{r}_{(m)}|,
\end{aligned}
$$

which are obtained from (79) and Proposition 4.13. We apply the operator $\partial_\eta^{s+1}\tilde{w}_{2(m+1)}\partial_\eta^{s+1}$ to (92), integrate with respect to η over one period and use the estimation techniques developed in the previous lemma; the result is

$$
\begin{aligned}
& \partial_\xi \mathcal{E}^\mathrm{e}_{s+1}(\tilde{w}_{(m+1)}) \\
& \quad \leq c\mu(\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}^\mathrm{e}_{s+1}(\tilde{w}_{(m+1)}) \\
& \qquad + c\mu(\mathrm{e}^{-c^\star/2\mu} + \mu^2 \mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}^\mathrm{e}_{s+1}(\tilde{w}_{(m+1)})^{1/2}(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}|^2 + |\tilde{r}_{(m)}|),
\end{aligned} \tag{93}
$$

where we have used the further calculations

$$
\begin{aligned}
& \left| \int \partial_\eta^{s+1}\tilde{w}_{2(m+1)}\partial_\eta^{s+1} P_\mathrm{c}((d_{1(m)} - d_{1(m-1)})w_{1(m)}) \, \mathrm{d}\eta \right| \\
& \qquad \leq \mathcal{E}_{s+1}(\tilde{w}_{(m+1)})^{1/2}\mathcal{E}_{s+1}(w_{(m)})^{1/2}\|d_{1(m)} - d_{1(m-1)}\|_{s+1} \\
& \qquad \leq c\mathrm{e}^{-c^\star/2\mu}\mathcal{E}_{s+1}(\tilde{w}_{(m+1)})|\tilde{r}_{(m)}|, \\
& \left| \int \partial_\eta^{s+1}\tilde{w}_{2(m+1)}\partial_\eta^{s+1} P_\mathrm{c}((\hat{g}^\mu_{3(m)} - \hat{g}^\mu_{3(m-1)})\partial_\eta^2 w_{1(m)}) \, \mathrm{d}\eta \right| \\
& \qquad \leq \mathcal{E}_{s+1}(\tilde{w}_{(m+1)})^{1/2}\mathcal{E}_{s+2}(w_{(m)})^{1/2}\|\hat{g}^\mu_{3(m)} - \hat{g}^\mu_{3(m-1)}\|_{s+1} \\
& \qquad \leq c\mu \mathrm{e}^{-c^\star/2\mu}\mathcal{E}_{s+1}(\tilde{w}_{(m+1)})(\mathcal{E}_{s+1}(\tilde{w}_{(m)})^{1/2} + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}|).
\end{aligned}
$$

(The estimate $\mathcal{E}_{s+2}(w_{(m)}) \leq \mathrm{e}^{-c^\star/\mu}$ is obtained by repeating Lemma 5.3 with s replaced by $s+1$ (and requires the stronger condition $\|w^0\|_{\mathcal{X}_\mathrm{c}^{s+2}} \leq \mu \mathrm{e}^{c^\star/2\mu}$). Using the two-step method in the proof of the previous lemma, one finds from (93) that

$$\sup_{\tau \in [0, \mathrm{e}^{c^\star/2\mu}]} \mathcal{E}_{s+1}(\tilde{w}_{(m+1)}(\tau)) \leq c\mu |\log \mu|^2 \sup_{\tau \in [0, \mathrm{e}^{c^\star/2\mu}]} \left(\mathcal{E}_{s+1}(\tilde{w}_{(m)}(\tau)) + |\tilde{Z}_{(m)}| + |\tilde{r}_{(m)}(\tau)|^2 \right). \quad \square$$

The following convergence result is a direct consequence of the above lemmata.

Theorem 5.5 *For each Z^0, r_1^0, r_2^0 and w^0 with*

$$|Z^0| \leq \mu \mathrm{e}^{-c^\star/2\mu}, \quad |r_1^0| \leq \mu \mathrm{e}^{-c^\star/2\mu}, \quad |r_2^0| \leq \mu \mathrm{e}^{-c^\star/2\mu}, \quad \|w^0\|_{\mathcal{X}_\mathrm{c}^{s+2}} \leq \mu \mathrm{e}^{-c^\star/2\mu}$$

the sequence $(Z_{(m)}, r_{(m)}, w_{(m)})_{m \in \mathbb{N}_0}$ converges in $C([0, \mathrm{e}^{c^\star/2\mu}], \mathcal{X}^{s+1})$ to a limit $(Z_\star, r_\star, w_\star)$ which satisfies the estimate

$$\sup_{\xi \in [0, \mathrm{e}^{c^\star/2\mu}]} \|(Z_\star(\xi), r_\star(\xi), w_\star(\xi))\|_{\mathcal{X}^{s+1}} \leq \mathrm{e}^{-c^\star/2\mu}$$

and solves equations (72)–(75).

We now use the above results to define a local centre-stable manifold at time $\xi = 0$ for the nonautonomous equations (72)–(75). According to Lemmata 5.3 and 5.4 the solutions defining this manifold are available under the hypothesis that $\|w^0\|_{\mathcal{X}_\mathrm{c}^{s+2}} \leq \mu \mathrm{e}^{-c^\star/2\mu}$; to ensure its differentiability one however requires the stronger hypothesis that $\|w^0\|_{\mathcal{X}_\mathrm{c}^{s+4}} \leq \mu \mathrm{e}^{-c^\star/2\mu}$ (see Section 6 below), and we therefore make this hypothesis from the outset.

Definition 5.6 *The set of points*

$$W_{\mathrm{loc}}^{\mathrm{cs}} = \bigcup \{(Z_\star(0), r_\star(0), w_\star(0))\},$$

in which the union is taken over the set of Z^0, r_1^0, r_2^0 and w^0 such that

$$|Z^0| \leq \mu \mathrm{e}^{-c^\star/2\mu}, \quad |r_1^0| \leq \mu \mathrm{e}^{-c^\star/2\mu}, \quad |r_2^0| \leq \mu \mathrm{e}^{-c^\star/2\mu}, \quad \|w^0\|_{\mathcal{X}_\mathrm{c}^{s+4}} \leq \mu \mathrm{e}^{-c^\star/2\mu},$$

is called the local centre-stable manifold for solutions to (72)–(75) at time $\xi = 0$.

