Pulse solutions to a class of quasilinear symmetrisable hyperbolic systems with applications to nonlinear wave equations

A dissertation submitted towards the degree

Doctor of Natural Sciences

of the Faculty of Mathematics and Computer Science of Saarland University

by

Nils Gutheil

Saarbrücken, 2025

Datum des Kolloquiums - Date of Defense

April 30, 2026

Dekan - Dean

Prof. Dr. Jan Reineke

Prüfungsausschuss - Examination Board

Prof. Dr. Michael Bildhauer
(Vorsitzender - Chairman)

Prof. Dr. Mark Groves
(Betreuer - Thesis supervisor)

Prof. Sylvie Benzoni-Gavage
(Gutachterin - Reviewer)

Prof. Tom Bridges
(Gutachter - Reviewer)

Dr. Christian Steinhart
(Akademischer Mitarbeiter - Academic Staff)

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Declaration of original authorship

I hereby declare that this dissertation is my own original work except where otherwise indicated. All data or concepts drawn directly or indirectly from other sources have been correctly acknowledged. This dissertation has not been submitted in its present or similar form to any other academic institution either in Germany or abroad for the award of any other degree.

Saarbrücken, ***Dezember 2025***

gez. / signed

(*Nils Gutheil*)

Abstract

In this thesis we present an existence theory for pulse solutions to a class of reversible, analytic, hyperbolic evolutionary equations of the form

$$v_\xi = A^\varepsilon(v)v_\tau + B^\varepsilon(v), \tag{1}$$

where $v(\xi,\tau) \in \mathbb{R}^n$ is 2π-periodic in its second argument. Combining methods from the theory of (infinite-dimensional) dynamical systems (in particular normal forms and invariant manifolds) and symmetrisable, quasilinear hyperbolic systems, we construct symmetric generalised pulse solutions satisfying

$$v(\xi,\tau+2\pi) = v(\xi,\tau), \quad |v(\xi,\tau) - p^\varepsilon(\xi,\tau)| \le \mathrm{e}^{-c^\star/2\sqrt{\varepsilon}} \tag{2}$$

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

$$p^\varepsilon(\xi,\tau) \sim \pm\varepsilon\,\mathrm{sech}(\varepsilon\xi)u^\star\cos(\tau),$$

$c^\star > 0$ and $u^\star$ is a constant vector. In the special case where $A^\varepsilon(v)$ is a constant function of v and the system (1) admits a suitable conserved quantity, the estimate in (2) holds for all $\xi \in \mathbb{R}$. As an application of our theory we discuss how a wide class of quadratic, quasilinear Klein-Gordon equations fits into the above framework.

Zuammenfassung

In dieser Arbeit präsentieren wir eine Existenztheorie für Puls Lösungen zu einer Klasse reversibler, analytischer, hyperbolischer Evolutionsgleichungen der Form

$$v_\xi = A^\varepsilon(v)v_\tau + B^\varepsilon(v), \tag{1}$$

wobei $v(\xi,\tau) \in \mathbb{R}^n$ 2π-periodisch im zweiten Argument ist. Methoden aus der Theorie der (unendlichdimesionalen) dynamischen Systeme (insbesondere Normalformen und invariante Mannigfaltigkeiten) und der symmetrisierbaren, quasilinearen hyperbolischen Systeme kombinierend, konstruieren wir verallgemeinerte Puls Lösungen mit

$$v(\xi,\tau+2\pi) = v(\xi,\tau), \quad |v(\xi,\tau) - p^\varepsilon(\xi,\tau)| \le \mathrm{e}^{-c^\star/2\sqrt{\varepsilon}} \tag{2}$$

für alle $\tau \in \mathbb{R}$ und $\xi \in [-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]$, wobei

$$p^\varepsilon(\xi,\tau) \sim \pm\varepsilon\,\mathrm{sech}(\varepsilon\xi)u^\star\cos(\tau),$$

$c^\star > 0$ und $u^\star$ ein konstanter Vektor ist. Im Spezialfall wo $A^\varepsilon(v)$ eine konstante Funktion von v ist und das System (1) eine geeignete Erhaltungsgröße hat, gilt die Abschätzung in (2) für alle $\xi \in \mathbb{R}$. Als Anwendung unserer Theorie diskutieren wir wie eine weite Klasse quadratischer, quasilinearer Klein-Gordon Gleichungen in die obige Struktur passt.

Danksagung

Zuerst möchte ich mich bei meinem Betreuer Prof. Mark Groves bedanken, bei dem ich nun unglaubliche 13 Jahre Student und Doktorand sein durfte. Als ich im Wintersemester 2012 in der ersten Vorlesung und im Anschluss verzweifelt am ersten Übungsblatt Analysis 1 gesessen habe, hätte ich mir nicht erträumt, dass mein Studium im Jahr 2025 mit einer Promotion am Lehrstuhl Groves endet. Ich bedanke mich für eine außergewöhnliche Betreuung, die auch vor Projekten wie dem ersten Halbmarathon oder großartigen Forschungsreisen um die ganze Welt keinen Halt machte. Eine Betreuung, die hohes mathematisches Niveau, Gewissenhaftigkeit, Akribie und den Blick für Konsistenz zum einen gefordert, zum anderen gefördert hat, hat mich zu dem Studenten und wissenschaftlichen Mitarbeiter gemacht der ich heute bin. Unsere gemeinsame Zeit war geprägt von lockerem Humor und einer Prise Sportsgeist (natürlich niemals kompetitiv), aber definitiv nie von Langeweile. Unzählige Stunden, Tage, gar Nächte voller intensiver Betreuungsgespräche haben mich nie den Glauben verlieren lassen diese Arbeit beenden zu können. Ohne einen so hilfsbereiten Betreuer, der immer den Blick für das große Ganze hatte, wenn ich ihn mal verloren hatte, wäre dies nicht möglich gewesen. Lieber Mark, den Satz “Ich lasse euch nicht im Stich” hast du bis zuletzt vorgelebt und dafür danke ich dir von Herzen. Es war mir eine Ehre!

Ich möchte mich bedanken bei Herrn Prof. Michael Bildhauer, der durch seine freundliche, grenzenlos hilfsbereite und pragmatische Art über Jahre dazu beigetragen hat, dass ich meinen Weg als festangestellter wissenschaftlicher Mitarbeiter in der Studienkoordination weitergehen darf. Meine Idee von Lehre und deren Organisation wurde durch unsere fast sechsjährige Zusammenarbeit in den Höhere Mathematik Vorlesungen, deren AdressatInnenkreis mittlerweile weit über die Ingenieurswissenschaften hinausgeht, entscheidend geprägt. Eine Meinungsverschiedenheit hatten wir nie, das gemeinsame Konzipieren von Übungsblättern und Klausuren und nicht zuletzt die äußerst fruchtbaren Ideenaustausche am Whiteboard empfinde ich als Privileg und große Freude. Ich freue mich auf hoffentlich noch viele weitere gemeinsame Lehrveranstaltungen! Darüber hinaus muss ich mich ausdrücklich bedanken für die Hilfestellung in allen Verwaltungsfragen. Ob Telefonate am Wochenende, das Bereinigen einer Studiengangsdefinitionstabelle, Verträge, eine bloße Unterschrift oder das Erkunden des Prüfungsverwaltungssystems: wenn ich Hilfe oder den richtigen Kontakt brauche, weiss ich wo ich sie finde. Einen solchen Geschäftsführer kann sich jede Fachrichtung nur wünschen!

Bedanken möchte ich mich bei meinem langjährigen Wegbegleiter, Kollegen und besten Freund Rami Ahmad. Gemeinsam haben wir in den letzten 13 Jahren eine Reise unternommen wie sie kein Drehbuch hätte vorschreiben können. Von der zur Abiturzeit entwickelten Idee Mathematik zu studieren bis zur gemeinsamen Festanstellung am heutigen Tag haben wir jede Hürde, die das Studium und unsere Tätigkeit als wissenschaftliche Mitarbeiter mit sich brachten, zusammen bewältigt. In jeder Vorlesung, Übung, Klausur, bei jeder Tutorentätigkeit und während der Bachelor- und Masterarbeit seinen Partner im Büro neben sich zu wissen, auf den man sich verlassen kann und der immer ein offenes Ohr hat, hat mich durch das Studium getragen, das ich alleine nicht geschafft hätte. Dafür danke ich dir. Auf viele weitere Jahre Freundschaft und Zusammenarbeit, unser Drehbuch hält noch unendlich viele Seiten parat!

Ich bedanke mich bei Benedikt Hewer, der mir seit Anbeginn meines Studiums als Tutor, Kommilitone, Kollege und Freund immer mit Rat und Tat zur Seite steht. Sei es Kombinatorik, IT oder LaTeX, immer konnte ich bei dir Hilfe finden. Unsere stets amüsanten Gespräche über gemeinsame Erinnerungen und "Highlights", die wir in der Fachrichtung erleben durften und als entfristete wissenschaftliche Mitarbeiter weiter dürfen, füllen Abende. Dafür bin ich dankbar und freue mich darauf, dass künftig noch viele dazu kommen werden!

I would like to thank my colleague Dan Hill, who helped me find my way during my first research stay in England. As a former member of our working group you always had clever advice for me on all mathematical and linguistic questions. Apart from that you taught me what a 'parliament of owls' is, among other things. Thank you for three and a half great years together in our working group in Saarbrücken!

Ich bedanke mich bei Christian Steinhart, der mit seiner heiteren und humorvollen Persönlichkeit immer gute Laune an meinen Schreibtisch gebracht hat, wenn es notwendig war. Danke dass du mein Portfolio an mathematischen Witzen erweitert hast, auch wenn wir den Humor diesbezüglich nicht immer teilen. Die Schildkröte wird mir im Gedächtnis bleiben. So schlimm außerdem das Laufen ist, wenn du dabei bist ist es erträglicher. Danke dafür!

Danke an meinen Kollegen Johannes Hoffmann, der mit seiner besonnenen und aufgeschlossenen Art eine Bereicherung für jedes Gespräch ist und immer eine schlaue Idee hat, wenn es gerade nötig ist. Danke für den Moodle-Support und das spontane Einrichten der Task-Force LaTeX gemeinsam mit Benedikt Hewer. Das Lineal am Bildschirm wird in guter Erinnerung bleiben!

Ich bedanke mich bei meinem Freund und ehemaligen Tutor Daniel Kraemer, dessen akribische und gewissenhafte Arbeitsweise mich während meines Studiums inspiriert hat. Vielen Dank für deine Übungsblätter, Materialien von solcher Qualität haben mir meinen Einstieg in die Assistenz für die Höhere Mathematik Vorlesungen ungemein erleichtert!

Vielen Dank an meinen ehemaligen Kommilitonen Felix Leid, mit dem ich Tage und Nächte meines Grundstudiums verbracht habe. In der gemeinsamen Zeit im CIP-Pool eine mentale Stütze und einen verlässlichen Übungspartner zu haben, mit dem man außerdem eine gepflegte Runde K-Blocks spielen konnte, hat das Studium studierenswerter gemacht. Danke dafür!

Ich bedanke mich bei meiner Frau Anna Gutheil für unzählige Tage, Abende, Nächte, Wochenenden, die sie ohne mich auskommen musste. Ich bedanke mich dafür dass du mich und meine Launen ertragen hast in den schwierigen Phasen die das Fertigstellen dieser Arbeit so oft mit sich brachte. Die gemeinsame Zeit macht diese Phasen schnell vergessen, vielen Dank dass du mich immer aufgebaut und erheitert hast, wenn es mal wieder nötig war. So oft hast du mir den Rücken freigehalten und mich zuhause unterstützt. Auch dank dir konnte ich diese Arbeit überhaupt fertigstellen!

Ich danke meiner Familie, die mir während meines gesamten Studiums immer ein sicherer Hafen gewesen ist und mich unterstützt und bestärkt hat. Bei allen Projekten in meinem Leben konnte ich mir eurer Unterstützung sicher sein und ohne ein solches familiäres Umfeld wäre ich nicht in der Lage gewesen, mein Studium und meine Promotion zu beenden. Dafür bin ich dankbar!

Special thanks go to my proofreaders Johannes Hoffmann, Dan Hill, Benedikt Hewer and Daniel Kraemer for their many great suggestions and comments which have contributed significantly to improving this work!

Contents

1 Introduction

1.1 Homoclinic solutions and pulses

A *homoclinic solution* to a dynamical system

$$u_\xi = f(u)$$

for $u\colon \mathbb{R} \to \mathbb{R}^n$ is a nontrivial solution which satisfies $|u(\xi)| \to 0$ as $|\xi| \to \infty$. The canonical example of a two-dimensional system with a cubic nonlinearity admitting homoclinic solutions is

$$x_{1\xi} = y_1, \tag{1.1}$$

$$y_{1\xi} = x_1 - \check{C}x_1^3, \tag{1.2}$$

where $\check{C} > 0$. These homoclinic solutions are given by

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \pm \begin{pmatrix} H \\ H_\xi \end{pmatrix},$$

where

$$H(\xi) = \sqrt{\frac{2}{\check{C}}}\,\mathrm{sech}(\xi).$$

The dynamics of equations (1.1), (1.2) are shown in the phase portrait in Figure 1.1. The four-dimensional system

$$x_{1\xi} = y_1, \tag{1.3}$$

$$y_{1\xi} = x_1 - \check{C}x_1(x_1^2 + x_2^2), \tag{1.4}$$

$$x_{2\xi} = y_2, \tag{1.5}$$

$$y_{2\xi} = x_2 - \check{C}x_2(x_1^2 + x_2^2), \tag{1.6}$$

also has homoclinic solutions since it admits the invariant plane $\{(x_1, y_1)\}$, the flow in which is governed by equations (1.1), (1.2).

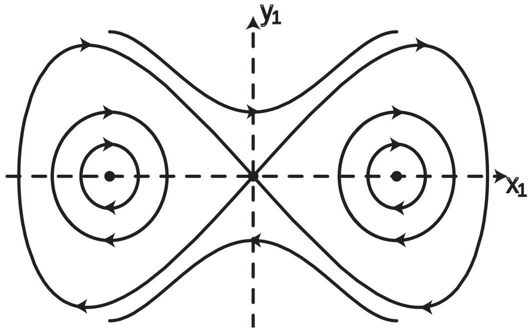

Figure 1.1: Phase portrait of the invariant plane $\{(x_1, y_1)\}$ of equations (1.3)–(1.6) with two homoclinic solutions.

Perturbations of equations (1.3)–(1.6) of the form

$$\dot{x}_1 = y_1 + \varepsilon \mathcal{R}^{\varepsilon}_{1,1}(x_1, y_1, x_2, y_2), \tag{1.7}$$

$$\dot{y}_1 = x_1 - \check{C} x_1 (x_1^2 + x_2^2) + \varepsilon \mathcal{R}^{\varepsilon}_{1,2}(x_1, y_1, x_2, y_2), \tag{1.8}$$

$$\dot{x}_2 = y_2 + \varepsilon \mathcal{R}^{\varepsilon}_{2,1}(x_1, y_1, x_2, y_2), \tag{1.9}$$

$$\dot{y}_2 = x_2 - \check{C} x_2 (x_1^2 + x_2^2) + \varepsilon \mathcal{R}^{\varepsilon}_{2,2}(x_1, y_1, x_2, y_2), \tag{1.10}$$

where

$$R^{\varepsilon}_{i,j}(x_1, y_1, x_2, y_2) = \mathcal{O}(|(x_1, y_1, x_2, y_2)|^3)$$

and ε is a small parameter, in general do not have homoclinic solutions. Each point on a homoclinic solution is a point of intersection of the stable and unstable manifolds $W^{\varepsilon}_{\mathrm{s}}$ and $W^{\varepsilon}_{\mathrm{u}}$ which are both two-dimensional ($\dim W^{\varepsilon}_{\mathrm{s}}$ and $\dim W^{\varepsilon}_{u}$ are the number of eigenvalues of the linear operator on the right-hand side of equations (1.7)–(1.10) with respectively negative and positive real part), and in general two two-dimensional manifolds do not intersect in $\mathbb{R}^4$. If we additionally assume that equations (1.7)–(1.10) are *reversible*, i.e. invariant under the transformation $\xi \mapsto -\xi$, $(x_1, y_1, x_2, y_2) \mapsto (x_1, -y_1, -x_2, y_2)$, we are still able to construct symmetric homoclinic solutions by showing that the stable manifold $W^{\varepsilon}_{\mathrm{s}}$ intersects the *symmetric section* $\{y_1 = 0, x_2 = 0\}$. Indeed

$$W^0_{\mathrm{s}} = \left\{ \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = R_\theta \begin{pmatrix} H(\xi) \\ 0 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = R_\theta \begin{pmatrix} H_\xi(\xi) \\ 0 \end{pmatrix}, \ \xi \in \mathbb{R}, \ \theta \in [0, 2\pi) \right\},$$

where R_θ is a rotation through θ, and this manifold intersects the symmetric section transversally at the points $(\pm H(0), 0, 0, 0)^\top$. We thus obtain the existence of two symmetric homoclinic solutions for small positive values of ε (see Lemma 2.13 for a functional-analytic proof of this result).

Equations (1.7)–(1.10) arise from the system given in complex coordinates by

$$z_{1\xi} = \mathrm{i}(\omega + \nu^{\varepsilon}) z_1 + z_2 + \dots, \tag{1.11}$$

$$z_{2\xi} = (\lambda^{\varepsilon})^2 z_1 + \mathrm{i}(\omega + \nu^{\varepsilon}) z_2 - \check{C} z_2 |z_1|^2 + \dots, \tag{1.12}$$

where $\lambda^\varepsilon = \varepsilon + \mathcal{O}(\varepsilon^2)$, $\nu^\varepsilon = \mathcal{O}(\varepsilon)$, by the scaling

$$\check{\xi} = \varepsilon\xi, \qquad \begin{pmatrix} z_1(\xi) \\ z_2(\xi) \end{pmatrix} = \mathrm{e}^{\mathrm{i}(\omega+\nu^\varepsilon)\xi} \begin{pmatrix} \lambda^\varepsilon\big(\check{x}_1(\check{\xi}) + \mathrm{i}\check{x}_2(\check{\xi})\big) \\ (\lambda^\varepsilon)^2\big(\check{y}_1(\check{\xi}) + \mathrm{i}\check{y}_2(\check{\xi})\big) \end{pmatrix}.$$

This system exhibits homoclinic bifurcation associated with a reversible $(\mathrm{i}\omega)^2$ *resonance*: as $\varepsilon \downarrow 0$ two pairs of simple complex eigenvalues collide to form geometrically simple and algebraically double eigenvalues $\pm\mathrm{i}\omega$ at $\varepsilon = 0$ (see Figure 1.2).

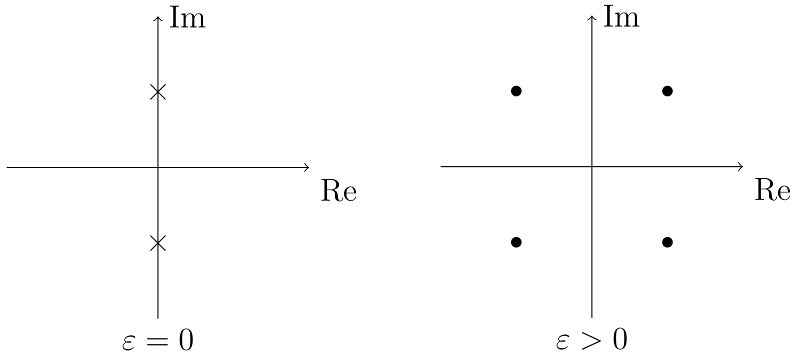

Figure 1.2: A reversible $(\mathrm{i}\omega)^2$ resonance takes place at $\varepsilon = 0$.