6 Existence theory for symmetric modulating pulses

In this section we identify solutions $(Z_\star, r_\star, w_\star)$ to equations (72)–(75) on the interval $[0, \mathrm{e}^{c^\star/2\mu}]$ whose initial data $(Z_\star(0), r_\star(0), w_\star(0))$ lies on $W_{\mathrm{loc}}^{\mathrm{cs}}$ and which can be extended to solutions that remain $\mathcal{O}(\mathrm{e}^{-c^\star/2\mu})$ on $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$. The idea is to exploit the reversibility of equations (72)–(75) (see Section 2); in particular, solutions with the property that $(Z_\star(0), r_\star(0), w_\star(0))$ lies on the *symmetric section*

$$\Sigma := \mathrm{Fix}\, S = \mathcal{X}^{s+1} \cap \{(v_{1,\mathrm{o}}, v_{2,\mathrm{e}}) = (0,0)\}$$

can be extended to symmetric solutions on $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$. Because $w_\star(0) = w^0$ we have that $w_\star(0) \in \Sigma_\mathrm{c} := P_\mathrm{c}\Sigma$ whenever $w^0 \in \Sigma_\mathrm{c}$ and our task is reduced to that of finding a criterion on (Z^0, r_1^0, r_2^0) which guarantees that $(Z_\star(0), r_\star(0)) \in \Sigma_\mathrm{h} := P_\mathrm{h}\Sigma$.

Our first step is to introduce an artificial parameter by replacing F_1^μ and $\hat{h}^\mu$ in equations (72)–(75) by ρF_1^μ and $\rho\hat{h}^\mu$; the construction of $W_\mathrm{cs}^\mathrm{loc}$ undertaken in Section 5 above clearly remains valid for all values of $\rho \in [0, 1]$. Observe that $\rho = 1$ yields the original equations while $\rho = 0$ yields the system considered in Section 3, in which $\{(Z, w) = (0, 0)\}$ is an invariant subspace containing the homoclinic solution p^μ (generated by the solution $(Z, r, w) = (0, 0, 0)$ in the present coordinates). We consider a solution $(Z_\star, r_\star, w_\star)$ with $(Z_\star(0), r_\star(0), w_\star(0)) \in W_\mathrm{loc}^\mathrm{cs}$ as a function of Z^0, r_1^0, r_2^0 which depends upon $\rho \in \mathbb{R}$ and $w^0 \in \Sigma_\mathrm{c}$ as parameters (with $\rho \in [0, 1]$, $\|w^0\|_{\mathcal{X}_\mathrm{c}^{s+4}} \leq \mu\mathrm{e}^{-c^\star/2\mu}$) and therefore write $(Z_\star, r_\star, w_\star)$ as $(Z_{\rho,w^0}, r_{\rho,w^0}, w_{\rho,w^0})(Z^0, r_1^0, r_2^0)$ in the following analysis. Notice that $(Z_{\rho,w^0}, r_{\rho,w^0}, w_{\rho,w^0})(Z^0, r_1^0, r_2^0)|_{\xi=0} \in \Sigma$ whenever (Z^0, r_1^0, r_2^0) is a solution of the equation

$$J_{\rho,w^0}(Z^0, r_1^0, r_2^0) = 0, \tag{94}$$

where $J_{\rho,w^0} : \bar{B}_{\mu\mathrm{e}^{-c^\star/2\mu}}(0) \subset \mathbb{R}^3 \to \mathbb{R}^3$ is defined by

$$J_{\rho,w^0}(Z^0, r_1^0, r_2^0) = \left(\begin{array}{c} (I - S_\mathrm{sh})Z_{\rho,w^0}(Z^0, r_1^0, r_2^0)|_{\xi=0} \\ (I - S_\mathrm{wh})r_{\rho,w^0}(Z^0, r_1^0, r_2^0)|_{\xi=0} \end{array} \right).$$

(The right-hand side of this equation is a vector in $\mathcal{X}_\mathrm{h}^{s+1}$ with only three nonzero entries, namely its Z_2, $z_{1,\mathrm{o}}$ and $z_{2,\mathrm{e}}$ components, and is therefore identified with a triplet of real numbers.) Equation (94) has the solution $(Z^0, r_1^0, r_2^0) = (0, 0, 0)$ at $(\rho, w^0) = (0, 0)$ since the unique solution of (72)–(75) with $(\rho, w^0) = (0, 0)$ is $(Z, r, w) = (0, 0, 0)$. We therefore seek a solution of (72)–(75) near this known solution for parameter values (ρ, w^0) near $(0, 0)$, and it seems natural to apply the implicit-function theorem; notice, however, that we are forced to work from first principles (by applying the contraction mapping principle) since we require precise information concerning the parameter-dependence of the solutions, in particular that the solution exists for values of ρ up to one.

In order to carry out the above programme it is necessary to show that J_{ρ,w^0} is differentiable with respect to Z^0, r_1^0, r_2^0 and obtain some estimates on its derivatives. We therefore need to show that the solutions $(Z_{\rho,w^0}, r_{\rho,w^0}, w_{\rho,w^0})$ described above are differentiable with respect to Z^0, r_1^0, r_2^0 and obtain some estimates on their derivatives. To this end we formally differentiate equations (81)–(84) with respect to Z^0 and use a dot to denote ∂_{Z^0}; we treat the resulting linear equations for $\dot{Z}$, $\dot{r}$, $\dot{w}$ with the iteration scheme

$$\begin{aligned}\dot{Z}_{(m+1)}(\xi) = s_0(\xi) + \int_0^\xi &\langle (\mathrm{d}F_0^\mu[Z, p^\mu + r, w](\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)}) \\ &\qquad + \rho\mathrm{d}F_1^\mu[p^\mu + r](\dot{r}_{(m)}))(\tau), s_0^*\mathrm{e}^{\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau \, s_0\mathrm{e}^{-\lambda_{0,\mu}\xi} \\ - \int_\xi^{\mathrm{e}^{c^\star/2\mu}} &\langle (\mathrm{d}F_0^\mu[Z, p^\mu + r, w](\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)}) \\ &\qquad + \rho\mathrm{d}F_1^\mu[p^\mu + r](\dot{r}_{(m)}))(\tau), u_0^*\mathrm{e}^{-\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau \, u_0\mathrm{e}^{\lambda_{0,\mu}\xi},\end{aligned} \tag{95}$$

$$
\begin{aligned}
\dot r_{(m+1)}(\xi) &= \sum_{j=1}^{2}\int_0^{\xi}\langle \mathrm{d}N^\mu(Z,r,w)(\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})(\tau), s^*_{1,j}(\tau)\rangle \,\mathrm{d}\tau\, s_{1,j}(\xi) \\
&\quad - \sum_{j=1}^{2}\int_{\xi}^{\mathrm{e}^{c^\star/2\mu}}\langle \mathrm{d}N^\mu(Z,r,w)(\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})(\tau), u^*_{1,j}(\tau)\rangle \,\mathrm{d}\tau\, u_{1,j}(\xi),
\end{aligned} \tag{96}
$$

$$
\partial_\xi \dot w_{1(m+1)} = \dot w_{2(m+1)} + \mathrm{d}\hat g_6^\mu[Z,p^\mu+r,w](\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)}) + \rho\mathrm{d}\hat h_1^\mu[p^\mu+r](\dot r_{(m)}), \tag{97}
$$

$$
\begin{aligned}
\partial_\xi \dot w_{2(m+1)} &= -c_3^\mu k_0^2\partial_\eta^2 \dot w_{1(m+1)} - c_4^\mu \dot w_{1(m+1)} \\
&\quad + \mu P_\mathrm{c}(d_1(p^\mu+r)\dot w_{1(m+1)}) + \mu P_\mathrm{c}(d_2(p^\mu+r)\partial_\eta \dot w_{1(m+1)}) + \mu P_\mathrm{c}(d_3(p^\mu+r)\dot w_{2(m+1)}) \\
&\quad + \mu P_\mathrm{c}(d_1(\dot r_{(m)})w_1) + \mu P_\mathrm{c}(d_2(\dot r_{(m)})\partial_\eta w_1) + \mu P_\mathrm{c}(d_3(\dot r_{(m)})w_2) \\
&\quad + P_\mathrm{c}(\hat g_3^\mu(Z,p^\mu+r,w)\partial_\eta^2 \dot w_{1(m+1)}) + P_\mathrm{c}(\mathrm{d}\hat g_3^\mu[Z,p^\mu+r,w](\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})\partial_\eta^2 w_1) \\
&\quad + P_\mathrm{c}(\hat g_5^\mu(Z,p^\mu+r,w)\partial_\eta \dot w_{2(m+1)}) + P_\mathrm{c}(\mathrm{d}\hat g_5^\mu[Z,p^\mu+r,w](\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})\partial_\eta w_2) \\
&\quad + \mathrm{d}\hat g_4^\mu[Z,p^\mu+r,w](\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)}) + \rho\mathrm{d}\hat h_2^\mu[p^\mu+r](\dot r_{(m)}).
\end{aligned} \tag{98}
$$

Let us now choose Z, r, w which satisfy $|Z(\xi)|$, $|r(\xi)|$, $\|w(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+3}} \le \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0,\mathrm{e}^{c^\star/2\mu}]$, take $\dot Z_{(0)} = 0$, $\dot r_{(0)} = 0$, $\dot w_{(0)} = 0$, and for $m = 0,1,2,\ldots$ define $\dot Z_{(m+1)} \in C([0,\mathrm{e}^{c^\star/2\mu}],\mathbb{R}^2)$, $\dot r_{(m+1)} \in C([0,\mathrm{e}^{c^\star/2\mu}],\mathbb{R}^4)$ by formulae (95), (96) and let $\dot w_{(m+1)} \in C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}_\mathrm{c}^{s+1})$ be the solution of (97), (98) with initial data $(\dot w_1,\dot w_2)|_{\xi=0} = (0,0)$.

Lemma 6.1

(i) The estimate

$$
\begin{aligned}
&\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]}\left(|\bar Z_{(m+1)}(\xi)| + |\bar r_{(m+1)}(\xi)| + \|\bar w_{(m+1)}(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+1}}\right) \\
&\qquad \le \frac{1}{2}\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]}\left(|\bar Z_{(m)}(\xi)| + |\bar r_{(m)}(\xi)| + \|\bar w_{(m)}(\xi)\|_{\mathcal{X}_\mathrm{sh,c}^{s+1}}\right), \qquad m \in \mathbb{N}
\end{aligned}
$$

holds uniformly over the set of (Z,r,w) which satisfy $|Z(\xi)|$, $|r(\xi)|$, $\|w(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+2}} \le \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0,\mathrm{e}^{c^\star/2\mu}]$, where $\bar R_{(m+1)} = \dot R_{(m+1)} - \dot R_{(m)}$, $\bar w_{(m+1)} = \dot w_{(m+1)} - \dot w_{(m)}$. Under these hypotheses the sequence $\{(\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})\}$ is bounded independently of (Z,r,w) in $C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+1})$.

(ii) Suppose additionally that $\|w(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+3}} \le \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0,\mathrm{e}^{c^\star/2\mu}]$. For each fixed value of $(\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})$ the iterate $(\dot Z_{(m+1)},\dot r_{(m+1)},\dot w_{(m+1)}) \in C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+1})$ depends Lipschitz-continuously on $(Z,r,w) \in C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+2})$; the Lipschitz constant is an affine function of the norm $\|(\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})\|_{C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+1})}$.