It is convenient to write the system (1.11), (1.12) as

$$z_\xi = L^\varepsilon_{\mathrm{wh}} z + f^\varepsilon(z), \tag{1.13}$$

where $z \in \mathbb{C}^4$, $L^\varepsilon_{\mathrm{wh}}$ is the matrix associated with the reversible $(\mathrm{i}\omega)^2$ resonance and $f^\varepsilon(z)$ is the (parameter-dependent) nonlinearity. Let us now couple equation (1.13) with a second equation to arrive at the system

$$z_\xi = L^\varepsilon_{\mathrm{wh}} z + f^\varepsilon_1(z) + f^\varepsilon_2(z, w), \tag{1.14}$$
$$w_\xi = L^\varepsilon_{\mathrm{c}} w + g^\varepsilon(z, w) + h^\varepsilon(z), \tag{1.15}$$

where $w \in \mathbb{R}^{2d}$ and $L^\varepsilon_{\mathrm{c}}$ is a matrix with d pairs of purely imaginary eigenvalues $\pm\mathrm{i}\nu^\varepsilon_1, \ldots, \pm\mathrm{i}\nu^\varepsilon_d$ (see Figure 1.3), and $f^\varepsilon_1(z)$, $f^\varepsilon_2(z, w)$, $g^\varepsilon(z, w)$, $h^\varepsilon(z)$ are nonlinearities with $f^\varepsilon_2(z, 0) = 0$, $g^\varepsilon(z, 0) = 0$. In general the coupled system (1.14), (1.15) does not have homoclinic solutions, since $W^\varepsilon_{\mathrm{s}}$ and $W^\varepsilon_{\mathrm{u}}$ are both two-dimensional, and in general two two-dimensional manifolds do not intersect in $\mathbb{R}^{4+2d}$.

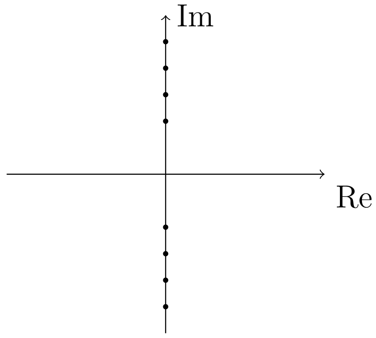

Figure 1.3: The spectrum of L_c^ε consists of d pairs of purely imaginary eigenvalues.

A strategy for treating equations of the form (1.14), (1.15) was devised in a series of papers by Groves & Schneider [8, 9, 10]. In the special case that $h^\varepsilon = 0$ the coupled system (1.14), (1.15) has the invariant subspace $\{w = 0\}$, the flow in which is determined by (1.13), and one can find a homoclinic solution p^ε to (1.13), and hence a homoclinic solution $(p^\varepsilon, 0)$ to (1.14), (1.15), using the methods described above. Groves & Schneider showed that (1.14), (1.15) has reversible solutions (z, w) which satisfy

$$|(z(\xi) - p^\varepsilon(\xi), w(\xi))| = O(\varepsilon^N), \quad \xi \in [-\varepsilon^{-N}, \varepsilon^{-N}],$$

if $h^\varepsilon(z) = O(|z|^N)$. Furthermore, a sequence of near-identity 'normal-form' transformations can be used to successively remove terms in the Maclaurin expansion of $h^\varepsilon(z)$ and hence achieve the condition $h^\varepsilon(z) = O(|z|^N)$.

This result is considerably improved if the nonlinearities are analytic: one can find an optimal value of N so that $h^\varepsilon(z)$ is exponentially small, that is $h^\varepsilon(z) = O(\mathrm{e}^{-c^\star/\sqrt{\varepsilon}})$, for values of (z, ε) in an appropriately chosen neighbourhood of the origin. In this case (1.14), (1.15) has reversible solutions (z, w) which satisfy

$$|(z(\xi) - p^\varepsilon(\xi), w(\xi))| = O(\mathrm{e}^{-c^\star/2\sqrt{\varepsilon}}), \quad \xi \in [-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}].$$

A different improvement is achieved if (1.14), (1.15) has a conserved quantity which serves as a Lyapunov functional in a suitable sense. In this setting (1.14), (1.15) has reversible solutions which satisfy

$$|(z(\xi) - p^\varepsilon(\xi), w(\xi))| = O(\varepsilon^N), \quad \xi \in \mathbb{R}; \tag{1.16}$$

note however that $p^\varepsilon(\xi)$ does not necessarily tend to zero as $|\xi| \to \infty$. We refer to these solutions as (*generalised*) *pulse solutions* since they track homoclinic solutions over large intervals of values for ξ.

Groves & Schneider studied systems in which w has infinitely many components, concentrating in particular on wave equations. In this thesis we show how their method applies to general analytic, symmetrisable, quasilinear, hyperbolic systems of equations and improve the estimate (1.16) to

$$|(z(\xi) - p^\varepsilon(\xi), w(\xi))| = O(\mathrm{e}^{-c^\star\sqrt{\varepsilon}}), \quad \xi \in \mathbb{R},$$

if the system is semilinear and has a Lyapunov functional. The thesis builds upon preliminary results for quasilinear systems in the author's BSc and MSc theses [11, 12] and makes heavy use of the theory of analytic functions presented by Buffoni and Toland [4, Chapter 4], to which we defer for all definitions and basic facts. Throughout the subsequent theory we use the notation

$$\|\mathcal{R}(x_1,\dots,x_l)\|_Y = \mathcal{O}(\|x_1\|_{X_1}^{j_1}\cdot\dots\cdot\|x_l\|_{X_l}^{j_l})$$

for a function $\mathcal{R}\colon X_1\times\dots\times X_l\to Y$ between Banach spaces $X_1,\dots,X_l,Y$ which is analytic at the origin with a Maclaurin series whose leading-order term is homogeneous of degree at least $j_1,\dots,j_l$ in respectively $x_1,\dots,x_l$.

1.2 The main results

Framework

In this thesis we construct generalised pulse solutions to a class of nonlinear hyperbolic evolutionary equations of the form

$$v_\xi = A^\varepsilon(v)v_\tau + B^\varepsilon(v) \tag{1.17}$$

where $v(\xi,\tau)\in\mathbb{R}^n$ is 2π-periodic in its second argument and A and B are respectively matrix- and vector-valued functions of (ε,v) defined in a neighbourhood of the origin in $\mathbb{R}\times\mathbb{R}^n$. We refer to equation (1.17) as being *quasilinear* if the matrix-valued function A^ε is non-constant and as being *semilinear* if A^ε is a constant function, that is $A^\varepsilon(x)=A^\varepsilon(0)$ for each $x\in\mathbb{R}^n$. In the latter case we just write A^ε instead of $A^\varepsilon(0)$. The hypotheses on the functions A and B are as follows.

Assumptions 1.

(i) The matrix-valued function $A\colon\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^{n\times n}$ is analytic at the origin.

The function $B\colon\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$ is analytic at the origin with $B^\varepsilon(0)=0$.

Both A and B are even functions of ε.

(ii) The matrix $A^\varepsilon(0)$ is strictly hyperbolic, that is it has n distinct real eigenvalues.

(iii) There exists a matrix-valued function $S\colon\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^{n\times n}$ such that S is analytic at the origin and $S^\varepsilon(x)$, $S^\varepsilon(x)A^\varepsilon(x)$ are symmetric with

$$|y|^2 \lesssim \langle S^\varepsilon(x)y,y\rangle \lesssim |y|^2, \quad y\in\mathbb{R}^n,$$

uniformly over (ε,x) in a neighbourhood of the origin in $\mathbb{R}\times\mathbb{R}^n$.

This assumption is not needed if A^ε is constant. In this case we set

$$S^\varepsilon(x)=S^\varepsilon(0)=I$$

for notational simplicity below.

(iv) There exists a matrix $T \in \mathbb{R}^{n\times n}$ such that $T^2 = I$ and

$$A^\varepsilon(Tx)T = TA^\varepsilon(x), \quad B^\varepsilon(Tx) = -TB^\varepsilon(x), \quad S^\varepsilon(0)T = TS^\varepsilon(0).$$

The existence of the matrix T guarantees that equation (1.17) is *reversible*, that is invariant under the transformation $\xi \mapsto -\xi$, $v \mapsto Rv$, where the *reverser* R is defined by the formula

$$(Rv)(\tau) = Tv(-\tau).$$

We begin with the change of variable $u = S^\varepsilon(0)^{\frac{1}{2}}v$ in equation (1.17) to obtain

$$u_\xi = \tilde{A}^\varepsilon(u)u_\tau + \tilde{B}^\varepsilon(u),$$

where

$$\tilde{A}^\varepsilon(u) = S^\varepsilon(0)^{\frac{1}{2}}A^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u)S^\varepsilon(0)^{-\frac{1}{2}}, \quad \tilde{B}^\varepsilon(u) = S^\varepsilon(0)^{\frac{1}{2}}B^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u);$$

note that $\tilde{A}$, $\tilde{B}$ satisfy Assumptions 1(i), 1(ii) and 1(iv) and $\tilde{A}^\varepsilon(0)$ is symmetric. Define

$$\mathcal{X}^t = H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C}^n)$$

for arbitrary $t \geq 0$, so that in particular the linear operator

$$L^\varepsilon = \tilde{A}^\varepsilon(0)\partial_\tau + \partial\tilde{B}^\varepsilon[0]$$

lies in $\mathcal{L}(\mathcal{X}^{t+1}, \mathcal{X}^t)$. Note that a function $u \in \mathcal{X}^t$ admits a Fourier series expansion

$$u(\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}} u_m \mathrm{e}^{\mathrm{i}m\tau}, \tag{1.18}$$

where the coefficient vectors $u_m \in \mathbb{C}^n$ satisfy the reality condition $\overline{u}_m = u_{-m}$ for each $m \in \mathbb{Z}$ if u is real-valued.

It follows from the expansion (1.18) that

$$(L^\varepsilon u)(\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}} L^\varepsilon_m(u_m\mathrm{e}^{\mathrm{i}m\tau}) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}}\Big(\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]\Big)u_m\mathrm{e}^{\mathrm{i}m\tau},$$

and we have that

$$\sigma(L^\varepsilon) = \bigcup_{m\in\mathbb{Z}}\sigma\Big(\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]\Big),$$

that is the spectrum of the operator L^ε is discrete. We make the following assumptions on $\sigma(L^\varepsilon)$.

Assumptions 2.

(i) The mode m eigenvalues, that is the eigenvalues of the matrix $\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]$ are geometrically simple for each $m \in \mathbb{Z}$.

(ii) There exists $\omega \geq 0$ such that a *non-semisimple collision* takes place in modes ± 1 for $\varepsilon = 0$ at $\pm \mathrm{i}\omega$, that is as $\varepsilon \downarrow 0$ two pairs of simple complex mode ± 1 eigenvalues of the operator L^ε collide to form geometrically simple and algebraically double eigenvalues $\pm \mathrm{i}\omega$, i.e. a Jordan block of the matrix $\pm \mathrm{i}\tilde{A}^0(0) + \partial \tilde{B}^0[0]$ (see Figure 1.2).

(iii) All further eigenvalues of L^ε are assumed to be semisimple. Moreover we suppose that $\pm \mathrm{i}\omega$ is not in resonance with any other eigenvalue of L^0 (that is no other eigenvalue of L^0 is an integer multiple of $\pm \mathrm{i}\omega$).

(iv) The number $-\omega$ is not an eigenvalue of the matrix $\tilde{A}^0(0)$.

Assumptions 1 and 2 imply the following result. The eigenvalue picture is summarised in Figure 1.4.

Proposition 1.1.

(i) The spectrum of L^ε is symmetric with respect to the real and imaginary axes.

(ii) A reversible $(\mathrm{i}\omega)^2$ resonance takes place in modes ± 1, that is there exist $u, v \in \mathbb{C}^n$ with

$$L^0 u \mathrm{e}^{\mathrm{i}\tau} = \mathrm{i}\omega u \mathrm{e}^{\mathrm{i}\tau}, \quad L^0 v \mathrm{e}^{\mathrm{i}\tau} = \mathrm{i}\omega v \mathrm{e}^{\mathrm{i}\tau} + u \mathrm{e}^{\mathrm{i}\tau}$$

and

$$L^0 \overline{u} \mathrm{e}^{-\mathrm{i}\tau} = -\mathrm{i}\omega \overline{u} \mathrm{e}^{-\mathrm{i}\tau}, \quad L^0 \overline{v} \mathrm{e}^{-\mathrm{i}\tau} = -\mathrm{i}\omega \overline{v} \mathrm{e}^{-\mathrm{i}\tau} + \overline{u} \mathrm{e}^{-\mathrm{i}\tau}.$$

(iii) There is a finite number of semisimple eigenvalues with non-zero real part and an infinite number of semisimple, purely imaginary eigenvalues.

(iv) All eigenvalues depend analytically upon ε.

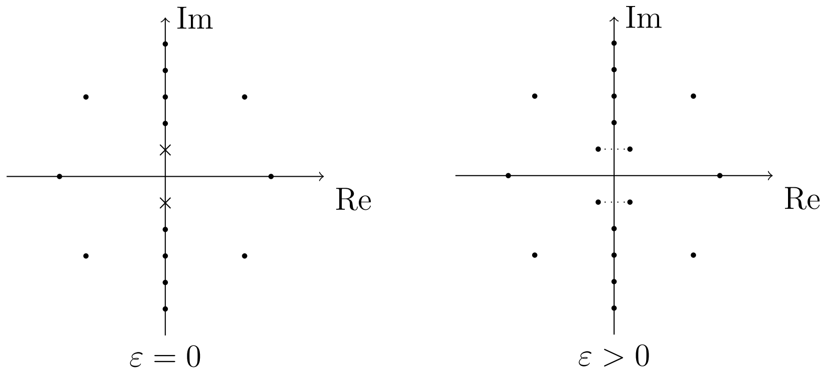

Figure 1.4: The spectrum of the linear operator L^ε consists of two geometrically double eigenvalues $\pm \mathrm{i}\omega$ which split into four geometrically simple eigenvalues for $\varepsilon > 0$, a finite number of semisimple eigenvalues with non-zero real part and an infinite number of semisimple, purely imaginary eigenvalues.

We call the semisimple analytic perturbations of the eigenvalues $\pm \mathrm{i}\omega$ of L^0 the *weakly hyperbolic eigenvalues*, since they are complex for $\varepsilon > 0$ and become purely imaginary for $\varepsilon = 0$. The remaining (finitely many) eigenvalues which have non-zero real part for all ε are called *strongly hyperbolic eigenvalues*; the (infinitely many) purely imaginary eigenvalues are called *centre eigenvalues*. We denote the spectral projections onto the corresponding *weakly hyperbolic subspace*, *strongly hyperbolic subspace* and *centre subspace* of L^ε by $P^\varepsilon_{\mathrm{wh}}$, $P^\varepsilon_{\mathrm{sh}}$ and $P^\varepsilon_{\mathrm{c}}$ (see Section 2.1 for precise definitions). We work under the assumption that the space $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t$ is independent of ε and henceforth write it as $\mathcal{X}_{\mathrm{wh}}$. The ε-dependence of the spaces $P^\varepsilon_{\mathrm{sh}}\mathcal{X}^t$ and $P^\varepsilon_{\mathrm{c}}\mathcal{X}^t$ is removed using a *Kato similarity transformation* $U^\varepsilon \in \mathcal{L}(\mathcal{X}^t)$ which has the property that

$$P^\varepsilon_{\mathrm{wh}} = U^\varepsilon P^0_{\mathrm{wh}}(U^\varepsilon)^{-1}, \quad P^\varepsilon_{\mathrm{sh}} = U^\varepsilon P^0_{\mathrm{sh}}(U^\varepsilon)^{-1}, \quad P^\varepsilon_{\mathrm{c}} = U^\varepsilon P^0_{\mathrm{c}}(U^\varepsilon)^{-1}$$

and thus maps the spaces $\mathcal{X}_{\mathrm{sh}} = P^0_{\mathrm{sh}}\mathcal{X}^t$ and $\mathcal{X}^t_{\mathrm{c}} = P^0_{\mathrm{c}}\mathcal{X}^t$ onto the spaces $P^\varepsilon_{\mathrm{sh}}\mathcal{X}^t$ and $P^\varepsilon_{\mathrm{c}}\mathcal{X}^t$ respectively (see Section 2.1). To simplify the notation we write P_{sh}, P_{c} instead of P^0_{sh}, P^0_{c} and also define $P_{\mathrm{sh,c}} = P_{\mathrm{sh}} + P_{\mathrm{c}}$.

Writing

$$z = P^\varepsilon_{\mathrm{wh}}u, \quad q = (U^\varepsilon)^{-1}(P^\varepsilon_{\mathrm{sh}} + P^\varepsilon_{\mathrm{c}})u,$$

we arrive at the coupled system

$$z_\xi = L^\varepsilon_{\mathrm{wh}}z + f^\varepsilon_1(z) + f^\varepsilon_2(z,q), \tag{1.19}$$

$$q_\xi = L^\varepsilon_{\mathrm{sh,c}}q + P_{\mathrm{sh,c}}\Big(g^\varepsilon_1(z,q)q_\tau\Big) + g^\varepsilon_2(z,q) + h^\varepsilon(z), \tag{1.20}$$

with

$$L^\varepsilon_{\mathrm{wh}} = L^\varepsilon|_{\mathcal{X}_{\mathrm{wh}}}, \qquad L^\varepsilon_{\mathrm{sh,c}} = (U^\varepsilon)^{-1}L^\varepsilon U^\varepsilon|_{\mathcal{X}^{t+1}_{\mathrm{sh,c}}} \quad (t \geq 0)$$

and the nonlinearities

$$\begin{aligned}
f^\varepsilon_1(z) &= f^\varepsilon(z,0),\\
f^\varepsilon_2(z,q) &= f^\varepsilon(z,q) - f^\varepsilon(z,0),\\
f^\varepsilon(z,q) &= P^\varepsilon_{\mathrm{wh}}\Big(\tilde{A}^\varepsilon(z+U^\varepsilon q)(z+U^\varepsilon q)_\tau - \tilde{A}^\varepsilon(0)z_\tau + \tilde{B}^\varepsilon(z+U^\varepsilon q) - \partial\tilde{B}^\varepsilon[0]z\Big),\\
g^\varepsilon_1(z,q) &= (U^\varepsilon)^{-1}\Big(\tilde{A}^\varepsilon(z+U^\varepsilon q) - \tilde{A}^\varepsilon(0)\Big)U^\varepsilon,\\
g^\varepsilon_2(z,q) &= P_{\mathrm{sh,c}}(U^\varepsilon)^{-1}\Big((\tilde{A}^\varepsilon(z+U^\varepsilon q) - \tilde{A}^\varepsilon(z))z_\tau + \tilde{B}^\varepsilon(z+U^\varepsilon q) - \tilde{B}^\varepsilon(z) - \partial\tilde{B}^\varepsilon[0]U^\varepsilon q\Big),\\
h^\varepsilon(z) &= P_{\mathrm{sh,c}}(U^\varepsilon)^{-1}\Big(\tilde{A}^\varepsilon(z)z_\tau + \tilde{B}^\varepsilon(z)\Big),
\end{aligned}$$

which are analytic at the origin as functions $\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^t_{\mathrm{sh,c}} \to \mathcal{X}_{\mathrm{wh}}$ or $\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^t_{\mathrm{sh,c}} \to \mathcal{X}^t_{\mathrm{sh,c}}$ for $t > \frac{1}{2}$.

For $\varepsilon > 0$ the operator L^ε has a pair of simple complex mode 1 eigenvalues $\mathrm{i}(\omega+\nu^\varepsilon)\pm\lambda^\varepsilon$ with the corresponding complex conjugate mode -1 eigenvalues $-\mathrm{i}(\omega+\nu^\varepsilon)\pm\lambda^\varepsilon$. Here λ^ε and ν^ε are (real-valued) analytic functions of ε. We make the following assumptions on the qualitative behaviour of λ^ε, ν^ε and the form of the mode ± 1 eigenvectors for $\varepsilon > 0$.