Proof. Equations for the difference $(\bar Z_{(m)},\bar r_{(m)},\bar w_{(m)})$ are obtained by replacing $(\dot Z_{(m)},\dot r_{(m)},\dot w_{(m)})$ by $(\bar Z_{(m)},\bar r_{(m)},\bar w_{(m)})$ and $\tilde Z_0$ by zero in equations (95)–(98). Observe that the equations for $(\bar Z_{(m)},\bar r_{(m)},\bar w_{(m)})$ are transformed into those for $(\tilde Z_{(m)},\tilde r_{(m)},\tilde w_{(m)})$ examined in Lemma 5.4 by replacing derivatives such as $\mathrm{d}F_0^\mu[Z,p^\mu+r,w](\bar Z_{(m)},\bar r_{(m)},\bar w_{(m)})$ with differences such as $|F^\mu_{0(m)} - F^\mu_{0(m-1)}|$, where the derivatives obey the same estimates as the differences. The first assertion in part (i) therefore follows from the conclusion of Lemma 5.4; the second assertion is a consequence of the first together with the linearity of the right-hand sides of equations (95)–(98) in $(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})$.

Turning to part (ii), note that

$$
\begin{aligned}
\hat{Z}(\xi) &= \int_0^\xi \langle (\mathrm{d}F^\mu_{0,1} - \mathrm{d}F^\mu_{0,2} + \rho\mathrm{d}F^\mu_{1,1} - \rho\mathrm{d}F^\mu_{1,2})(\tau), s_0^* \mathrm{e}^{\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, s_0 \mathrm{e}^{-\lambda_{0,\mu}\xi} \\
&\quad - \int_\xi^{\mathrm{e}^{c^\star/2\mu}} \langle (\mathrm{d}F^\mu_{0,1} - \mathrm{d}F^\mu_{0,2} + \rho\mathrm{d}F^\mu_{1,1} - \rho\mathrm{d}F^\mu_{1,2})(\tau), u_0^* \mathrm{e}^{-\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, u_0 \mathrm{e}^{\lambda_{0,\mu}\xi},
\end{aligned}
$$

$$
\begin{aligned}
\hat{r}(\xi) &= \sum_{j=1}^2 \int_0^\xi \langle (\mathrm{d}N_1^\mu - \mathrm{d}N_2^\mu)(\tau), s_j^*(\tau)\rangle \,\mathrm{d}\tau\, s_j(\xi) \\
&\quad - \sum_{j=1}^2 \int_\xi^{e^{c^\star/2\mu}} \langle (\mathrm{d}N_1^\mu - \mathrm{d}N_2^\mu)(\tau), u_j^*(\tau)\rangle \,\mathrm{d}\tau\, u_j(\xi),
\end{aligned}
$$

$$
\begin{aligned}
\partial_\xi \hat{w}_{1(m+1)} &= \hat{w}_{2(m+1)} + \mathrm{d}\hat{g}^\mu_{6,1} - \mathrm{d}\hat{g}^\mu_{6,2} + \rho\mathrm{d}\hat{h}^\mu_{1,1} - \rho\mathrm{d}\hat{h}^\mu_{2,1}, \\
\partial_\xi \hat{w}_{2(m+1)} &= -c_3^\mu k_0^2 \partial_\eta^2 \hat{w}_{1(m+1)} - c_4^\mu \hat{w}_{1(m+1)} \\
&\qquad + \mu P_\mathrm{c}(d_{1,1}\hat{w}_{1(m+1)}) + \mu P_\mathrm{c}(d_{2,1}\partial_\eta \hat{w}_{1(m+1)}) + \mu P_\mathrm{c}(d_{3,1}\hat{w}_{2(m+1)}) \\
&\qquad + \mu P_\mathrm{c}((d_{1,1}-d_{1,2})\dot{w}^2_{1(m+1)}) + \mu P_\mathrm{c}((d_{2,1}-d_{2,2})\partial_\eta \dot{w}^2_{1(m+1)}) + \mu P_\mathrm{c}((d_{3,1}-d_{3,2})\dot{w}^2_{2(m+1)}) \\
&\qquad + \mu P_\mathrm{c}(d_1(\dot{r}_{(m)})\tilde{w}_1) + \mu P_\mathrm{c}(d_2(\dot{r}_{(m)})\partial_\eta \tilde{w}_1) + \mu P_\mathrm{c}(d_3(\dot{r}_{(m)})\tilde{w}_2) \\
&\qquad + P_\mathrm{c}(\hat{g}^\mu_{3,1}\partial_\eta^2 \hat{w}_{1(m+1)}) + P_\mathrm{c}(\hat{g}^\mu_{5,1}\partial_\eta \hat{w}_{2(m+1)}) \\
&\qquad + P_\mathrm{c}((\hat{g}^\mu_{3,1}-\hat{g}^\mu_{3,2})\partial_\eta^2 \dot{w}^2_{1(m+1)}) + P_\mathrm{c}((\hat{g}^\mu_{5,1}-\hat{g}^\mu_{5,2})\partial_\eta \dot{w}^2_{2(m+1)}) \\
&\qquad + P_\mathrm{c}((\mathrm{d}\hat{g}^\mu_{3,1}-\mathrm{d}\hat{g}^\mu_{3,2})\partial_\eta^2 w_1^2) + P_\mathrm{c}((\mathrm{d}\hat{g}^\mu_{5,1}-\mathrm{d}\hat{g}^\mu_{5,2})\partial_\eta w_2^2) \\
&\qquad + P_\mathrm{c}(\mathrm{d}\hat{g}^\mu_{3,1}\partial_\eta^2 \tilde{w}_1) + P_\mathrm{c}(\mathrm{d}\hat{g}^\mu_{5,1}\partial_\eta \tilde{w}_2) + \mathrm{d}\hat{g}^\mu_{4,1} - \mathrm{d}\hat{g}^\mu_{4,2} + \rho\mathrm{d}\hat{h}^\mu_1 - \rho\mathrm{d}\hat{h}^\mu_2,
\end{aligned}
$$

in which $(\hat{Z}, \hat{r}, \hat{w})$ denotes the difference between the values $(\dot{Z}^1_{(m+1)}, \dot{r}^1_{(m+1)}, \dot{w}^1_{(m+1)})$ and $(\dot{Z}^2_{(m+1)}, \dot{r}^2_{(m+1)}, \dot{w}^2_{(m+1)})$ of $(\dot{Z}_{(m+1)}, \dot{r}_{(m+1)}, \dot{w}_{(m+1)})$ for $(Z, r, w) = (Z^1, r^1, w^1)$ and $(Z, r, w) = (Z^2, r^2, w^2)$, $\mathrm{d}F^\mu_j$ is an abbreviation for $\mathrm{d}F^\mu[Z^j, r^j, w^j](\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})$, $j = 1, 2$ (similar abbreviations are used for the other functions) and $(\tilde{Z}, \tilde{r}, \tilde{w}) = (Z^1, r^1, w^1) - (Z^2, r^2, w^2)$.

Employing the symbol c_μ to denote a constant which depends upon μ and estimating

$$
\|\mathrm{d}\hat{g}^\mu_{3,1} - \mathrm{d}\hat{g}^\mu_{3,2}\|_{s+1} \leq c_\mu \|(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})\|_{\mathcal{X}^{s+1}} \|(\tilde{Z}, \tilde{r}, \tilde{w})\|_{\mathcal{X}^{s+1}}
$$

together with similar estimates for the other terms involving differences of derivatives (see the remarks at the end of Section 4.4), we find that

$$
\begin{aligned}
&|\hat{Z}(\xi)| \\
&\quad \leq c_\mu \Bigg(\frac{1}{\lambda_{0,\mu}} \int_0^\xi \|(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})(\tau)\|_{\mathcal{X}^{s+1}} \|(\tilde{Z}, \tilde{r}, \tilde{w})(\tau)\|_{\mathcal{X}^{s+1}} \mathrm{e}^{\lambda_{0,\mu}\tau} \,\mathrm{d}\tau\, \mathrm{e}^{-\lambda_{0,\mu}\xi} \\
&\qquad + \frac{1}{\lambda_{0,\mu}} \int_\xi^{\mathrm{e}^{c^\star/2\mu}} \|(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})(\tau)\|_{\mathcal{X}^{s+1}} \|(\tilde{Z}, \tilde{r}, \tilde{w})(\tau)\|_{\mathcal{X}^{s+1}} \mathrm{e}^{-\lambda_{0,\mu}\tau} \,\mathrm{d}\tau\, \mathrm{e}^{\lambda_{0,\mu}\xi} \Bigg)
\end{aligned}
$$

for $\xi \le \mathrm{e}^{c^\star/2\mu}$, whence

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |\hat{Z}(\xi)| \le c_\mu \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|(\dot{Z}_{(m)},\dot{r}_{(m)},\dot{w}_{(m)})(\xi)\|_{\mathcal{X}^{s+1}} \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|(\tilde{Z},\tilde{r},\tilde{w})(\xi)\|_{\mathcal{X}^{s+1}}$$

and similarly

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |\hat{r}(\xi)| \le c_\mu \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|(\dot{Z}_{(m)},\dot{r}_{(m)},\dot{w}_{(m)})(\xi)\|_{\mathcal{X}^{s+1}} \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|(\tilde{Z},\tilde{r},\tilde{w})(\xi)\|_{\mathcal{X}^{s+1}}.$$

Furthermore, the usual energy estimates show that

$$\begin{aligned}
&\partial_\xi \mathcal{E}^{\mathrm{e}}_{s+1}(\hat{w}_{(m+1)})\\
&\quad\le\ c\mu(\mathrm{e}^{-c^\star/2\mu}+\mu\mathrm{e}^{-\mu^2\theta\xi})\mathcal{E}^{\mathrm{e}}_{s+1}(\hat{w}_{(m+1)})\\
&\qquad + c_\mu \mathcal{E}^{\mathrm{e}}_{s+1}(\hat{w}_{(m+1)})^{1/2}\mathcal{E}_{s+2}(\dot{w}^2_{(m+1)})^{1/2}(\mathcal{E}_{s+1}(\tilde{w})^{1/2}+|\tilde{Z}|+|\tilde{r}|)\\
&\qquad + c_\mu \mathcal{E}^{\mathrm{e}}_{s+1}(\hat{w}_{(m+1)})^{1/2}\mathcal{E}_{s+2}(w^2)^{1/2}(\mathcal{E}_{s+1}(\dot{w}_{(m)})^{1/2}+|\dot{Z}_{(m)}|+|\dot{r}_{(m)}|)(\mathcal{E}_{s+1}(\tilde{w})^{1/2}+|\tilde{Z}|+|\tilde{r}|)\\
&\qquad + c_\mu \mathcal{E}^{\mathrm{e}}_{s+1}(\hat{w}_{(m+1)})^{1/2}\mathcal{E}_{s+2}(\tilde{w})^{1/2}(\mathcal{E}_{s+1}(\dot{w}_{(m)})^{1/2}+|\dot{Z}_{(m)}|+|\dot{r}_{(m)}|)\\
&\qquad + c_\mu \mathcal{E}^{\mathrm{e}}_{s+1}(\hat{w}_{(m+1)})^{1/2}(\mathcal{E}_{s+1}(\tilde{w})^{1/2}+|\tilde{Z}|+|\tilde{r}|)(\mathcal{E}_{s+1}(\dot{w}_{(m)})^{1/2}+|\dot{Z}_{(m)}|+|\dot{r}_{(m)}|),
\end{aligned}$$

which in turn yields the estimate

$$\begin{aligned}
&\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|(\hat{Z}_{(m)},\hat{r}_{(m)},\hat{w}_{(m)})(\xi)\|_{\mathcal{X}^{s+1}}\\
&\qquad \le c_\mu \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]}\Big(\|(\dot{Z}_{(m)},\dot{r}_{(m)},\dot{w}_{(m)})(\xi)\|_{\mathcal{X}^{s+1}}+1\Big) \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|(\tilde{Z},\tilde{r},\tilde{w})(\xi)\|_{\mathcal{X}^{s+2}}
\end{aligned}$$

because

$$\mathcal{E}_{s+2}(\dot{w}^2_{(m+1)}) \le c, \qquad \mathcal{E}_{s+2}(w^2) \le c\mathrm{e}^{-c^\star/2\mu}$$

for $\xi \le \mathrm{e}^{c^\star/2\mu}$. (The first of the above inequalities follows from part (i) of this lemma and Lemma 5.3 with s replaced by $s+1$.) □