Assumptions 3. The functions λ^ε and ν^ε satisfy

$$\lambda^\varepsilon = \varepsilon + \mathcal{O}(\varepsilon^2), \quad \nu^\varepsilon = \mathcal{O}(\varepsilon).$$

For $\varepsilon > 0$ the vectors $(u+\lambda^\varepsilon v)\mathrm{e}^{\mathrm{i}\tau}$, $(u-\lambda^\varepsilon v)\mathrm{e}^{\mathrm{i}\tau}$ and $(\overline{u}+\lambda^\varepsilon \overline{v})\mathrm{e}^{-\mathrm{i}\tau}$, $(\overline{u}-\lambda^\varepsilon \overline{v})\mathrm{e}^{-\mathrm{i}\tau}$ are mode ± 1 eigenvectors of L^ε with corresponding eigenvalues $\mathrm{i}(\omega+\nu^\varepsilon) \pm \lambda^\varepsilon$ and $-\mathrm{i}(\omega+\nu^\varepsilon) \pm \lambda^\varepsilon$, that is

$$\begin{aligned} L^\varepsilon(u\mathrm{e}^{\mathrm{i}\tau} \pm \lambda^\varepsilon v \mathrm{e}^{\mathrm{i}\tau}) &= (\mathrm{i}(\omega+\nu^\varepsilon) \pm \lambda^\varepsilon)(u\mathrm{e}^{\mathrm{i}\tau} \pm \lambda^\varepsilon v \mathrm{e}^{\mathrm{i}\tau}), \\ L^\varepsilon(\overline{u}\mathrm{e}^{-\mathrm{i}\tau} \pm \lambda^\varepsilon \overline{v} \mathrm{e}^{-\mathrm{i}\tau}) &= (-\mathrm{i}(\omega+\nu^\varepsilon) \pm \lambda^\varepsilon)(\overline{u}\mathrm{e}^{-\mathrm{i}\tau} \pm \lambda^\varepsilon \overline{v} \mathrm{e}^{-\mathrm{i}\tau}). \end{aligned}$$

It follows that the matrix of the operator $L^\varepsilon_{\mathrm{wh}}$ for $\varepsilon > 0$ with respect to the ordered basis $\{u\mathrm{e}^{\mathrm{i}\tau}, v\mathrm{e}^{\mathrm{i}\tau}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}, \overline{v}\mathrm{e}^{-\mathrm{i}\tau}\}$ for $\mathcal{X}_{\mathrm{wh}}$ is given by

$$L^\varepsilon_{\mathrm{wh}} = \begin{pmatrix} \mathrm{i}(\omega+\nu^\varepsilon) & 1 & 0 & 0 \\ (\lambda^\varepsilon)^2 & \mathrm{i}(\omega+\nu^\varepsilon) & 0 & 0 \\ 0 & 0 & -\mathrm{i}(\omega+\nu^\varepsilon) & 1 \\ 0 & 0 & (\lambda^\varepsilon)^2 & -\mathrm{i}(\omega+\nu^\varepsilon) \end{pmatrix}.$$

Working in this coordinate system by making the *Ansatz*

$$z(\tau) = z_1 u\mathrm{e}^{\mathrm{i}\tau} + z_2 v\mathrm{e}^{\mathrm{i}\tau} + \overline{z}_1 \overline{u}\mathrm{e}^{-\mathrm{i}\tau} + \overline{z}_2 \overline{v}\mathrm{e}^{-\mathrm{i}\tau},$$

one finds that, after a near-identity change of variable and scaling, the equation

$$z_\xi = L^\varepsilon_{\mathrm{wh}} z + f^\varepsilon_1(z)$$

reduces to (1.11), (1.12).

Assumptions 4. The coefficient $\check{C}$ in equations (1.11), (1.12) is positive.

Results

Our main result is the following theorem, which is proved in Chapter 4 by energy methods for symmetrisable hyperbolic equations.

Theorem 1.2. *Let $s > \frac{1}{2}$, $0 < \eta < 1$, $Z = P_{\mathrm{sh}}q$ and $w = P_{\mathrm{c}}q$. There exist positive constants ε_0 and $c^\star$ with the property that for each $\varepsilon \in (0, \varepsilon_0]$ equations (1.19), (1.20) admit reversible solutions $(Z^\star_{w_0}, z^\star_{w_0}, w^\star_{w_0})$ in*

$$C([-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]; \mathcal{X}_{\mathrm{sh}}) \times C([-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]; \mathcal{X}_{\mathrm{wh}}) \times C([-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]; \mathcal{X}^{s+1}_{\mathrm{c}}),$$

parameterised by $w_0 \in \mathcal{X}^{s+2}_{\mathrm{c}}$ with $\|w_0\|_{\mathcal{X}^{s+2}_{\mathrm{c}}} \leq \varepsilon \mathrm{e}^{-c^\star/2\sqrt{\varepsilon}}$ and $Rw_0 = w_0$, such that

$$\begin{aligned} &\|Z^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]; \mathcal{X}_{\mathrm{sh}})}, \\ &\quad \|z^\star_{w_0} - p^\varepsilon\|_{C([-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]; \mathcal{X}_{\mathrm{wh}})}, \\ &\quad \|w^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]; \mathcal{X}^{s+1}_{\mathrm{c}})} \leq \mathrm{e}^{-c^\star/2\sqrt{\varepsilon}}, \end{aligned}$$

where

$$p^{\varepsilon}(\xi,\tau) = \pm\lambda^{\varepsilon}\left(\frac{2}{\check{C}}\right)^{1/2}\operatorname{sech}(\lambda^{\varepsilon}\xi)u\mathrm{e}^{\mathrm{i}\tau} + \text{c.c.} + O(\varepsilon^{3/2}\mathrm{e}^{-\eta\lambda^{\varepsilon}|\xi|}),$$

(so that $\lim\limits_{\xi\to\pm\infty} p^{\varepsilon}(\xi,\tau) = 0$ *uniformly in* $\tau\in\mathbb{R}$*). These solutions satisfy*

$$\begin{aligned}
\|Z^{\star}_{w_0}\|_{C([-\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}},\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}}];\mathcal{X}_{\mathrm{sh}})} &\lesssim \varepsilon\mathrm{e}^{-c^{\star}/2\sqrt{\varepsilon}},\\
\|z^{\star}_{w_0} - p^{\varepsilon}\|_{C([-\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}},\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}}];\mathcal{X}_{\mathrm{wh}})} &\lesssim \sqrt{\varepsilon}\mathrm{e}^{-c^{\star}/2\sqrt{\varepsilon}},\\
\|w^{\star}_{w_0}\|_{C([-\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}},\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}}];\mathcal{X}^{s+1}_{\mathrm{c}})} &\lesssim \varepsilon^{\delta}\mathrm{e}^{-c^{\star}/2\sqrt{\varepsilon}}
\end{aligned}$$

for arbitrary but fixed $\delta\in[0,\frac{1}{4})$*.*

Corollary 1.3. *For each* $\varepsilon\in(0,\varepsilon_0]$ *equation (1.17) admits an infinite-dimensional, continuous family of pulse solutions satisfying*

$$v(\xi,\tau) = Tv(-\xi,-\tau), \quad v(\xi,\tau+2\pi) = v(\xi,\tau), \quad |v(\xi,\tau) - p^{\varepsilon}(\xi,\tau)| \le \mathrm{e}^{-c^{\star}/2\sqrt{\varepsilon}}$$

for all $\tau\in\mathbb{R}$ *and* $\xi\in[-\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}},\mathrm{e}^{c^{\star}/2\sqrt{\varepsilon}}]$ *(see Figure 1.5).*

A global existence result can be obtained in the semilinear case, which corresponds to $g_1 = 0$ in equation (1.20), using semigroup methods. The result is proved in Chapter 5 under the following further assumption.

Assumptions 5. There exist $s \ge 0$ and a conserved quantity $\mathcal{I}^{\varepsilon}(z,q)$ for the system (1.19), (1.20) which is analytic at the origin in $\mathbb{R}\times\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^{s}_{\mathrm{sh,c}}$ and satisfies the estimate

$$\|\mathcal{I}^{\varepsilon}(z,q)\|_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^{s}_{\mathrm{sh,c}}} = \mathcal{O}(\|(z,q)\|^2_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^{s}_{\mathrm{sh,c}}}).$$

Furthermore,

$$\mathcal{I}^{0}(0,w) = Q(w) + \mathcal{O}(\|w\|^3_{\mathcal{X}^{s}_{\mathrm{c}}})$$

for some quadratic form $Q\colon \mathcal{X}^{s}_{\mathrm{c}} \to \mathbb{R}$ satisfying

$$\|w\|^2_{\mathcal{X}^{s}_{\mathrm{c}}} \lesssim \|Q(w)\|_{\mathcal{X}^{s}_{\mathrm{c}}} \lesssim \|w\|^2_{\mathcal{X}^{s}_{\mathrm{c}}},$$

where $w = P_{\mathrm{c}}q$.

Theorem 1.4. *Let* $0<\eta<1$*,* $Z = P_{\mathrm{sh}}q$*,* $w = P_{\mathrm{c}}q$ *and suppose that* $g_1 = 0$*. There exist positive constants* ε_0 *and* $c^{\star}$ *with the property that for each* $\varepsilon\in(0,\varepsilon_0]$ *equations (1.19), (1.20) admit reversible solutions* $(Z^{\star}_{w_0}, z^{\star}_{w_0}, w^{\star}_{w_0})$ *in*

$$C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})\times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})\times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}^{s}_{\mathrm{c}}),$$

parameterised by $w_0\in\mathcal{X}^{s}_{\mathrm{c}}$ *with* $\|w_0\|_{\mathcal{X}^{s}_{\mathrm{c}}} \le \varepsilon\mathrm{e}^{-c^{\star}/2\sqrt{\varepsilon}}$ *and* $Rw_0 = w_0$*, such that*

$$\|Z^{\star}_{w_0}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|z^{\star}_{w_0} - p^{\varepsilon}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})}, \|w^{\star}_{w_0}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}^{s}_{\mathrm{c}})} \le \mathrm{e}^{-c^{\star}/2\sqrt{\varepsilon}},$$

where

$$p^\varepsilon(\xi,\tau) = \pm\lambda^\varepsilon\left(\frac{2}{\check{C}}\right)^{1/2}\operatorname{sech}(\lambda^\varepsilon\xi)u\mathrm{e}^{\mathrm{i}\tau} + \text{c.c.} + O(\varepsilon^{3/2}\mathrm{e}^{-\eta\lambda^\varepsilon|\xi|}),$$

(so that $\lim_{\xi\to\pm\infty} p^\varepsilon(\xi,\tau) = 0$ *uniformly in* $\tau\in\mathbb{R}$*). These solutions satisfy*

$$\begin{aligned}\|Z^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} &\lesssim \varepsilon\mathrm{e}^{-c^\star/2\sqrt{\varepsilon}},\\ \|z^\star_{w_0} - p^\varepsilon\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})} &\lesssim \sqrt{\varepsilon}\mathrm{e}^{-c^\star/2\sqrt{\varepsilon}},\\ \|w^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}},\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}];\mathcal{X}_c^s)} &\lesssim \varepsilon^\delta\mathrm{e}^{-c^\star/2\sqrt{\varepsilon}}\end{aligned}$$

for arbitrary but fixed $\delta\in[0,\frac{1}{4})$*.*

Corollary 1.5. *Suppose that the matrix-valued function* A^ε *in equation (1.17) is constant. For each* $\varepsilon\in(0,\varepsilon_0]$ *equation (1.17) admits an infinite-dimensional, continuous family of pulse solutions satisfying*

$$v(\xi,\tau) = Tv(-\xi,-\tau), \quad v(\xi,\tau+2\pi) = v(\xi,\tau), \quad |v(\xi,\tau) - p^\varepsilon(\xi,\tau)| \le \mathrm{e}^{-c^\star/2\sqrt{\varepsilon}}$$

for all $\tau\in\mathbb{R}$ *and* $\xi\in\mathbb{R}$ *(see Figure 1.5).*

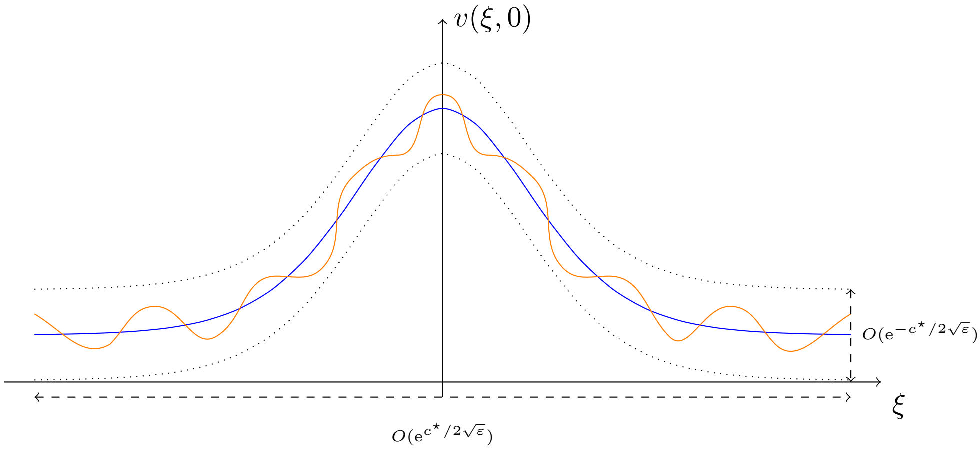

Figure 1.5: A reversible generalised pulse solution v (orange) to equation (1.17) lies in an exponentially small tubular neighbourhood of p^ε (blue) over an exponentially long interval of values for ξ. In the special case that A^ε is constant, this interval extends to the entire real line.

Application to Klein-Gordon equations

In Chapter 6 we apply our theory to the quasilinear Klein-Gordon equation

$$\phi_{tt} = \phi_{xx} - \phi + f_1(\phi,\phi_t,\phi_x)\phi_{xx} + f_2(\phi,\phi_t,\phi_x), \tag{1.21}$$

where $f_1, f_2 \colon \mathbb{R}^3 \to \mathbb{R}$ are analytic functions with $f_1(0)=0$, $f_2(0)=0$, $\mathrm{d}f_2[0]=0$ and

$$f_i(a,-b,-c) = f_i(a,b,c), \quad i=1,2.$$

We prove the existence of *generalised breather solutions* to equation (1.21) of the form

$$\phi(x,t) = v_1(\xi,\tau), \qquad \xi = x, \quad \tau = \sqrt{1-\varepsilon^2}t, \tag{1.22}$$

in which v_1 is 2π-periodic in its second argument (one cannot expect the existence of 'true' breathers with $v_1(\xi,\tau) \to 0$ as $|\xi| \to \infty$; it was for instance shown by Denzler [6] and Birnir, McKean & Weinstein [2] that no perturbation of the sine-Gordon equation up to rescalings admits true breathers).

Theorem 1.6. *There exist positive constants ε_0 and $c^\star$ with the property that for each $\varepsilon \in (0,\varepsilon_0]$ equation (1.21) admits an infinite-dimensional, continuous family of generalised breather solutions of the form*

$$\phi(x,t) = v_1(x,\omega t),$$

parameterised by the values $v_1(0,0)$. These solutions satisfy

$$v_1(\xi,\tau) = v_1(-\xi,\tau), \quad v_1(\xi,\tau+2\pi) = v_1(\xi,\tau), \quad |v(\xi,\tau) - p^\varepsilon(\xi,\tau)| \leq \mathrm{e}^{-c^\star/2\sqrt{\varepsilon}}$$

for all $\tau \in \mathbb{R}$ and respectively $\xi \in [-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]$ and $\xi \in \mathbb{R}$ if additionally $f_1 = 0$, $f_2(a,b,c) = f_2(a)$. Here

$$p^\varepsilon(\xi,\tau) = \pm 2\lambda^\varepsilon \left(\frac{2}{\check{C}}\right)^{1/2} \operatorname{sech}(\lambda^\varepsilon \xi)\cos(\tau) + O(\varepsilon^{3/2}\mathrm{e}^{-\varepsilon\eta|\xi|}), \quad 0 < \eta < 1$$

(so that $\lim_{\xi\to\pm\infty} p^\varepsilon(\xi,\tau) = 0$ uniformly in $\tau \in \mathbb{R}$).

Theorem 1.6 is obtained by applying Theorems 1.2 and 1.4 to a suitable formulation of the breather version of equation (1.21) as a first order system. We begin by writing equation (1.21) as a first order system by inserting our *Ansatz* (1.22) and writing $v_2 = v_{1\xi}$ and $v_3 = v_{1\tau}$, such that

$$v_{1\xi} = v_2, \quad v_{2\xi} = \frac{(1-\varepsilon^2)v_{3\tau} + v_1 - f_2(v_1, \sqrt{1-\varepsilon^2}v_3, v_2)}{1 + f_1(v_1, \sqrt{1-\varepsilon^2}v_3, v_2)}, \quad v_{3\xi} = v_{2\tau}.$$

This system satisfies Assumptions 1(i), 1(iii) and 1(iv) with

$$S^\varepsilon(v) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & f_3^\varepsilon(v) \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

and the change of variable $u = S^\varepsilon(0)^{\frac{1}{2}}v$ yields the system

$$u_\xi = \tilde{A}^\varepsilon(u)u_\tau + \tilde{B}^\varepsilon(u) = L^\varepsilon u + N^\varepsilon(u), \tag{1.23}$$

where

$$L^\varepsilon u = \begin{pmatrix} u_2 \\ \sqrt{1-\varepsilon^2}u_{3\tau} + u_1 \\ \sqrt{1-\varepsilon^2}u_{2\tau} \end{pmatrix}$$

and

$$N^\varepsilon(u) = \begin{pmatrix} 0 \\ \dfrac{f_3^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u) - f_3^\varepsilon(0)}{\sqrt{1-\varepsilon^2}} u_{3\tau} + f_4^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u) - u_1 \\ 0 \end{pmatrix}.$$

In the special case $f_1 = 0,\ f_2(a,b,c) = f_2(a)$ it follows that

$$N^\varepsilon(u) = \begin{pmatrix} 0 \\ f_2(u_1) \\ 0 \end{pmatrix},$$

so that the function

$$\mathcal{H}^\varepsilon(u) = \int_0^{2\pi} \left\{ \frac{1}{2}u_2^2 + \frac{1}{2}u_3^2 - \frac{1}{2}u_1^2 - F_2(u_1) \right\} \mathrm{d}\tau, \tag{1.24}$$

in which F_2 is an antiderivative of f_2, is a conserved quantity for the system (1.23).

Calculating the mode m eigenvalues of L^ε yields the following results.

$\boldsymbol{m=0}$: The matrix $\partial\tilde{B}^\varepsilon[0]$ has, independently of ε, the geometrically simple eigenvalues -1, 0 and 1.

$\boldsymbol{m=\pm 1}$: For $\varepsilon = 0$ the matrix $\pm \mathrm{i}\tilde{A}^0(0) + \partial\tilde{B}^0[0]$ has the algebraically triple, geometrically simple eigenvalue 0. For $\varepsilon > 0$ the matrix $\pm \mathrm{i}\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]$ has the geometrically simple eigenvalues $-\varepsilon$, 0 and ε.

$\boldsymbol{|m| > 1}$: The matrix $\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]$ has the geometrically simple eigenvalues $-\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}$, 0 and $\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}$.