Corollary 6.2 *Any solution $(Z_\star, r_\star, w_\star)$ to equations (72)–(75) whose initial data lies on $W^{\mathrm{cs}}_{\mathrm{loc}}$ is differentiable in the topology of $\mathcal{X}^{s+1}$ with respect to Z^0, r_1^0 and r_2^0.*

Proof. Let T be the operator which maps $(\dot{Z}_{(m)},\dot{r}_{(m)},\dot{w}_{(m)})$ to $(\dot{Z}_{(m+1)},\dot{r}_{(m+1)},\dot{w}_{(m+1)})$ in the iteration scheme (95)–(98), which may therefore be written as

$$(\dot{Z}_{(m+1)},\dot{r}_{(m+1)},\dot{w}_{(m+1)}) = T((Z,r,w),(\dot{Z}_{(m)},\dot{r}_{(m)},\dot{w}_{(m)})).$$

Consider the new iteration scheme

$$(\dot{Z}_{(m+1)},\dot{r}_{(m+1)},\dot{w}_{(m+1)}) = T((Z_{(m)},r_{(m)},w_{(m)}),(\dot{Z}_{(m)},\dot{r}_{(m)},\dot{w}_{(m)}))$$

with initial data $\dot{Z}_{(0)}=0$, $\dot{r}_{(0)}=0$, $\dot{w}_{(0)}=0$, which is obtained by differentiating (81)–(84) with respect to Z^0 and using the dot to denote ∂_{Z^0}. Let us write this iteration scheme as

$$(\dot{Z}_{(m+1)},\dot{r}_{(m+1)},\dot{w}_{(m+1)}) = T((Z_\star,r_\star,w_\star),(\dot{Z}_{(m)},\dot{r}_{(m)},\dot{w}_{(m)})) + \alpha_{(m)},$$

where

$$\alpha_{(m)} = T((Z_{(m)}, r_{(m)}, w_{(m)}), (\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})) - T((Z_\star, r_\star, w_\star), (\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})).$$

It follows from Lemma 6.1(i) that $T((R_\star, w_\star), \cdot) : C([0, \mathrm{e}^{c^\star/2\mu}], \mathcal{X}^{s+1}) \to C([0, \mathrm{e}^{c^\star/2\mu}], \mathcal{X}^{s+1})$ is a contraction whenever $\|w_\star(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+2}} \leq \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$, while Lemma 6.1(ii) and Theorem 5.5 with s replaced by $s+1$ show that

$$\begin{aligned}
&\|\alpha_{(m)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+1})} \\
&\quad \leq c_\mu \|(Z_{(m)} - Z_\star, r_{(m)} - r_\star, w_{(m)} - w_\star)\|_{C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+2})} \|(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})\|_{C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+1})} \\
&\qquad\quad + c_\mu \|(Z_{(m)} - Z_\star, r_{(m)} - r_\star, w_{(m)} - w_\star)\|_{C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+2})} \\
&\quad = o(1)\|(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})\|_{C([0,\mathrm{e}^{c^\star/2\mu}],\mathcal{X}^{s+1})} + o(1)
\end{aligned}$$

whenever $\|w_{(m)}(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+3}}, \|w_\star(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+3}} \leq \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$; according to Theorem 5.5 with s replaced by $s+3$ the hypothesis $\|w^0\|_{\mathcal{X}_\mathrm{c}^{s+4}} \leq \mu \mathrm{e}^{-c^\star/2\mu}$ in the definition of $W_\mathrm{loc}^\mathrm{cs}$ guarantees that these conditions are met. Elementary arguments show that $(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)})$ converges in $C([0, \mathrm{e}^{c^\star/2\mu}], \mathcal{X}^{s+1})$ to $(\dot{Z}_\star, \dot{r}_\star, \dot{w}_\star)$. By construction, one has that $(\dot{Z}_{(m)}, \dot{r}_{(m)}, \dot{w}_{(m)}) = (\partial_{Z^0} Z_{(m)}, \partial_{Z^0} r_{(m)}, \partial_{Z^0} w_{(m)})$ for each $m \in \mathbb{N}_0$, and a familiar uniform continuity argument asserts that

$$(\dot{Z}_\star, \dot{r}_\star, \dot{w}_\star) = (\partial_{Z^0} Z_\star, \partial_{Z^0} r_\star, \partial_{Z^0} w_\star).$$

A similar procedure yields the differentiability of $(Z_\star, r_\star, w_\star)$ with respect to r_1^0 and r_2^0. □

We now turn to the requisite estimates on the derivative of J_{ρ,w^0}.

Proposition 6.3

(i) The operator $\mathrm{d}J_{0,0}[0,0,0] : \mathbb{R}^3 \to \mathbb{R}^3$ *is a bijection and*

$$|\mathrm{d}J_{0,0}[0,0,0]^{-1}| \leq \frac{c}{\lambda_{1,\mu}}. \tag{99}$$

(ii) The operator $\mathrm{d}J_{\rho,w^0}[Z^0, r_1^0, r_2^0] : \mathbb{R}^3 \to \mathbb{R}^3$ *satisfies the estimate*

$$|\mathrm{d}J_{\rho,w^0}[Z^0, r_1^0, r_2^0] - \mathrm{d}J_{0,0}[0,0,0]| \leq c\mu^3. \tag{100}$$

Proof. Clearly

$$\begin{aligned}
\partial_1 J_{0,0}(0,0,0) &= \begin{pmatrix} (I - S_\mathrm{sh})\partial_{Z_0} Z_{0,0}(0,0,0)|_{\xi=0} \\ (I - S_\mathrm{wh})\partial_{Z_0} r_{0,0}(0,0,0)|_{\xi=0} \end{pmatrix}, \\
\partial_j J_{0,0}(0,0,0) &= \begin{pmatrix} (I - S_\mathrm{sh})\partial_{r_{j-1}^0} Z_{0,0}(0,0,0)|_{\xi=0} \\ (I - S_\mathrm{wh})\partial_{r_{j-1}^0} r_{0,0}(0,0,0)|_{\xi=0} \end{pmatrix}, \qquad j = 2,3
\end{aligned}$$

and

$$\langle \partial_{Z^0} Z_{0,0}(0,0,0)\tilde{Z}^0|_{\xi=0}, s_0^*\rangle = \tilde{Z}^0, \qquad \langle \partial_{Z^0} r_{0,0}(0,0,0)\tilde{Z}^0|_{\xi=0}, s_0^*\rangle = 0,$$

$$\langle \partial_{r_j^0} Z_{0,0}(0,0,0)\tilde{r}_j^0|_{\xi=0}, s_0^*\rangle = 0, \qquad \langle \partial_{r_j^0} r_{0,0}(0,0,0)\tilde{r}_j{}^0|_{\xi=0}, s_0^*\rangle = 0,$$

in which $\langle\cdot\,,\cdot\rangle$ is the inner product on $\mathcal{X}_\mathrm{h}^{s+1} \simeq \mathbb{R}^6$ and the vectors whose inner products are taken are regarded as elements of $\mathcal{X}_\mathrm{h}^{s+1}$. Using the fact that $S_\mathrm{h} : \mathcal{X}_\mathrm{h}^{s+1} \to \mathcal{X}_\mathrm{h}^{s+1}$ is a self-adjoint involution, we may rewrite the first of the above formulae as

$$\tilde{Z}^0 = \frac{1}{2}\langle (I - S_\mathrm{h})\partial_{Z^0} Z_{0,0}(0,0,0)\tilde{Z}^0|_{\xi=0}, (I + S_\mathrm{h})s_0^*\rangle$$

with corresponding expressions for the others, so that

$$\tilde{Z}^0 = \frac{1}{2}\langle \mathrm{d}J_{0,0}[0,0,0](\tilde{Z}^0, \tilde{r}_1^0, \tilde{r}_2^0), (I + S_\mathrm{h})s_0^*\rangle. \tag{101}$$

A similar argument shows that

$$\tilde{r}_{1,j}^0 = \frac{1}{2}\langle \mathrm{d}J_{0,0}[0,0,0](\tilde{Z}^0, \tilde{r}_1^0, \tilde{r}_2^0), (I + S_\mathrm{h})s_{1,j}^*(0)\rangle, \qquad j = 1, 2, \tag{102}$$

and the first assertion is a direct consequence of (101), (102).

Define $Z^1 = Z_{\rho,w^0}(Z^0, r_1^0, r_2^0)$, $Z^2 = Z_{0,0}(0,0,0)$, $r^1 = r_{\rho,w^0}(Z^0, r_1^0, r_2^0)$, $r^2 = r_{0,0}(0,0,0)$, $\dot{Z}^1 = \mathrm{d}Z_{\rho,w^0}[Z^0, r_1^0, r_2^0]$, $\dot{Z}^2 = \mathrm{d}Z_{0,0}[0,0,0]$ and $\dot{r}^1 = \mathrm{d}r_{\rho,w^0}[Z^0, r_1^0, r_2^0]$, $\dot{r}^2 = \mathrm{d}r_{0,0}[0,0,0]$. By construction we have that

$$\begin{aligned}
&(\dot{Z}_1 - \dot{Z}_2)(\xi) \\
&\quad = \int_0^\xi \langle (\partial_1 F_0^\mu(Z_1, p^\mu + r_1, w_1)\dot{Z}_1 - \partial_1 F_0^\mu(0, p^\mu, 0)\dot{Z}_2 \\
&\qquad\qquad + \partial_2 F_0^\mu(Z_1, p^\mu + r_1, w_1)\dot{r}_1 - \partial_2 F_0^\mu(0, p^\mu, 0)\dot{r}_2 \\
&\qquad\qquad + \partial_3 F_0^\mu(Z_1, p^\mu + r_1, w_1)\dot{w}_1 - \partial_3 F_0^\mu(0, p^\mu, 0)\dot{w}_2 \\
&\qquad\qquad + \rho\partial F_1^\mu(p^\mu + r_1)\dot{r}_1 - \rho\partial F_1^\mu(p^\mu)\dot{r}_1)(\tau), s_0^* \mathrm{e}^{\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, s_0 \mathrm{e}^{-\lambda_{0,\mu}\xi} \\
&\qquad - \int_\xi^{\mathrm{e}^{c^\star/2\mu}} \langle (\partial_1 F_0^\mu(Z_1, p^\mu + r_1, w_1)\dot{Z}_1 - \partial_1 F_0^\mu(0, p^\mu, 0)\dot{Z}_2 \\
&\qquad\qquad + \partial_2 F_0^\mu(Z_1, p^\mu + r_1, w_1)\dot{r}_1 - \partial_2 F_0^\mu(0, p^\mu, 0)\dot{r}_2 \\
&\qquad\qquad + \partial_3 F_0^\mu(Z_1, p^\mu + r_1, w_1)\dot{w}_1 - \partial_3 F_0^\mu(0, p^\mu, 0)\dot{w}_2 \\
&\qquad\qquad + \rho\partial F_1^\mu(p^\mu + r_1)\dot{r}_1 - \rho\partial F_1^\mu(p^\mu)\dot{r}_1)(\tau), u_0^* \mathrm{e}^{-\lambda_{0,\mu}\tau}\rangle \,\mathrm{d}\tau\, u_0 \mathrm{e}^{\lambda_{0,\mu}\xi},
\end{aligned}$$

whence

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |(\dot{Z}_1 - \dot{Z}_2)(\xi)| \le c\mu^3 \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \sum_{j=1}^2 (|\dot{Z}_j(\xi)| + |\dot{r}_j(\xi)| + \|\dot{w}_j(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+1}}).$$