Note that the simple zero eigenvalues in modes $m \in \mathbb{Z} \setminus \{\pm 1\}$ remain simple zero eigenvalues for all values of ε, and we therefore seperate them from the eigenvalues undergoing a non-semisimple collision in modes ± 1. We find that the spectral projections $P_{\mathrm{wh}}^\varepsilon$, $P_{\mathrm{sh}}^\varepsilon$, $P_{\mathrm{c}}^\varepsilon$ and P_0^ε associated with the eigenvalues $-\varepsilon$, 0, ε in modes ± 1, the (finitely many) remaining eigenvalues with non-zero real part, the purely imaginary eigenvalues and the remaining zero eigenvalues are given by respectively

$$\begin{aligned}
P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t &= E_{-1} \oplus E_1,\\
P^\varepsilon_{\mathrm{sh}}\mathcal{X}^t &= \left\langle \begin{pmatrix} 1\\ -1\\ 0\end{pmatrix}, \begin{pmatrix} 1\\ 1\\ 0\end{pmatrix} \right\rangle \subseteq E_0,\\
P^\varepsilon_{\mathrm{c}}\mathcal{X}^t &= \left\{ \begin{pmatrix} 1\\ \mathrm{i}\sqrt{(1-\varepsilon^2)D^2-1}\\ \sqrt{1-\varepsilon^2}\partial_\tau\end{pmatrix} u_{1,1} + \begin{pmatrix} 1\\ -\mathrm{i}\sqrt{(1-\varepsilon^2)D^2-1}\\ \sqrt{1-\varepsilon^2}\partial_\tau\end{pmatrix} u_{1,2} \colon \right.\\
&\qquad \left. u_{1,1}, u_{1,2} \in H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{C}) \right\},\\
P^\varepsilon_{0}\mathcal{X}^t &= \left\{ \begin{pmatrix} -\partial_\tau\\ 0\\ 1\end{pmatrix} u_3 \colon u_3 \in H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{C}) \right\}, \qquad \varepsilon \geq 0.
\end{aligned}$$

It follows from this observation, and the fact that a function $u \in \mathcal{X}^t$ of the form

$$u = \begin{pmatrix} 0\\ u_2\\ 0\end{pmatrix}$$

has by definition no components in the space $P^\varepsilon_0\mathcal{X}^t$, that the nonlinearity N^ε maps $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^{t+1} \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^{t+1}$ analytically into $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^t$ for each $t > \frac{1}{2}$ and each $\varepsilon \geq 0$. In the special case $f_1 = 0$, $f_2(a,b,c) = f_2(a)$, where

$$N^\varepsilon(u) = \begin{pmatrix} 0\\ f_2(u_1)\\ 0\end{pmatrix}$$

we further observe that N^ε maps $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^t$ analytically into itself (in fact into $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^{t+1} \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^{t+1}$) for each $t \geq 0$ and each $\varepsilon \geq 0$. These facts allow us to study equation (1.23) in respectively the phase space $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^s \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^s$ for fixed $s > \frac{1}{2}$ and the phase space $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^0 \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^0$ in the special case.

We proceed by writing $z = P^\varepsilon_{\mathrm{wh}}u$ and $q = (U^\varepsilon)^{-1}P^\varepsilon_{\mathrm{sh,c}}u$ in the usual fashion (see above). Equation (1.23) is however still not of the required form due to the presence of an additional zero eigenvalue in modes ± 1. We therefore obtain a six-dimensional rather than a four-dimensional weakly hyperbolic subspace $\mathcal{X}_{\mathrm{wh}}$. Because of the special structure of the nonlinearity the equations for the components z_1 and z_3 are in fact

$$\begin{aligned}
z_{1\xi} &= \sqrt{1-\varepsilon^2}\, z_2,\\
z_{3\xi} &= (1-\sqrt{1-\varepsilon^2})z_2.
\end{aligned}$$

The quantity

$$I^\varepsilon_{\mathrm{wh}}(z) = z_3 - \frac{1-\sqrt{1-\varepsilon^2}}{\sqrt{1-\varepsilon^2}} z_1$$

is therefore conserved, so that we can study equation (1.23) in the invariant subspace $\{I^{\varepsilon}_{\mathrm{wh}}(z) = 0\}$. After 'replacing' the weakly hyperbolic component of (1.23) with

$$z_{\xi} = \check{L}^{\varepsilon}_{\mathrm{wh}} z + \check{N}^{\varepsilon}_{\mathrm{wh}}(z, q),$$

identifying $\mathcal{X}_{\mathrm{wh}} \cap \{I^{\varepsilon}_{\mathrm{wh}}(z) = 0\}$ with $\mathbb{C}^4$ and introducing the further scaled variables $z_1 = \sqrt{1-\varepsilon^2}\check{z}_1$, $z_2 = \check{z}_2$ we arrive at the system

$$\check{z}_{\xi} = \check{L}^{\varepsilon}_{\mathrm{wh}} \check{z} + f^{\varepsilon}(\check{z}, q), \tag{1.25}$$

$$q_{\xi} = L^{\varepsilon}_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}\Big(g^{\varepsilon}_1(\check{z}, q) q_{\tau}\Big) + g^{\varepsilon}_2(\check{z}, q) + h^{\varepsilon}(\check{z}), \tag{1.26}$$

which satisfies the hypotheses of Theorem 1.2 (Assumptions 1–4).

Finally, recall that

$$\mathcal{H}^{\varepsilon}(u) = \int_0^{2\pi} \left\{ \frac{1}{2}u_2^2 + \frac{1}{2}u_3^2 - \frac{1}{2}u_1^2 - F_2(u_1) \right\} \mathrm{d}\tau$$

is a conserved quantity for the system (1.23) in the special case $f_1 = 0$, $f_2(a,b,c) = f_2(a)$. The formula

$$\mathcal{I}^{\varepsilon}(z, q) = \mathcal{H}^{\varepsilon}(z + U^{\varepsilon} q)$$

thus defines a conserved quantity for the system (1.25), (1.26) in $\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^0_{\mathrm{sh,c}}$. Clearly

$$|\mathcal{I}^{\varepsilon}(z, q)| = \mathcal{O}(\|(z, q)\|^2_{\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^0_{\mathrm{sh,c}}})$$

and taking $w \in \mathcal{X}^0_{\mathrm{c}} \subseteq H^1_{\mathrm{per}}(0, 2\pi; \mathbb{C}) \times L^2_{\mathrm{per}}(0, 2\pi; \mathbb{C}) \times L^2_{\mathrm{per}}(0, 2\pi; \mathbb{C})$, so that

$$w(\tau) = \sum_{m \in \mathbb{Z} \setminus \{-1,0,1\}} w_{m,1} \begin{pmatrix} 1 \\ -\mathrm{i}\sqrt{m^2-1} \\ \mathrm{i}m \end{pmatrix} \mathrm{e}^{\mathrm{i}m\tau} + w_{m,2} \begin{pmatrix} 1 \\ \mathrm{i}\sqrt{m^2-1} \\ \mathrm{i}m \end{pmatrix} \mathrm{e}^{\mathrm{i}m\tau},$$

we find by a direct computation that

$$\mathcal{I}^0(0, w) = Q(w) + \mathcal{O}(\|w\|^3_{\mathcal{X}^0_{\mathrm{c}}})$$

and

$$\|w\|^2_{\mathcal{X}^0_{\mathrm{c}}} \lesssim Q(w) \lesssim \|w\|^2_{\mathcal{X}^0_{\mathrm{c}}}.$$

The remaining hypothesis of Theorem 1.4 (Assumptions 5) is therefore satistied.

1.3 Methodology

Normal-form-theory

In Chapter 3 we adapt the normal-form theory given by Groves & Schneider [10] on the basis of the earlier work by Iooss & Lombardi [14] to construct a sequence of near-identity changes of variable which systematically remove the k-th order terms of the

Maclaurin expansion of $h^\varepsilon(z)$ for $k \in \{1, \dots, p\}$ while preserving the overall structure of the system. In general it is not possible to remove $h^\varepsilon(z)$ completely but we can at least make an optimal choice of p so that the remaining terms are exponentially small in comparison to ε in a suitable neighbourhood of the origin.

A central requirement of the normal-form theory is that the parameter-independent part in the dynamical system (1.19) for z should be diagonalisable. Since (the matrix of) L^0_{wh} on the right-hand side of equation (1.19) is evidently not diagonalisable, we use a change of parameter to 'replace' it by a diagonalisable matrix. We introduce a new parameter μ by writing $\varepsilon = \mu^2$ and use the scaled variables

$$z_1' = (\lambda^\mu)^{-\frac{1}{2}} z_1, \quad z_2' = (\lambda^\mu)^{-\frac{3}{2}} z_2, \quad q' = (\lambda^\mu)^{-1} q, \tag{1.27}$$

where we have abused the notation to abbreviate $\lambda^\mu = \lambda^\varepsilon|_{\varepsilon=\mu^2}$ and $\nu^\mu = \nu^\varepsilon|_{\varepsilon=\mu^2}$. Omitting the primes for notational simplicity, equations (1.19), (1.20) are transformed into

$$z_\xi = L^\mu_{\mathrm{wh}} z + \check{f}^\mu_1(z) + \check{f}^\mu_2(z, q), \tag{1.28}$$

$$q_\xi = L^\mu_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}\Big(\check{g}^\mu_1(z, q) q_\tau\Big) + \check{g}^\mu_2(z, q) + \check{h}^\mu(z), \tag{1.29}$$

where we have (with a further abuse of notation) abbreviated $L^\varepsilon_{\mathrm{sh,c}}|_{\varepsilon=\mu^2}$ to $L^\mu_{\mathrm{sh,c}}$ and

$$L^\mu_{\mathrm{wh}} = \begin{pmatrix} \mathrm{i}(\omega+\nu^\mu) & \lambda^\mu & 0 & 0 \\ \lambda^\mu & \mathrm{i}(\omega+\nu^\mu) & 0 & 0 \\ 0 & 0 & -\mathrm{i}(\omega+\nu^\mu) & \lambda^\mu \\ 0 & 0 & \lambda^\mu & -\mathrm{i}(\omega+\nu^\mu) \end{pmatrix};$$

the nonlinearities $\check{f}^\mu_1$, $\check{f}^\mu_2$, $\check{g}^\mu_1$, $\check{g}^\mu_2$, $\check{h}^\mu(z)$ are obtained by applying the scaling (1.27) to $\tilde{f}^\varepsilon_1$, $\tilde{f}^\varepsilon_2$, $\tilde{g}^\varepsilon_1$, $\tilde{g}^\varepsilon_2$, $\tilde{h}^\varepsilon(z)$ in the obvious way. We use the analogous abbreviations L^μ_{sh} and L^μ_{c} for respectively $L^\varepsilon_{\mathrm{sh}}|_{\varepsilon=\mu^2}$ and $L^\varepsilon_{\mathrm{c}}|_{\varepsilon=\mu^2}$. The result of the theory in Chapter 3 is summarised in the following theorem.

Theorem 1.7. *There exist a constant $c^\star > 0$ and a near-identity, finite-dimensional change of coordinates of the form $q \mapsto q + \Phi^\mu(z)$, where Φ is a polynomial of degree p in (μ, z) with only finitely many nonvanishing Fourier components, which transforms the coupled system (1.28), (1.29) into*

$$z_\xi = L^\mu_{\mathrm{wh}} z + \tilde{f}^\mu_1(z) + \tilde{f}^\mu_2(z, q), \tag{1.30}$$

$$q_\xi = L^\mu_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}\Big(\tilde{g}^\mu_1(z, q) q_\tau\Big) + \tilde{g}^\mu_2(z, q) + \tilde{h}^\mu(z), \tag{1.31}$$

and preserves the reversibility. The transformed nonlinearities $\tilde{f}^\mu_1$, $\tilde{f}^\mu_2$ and $\tilde{g}^\mu_1$ satisfy the estimates

$$|\tilde{f}^\mu_1(z)| = \mathcal{O}(|z|^3)$$

with

$$|\tilde{f}^\mu_1(z) - \check{f}^\mu_1(z)| = \mathcal{O}(|z|^3|(z, \mu)|),$$

and

$$|\tilde{f}_2^\mu(z,q)| = \mathcal{O}(|z||(z,\mu)|^2\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}}),$$
$$\|\tilde{g}_1^\mu(z,q)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(\mu\|(z,q)\|_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^t_{\mathrm{sh,c}}}),$$

while

$$\|\tilde{g}_2^\mu(z,q)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(\mu|z|\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu^2\|q\|^2_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu\|q\|^3_{\mathcal{X}^t_{\mathrm{sh,c}}} + |z|^3\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}})$$

for $t > \frac{1}{2}$*. Here*

$$P_{\mathrm{c}}\Big(\tilde{g}_1^{1;1,0}(z,0)Z_\tau\Big) = P_{\mathrm{c}}\tilde{g}_2^{1;1,1}(z,Z) = 0,$$

in which $Z = P_{\mathrm{sh}}q$ *and* $\tilde{g}_i^{j;k,l}(z,q)$ *denotes the term in the Maclaurin expansion of* $\tilde{g}_i^\varepsilon(z,q)$ *which is homogeneous of degree* k*,* j *and* l *in respectively* ε*,* z *and* q*. Furthermore there exists a constant* $\rho^\star$ *such that*

$$\|\tilde{h}^\mu(z)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \lesssim \mu^2\mathrm{e}^{-c^\star/\mu}, \quad \|\mathrm{d}\tilde{h}^\mu[z]\|_{\mathcal{L}(\mathcal{X}_{\mathrm{wh}},\mathcal{X}^t_{\mathrm{sh,c}})} \lesssim \mu\mathrm{e}^{-c^\star/\mu}$$

for $(z,\mu) \in B_\delta(0) \subseteq \mathbb{C}^5$ *whenever* $\delta p \leq \rho^\star$*.*

Examining the equation

$$z_\xi = L^\mu_{\mathrm{wh}}z + \check{f}_1^\mu(z)$$

yields homoclinic solutions $p^{\mu\pm}$ (we henceforth write p^μ to denote either of these solutions) which can be used in a perturbation argument as approximations to genuine pulse solutions. Note that the normal-form transformations do not affect the leading-order terms in the scaled system (1.28), (1.29), so that the leading-order term in the asymptotic formula for p^μ is also unaffected.

Existence theory

In Chapter 4 we prove the existence result in Theorem 1.2: we construct reversible solutions to the transformed equations (1.30), (1.31) whose pointwise distance to the approximate pulse p^μ does not exceed $\mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$ (see Figure 1.6). To this end we write z as a perturbation $z = p^\mu + r$ of p^μ in equation (1.30) and separate equation (1.31) into its finite-dimensional strongly hyperbolic part and its infinite-dimensional centre part by writing $Z = P_{\mathrm{sh}}q$, $w = P_{\mathrm{c}}q$. We thus obtain

$$Z_\xi = L^\mu_{\mathrm{sh}}Z + G_1^\mu(Z,z,w) + G_2^\mu(z), \tag{1.32}$$
$$r_\xi = L^\mu_{\mathrm{wh}}z + \mathcal{L}^\mu r + f_3^\mu(r) + f_4^\mu(Z,r,w), \tag{1.33}$$
$$w_\xi = L^\mu_{\mathrm{c}}w + P_{\mathrm{c}}\Big(\hat{g}_1^\mu(Z,z,w)w_\tau\Big) + \hat{g}_2^\mu(Z,z,w) + \hat{h}^\mu(z), \tag{1.34}$$

where

$$
\begin{aligned}
G_1^\mu(Z,z,w) &= P_{\mathrm{sh}}\Big(\tilde{g}_1(z,Z+w)(Z+w)_\tau + \tilde{g}_2^\mu(z,Z+w)\Big),\\
G_2^\mu(z) &= P_{\mathrm{sh}}\tilde{h}^\mu(z),\\
\mathcal{L}^\mu &= \mathrm{d}\tilde{f}_1^\mu[p^\mu],\\
f_3^\mu(r) &= \tilde{f}_1^\mu(p^\mu + r) - \tilde{f}_1^\mu(p^\mu) - \mathrm{d}\tilde{f}_1^\mu[p^\mu](r),\\
f_4^\mu(Z,r,w) &= \tilde{f}_2^\mu(p^\mu + r, Z+w),\\
\hat{g}_1^\mu(Z,z,w) &= \tilde{g}_1^\mu(z,Z+w),\\
\hat{g}_2^\mu(Z,z,w) &= P_{\mathrm{c}}\Big(\tilde{g}_2^\mu(z,Z+w) + \tilde{g}_1^\mu(Z,z,w)Z_\tau\Big),\\
\hat{h}^\mu(z) &= P_{\mathrm{c}}\tilde{h}^\mu(z).
\end{aligned}
$$

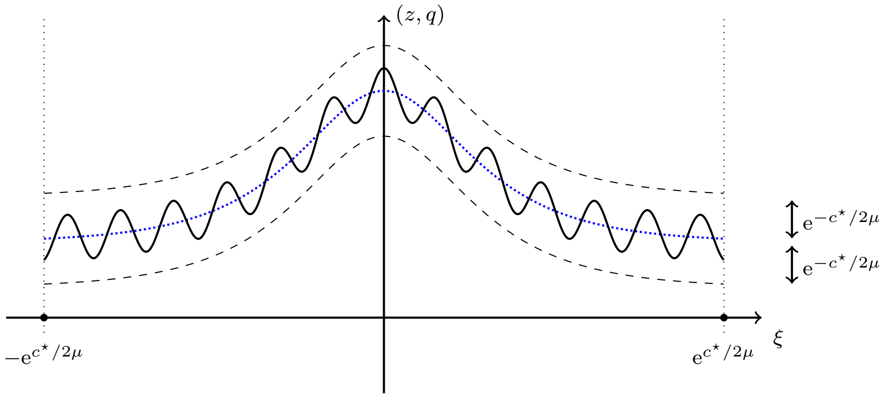

Figure 1.6: A generalised pulse solution to equations (1.30), (1.31) lies in an exponentially thin tubular neighbourhood of $(p^\mu,0)$ on the exponentially large interval $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$.

The proof of Theorem 1.2 is based upon the following iteration scheme for equations (1.32)–(1.34). Suppose that $Z_{(0)}$, $r_{(0)}$ and $w_{(0)}$ are appropriately chosen initial values (see below) and let

$$
\begin{aligned}
Z_{(j+1)}(\xi) = \int_{-\mathrm{e}^{-c^\star/2\mu}}^{\xi} &\mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{ss}}\Big(G_1^\mu(Z_{(j)}, p^\mu + \mathcal{E}_{\mathrm{wh}} r_{(j)}, \mathcal{E}_{\mathrm{c}} w_{(j)})(t) + G_2^\mu(\mathcal{E}_{\mathrm{wh}} r_{(j)})(t)\Big)\,\mathrm{d}t\\
&- \int_{\xi}^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{su}}\Big(G_1^\mu(Z_{(j)}, p^\mu + \mathcal{E}_{\mathrm{wh}} r_{(j)}, \mathcal{E}_{\mathrm{c}} w_{(j)})(t) + G_2^\mu(\mathcal{E}_{\mathrm{wh}} r_{(j)})(t)\Big)\,\mathrm{d}t,
\end{aligned}
$$

$$
\begin{aligned}
r_{(j+1)}(\xi) = \sum_{i=1}^{2} \bigg(&\int_0^\xi \Big\langle \Big(f_3^\mu(r_{(j)})(t) + f_4^\mu(Z_{(j)}, r_{(j)}, w_{(j)})(t)\Big), S_i^*(t)\Big\rangle\,\mathrm{d}t\bigg) S_i(\xi)\\
&- \bigg(\int_\xi^\infty \Big\langle \Big(f_3^\mu(r_{(j)})(t) + f_4^\mu(Z_{(j)}, r_{(j)}, w_{(j)})(t)\Big), U_i^*(t)\Big\rangle\,\mathrm{d}t\bigg) U_i(\xi),
\end{aligned}
$$

and $w_{(j+1)}$ be the unique solution of the initial value problem

$$\begin{aligned}\partial_\xi w_{(j+1)} &= L_{\mathrm{c}}^\mu w_{(j)} + P_{\mathrm{c}}\Big(\hat{g}_1^\mu(Z_{(j)}, p^\mu + r_{(j)}, w_{(j)})\partial_\tau w_{(j+1)}\Big) + \mu b(p^\mu + r_{(j)}, w_{(j)}) \\ &\qquad + \hat{g}_2^\mu(Z_{(j)}, p^\mu + r_{(j)}, w_{(j)}) + \hat{h}^\mu(r_{(j)}), \\ w_{(j+1)}(0) &= w_{(0)}(0)\end{aligned}$$

for fixed $j \in \mathbb{N}_0$ (see below). Here P_{ss} and P_{su} denote the projections onto the *stable subspace* of L_{sh}^μ spanned by the eigenvectors of L_{sh}^μ whose corresponding eigenvalues have negative real part, and the *unstable subspace* of L_{sh}^μ spanned by the eigenvectors of L_{sh}^μ whose corresponding eigenvalues have positive real part. The mappings $\mathcal{E}_{\mathrm{wh}}$ and $\mathcal{E}_{\mathrm{c}}$ are the extensions by reversibility defined by

$$(\mathcal{E}_{\mathrm{wh}}r)(\xi) = \begin{cases} r(\xi), & \xi \geq 0, \\ R_{\mathrm{wh}}r(-\xi), & \xi < 0, \end{cases} \qquad (\mathcal{E}_{\mathrm{c}}w)(\xi) = \begin{cases} w(\xi), & \xi \geq 0, \\ R_{\mathrm{c}}w(-\xi), & \xi < 0. \end{cases}$$

A solution $r(\xi)$ to equation (1.33) which is defined for $\xi \geq 0$ and satisfies $R_{\mathrm{wh}}r(0) = r(0)$ automatically gives rise to a reversible solution $r(\xi)$ to equation (1.33) due to extension by reversibility. The reversibility of a solution Z to equation (1.32) defined for $\xi \in \mathbb{R}$ is verified *a posteriori* by observing that $R_{\mathrm{sh}}Z(-\xi)$ solves equation (1.32) whenever $Z(\xi)$ is a solution and using uniqueness. We thus obtain a reversible solution to equation (1.32). Finally $\{S_1^\mu, S_2^\mu, U_1^\mu, U_2^\mu\}$ is a fundamental solution set consisting of two anti-reversible solutions S_1^μ, S_2^μ and two reversible solutions U_1^μ, U_2^μ for the linearised equation

$$r_\xi = L_{\mathrm{wh}}^\mu z + \mathcal{L}^\mu r$$

for r (see Lemma 2.19 for the precise construction).