A similar argument yields

$$\sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} |(\dot{r}_1 - \dot{r}_2)(\xi)| \le c\frac{\mathrm{e}^{-c^\star/2\mu}}{\mu^2\lambda_{1,\mu}^2} \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \sum_{j=1}^2 (|\dot{Z}_j(\xi)| + |\dot{r}_j(\xi)| + \|\dot{w}_j(\xi)\|_{\mathcal{X}_\mathrm{c}^{s+1}}),$$

and these inequalities imply the second assertion. □

We now study the solution set of the equation

$$J_{\rho,w^0}(Z^0,r_1^0,r_2^0)=0$$

near the known solution $(Z^0,r_1^0,r_2^0)=(0,0,0)$ at $(\rho,w^0)=(0,0)$ by writing it as

$$(Z^0,r_1^0,r_2^0)=(Z^0,r_1^0,r_2^0)-\mathrm{d}J_{0,0}[0,0,0]^{-1}J_{\rho,w^0}(Z^0,r_1^0,r_2^0) \tag{103}$$

and examining this fixed point problem. According to a standard argument in nonlinear analysis the fixed-point problem (103) has a unique solution $(Z^0,r_1^0,r_2^0)=(Z^0,r_1^0,r_2^0)(\rho,w^0)$ in $\bar{B}_\eta(0)\subset\mathbb{R}^3$ whenever

$$\begin{aligned}|\mathrm{d}J_{0,0}[0,0,0]^{-1}||J_{\rho,w^0}(0,0,0)| &\le \frac{\eta}{2},\\ |\mathrm{d}J_{0,0}[0,0,0]^{-1}||\mathrm{d}J_{\rho,w^0}[Z^0,r_1^0,r_2^0]-\mathrm{d}J_{0,0}[0,0,0]| &\le \frac{1}{2}, \qquad (Z^0,r_1^0,r_2^0)\in\bar{B}_\eta(0).\end{aligned}$$

The estimates (99), (100) and

$$|J_{\rho,w^0}(0,0,0)|\ \le\ c\left|\begin{pmatrix} Z_{\rho,w^0}(0,0,0)|_{\xi=0}\\ r_{\rho,w^0}(0,0,0)|_{\xi=0}\end{pmatrix}\right|\ \le\ c\mu^3\mathrm{e}^{-c^\star/2\mu}$$

(see formulae (85), (86)) show that we can take $\eta=\mu^2\mathrm{e}^{-c^\star/2\mu}$.

We have therefore constructed a family of symmetric solutions $(Z_{w^0},r_{w^0},w_{w^0})$ to (72)–(75) on $[-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}]$ which are parameterised by $w^0\in\Sigma_\mathrm{c}$ with $\|w^0\|_{\mathcal{X}_\mathrm{c}^{s+4}}\le\mu\mathrm{e}^{-c^\star/2\mu}$ and satisfy $\|(Z_{w^0}(\xi),r_{w^0}(\xi),w_{w^0}(\xi))\|_{\mathcal{X}^{s+1}}\le\mathrm{e}^{-c^\star/2\mu}$ for each $\xi\in[-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}]$. The formula

$$z_{w^0}(\xi)=p^\mu(\xi)+r_{w^0}(\xi), \qquad \xi\in[-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}]$$

defines a family of modulating pulse solutions to the coupled system (65)–(67) which was obtained from the original spatial dynamics formulation of the problem by the normal-form theory in Section 4; these solutions are parameterised by $w^0\in\Sigma_\mathrm{c}$, that is by $w_{1,\mathrm{e}}^0$ and $w_{2,\mathrm{o}}^0$. Notice that $p^\mu(0),r_{w^0}(0)\in\Sigma_\mathrm{wh}$, so that $z_{w^0}(0)\in\Sigma_\mathrm{wh}$, and by construction $Z_{w^0}(0)\in\Sigma_\mathrm{sh}$, $w_{w^0}(0)\in\Sigma_\mathrm{c}$. The existence result for modulating pulses stated in Section 1 (Theorem 1.1) follows by tracing the coordinate transformations back to the original variable $v(\xi,\eta)$ and replacing μ with $\sqrt{\varepsilon}$.

References

[1] BIRNIR, B., MCKEAN, H. & WEINSTEIN, A. 1994 The rigidity of sine-Gordon breathers. *Commun. Pure Appl. Math.* **47**, 1043–1051.

[2] DENZLER, J. 1993 Nonpersistence of breather families for the perturbed sine Gordon equation. *Commun. Math. Phys.* **158**, 397–430.

[3] GROVES, M. D. & MIELKE, A. 2001 A spatial dynamics approach to three-dimensional gravity-capillary steady water waves. *Proc. Roy. Soc. Edin. A* **131**, 83–136.

[4] GROVES, M. D. & SCHNEIDER, G. 2001 Modulating pulse solutions for a class of nonlinear wave equations. *Commun. Math. Phys.* **219**, 489–522.

[5] GROVES, M. D. & SCHNEIDER, G. 2005 Modulating pulse solutions for quasilinear wave equations. *J. Diff. Eqns.* **219**, 221–258.

[6] IOOSS, G. & LOMBARDI, E. 2004 Normal forms with exponentially small remainder: application to homoclinic connections for the reversible $0^2\mathrm{i}\omega$ resonance. *C. R. Acad. Sci. Paris, Sér. 1* **339**, 831–838.

[7] IOOSS, G. & LOMBARDI, E. 2005 Polynomial normal forms with exponentially small remainder for analytic vector fields. *J. Diff. Eqns.* **212**, 1–61.

[8] KALYAKIN, L. A. 1988 Asymptotic decay of a one-dimensional wave packet in a nonlinear dispersive medium. *Mat. Sb. (N.S.)* **132(174)**, 470–495. (English translation *Math. USSR-Sb.* **60**, 457–483.)

[9] KATO, T. 1975 Quasi-linear equations of evolution, with applications to partial differential equations. In *Lecture Notes in Mathematics* **448** — *Spectral Theory and Differental Equations, Dundee 1974*, pages 25–70. Berlin: Springer-Verlag.

[10] MCLAUGHLIN, D. W. & SHATAH, J. 1998 Homoclinic orbits for PDEs. *AMS Proc. Symp. Pure Math.* **54**, 281–299.

[11] PÖSCHEL, J. 1998 Nonlinear partial differential equations, Birkhoff normal forms and KAM theory. In *European Congress of Mathematics, Vol. II (Budapest 1996). Prog. Math.* **169**, 167–186.

[12] SCHNEIDER, G. 1998 Justification of modulation equations for hyperbolic systems via normal forms. *Nonlinear Differential Equations and Applications (NODEA)* **5**, 69–82.

[13] SHATAH, J. 1985 Normal forms and quadratic nonlinear Klein-Gordon equations. *Commun. Pure Appl. Math.* **38**, 685–696.