We establish that the iteration scheme has a unique limit $(Z^\star, r^\star, w^\star)$ in

$$\begin{aligned}\Big\{&(Z, r, w) \in C\Big([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{sh}}\Big) \times C\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{wh}}\Big) \times C\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{c}}^{s+1}\Big) : \\ &\|Z\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{sh}})}, \|r\|_{C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{wh}})}, \|w\|_{C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{c}}^{s+1})} \leq \mathrm{e}^{-c^\star/2\mu}, \\ &R_{\mathrm{wh}}r(0) = r(0),\ R_{\mathrm{c}}w(0) = w(0)\Big\}\end{aligned}$$

provided that the initial values lie in the set

$$W_{\mathrm{loc}}^{\mathrm{cs}} = \Big\{\Big(Z_{w_0}^\star(0), r_{w_0}^\star(0), w_{w_0}^\star(0)\Big) : w_0 \in \mathcal{X}_{\mathrm{c}}^{s+2},\ R_{\mathrm{c}}w_0 = w_0,\ \|w_0\|_{s+2} \leq \mu^2 \mathrm{e}^{-c^\star/2\mu}\Big\}$$

which we refer to as a *local centre-stable manifold* $W_{\mathrm{loc}}^{\mathrm{cs}}$ for equations (1.32)–(1.34); here we make use of familiar dynamical systems terminology regarding invariant manifolds (see Kelley [16]). An exponentially small solution to the reformulated system (1.32)–(1.34) corresponds to a solution of equations (1.30), (1.31) which is exponentially close to p^μ (see above).

The above definition of the iterate $w_{(j+1)}$ requires the solvability of an associated linear Cauchy problem. The relevant theory is given in Section 4.1 and uses methods for symmetrisable hyperbolic systems as presented by Benzoni-Gavage & Serre

[1, Chapter 2], who consider the class of equations

$$\partial_\xi u + A(\xi,\tau)\partial_\tau u = B(\xi,\tau)u + f(\xi,\tau)$$

with coefficients A and B which depend smoothly on ξ and τ. It is necessary to extend their methods by treating coefficients with values in finite-order Sobolev spaces and replacing the function B by a general linear operator acting on the function u. In order to prove the convergence of the sequence $\{w_{(j)}\}_{j\in\mathbb{N}}$ we use the method given by Benzoni-Gavage & Serre [1, Chapter 10] and construct *a priori* energy estimates, the proof of which requires the symmetrisation of the matrix-valued function g_1^μ via multiplication by the matrix $S^\varepsilon(z+Z+w)|_{\varepsilon=\mu^2}$ (see Assumptions 1(iii)). This symmetrisation clearly changes the spectrum of the operator L_c^μ on the right-hand side of equation (1.34), the structure of which determines the basis for the space $\mathcal{X}_\mathrm{c}^t$ (see above). The symmetrisation of the linear part is therefore carried out as a preliminary step (see the change of variable below Assumptions 1); a further near-identity symmetrisation for the nonlinear part yields the desired *a priori* energy estimates.

A major difficulty in obtaining the necessary estimates for $w_{(j)}$ lies in the presence of quadratic terms in our nonlinearities. Specifically, one arrives at the differential inequality

$$\partial_\xi\|w_{(j)}(\xi)\|^2 \lesssim \mu(\mathrm{e}^{-c^\star/\mu} + \mu^2\mathrm{e}^{-\mu^2\eta\xi}\mathrm{e}^{-c^\star/2\mu})\|w_{(j)}(\xi)\| + \mu(\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-\mu^2\eta\xi})\|w_{(j)}(\xi)\|^2 \tag{1.35}$$

for $\|w_{(j)}(\xi)\|$, and it is necessary to deduce that

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

The required estimate for $\|w_{(j)}(\xi)\|$ follows from inequality (1.35) by a two-step estimation technique devised by Groves and Schneider [10, §5]. In the first step we define $\xi^\star$ so that $\mathrm{e}^{-\mu^2\eta\xi^\star}=\mu^\alpha$, where α is an appropriately chosen positive constant; a straightforward application of Gronwall's inequality shows that $\|w_{(j)}(\xi)\|^2 \lesssim \mu|\log\mu|\mathrm{e}^{-c^\star/\mu}$ for $\xi\in[0,\xi^\star]$ whenever $\|w_{(j)}(0)\|\leq\mu\mathrm{e}^{-c^\star/2\mu}$. In the second step we integrate (1.35) over $[0,\mathrm{e}^{c^\star/2\mu}]$ and split the range of integration into $[0,\xi^\star]$ and $[\xi^\star,\mathrm{e}^{c^\star/2\mu}]$. Satisfactory estimates for the integrals over $[\xi^\star,\mathrm{e}^{c^\star/2\mu}]$ 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 $\|w_{(j)}(\xi)\|\leq c\mu^{1/2}|\log\mu|\mathrm{e}^{-c^\star/2\mu}$ for $\xi\in[0,\mathrm{e}^{c^\star/2\mu}]$.

Existence theory for the semilinear problem

In Chapter 5 we prove the global existence result in Theorem 1.4. We separate equation (1.31) into its finite-dimensional strongly hyperbolic part and its infinite-dimensional (semilinear) centre part by writing $Z=P_\mathrm{sh}q$, $w=P_\mathrm{c}q$ to obtain

$$Z_\xi = L_\mathrm{sh}^\mu Z + G_1^\mu(Z,z,w) + G_2^\mu(z), \tag{1.36}$$

$$z_\xi = L_\mathrm{wh}^\mu z + \hat{f}_1^\mu(z) + \hat{f}_2^\mu(Z,z,w), \tag{1.37}$$

$$w_\xi = L_\mathrm{c}^\mu w + \hat{g}^\mu(Z,z,w) + \hat{h}^\mu(z), \tag{1.38}$$

where

$$\begin{aligned}
G_1^\mu(Z,z,w) &= P_{\mathrm{sh}}\tilde{g}_2^\mu(z,Z+w),\\
G_2^\mu(z) &= P_{\mathrm{sh}}\tilde{h}^\mu(z),\\
\hat{f}_1^\mu(z) &= \tilde{f}_1^\mu(z),\\
\hat{f}_2^\mu(Z,z,w) &= \tilde{f}_2^\mu(z,Z+w),\\
\hat{g}^\mu(Z,z,w) &= P_{\mathrm{c}}\tilde{g}_2^\mu(z,Z+w),\\
\hat{h}^\mu(z) &= P_{\mathrm{c}}\tilde{h}^\mu(z)
\end{aligned}$$

(recall that $\tilde{g}_1 = 0$ in equation (1.31) in the semilinear case). The proof of Theorem 1.4 has three main steps which are summarised in Theorems 1.8, 1.9 and 1.11.

We begin by writing z as a perturbation $p^\mu + r$ of p^μ and introduce an artificial bound on the infinite-dimensional w-component of the system by applying a cut-off function $\phi\colon \mathcal{X}_{\mathrm{c}}^s \to \mathcal{X}_{\mathrm{c}}^s$ defined as $\phi(w) = \psi(\mathrm{e}^{c^\star/2\mu}\|w\|_s)w$, in which $\psi \in C^\infty(\mathbb{R})$ satisfies

$$\psi(x) = \begin{cases} 1, & |x| \le 1,\\ 0, & |x| \ge 2,\end{cases}$$

and $|\psi^{(l)}(x)| \le 2^l$ for all $x \in \mathbb{R}$, $l \in \mathbb{N}_0$, to the w-component of the nonlinearities. The modification via the cut-off function ϕ enables us to control the linear growth of the centre part w on the entire real line in the estimates below (see inequality (5.33)). The resulting system reads

$$Z_\xi = L_{\mathrm{sh}}^\mu Z + G_1^\mu(Z, p^\mu + r, \phi(w)) + G_2^\mu(r), \tag{1.39}$$
$$r_\xi = L_{\mathrm{wh}}^\mu r + \mathcal{L}^\mu r + f_3^\mu(r) + f_4^\mu(Z, r, \phi(w)), \tag{1.40}$$
$$w_\xi = L_{\mathrm{c}}^\mu w + \hat{g}^\mu(Z, p^\mu + r, \phi(w)) + \hat{h}^\mu(r), \tag{1.41}$$

where

$$\begin{aligned}
\mathcal{L}^\mu &= \mathrm{d}\tilde{f}_1^\mu[p^\mu],\\
f_3^\mu(r) &= \tilde{f}_1^\mu(p^\mu + r) - \tilde{f}_1^\mu(p^\mu) - \mathrm{d}\tilde{f}_1^\mu[p^\mu](r),\\
f_4^\mu(Z,r,w) &= \tilde{f}_2^\mu(p^\mu + r, Z + w).
\end{aligned}$$

Using semigroup methods we construct solutions to the modified equations (1.39)–(1.41) which exist globally and are exponentially close to $(0, p^\mu, 0)$ on the exponentially large interval $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$, where they in particular solve equations (1.36)–(1.38) (cf. Figure 1.6). The result is summarised in the following theorem.

Theorem 1.8. *Equations (1.39)–(1.41) admit reversible solutions* $(Z_{w_0}^\star, r_{w_0}^\star, w_{w_0}^\star)$ *in*

$$C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_{\mathrm{c}}^s),$$

parameterised by $w_0 \in \mathcal{X}_{\mathrm{c}}^s$ *with* $\|w_0\|_s \leq \mu \mathrm{e}^{-c^\star/2\mu}$, $R_{\mathrm{c}} w_0 = w_0$, *such that*

$$\|Z^\star_{w_0}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|r^\star_{w_0}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})}, \|w^\star_{w_0}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \leq \mathrm{e}^{-c^\star/2\mu}$$

and

$$\|w^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{c}}^s)} \leq \mathrm{e}^{-c^\star/2\mu}.$$

The solutions $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ obtained by Theorem 1.8 are parameterised by the initial values w_0. We therefore define

$$W^{\mathrm{cs}}_{\mathrm{loc}} = \left\{ \left(Z^\star_{w_0}(0), r^\star_{w_0}(0), w^\star_{w_0}(0) \right) : w_0 \in \mathcal{X}_{\mathrm{c}}^s,\ R_{\mathrm{c}} w_0 = w_0,\ \|w_0\|_s \leq \mu \mathrm{e}^{-c^\star/2\mu} \right\}$$

and refer to $W^{\mathrm{cs}}_{\mathrm{loc}}$ as a *local centre-stable manifold for reversible solutions to equations (1.39)–(1.41) at* $\xi = 0$. Solutions $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ to equations (1.39)–(1.41) whose initial values lie on $W^{\mathrm{cs}}_{\mathrm{loc}}$ have exponentially small hyperbolic part (see Figure 1.7).

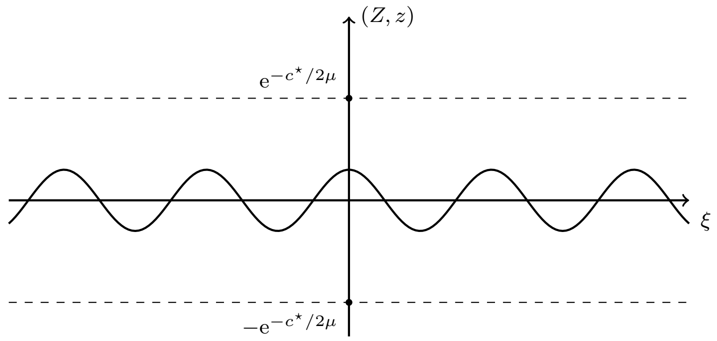

Figure 1.7: The hyperbolic part $(Z^\star, r^\star)$ of solutions to equations (1.39)–(1.41) with initial values on $W^{\mathrm{cs}}_{\mathrm{loc}}$ satisfies $\|(Z^\star(\xi), r^\star(\xi))\|_{\mathcal{X}_{\mathrm{sh}} \times \mathcal{X}_{\mathrm{wh}}} \leq \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$.

The proof of Theorem 1.4 is completed by showing that in fact

$$\|w^\star_{w_0}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s)} \leq \mathrm{e}^{-c^\star/2\mu},$$

from which it follows that $W^{\mathrm{cs}}_{\mathrm{loc}}$ is actually a *global* centre-stable manifold, meaning that its elements are initial values of generalised pulse solutions which exist on the entire real line. We use the following steps.

In Section 5.2 we consider the modified equations

$$Z_\xi = L^\mu_{\mathrm{sh}} Z + G^\mu_1(Z, z, \phi(w)) + G^\mu_2(z), \tag{1.42}$$

$$z_\xi = L^\mu_{\mathrm{wh}} z + \hat{f}^\mu_1(r) + \hat{f}^\mu_2(Z, z, \phi(w)), \tag{1.43}$$

$$w_\xi = L^\mu_{\mathrm{c}} w + \hat{g}^\mu(Z, z, \phi(w)) + \hat{h}^\mu(z), \tag{1.44}$$

which are obtained from the original equations (1.36)–(1.38) by applying the cut-off function ϕ (see above) to the w-component of the nonlinearities. Note that solutions

(Z,z,w) to equations (1.42)–(1.44) with $\|w(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \leq \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$ correspond to solutions (Z,z,w) of equations (1.36)–(1.38), where the cut-off function is not used. The construction of a suitable *centre manifold* for equations (1.42)–(1.44) below requires the existence of solutions with exponentially small hyperbolic part.

Theorem 1.9. *For each $w_0 \in \mathcal{X}_\mathrm{c}^s$ equations (1.42)–(1.44) admit a unique solution*

$$(Z_{w_0}^{\star\star}, z_{w_0}^{\star\star}, w_{w_0}^{\star\star}) \in C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh}) \times C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{c}^s)$$

satisfying

$$\|Z_{w_0}^{\star\star}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})}, \|z_{w_0}^{\star\star}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

In the context of Theorem 1.9 we define the mapping

$$\Psi\colon \mathcal{X}_\mathrm{c}^s \to \mathcal{X}_\mathrm{sh} \times \mathcal{X}_\mathrm{wh}, \quad \Psi(w_0) = \big(Z_{w_0}^{\star\star}(0), z_{w_0}^{\star\star}(0)\big)$$

and observe that any solution

$$(Z,z,w) \in C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh}) \times C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh}) \times C(\mathbb{R};\mathcal{X}_\mathrm{c}^s)$$

of equations (1.42)–(1.44) with

$$\|Z\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})}, \|z\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})} \leq \mathrm{e}^{-c^\star/2\mu}$$

lies in

$$C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh}) \times C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{c}^s)$$

(see the explanation below Theorem 5.8). Theorem 1.9 therefore implies that $(Z(0),z(0)) = \Psi(w(0))$ and we refer to the set

$$W^\mathrm{c} = \left\{\big(Z_{w_0}^{\star\star}(0), z_{w_0}^{\star\star}(0), w_{w_0}^{\star\star}(0)\big)\colon w_0 \in \mathcal{X}_\mathrm{c}^s\right\} = \left\{\big(\Psi(w_0), w_0\big)\colon w_0 \in \mathcal{X}_\mathrm{c}^s\right\}$$

as a *centre manifold for reversible solutions to equations (1.42)–(1.44) at $\xi = 0$*. Due to the autonomy of equations (1.42)–(1.44) the manifold W^c is an invariant manifold, that is $(Z(\xi), z(\xi), w(\xi)) \in W^\mathrm{c}$ and in particular

$$|Z(\xi)|, |z(\xi)| \leq \mathrm{e}^{-c^\star/2\mu}$$

for all $\xi \in \mathbb{R}$ whenever $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W^\mathrm{c}$ for some $\xi_0 \in \mathbb{R}$. This global invariance is not inherited by the *local centre manifold*

$$W_\mathrm{loc}^\mathrm{c} = \{(Z,z,w) \in \mathcal{X}^s\colon (Z,z) = \Psi(w),\ w \in \mathcal{X}_\mathrm{c}^s,\ \|w\|_s \leq \mathrm{e}^{-c^\star/2\mu}\} \subseteq W^\mathrm{c}$$

consisting of those points on W^c which have exponentially small centre part. However we observe that each solution (Z,z,w) to (1.36)–(1.38)

(i) with
$$\|Z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})}, \|w\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}^s_{\mathrm{c}})} \le \mathrm{e}^{-c^\star/2\mu}$$
satisfies $(Z(\xi), z(\xi), w(\xi)) \in W^{\mathrm{c}}_{\mathrm{loc}}$ for all $\xi \in \mathbb{R}$,

(ii) with $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W^{\mathrm{c}}_{\mathrm{loc}}$ for some $\xi_0 \in \mathbb{R}$ remains on $W^{\mathrm{c}}_{\mathrm{loc}}$ as long as
$$|Z(\xi)|, |z(\xi)|, \|w(\xi)\|_{\mathcal{X}^s_{\mathrm{c}}} \le \mathrm{e}^{-c^\star/2\mu}.$$

Using a suitable estimate for the function Ψ and applying the estimate for the Lyapunov functional $\mathcal{I}^\mu(Z,z,w) = \mathcal{I}^\varepsilon(z, Z+w)|_{\varepsilon=\mu^2}$ (see Assumptions 5) we conclude the following invariance result for $W^{\mathrm{c}}_{\mathrm{loc}}$.

Lemma 1.10. *Suppose that*
$$(Z,z,w) \in C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}^s_{\mathrm{c}})$$
is a solution of equations (1.42)–(1.44) and $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W^{\mathrm{c}}$ *with*
$$\|w(\xi_0)\|_s \le \tfrac{1}{2}\mathrm{e}^{-c^\star/2\mu}$$
(so that $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W^{\mathrm{c}}_{\mathrm{loc}}$*). It follows that*
$$\|w(\xi)\|_s \le \tfrac{7}{8}\mathrm{e}^{-c^\star/2\mu}$$
(and therefore $(Z(\xi), z(\xi), w(\xi)) \in W^{\mathrm{c}}_{\mathrm{loc}}$*) for all* $\xi \ge \xi_0$.

The final step in the proof of Theorem 1.4 is to use the results in Lemma 1.10 to establish that a solution $(Z^\star, p^\mu + r^\star, w^\star)$ to equations (1.36)–(1.38) with $(Z^\star(0), r^\star(0), w^\star(0)) \in W^{\mathrm{cs}}_{\mathrm{loc}}$ converges exponentially fast to a solution (Z,z,w) to equations (1.42)–(1.44) with $(Z(\xi), z(\xi), w(\xi)) \in W^{\mathrm{c}}$ (that is $\|(Z(\xi), z(\xi))\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{wh}}} \le \mathrm{e}^{-c^\star/2\mu}$) for all $\xi \in \mathbb{R}$ (see Figure 1.8).

Theorem 1.11. *Let* $0 < \eta < 1$ *and suppose that* $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ *is a solution of the modified equations (1.39)–(1.41). It follows that* $(Z^\star_{w_0}, p^\mu + r^\star_{w_0}, w^\star_{w_0})$ *converges exponentially to a solution* (Z,z,w) *of equations (1.42)–(1.44) on* W^{c}*, that is*
$$\lim_{\xi\to\infty} \mathrm{e}^{\eta\lambda^\mu\xi}\|(Z^\star_{w_0}(\xi), p^\mu(\xi) + r^\star_{w_0}(\xi), w^\star_{w_0}(\xi)) - (Z(\xi), z(\xi), w(\xi))\|_s = 0.$$
In particular there exists $\xi^\star = \xi^\star(\mu) \in (0, \mathrm{e}^{c^\star/2\mu}]$ *such that*
$$\sup_{\xi\in[\xi^\star,\infty)} \|w^\star_{w_0}(\xi) - w(\xi)\|_s \lesssim \mu\mathrm{e}^{-c^\star/2\mu}.$$

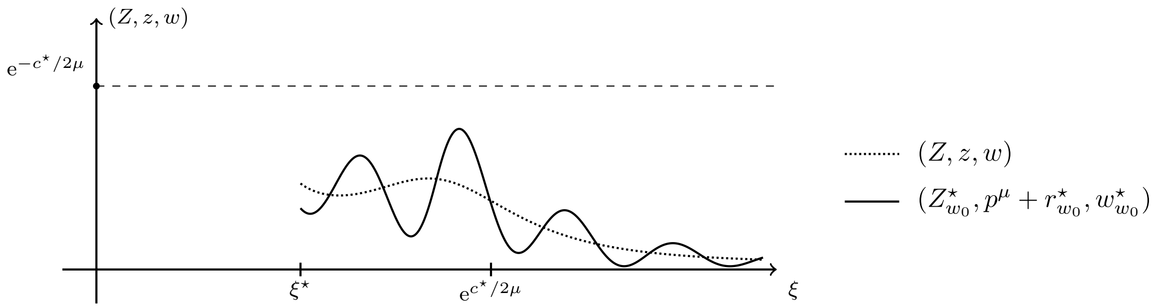

Figure 1.8: A solution $(Z^\star_{w_0}, p^\mu + r^\star_{w_0}, w^\star_{w_0})$ with initial values on $W^{\mathrm{cs}}_{\mathrm{loc}}$ converges exponentially fast to a solution (Z, z, w) on W^{c}. Observe that $(Z^\star_{w_0}, p^\mu + r^\star_{w_0}, w^\star_{w_0})$ is (i) exponentially small, (ii) exponentially close to (Z, z, w) on $[\xi^\star, \mathrm{e}^{c^\star/2\mu})$. The (exponential) convergence of $(Z^\star_{w_0}, p^\mu + r^\star_{w_0}, w^\star_{w_0})$ to (Z, z, w) and the fact that (Z, z, w) remains on $W^{\mathrm{c}}_{\mathrm{loc}}$ with for all $\xi \geq \xi^\star$ imply that $(Z^\star_{w_0}, p^\mu + r^\star_{w_0}, w^\star_{w_0})$ remains exponentially small for all $\xi \geq \xi^\star$.

Since by construction $\|w^\star_{w_0}(\xi^\star)\|_{\mathcal{X}^s_{\mathrm{c}}} \lesssim \mu^\delta \mathrm{e}^{-c^\star/2\mu}$, it follows that $\|w(\xi^\star)\|_{\mathcal{X}^s_{\mathrm{c}}} \leq \frac{1}{2}\mathrm{e}^{-c^\star/2\mu}$, and since $(Z(\xi^\star), z(\xi^\star), w(\xi^\star)) \in W^{\mathrm{c}}$ we conclude that $(Z(\xi), z(\xi), w(\xi)) \in W^{\mathrm{c}}_{\mathrm{loc}}$ for all $\xi \geq \xi^\star$ (see Lemma 1.10). This observation finally implies that $\phi(w^\star_{w_0}(\xi)) = w^\star_{w_0}(\xi)$ for all $\xi \geq \xi^\star$, that is

$$\|w^\star_{w_0}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}^s_{\mathrm{c}})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

The statement in Theorem 1.4 is now obtained by using the rescaling $\mu = \sqrt{\varepsilon}$.

The theory in Chapter 5 actually applies to more general systems of the form (1.36)–(1.38) in which the spectral hypotheses are replaced by the statements in Proposition 1.1 (the specific structure (1.17) is not required). We can also incorporate the case where the underlying strongly hyperbolic subspace $\mathcal{X}_{\mathrm{sh}}$ is infinite-dimensional (and thus written as $\mathcal{X}^s_{\mathrm{sh}}$) under two additional hypotheses.

The identity

$$P_{\mathrm{c}}\tilde{g}^{1;1,1}(z, Z) = 0$$

for the transformed nonlinearity $\tilde{g}$ (recall that $g_1 = 0$ and we write g_2 as g in the semilinear case) given by Theorem 1.7 guarantees the absence of quadratic terms with an insufficient power of μ in the estimates for our existence theory. As $\mathcal{X}^s_{\mathrm{sh}}$ is infinite-dimensional this criterion cannot be achieved via a finite-dimensional change of coordinates. It is thus necessary to work on the assumption that there exists an additional infinite-dimensional, reversibility preserving change of coordinates which removes these problematic terms.

We also implicitly use the fact that the ordinary differential equation

$$Z_\xi = L^\mu_{\mathrm{sh}} Z + G^\mu(Z^\dagger, z^\dagger, w^\dagger), \tag{1.45}$$

where $G^\mu \colon \mathcal{X}^s \to \mathcal{X}^s_{\mathrm{sh}}$ is defined by

$$G^\mu(Z, z, w) = G^\mu_1(Z, z, w) + G^\mu_2(z)$$

(see equation (1.36)), is solvable for all $(Z^\dagger, z^\dagger, w^\dagger)\colon \mathbb{R} \to \mathcal{X}^s$. This requirement is an additional assumption when $\mathcal{X}^s_{\mathrm{sh}}$ is infinite-dimensional: we suppose that for each $(Z^\dagger, z^\dagger, w^\dagger) \in C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}^s)$ equation (1.45) has a unique solution

$$\mathcal{K}^\mu(Z^\dagger, z^\dagger, w^\dagger) \in C^1_{\mathrm{b}}(\mathbb{R}; \mathcal{X}^s_{\mathrm{sh}}) \cap C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}^{s+1}_{\mathrm{sh}})$$

satisfying

$$\|\mathcal{K}^\mu(Z^\dagger, z^\dagger, w^\dagger)\|_{C^1_{\mathrm{b}}(\mathbb{R};\mathcal{X}^s_{\mathrm{sh}})} + \|\mathcal{K}^\mu(Z^\dagger, z^\dagger, w^\dagger)\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}^s_{\mathrm{sh}})} \lesssim \|(Z^\dagger, z^\dagger, w^\dagger)\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}^s)}$$

uniformly over (sufficiently small values of) μ.

Remark 1.12. The above methods have been used in a similar fashion by Hewer [13, Chapter 5] for quasilinear problems with finitely many centre and infinitely many strongly hyperbolic eigenvalues.

1.4 Function spaces

We complete the introduction with a glossary of the function spaces used throughout this thesis.

Spaces of continuous functions

Definition 1.13. Suppose that X is a complex Banach space, I an interval, $k \in \mathbb{N}_0$ and $\theta \in \mathbb{R}$.

(i) We denote by $C^k_{\mathrm{b}}(I; X)$ the Banach space of all k-times continuously differentiable, bounded functions $I \to X$ and equip it with the norm

$$\|f\|_{C^k_{\mathrm{b}}(I;X)} = \sup_{\xi\in I} \sum_{j=0}^{k} \|f^{(j)}(\xi)\|_X.$$

In the case $k = 0$ we just write $C_{\mathrm{b}}(I; X)$. If I is a compact interval we drop the subscript b and write $C^k(I; X)$. If $X = \mathbb{R}$ we abbreviate the space $C_{\mathrm{b}}(I; X)$ to $C^k_{\mathrm{b}}(I)$. As usual we define

$$C^\infty_{\mathrm{b}}(I; X) = \bigcap_{k=0}^{\infty} C^k_{\mathrm{b}}(I; X).$$

(ii) We denote by

$$C^k_\theta(I; X) = \{f \in C^k(I; X)\colon \|f\|_{k,\theta} < \infty\},$$

where

$$\|f\|_{k,\theta} = \sup_{\xi\in I} \sum_{j=0}^{k} \mathrm{e}^{-\theta|\xi|}|f^{(j)}(\xi)|,$$

the Banach space of all k-times continuously differentiable, exponentially weighted functions with weight $\mathrm{e}^{-\theta|\cdot|}$. Moreover, we define the subspaces

$$C^k_{\theta,\mathrm{e}}(I;X) = \{f \in C^k_\theta(I;X)\colon f(-\xi) = f(\xi),\ \xi \in I\}$$

and

$$C^k_{\theta,\mathrm{o}}(I;X) = \{f \in C^k_\theta(I)\colon f(-\xi) = -f(\xi),\ \xi \in I\}.$$

In the case $k = 0$ we just write $C_\theta(I;X)$, $C_{\theta,\mathrm{e}}(I;X)$ and $C_{\theta,\mathrm{o}}(I;X)$. If $X = \mathbb{R}$ we abbreviate the spaces $C^k_\theta(I;X)$, $C^k_{\theta,\mathrm{e}}(I;X)$, $C^k_{\theta,\mathrm{o}}(I;X)$ to respectively $C^k_\theta(I)$, $C^k_{\theta,\mathrm{e}}(I)$, $C^k_{\theta,\mathrm{o}}(I)$.

(iii) We denote by $C^k_{\mathrm{per}}([0,2\pi];X)$ the Banach space of k-times continuously differentiable, 2π-periodic functions $\mathbb{R} \to X$ and equip it with the norm

$$\|f\|_{C^k_{\mathrm{per}}([0,2\pi];X)} = \sum_{j=0}^{k} \sup_{\tau\in[0,2\pi]} \|f^{(j)}(\tau)\|_X.$$

In the case $k = 0$ we just write $C_{\mathrm{per}}([0,2\pi];X)$. As usual we define

$$C^\infty_{\mathrm{per}}([0,2\pi];X) = \bigcap_{k=0}^{\infty} C^k_{\mathrm{per}}([0,2\pi];X).$$

Lebesgue L^p-spaces

Definition 1.14. Suppose that X is a complex Banach space, I an interval and $p \geq 1$.

(i) We denote by

$$L^p(I;X) = \{f\colon I \to X\colon \|f\|_{L^p} < \infty\},$$

where

$$\|f\|_{L^p} = \Big(\int_I \|f(\xi)\|_X^p \,\mathrm{d}\xi\Big)^{\frac{1}{p}},$$

the Banach space of (equivalence classes of) p-integrable functions $I \to X$. Similarly, we denote the space of all 2π-periodic functions $\mathbb{R} \to X$ which belong to $L^p(0,2\pi;X)$ by $L^p_{\mathrm{per}}(0,2\pi;X)$; it is a Banach space with respect to the $L^p(0,2\pi;X)$-norm.

(ii) We denote by

$$L^\infty(I;X) = \{f\colon I \to X\colon \|f\|_{L^\infty} < \infty\},$$

where

$$\|f\|_{L^\infty} = \operatorname*{ess\,sup}_{\xi\in I} |f(\xi)|,$$

the Banach space of (equivalence classes of) essentially bounded functions $I \to X$. Similarly, we denote the space of all 2π-periodic functions $\mathbb{R} \to X$ which belong to $L^\infty(0,2\pi;X)$ by $L^\infty_{\mathrm{per}}(0,2\pi;X)$; it is a Banach space with respect to the $L^\infty(0,2\pi;X)$-norm.

Remark 1.15. The space $L^2_{\text{per}}(0, 2\pi; \mathbb{C})$ is a Hilbert space with respect to the inner product

$$\langle u, v\rangle_{L^2_{\text{per}}} = \int_0^{2\pi} u(\tau)\overline{v(\tau)}\, \mathrm{d}\tau.$$

The Fourier series expansion

$$u(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} u_m \mathrm{e}^{\mathrm{i}m\tau}, \quad u_m = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} u(\tau)\mathrm{e}^{-\mathrm{i}m\tau}\, \mathrm{d}\tau,$$

of a function $u \in L^2_{\text{per}}(0, 2\pi; \mathbb{C})$ satisfies

$$\|u\|^2_{L^2_{\text{per}}} = \sum_{m\in\mathbb{Z}} |u_m|^2$$

(Parseval's identity).

Sobolev spaces of periodic functions

We consider Sobolev spaces of periodic functions in one variable and briefly summarise some useful results concerning their norms and topologies (see Kress [18, Chapter 8]).

Definition 1.16. Let $t \in \mathbb{R}$. We denote by

$$H^t_{\text{per}}(0, 2\pi; \mathbb{C}) = \{u \in L^2_{\text{per}}(0, 2\pi; \mathbb{C})\colon \|u\|^2_t < \infty\},$$

where

$$\|u\|^2_t = \sum_{m\in\mathbb{Z}} (1 + m^2)^t |u_m|^2, \quad u(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} u_m \mathrm{e}^{\mathrm{i}m\tau},$$

the Sobolev space of order t of 2π-periodic functions. The (real) subspace of real-valued functions in $H^t_{\text{per}}(0, 2\pi; \mathbb{C})$ consists of the functions u which additionally satisfy the reality condition $u_{-m} = \overline{u}_m$ for each $m \in \mathbb{Z}$. In the case $t = 0$ we note that $H^0_{\text{per}}(0, 2\pi; \mathbb{C}) = L^2_{\text{per}}(0, 2\pi; \mathbb{C})$.

Proposition 1.17.

(i) The space $H^t_{\text{per}}(0, 2\pi; \mathbb{C})$ is a Banach algebra for $t > \frac{1}{2}$.

(ii) Let $t \in \mathbb{R}$. The space $H^t_{\text{per}}(0, 2\pi; \mathbb{C})$ is a Hilbert space with respect to the inner product

$$\langle u, v\rangle_t = \sum_{m\in\mathbb{Z}} (1 + m^2)^t u_m \overline{v}_m.$$

Its dual space is identified with $H^{-t}_{\text{per}}(0, 2\pi; \mathbb{C})$ by the duality pairing

$$\langle u, v\rangle_{H^t_{\text{per}}, H^{-t}_{\text{per}}} = \frac{1}{2\pi} \int_0^{2\pi} u(\tau)\overline{v(\tau)}\, \mathrm{d}\tau.$$

Remark 1.18. For $t \in \mathbb{N}_0$ the space $H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C})$ coincides with the completion of $C^\infty_{\mathrm{per}}([0, 2\pi]; \mathbb{C})$ with respect to the norm

$$\|u\|_t = \Big(\sum_{k=0}^{t} \|\partial_\tau^k u\|_0^2 \Big)^{\frac{1}{2}}.$$

This observation shows in particular that ∂_τ is a bounded linear operator $H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C}) \to H^{t-1}_{\mathrm{per}}(0, 2\pi; \mathbb{C})$ for $t \in \mathbb{N}$.

Definition 1.19. Let $t \in \mathbb{R}$ and $n \in \mathbb{N}$. The space $H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C}^n)$ is defined by componentwise extension of the above definitions .

The following technical result involving both classical and Sobolev spaces is used in Section 4.1.

Proposition 1.20. *Suppose that* $u \in C^k\big([0, T]; H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C})\big)$ *is given by*

$$u(\xi, \tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} u_m(\xi) \mathrm{e}^{\mathrm{i}m\tau}.$$

It follows that $u_m \in C^k([0, T]; \mathbb{C})$ *and* $\partial_\xi^l u \in C^{k-l}\big([0, T]; H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C})\big)$ *with*

$$\partial_\xi^l u(\xi, \tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} u_m^{(l)}(\xi) \mathrm{e}^{\mathrm{i}m\tau}$$

for each $l = 1, \ldots, k$.

Proof. Note that the functions $u_m \colon [0, T] \to \mathbb{C}$ are given by

$$u_m(\xi) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} u(\xi, \tau) \mathrm{e}^{-\mathrm{i}m\tau} \, \mathrm{d}\tau = \big\langle u(\xi), \mathrm{e}^{\mathrm{i}m(\cdot)} \big\rangle_{L^2_{\mathrm{per}}}$$

The continuity of the inner product on the right-hand side implies that we can interchange limits with respect to ξ and the integration with respect to τ. In particular we have that $u_m \in C^k([0, T]; \mathbb{C})$ with

$$u_m'(\xi) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \partial_\xi u(\xi, \tau) \mathrm{e}^{-\mathrm{i}m\tau} \, \mathrm{d}\tau = \big\langle \partial_\xi u(\xi), \mathrm{e}^{\mathrm{i}m(\cdot)} \big\rangle_{L^2_{\mathrm{per}}}.$$

On the other hand, since $\partial_\xi u \in C\big([0, T]; H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C})\big)$ we have that

$$\partial_\xi u(\xi, \tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} v_m(\xi) \mathrm{e}^{\mathrm{i}m\tau},$$

where $v_m \colon [0, T] \to \mathbb{C}$ is given by

$$v_m(\xi) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \partial_\xi u(\xi, \tau) \mathrm{e}^{-\mathrm{i}m\tau} \, \mathrm{d}\tau = \big\langle \partial_\xi u(\xi), \mathrm{e}^{\mathrm{i}m(\cdot)} \big\rangle_{L^2_{\mathrm{per}}} = u_m'(\xi).$$

It follows that

$$\partial_\xi u(\xi,\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} u_m'(\xi)\mathrm{e}^{\mathrm{i}m\tau}$$

and thus inductively

$$\partial_\xi^l u(\xi,\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} u_m^{(l)}(\xi)\mathrm{e}^{\mathrm{i}m\tau}, \quad l = 1,\dots,k.$$

■

Definition 1.21. Suppose that X is a complex Banach space, I an interval and $t \in \mathbb{N}_0$. The space $H^t(I;X)$ is the completion of $C^\infty(I;X)$ with respect to the norm

$$\|u\|_{H^t(I;X)} = \left(\sum_{k=0}^{t} \int_I \|\partial_\tau^k u\|_X^2\right)^{\frac{1}{2}}.$$

2 Mathematical formulation

We consider the equation

$$v_\xi = A^\varepsilon(v)v_\tau + B^\varepsilon(v), \tag{2.1}$$

where $v(\xi,\tau) \in \mathbb{R}^n$ is 2π-periodic in its second argument and A, B are respectively matrix- and vector-valued functions of (ε, v) defined in a neighbourhood of the origin in $\mathbb{R}\times\mathbb{R}^n$. We refer to equation (2.1) as *semilinear* if $A^\varepsilon(x)$ is a constant function of x, that is $A^\varepsilon(x) = A^\varepsilon(0)$, and *quasilinear* otherwise. The hypotheses on the functions A and B are as follows.

Assumptions 1.

(i) The matrix-valued function $A\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^{n\times n}$ is analytic at the origin.

The function $B\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ is analytic at the origin with $B^\varepsilon(0) = 0$.

Both A and B are even functions of ε.

(ii) The matrix $A^\varepsilon(0)$ is strictly hyperbolic, that is it has n distinct real eigenvalues.

(iii) There exists a matrix-valued function $S\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^{n\times n}$ such that S is analytic at the origin and $S^\varepsilon(x)$, $S^\varepsilon(x)A^\varepsilon(x)$ are symmetric with

$$|y|^2 \lesssim \langle S^\varepsilon(x)y, y\rangle \lesssim |y|^2, \quad y \in \mathbb{R}^n,$$

uniformly over (ε, x) in a neighbourhood of the origin in $\mathbb{R}\times\mathbb{R}^n$.

This assumption is not needed if A^ε is constant. In this case we set

$$S^\varepsilon(x) = S^\varepsilon(0) = I$$

for notational simplicity below.

(iv) There exists a matrix $T \in \mathbb{R}^{n\times n}$ such that $T^2 = I$ and

$$A^\varepsilon(Tx)T = TA^\varepsilon(x), \quad B^\varepsilon(Tx) = -TB^\varepsilon(x), \quad S^\varepsilon(0)T = TS^\varepsilon(0).$$

The existence of the matrix T guarantees that equation (2.1) is *reversible*, that is invariant under the transformation $\xi \mapsto -\xi$, $v \mapsto Rv$, where the *reverser* R is defined by the formula

$$(Rv)(\tau) = Tv(-\tau).$$

A standard approach to find solutions to the quasilinear system (2.1) is the construction of *a priori* energy estimates, the proof of which requires the symmetrisation of the matrix $A^\varepsilon(v)$ via multiplication by the matrix $S^\varepsilon(v)$ (see Benzoni-Gavage & Serre [1, Chapter 10]). The symmetrisation clearly changes the spectrum of the linearisation of the right-hand side of equation (2.1), the structure of which plays an important role in our analysis. We therefore symmetrise the linear part of the equation as a preliminary step and later use a further near-identity symmetrisation.

We use the change of variable $u = S^\varepsilon(0)^{\frac{1}{2}}v$ in equation (2.1) and write the transformed equation in terms of its linear and nonlinear parts as

$$u_\xi = L^\varepsilon u + N^\varepsilon(u) \tag{2.2}$$

with

$$L^\varepsilon u = S^\varepsilon(0)^{\frac{1}{2}}A^\varepsilon(0)S^\varepsilon(0)^{-\frac{1}{2}}u_\tau + S^\varepsilon(0)^{\frac{1}{2}}\partial B^\varepsilon[0]S^\varepsilon(0)^{-\frac{1}{2}}u$$

and

$$\begin{aligned} N^\varepsilon(u) = S^\varepsilon(0)^{\frac{1}{2}}\Big(A^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u) - A^\varepsilon(0)\Big)S^\varepsilon(0)^{-\frac{1}{2}}u_\tau \\ + S^\varepsilon(0)^{\frac{1}{2}}\Big(B^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u) - \partial B^\varepsilon[0]S^\varepsilon(0)^{-\frac{1}{2}}u\Big), \end{aligned}$$

where $\partial B^\varepsilon[\cdot]$ denotes the partial derivative of B with respect to u. For notational simplicity we define

$$\tilde{A}^\varepsilon(u) = S^\varepsilon(0)^{\frac{1}{2}}A^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u)S^\varepsilon(0)^{-\frac{1}{2}}, \quad \tilde{B}^\varepsilon(u) = S^\varepsilon(0)^{\frac{1}{2}}B^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u) \tag{2.3}$$

so that

$$L^\varepsilon u = \tilde{A}^\varepsilon(0)u_\tau + \partial\tilde{B}^\varepsilon[0]u.$$

Remark 2.1.

(i) Due to Assumptions 1(ii) the matrix $\tilde{A}^\varepsilon(0) = S^\varepsilon(0)^{\frac{1}{2}}A^\varepsilon(0)S^\varepsilon(0)^{-\frac{1}{2}}$ is strictly hyperbolic since it is similar to the matrix $A^\varepsilon(0)$. Moreover $\tilde{A}^\varepsilon(0)$ is symmetric since $S^\varepsilon(0)A^\varepsilon(0)$, $S^\varepsilon(0)^{-\frac{1}{2}}$ are symmetric and

$$\tilde{A}^\varepsilon(0) = S^\varepsilon(0)^{\frac{1}{2}}A^\varepsilon(0)S^\varepsilon(0)^{-\frac{1}{2}} = S^\varepsilon(0)^{-\frac{1}{2}}S^\varepsilon(0)A^\varepsilon(0)S^\varepsilon(0)^{-\frac{1}{2}}.$$

(ii) Since the matrices T and $S^\varepsilon(0)$ commute, the transformed functions $\tilde{A}$ and $\tilde{B}$ satisfy

$$\tilde{A}^\varepsilon(Tx)T = T\tilde{A}^\varepsilon(x), \quad \tilde{B}^\varepsilon(Tx) = -T\tilde{B}^\varepsilon(x).$$

We now specify the relevant function spaces for the system (2.2). Define

$$\mathcal{X}^t = H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C}^n)$$

for arbitrary $t \geq 0$, so that in particular $L^\varepsilon \in \mathcal{L}(\mathcal{X}^{t+1}, \mathcal{X}^t)$. Note that a function $u \in \mathcal{X}^t$ admits a Fourier series expansion

$$u(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} u_m \mathrm{e}^{\mathrm{i}m\tau}, \tag{2.4}$$

where the coefficient vectors $u_m \in \mathbb{C}^n$ satisfy the reality condition $\overline{u}_m = u_{-m}$ for each $m \in \mathbb{Z}$ if u is real-valued. We therefore have that

$$\mathcal{X}^t = \bigoplus_{m \in \mathbb{Z}} E_m,$$

where

$$E_m = \{u\mathrm{e}^{\mathrm{i}m\tau} \colon u \in \mathbb{C}^n\}, \quad m \in \mathbb{Z}.$$

For later use, we also define $P_k \colon \mathcal{X}^t \to \mathcal{X}^t$, $k \in \mathbb{Z}$, to be the natural projection onto the subspace E_k, that is

$$P_k\Big(\frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} u_m \mathrm{e}^{\mathrm{i}m\tau}\Big) = \frac{1}{\sqrt{2\pi}} u_k \mathrm{e}^{\mathrm{i}k\tau}.$$

We study the nonlinear system (2.2) in the real subspace of real-valued functions in the Banach algebra $\mathcal{X}^s$ for a fixed integer $s > \frac{1}{2}$, so that the domain of the vector field on its right-hand side is a neighbourhood of the origin in the real subspace of real-valued functions in $\mathcal{X}^{s+1}$.

2.1 Spectral analysis

Spectrum of the linearised system

We first examine the spectrum of the linear operator L^ε on the right-hand side of equation (2.2). It follows from the expansion (2.4) that

$$(L^\varepsilon u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} L^\varepsilon_m (u_m \mathrm{e}^{\mathrm{i}m\tau}) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \Big(\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]\Big) u_m \mathrm{e}^{\mathrm{i}m\tau},$$

and we have that

$$\sigma(L^\varepsilon) = \bigcup_{m \in \mathbb{Z}} \sigma\Big(\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]\Big),$$

that is the spectrum of the operator L^ε is discrete. We refer to $(\lambda, u) \in \mathbb{C} \times \mathbb{C}^n$ as a *mode m eigenpair* of L^ε if

$$L^\varepsilon u\mathrm{e}^{\mathrm{i}m\tau} = \lambda u\mathrm{e}^{\mathrm{i}m\tau},$$

so that (λ, u) is an eigenpair of the matrix $\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]$. We make the following assumptions concerning the mode m eigenvalues.

Assumptions 2.

(i) The mode m eigenvalues, that is the eigenvalues of the matrix $\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]$, are geometrically simple for each $m \in \mathbb{Z}$.

(ii) There exists $\omega \geq 0$ such that a *non-semisimple collision* takes place in modes ± 1 for $\varepsilon = 0$ at $\pm\mathrm{i}\omega$, that is as $\varepsilon \downarrow 0$ two pairs of simple complex mode ± 1 eigenvalues of the operator L^ε collide to form geometrically simple and algebraically double eigenvalues $\pm\mathrm{i}\omega$, i.e. a Jordan block of the matrix $\pm\mathrm{i}\tilde{A}^0(0)+\partial\tilde{B}^0[0]$ (see Figure 2.1).

(iii) All further eigenvalues of L^ε are assumed to be semisimple. Moreover we suppose that $\pm\mathrm{i}\omega$ is not in resonance with any other eigenvalue of L^0 (that is no other eigenvalue of L^0 is an integer multiple of $\pm\mathrm{i}\omega$).

(iv) The number $-\omega$ is not an eigenvalue of the matrix $\tilde{A}^0(0)$.

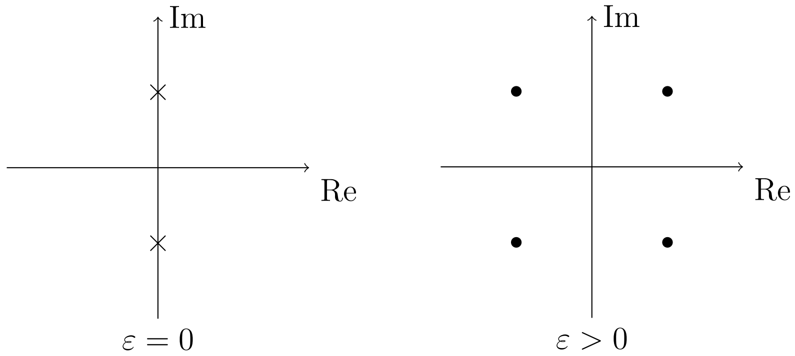

Figure 2.1: A reversible $(\mathrm{i}\omega)^2$ resonance takes place in modes ± 1 at $\varepsilon = 0$.

Differentiating the identity

$$\tilde{B}^\varepsilon(Tu) = -T\tilde{B}^\varepsilon(u)$$

with respect to u at 0 yields

$$\partial\tilde{B}^\varepsilon[0]T = -T\partial\tilde{B}^\varepsilon[0]$$

so that

$$(\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0])T = -T(-\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]).$$

For a mode m eigenpair (λ, u) we thus find that

$$L^\varepsilon u\mathrm{e}^{\mathrm{i}m\tau} = \lambda u\mathrm{e}^{\mathrm{i}m\tau}, \quad L^\varepsilon\overline{u}\mathrm{e}^{-\mathrm{i}m\tau} = \overline{\lambda}\overline{u}\mathrm{e}^{-\mathrm{i}m\tau}, \tag{2.5}$$

$$L^\varepsilon Tu\mathrm{e}^{-\mathrm{i}m\tau} = -\lambda Tu\mathrm{e}^{-\mathrm{i}m\tau}, \quad L^\varepsilon T\overline{u}\mathrm{e}^{\mathrm{i}m\tau} = -\overline{\lambda}T\overline{u}\mathrm{e}^{\mathrm{i}m\tau}. \tag{2.6}$$

In fact, if (λ, u) is a mode m eigenpair then $(-\overline{\lambda}, T\overline{u})$ is a mode m eigenpair and $(\overline{\lambda}, \overline{u})$, $(-\lambda, Tu)$ are mode $-m$ eigenpairs, so that the spectrum of L^ε is symmetric with respect to the real and imaginary axes as shown in Figure 2.2. Special cases arise when λ is real or purely imaginary. If $\lambda = \mathrm{i}\nu$ for some $\nu \in \mathbb{R}$ then $\mathrm{i}\nu$ is a mode m eigenvalue, $-\mathrm{i}\nu$ is a mode $-m$ eigenvalue and $\pm\mathrm{i}\nu$ are geometrically simple eigenvalues of L^ε.

If λ is real then λ and $-\lambda$ are both mode m and mode $-m$ eigenvalues and therefore geometrically double eigenvalues of L^ε if $m \neq 0$ and geometrically simple eigenvalues of L^ε if $m = 0$.

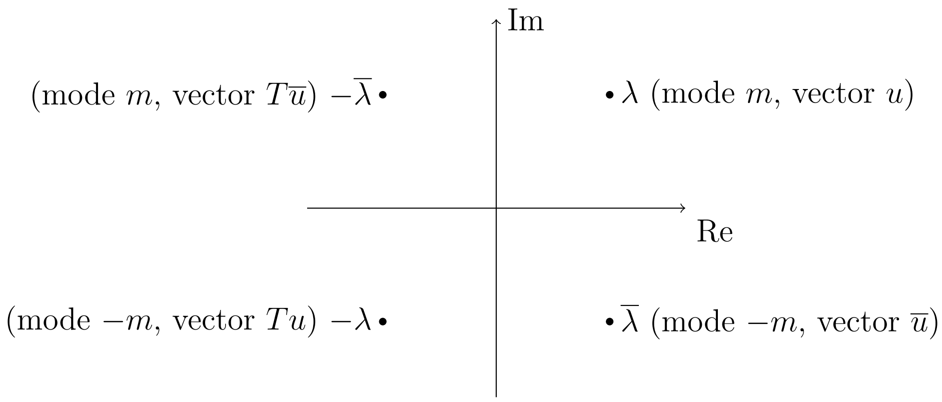

Figure 2.2: Symmetry in the spectrum of the operator L^ε.

In the next step we conclude from the above symmetry that the mode m eigenvalues of L^ε are purely imaginary for sufficiently large $|m|$ independently of ε. To this end we need the following result which states that the eigenvalues and eigenvectors of perturbed matrices depend analytically upon the perturbation parameter.

Proposition 2.2. *Suppose that $P^\varepsilon, Q^\varepsilon \in \mathbb{C}^{d\times d}$ depend analytically upon ε in a neighbourhood of the origin and the matrix P^0 has d simple eigenvalues $\lambda^0_{0,1}, \dots, \lambda^0_{0,d}$.*

(i) For all values of (ε, δ) in a sufficiently small neighbourhood Λ of the origin in $\mathbb{R}^2$ the matrix $K^\varepsilon_\delta = P^\varepsilon + \delta Q^\varepsilon$ has d simple eigenvalues $\lambda^\varepsilon_{\delta,1}, \dots, \lambda^\varepsilon_{\delta,d}$ which depend analytically upon $(\varepsilon, \delta) \in \Lambda$. Furthermore, the corresponding normalised eigenvectors $v^\varepsilon_{\delta,1}, \dots, v^\varepsilon_{\delta,d}$ also depend analytically upon $(\varepsilon, \delta) \in \Lambda$, so that the matrix

$$Q^\varepsilon_\delta = (v^\varepsilon_{\delta,1}|\cdots|v^\varepsilon_{\delta,d})$$

depends analytically upon $(\varepsilon, \delta) \in \Lambda$ as an element of $\mathcal{L}(\mathbb{C}^d)$.

(ii) The matrix Q^ε_δ is invertible and its inverse $(Q^\varepsilon_\delta)^{-1} \in \mathcal{L}(\mathbb{C}^d)$ depends analytically upon $(\varepsilon, \delta) \in \Lambda$.

Proof.

(i) Denote the normalised eigenvectors corresponding to the simple eigenvalues $\lambda^0_{0,1}, \dots, \lambda^0_{0,d}$ of P^0 by $v^0_{0,1}, \dots, v^0_{0,d}$ and define $F\colon \mathbb{C}^{d+1} \times \mathbb{R}^2 \to \mathbb{C}^d \times \mathbb{R}$ by the formula

$$F(v, \lambda, \delta, \varepsilon) = \begin{pmatrix} K^\varepsilon_\delta v - \lambda v \\ |v|^2 - 1 \end{pmatrix}.$$

Note that for fixed $i \in \{1, \dots, d\}$ we have that

$$F(v^0_{0,i}, \lambda^0_{0,i}, 0, 0) = 0$$

and

$$\mathrm{d}_{1,2}F[v,\lambda,\delta,\varepsilon]\begin{pmatrix} w \\ \eta \end{pmatrix} = \begin{pmatrix} K^\varepsilon_\delta w - \lambda w - \eta v \\ 2\operatorname{Re}\langle v, w\rangle \end{pmatrix},$$

so that in particular

$$\mathrm{d}_{1,2}F[v^0_{0,i},\lambda^0_{0,i},0,0] = \begin{pmatrix} K^0_0 - \lambda^0_{0,i} I & -v^0_{0,i} \\ 2\operatorname{Re}\langle v^0_{0,i}, \cdot\rangle & 0 \end{pmatrix}.$$

We show that this matrix is invertible by contradiction. Suppose its kernel is non-trivial, i.e.

$$\mathrm{d}_{1,2}F[v^0_{0,i},\lambda^0_{0,i},0,0]\begin{pmatrix} w \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

for some $(w,\eta) \neq (0,0)$. It follows that

$$(K^0_0 - \lambda^0_{0,i} I)w = \eta v^0_{0,i}, \quad \operatorname{Re}\langle v^0_{0,i}, w\rangle = 0,$$

so that $w \neq 0$ (since $w = 0$ implies $\eta = 0$). Hence w is a generalised eigenvector to the eigenvalue $\lambda^0_{0,i}$ of $K^0_0 = P^0$, which contradicts the fact that $\lambda^0_{0,i}$ is a simple eigenvalue.

The existence and analytic dependence upon δ and ε of $\lambda^\varepsilon_{\delta,1},\ldots,\lambda^\varepsilon_{\delta,d}$ and $v^\varepsilon_{\delta,1},\ldots,v^\varepsilon_{\delta,d}$ now follow from the analytic implicit function theorem. Note that (if Λ is chosen sufficiently small) $\lambda^\varepsilon_{\delta,1},\ldots,\lambda^\varepsilon_{\delta,d}$ are distinct (and hence simple) since $\lambda^0_{0,1},\ldots,\lambda^0_{0,d}$ are distinct.

(ii) First we note that Q^ε_δ is invertible since $v^\varepsilon_{\delta,1},\ldots,v^\varepsilon_{\delta,d}$ are linearly independent. Define $G\colon \mathcal{L}(\mathbb{C}^d)\times\Lambda \to \mathcal{L}(\mathbb{C}^d)$ by

$$G(Q,\varepsilon,\delta) = Q^\varepsilon_\delta Q - I,$$

so that

$$G((Q^0_0)^{-1},0,0) = Q^0_0(Q^0_0)^{-1} - I = 0$$

and

$$\mathrm{d}_1 G[(Q^0_0)^{-1},0,0](\tilde{Q}) = Q^0_0\tilde{Q},$$

that is $\mathrm{d}_1 G[(Q^0_0)^{-1},0,0]\colon \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^d)$ is invertible since Q^0_0 is invertible. By the analytic implicit function theorem the equation

$$G(Q,\varepsilon,\delta) = 0$$

has a unique solution $Q = Q(\varepsilon,\delta)$ which depends analytically upon $(\varepsilon,\delta)\in\Lambda$ (if Λ is chosen sufficiently small). However the solution to this equation is $Q = (Q^\varepsilon_\delta)^{-1}$.

■

In particular it follows from Proposition 2.2 that for sufficiently large values of $|m|$ (independently of ε) the eigenvalues and normalised eigenvectors of the matrix $\mathrm{i}m\tilde{A}^\varepsilon(0)+\partial\tilde{B}^\varepsilon[0]$ depend analytically upon ε in a neighbourhood of the origin which is independent of m. For the remaining finitely many values of m the same statement follows due to the perturbation theory given by Kato [15, II-§1]. The eigenvalues and normalised eigenvectors of the operator L^ε therefore also depend analytically upon ε in a neighbourhood of the origin.

Proposition 2.3. *The number of eigenvalues of the operator L^ε with non-zero real part, excluding the analytic perturbations of $\pm\mathrm{i}\omega$ in modes ± 1 (see Assumptions 2(ii)), is finite and does not depend upon ε.*

Proof. Denote by respectively $\lambda_1^\varepsilon(m),\dots,\lambda_n^\varepsilon(m)$ and $\mathrm{i}\mu_1^\varepsilon,\dots,\mathrm{i}\mu_n^\varepsilon$ the eigenvalues of the matrices $\mathrm{i}\tilde{A}^\varepsilon(0)+\frac{1}{m}\partial\tilde{B}^\varepsilon[0]$ and $\mathrm{i}\tilde{A}^\varepsilon(0)$. It follows from Proposition 2.2 that

$$\lambda_j^\varepsilon(m)\to\mathrm{i}\mu_j^\varepsilon,\quad j=1,\dots,n,$$

as $|m|\to\infty$ uniformly over ε. Recall that $\operatorname{Im}(\mu_j^\varepsilon)=0$ for $j=1,\dots,n$ (Assumptions 1(ii)) and that $\mu_j^\varepsilon\neq\mu_k^\varepsilon$ if $j\neq k$ for all values of ε (see Remark 2.1(i)). Hence there exist $M\in\mathbb{N}$ and $\kappa>0$, both independent of ε, with the properties that

$$B_\kappa(\mathrm{i}\mu_j^\varepsilon)\cap B_\kappa(\mathrm{i}\mu_k^\varepsilon)=\emptyset,\quad j\neq k,$$

and $\lambda_j^\varepsilon(m)\in B_\kappa(\mathrm{i}\mu_j^\varepsilon)$ for $|m|\geq M$. The matrix $\mathrm{i}\tilde{A}^\varepsilon(0)+\frac{1}{m}\partial\tilde{B}^\varepsilon[0]$ therefore has n distinct eigenvalues $\lambda_1^\varepsilon(m),\dots,\lambda_n^\varepsilon(m)$ for $m\geq M$. We know that for fixed $m\in\mathbb{Z}\setminus\{0\}$ the spectrum of the matrix $\mathrm{i}m\tilde{A}^\varepsilon(0)+\partial\tilde{B}^\varepsilon[0]$, and hence of $\mathrm{i}\tilde{A}^\varepsilon(0)+\frac{1}{m}\partial\tilde{B}^\varepsilon[0]$, is symmetric with respect to the imaginary axis. We conclude that $\operatorname{Re}(\lambda_j^\varepsilon(m))=0$, for otherwise the symmetry of the spectrum would imply that $-\overline{\lambda}_j^\varepsilon(m)$ is a second eigenvalue of $\mathrm{i}\tilde{A}^\varepsilon(0)+\frac{1}{m}\partial\tilde{B}^\varepsilon[0]$ in $B_\kappa(\mathrm{i}\mu_j^\varepsilon)$ which is a contradiction.

Because of

$$\sigma\Big(\mathrm{i}m\tilde{A}^\varepsilon(0)+\partial\tilde{B}^\varepsilon[0]\Big)=m\sigma\Big(\mathrm{i}\tilde{A}^\varepsilon(0)+\frac{1}{m}\partial\tilde{B}^\varepsilon[0]\Big),$$

we conclude that

$$\bigcup_{|m|>M}\sigma\Big(\mathrm{i}m\tilde{A}^\varepsilon(0)+\partial\tilde{B}^\varepsilon[0]\Big)$$

contains only purely imaginary eigenvalues. Thus

$$\sigma(L^\varepsilon)=\bigcup_{m\in\mathbb{Z}}\sigma\Big(\mathrm{i}m\tilde{A}^\varepsilon(0)+\partial\tilde{B}^\varepsilon[0]\Big)$$

contains only finitely many eigenvalues with non-zero real part.

Since all eigenvalues, with the exception of $\pm\mathrm{i}\omega$ in modes ±1 at $\varepsilon=0$, are algebraically simple (Assumptions 2(i) and 2(iii)), there are no additional (necessarily non-semisimple) eigenvalue collisions on the imaginary axis as ε is varied. The number of eigenvalues of L^ε with non-zero real part, excluding the analytic perturbations of $\pm\mathrm{i}\omega$, is therefore independent of ε. ■

The operator L^ε is a 'direct sum' of mode m operators L^ε_m defined on E_m by

$$L^\varepsilon_m(u_m \mathrm{e}^{\mathrm{i}m\tau}) = \Big(\mathrm{i}m\tilde{A}^\varepsilon(0) + \partial\tilde{B}^\varepsilon[0]\Big)u_m \mathrm{e}^{\mathrm{i}m\tau},$$

where L^ε and L^ε_m depend analytically upon ε due to Assumptions 1(i). In the spectral analysis below we construct the spectral projections for L^ε as a similar direct sum of the spectral projections for L^ε_m. To establish their analytic dependence upon ε we use the following result whose proof is given in Section 2.4.

Lemma 2.4. *Suppose that $T^\varepsilon_m \in \mathcal{L}(\mathbb{C}^d)$ depends analytically upon ε in a neighbourhood of the origin uniformly over $m \in \mathbb{Z}$. The formula*

$$(T^\varepsilon u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} (T^\varepsilon_m u_m)\mathrm{e}^{\mathrm{i}m\tau}$$

defines an operator $T^\varepsilon \in \mathcal{L}(\mathcal{X}^t)$ for all $t \geq 0$ which depends analytically upon ε in the same neighbourhood of the origin.

Formulation as a coupled system

We conclude from Proposition 2.3 that the spectrum of L^ε decomposes into a finite subset of *hyperbolic eigenvalues* with non-zero real part and an infinite subset of purely imaginary *centre eigenvalues*. Our aim is to write equation (2.2) as a coupled system by applying the corresponding spectral projections.

We call the semisimple analytic perturbations of the eigenvalues $\mathrm{i}\omega$ of $L^0|_{E_1}$ and $-\mathrm{i}\omega$ of $L^0|_{E_{-1}}$ (see Assumptions 2(ii)) the *weakly hyperbolic eigenvalues*, since they are complex for $\varepsilon > 0$ and become purely imaginary for $\varepsilon = 0$. The weakly hyperbolic eigenvalues form respectively an $(\mathrm{i}\omega)$-group and a $(-\mathrm{i}\omega)$-group in the sense of Kato [15, pp. 65–66], so that the corresponding spectral projections $P^{1,\varepsilon}_{\mathrm{wh}}$ and $P^{-1,\varepsilon}_{\mathrm{wh}}$ are analytic at $\varepsilon = 0$. The remaining (finitely many) eigenvalues which have non-zero real part for all ε are called *strongly hyperbolic eigenvalues*. For $m \in \mathbb{Z}$ the respective spectral projections $P^{m,\varepsilon}_{\mathrm{c}}\colon E_m \to E_m$ and $P^{m,\varepsilon}_{\mathrm{sh}}\colon E_m \to E_m$ associated with the centre eigenvalues of $L^\varepsilon|_{E_m}$ and the strongly hyperbolic eigenvalues of $L^\varepsilon|_{E_m}$ depend analytically upon ε. Since Proposition 2.3 implies that $P^{m,\varepsilon}_{\mathrm{sh}} = P^{m,0}_{\mathrm{sh}} = 0$ and $P^{m,\varepsilon}_{\mathrm{c}} = P^{m,0}_{\mathrm{c}} = I$ for sufficiently large $|m|$, the analytic dependence upon ε is uniform over $m \in \mathbb{Z}$ and the formulae

$$(P^\varepsilon_{\mathrm{wh}}u)(\tau) = \frac{1}{\sqrt{2\pi}}\Big(P^{-1,\varepsilon}_{\mathrm{wh}}(u_{-1}\mathrm{e}^{-\mathrm{i}\tau}) + P^{1,\varepsilon}_{\mathrm{wh}}(u_1\mathrm{e}^{\mathrm{i}\tau})\Big),$$

$$(P^\varepsilon_{\mathrm{sh}}u)(\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}} P^{m,\varepsilon}_{\mathrm{sh}}(u_m\mathrm{e}^{\mathrm{i}m\tau}), \quad (P^\varepsilon_{\mathrm{c}}u)(\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}} P^{m,\varepsilon}_{\mathrm{c}}(u_m\mathrm{e}^{\mathrm{i}m\tau}) \tag{2.7}$$

define projections $P^\varepsilon_{\mathrm{wh}}, P^\varepsilon_{\mathrm{sh}}, P^\varepsilon_{\mathrm{c}}\colon \mathcal{X}^t \to \mathcal{X}^t$ for all $t \geq 0$ which depend analytically upon ε due to Lemma 2.4 and satisfy $P^\varepsilon_{\mathrm{wh}} + P^\varepsilon_{\mathrm{sh}} + P^\varepsilon_{\mathrm{c}} = I$. A direct calculation using the symmetry in equations (2.5), (2.6) shows that $P^{-1,\varepsilon}_{\mathrm{wh}} + P^{1,\varepsilon}_{\mathrm{wh}}$, $P^{-m,\varepsilon}_{\mathrm{sh}} + P^{m,\varepsilon}_{\mathrm{sh}}$ and $P^{-m,\varepsilon}_{\mathrm{c}} + P^{m,\varepsilon}_{\mathrm{c}}$ commute with the reverser R so that the same is true for $P^\varepsilon_{\mathrm{wh}}$, $P^\varepsilon_{\mathrm{sh}}$ and $P^\varepsilon_{\mathrm{c}}$.

We apply the projections $P^\varepsilon_{\mathrm{wh}}$ and $P^\varepsilon_{\mathrm{sh,c}} = P^\varepsilon_{\mathrm{sh}} + P^\varepsilon_{\mathrm{c}}$ to equation (2.2) and write

$$z^\varepsilon = P^\varepsilon_{\mathrm{wh}} u, \quad q^\varepsilon = P^\varepsilon_{\mathrm{sh,c}} u$$

to obtain the reversible coupled system

$$z^\varepsilon_\xi = L^\varepsilon z^\varepsilon + P^\varepsilon_{\mathrm{wh}} N^\varepsilon(z^\varepsilon + q^\varepsilon), \tag{2.8}$$

$$q^\varepsilon_\xi = L^\varepsilon q^\varepsilon + P^\varepsilon_{\mathrm{sh,c}} N^\varepsilon(z^\varepsilon + q^\varepsilon). \tag{2.9}$$

This formulation has the disadvantage that the spaces $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t$ and $P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^t$ depend upon the parameter ε. The ε-dependence is removed by applying a *Kato similarity transformation* U^ε with the property that

$$U^\varepsilon P^0_{\mathrm{wh}}(U^\varepsilon)^{-1} = P^\varepsilon_{\mathrm{wh}}, \quad U^\varepsilon P^0_{\mathrm{sh}}(U^\varepsilon)^{-1} = P^\varepsilon_{\mathrm{sh}}, \quad U^\varepsilon P^0_{\mathrm{c}}(U^\varepsilon)^{-1} = P^\varepsilon_{\mathrm{c}}.$$

The operator U^ε is constructed as the direct sum of mode m similarity transformations $U^\varepsilon_m \in \mathcal{L}(E_m)$ such that

$$P^{\pm1,\varepsilon}_{\mathrm{wh}} = U^\varepsilon_{\pm1} P^{\pm1,0}_{\mathrm{wh}} (U^\varepsilon_{\pm1})^{-1}, \quad P^{m,\varepsilon}_{\mathrm{sh}} = U^\varepsilon_m P^{m,0}_{\mathrm{sh}} (U^\varepsilon_m)^{-1}, \quad P^{m,\varepsilon}_{\mathrm{c}} = U^\varepsilon_m P^{m,0}_{c} (U^\varepsilon_m)^{-1}.$$

The analytic dependence of U^ε upon ε is deduced from Proposition 2.2. The construction of the similarity transformations U^ε_m is based on the following result deduced from Kato [15, II-§4]. Its proof is presented in Section 2.4.

Lemma 2.5. *Let X be a Banach space, $J \in \mathcal{L}(X)$ with the property that $J^2 = I$, and $P^\varepsilon_1, \dots, P^\varepsilon_l \in \mathcal{L}(X)$ be projections which depend analytically upon ε in a neighbourhood of the origin and satisfy*

$$P^\varepsilon_i J = J P^\varepsilon_i, \quad P^\varepsilon_i P^\varepsilon_j = \delta_{ij} P^\varepsilon_i, \qquad i, j = 1, \dots, l,$$

and

$$\sum_{i=1}^{l} P^\varepsilon_i = I.$$

There exists $U^\varepsilon \in \mathcal{L}(X)$ which is invertible such that

$$U^\varepsilon P^0_i (U^\varepsilon)^{-1} = P^\varepsilon_i, \quad i = 1, \dots, l.$$

Furthermore, U^ε, $(U^\varepsilon)^{-1}$ depend analytically upon ε and $U^\varepsilon J = J U^\varepsilon$.

For each $m \in \mathbb{Z}$ Lemma 2.5 guarantees the existence of $U^\varepsilon_m \in \mathcal{L}(E_m)$ which depends analytically upon ε, commutes with the reverser R, and satisfies

$$P^{\pm1,\varepsilon}_{\mathrm{wh}} = U^\varepsilon_{\pm1} P^{\pm1,0}_{\mathrm{wh}} (U^\varepsilon_{\pm1})^{-1}, \quad P^{m,\varepsilon}_{\mathrm{sh}} = U^\varepsilon_m P^{m,0}_{\mathrm{sh}} (U^\varepsilon_m)^{-1}, \quad P^{m,\varepsilon}_{\mathrm{c}} = U^\varepsilon_m P^{m,0}_{c} (U^\varepsilon_m)^{-1},$$

where the inverse $(U^\varepsilon_m)^{-1}$ also depends analytically upon ε (in the same neighbourhood of the origin as $P^{\pm1,\varepsilon}_{\mathrm{wh}}$, $P^{m,\varepsilon}_{\mathrm{sh}}$ and $P^{m,\varepsilon}_{\mathrm{c}}$). Note that the analytic dependence is uniform over $m \in \mathbb{Z}$, since $P_{\mathrm{sh}}^{m,\varepsilon} = P_{\mathrm{sh}}^{m,0} = 0$ and $P_{\mathrm{c}}^{m,\varepsilon} = P_{\mathrm{c}}^{m,0} = I$ for sufficiently large values of $|m|$. By Lemma 2.4 the formula

$$U^\varepsilon v = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} U_m^\varepsilon(v_m \mathrm{e}^{\mathrm{i}m\tau})$$

defines an invertible linear operator $U^\varepsilon \in \mathcal{L}(\mathcal{X}^t)$ for all $t \geq 0$ which depends analytically upon ε, commutes with the reverser R, and whose inverse is given by

$$(U^\varepsilon)^{-1} w = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} (U_m^\varepsilon)^{-1}(w_m \mathrm{e}^{\mathrm{i}m\tau}).$$

We conclude that $P_{\mathrm{wh}}^\varepsilon = U^\varepsilon P_{\mathrm{wh}}^0 (U^\varepsilon)^{-1}$, $P_{\mathrm{sh}}^\varepsilon = U^\varepsilon P_{\mathrm{sh}}^0 (U^\varepsilon)^{-1}$ and $P_{\mathrm{c}}^\varepsilon = U^\varepsilon P_{\mathrm{c}}^0 (U^\varepsilon)^{-1}$. To simplify the notation we write P_{wh}, P_{sh}, P_{c} instead of P_{wh}^0, P_{sh}^0, P_{c}^0. The finite-dimensional subspaces $\mathcal{X}_{\mathrm{wh}} = P_{\mathrm{wh}}^0 \mathcal{X}^t$ and $\mathcal{X}_{\mathrm{sh}} = P_{\mathrm{sh}}^0 \mathcal{X}^t$ (which are equal for all values of $t \geq 0$) are called respectively the *weakly hyperbolic subspace* and the *strongly hyperbolic subspace* of L^0. The subspace $\mathcal{X}_{\mathrm{c}}^t = P_{\mathrm{c}}^0 \mathcal{X}^t$ (of regularity $t \geq 0$) is called the *centre subspace* of L^0. Moreover we define $\mathcal{X}_{\mathrm{h}} = \mathcal{X}_{\mathrm{wh}} \oplus \mathcal{X}_{\mathrm{sh}}$, $\mathcal{X}_{\mathrm{sh,c}}^t = \mathcal{X}_{\mathrm{sh}} \oplus \mathcal{X}_{\mathrm{c}}^t$ as well as $P_{\mathrm{h}}^\varepsilon = P_{\mathrm{wh}}^\varepsilon + P_{\mathrm{sh}}^\varepsilon$, $P_{\mathrm{sh,c}}^\varepsilon = P_{\mathrm{sh}}^\varepsilon + P_{\mathrm{c}}^\varepsilon$ and analogously write P_{h}, $P_{\mathrm{sh,c}}$ instead of P_{h}^0, $P_{\mathrm{sh,c}}^0$.

Remark 2.6. The spectral projection $P_{\mathrm{c}}^{m,0}$ is given by

$$P_{\mathrm{c}}^{m,0} v = \frac{1}{2\pi\mathrm{i}} \int_{\gamma_{\mathrm{c}}^m} (\lambda I - L^0)^{-1}\, \mathrm{d}\lambda,$$

where γ_{c}^m is a closed, positively-orientated curve enclosing the purely imaginary eigenvalues of the operator L_m^0 (see Kato [15, II-§1.4]). Its image is given by the linear span of the corresponding eigenvectors. It thus follows from the definition (2.7) that the space $\mathcal{X}_{\mathrm{c}}^t$ is the completion (with respect to the H_{per}^t-norm) of the set of all eigenvectors of the operators L_m^0, for all $m \in \mathbb{Z}$, which correspond to purely imaginary eigenvalues.

In the theory below we work under the additional assumption that the four-dimensional space $P_{\mathrm{wh}}^\varepsilon \mathcal{X}^t$ is independent of ε, that is $P_{\mathrm{wh}}^\varepsilon \mathcal{X}^t = P_{\mathrm{wh}}^0 \mathcal{X}^t = \mathcal{X}_{\mathrm{wh}}$ for all ε (see Assumptions 3), so that equation (2.8) is posed in the fixed, ε-independent space $\mathcal{X}_{\mathrm{wh}}$. To further formulate equation (2.9) in the fixed space $\mathcal{X}_{\mathrm{sh,c}}^s$ we write $q^\varepsilon = U^\varepsilon q$ with $q \in \mathcal{X}_{\mathrm{sh,c}}^{s+1}$. Writing z^ε as z for notational simplicity, we thus obtain the reversible coupled system

$$z_\xi = L_{\mathrm{wh}}^\varepsilon z + f^\varepsilon(z,q), \tag{2.10}$$

$$q_\xi = L_{\mathrm{sh,c}}^\varepsilon q + P_{\mathrm{sh,c}}\Big(g_1^\varepsilon(z,q) q_\tau\Big) + g_2^\varepsilon(z,q) + h^\varepsilon(z), \tag{2.11}$$

with

$$L_{\mathrm{wh}}^\varepsilon = L^\varepsilon|_{\mathcal{X}_{\mathrm{wh}}}, \qquad L_{\mathrm{sh,c}}^\varepsilon = (U^\varepsilon)^{-1} L^\varepsilon U^\varepsilon|_{\mathcal{X}_{\mathrm{sh,c}}^{t+1}} \quad (t \geq 0)$$

and the nonlinearities

$$f^\varepsilon(z,q) = P_{\mathrm{wh}}^\varepsilon\Big(\tilde{A}^\varepsilon(z+U^\varepsilon q)(z+U^\varepsilon q)_\tau - \tilde{A}^\varepsilon(0) z_\tau + \tilde{B}^\varepsilon(z+U^\varepsilon q) - \partial\tilde{B}^\varepsilon[0]z\Big),$$

$$g_1^\varepsilon(z,q) = (U^\varepsilon)^{-1}\Big(\tilde{A}^\varepsilon(z+U^\varepsilon q) - \tilde{A}^\varepsilon(0)\Big)U^\varepsilon,$$

$$g_2^\varepsilon(z,q) = P_{\mathrm{sh,c}}(U^\varepsilon)^{-1}\Big((\tilde{A}^\varepsilon(z+U^\varepsilon q) - \tilde{A}^\varepsilon(z))z_\tau + \tilde{B}^\varepsilon(z+U^\varepsilon q) - \tilde{B}^\varepsilon(z) - \partial\tilde{B}^\varepsilon[0]U^\varepsilon q\Big),$$

$$h^\varepsilon(z) = P_{\mathrm{sh,c}}(U^\varepsilon)^{-1}\Big(\tilde{A}^\varepsilon(z) z_\tau + \tilde{B}^\varepsilon(z)\Big).$$

Note that equation (2.11) is stated in the fixed, ε-independent space $\mathcal{X}^s_{\mathrm{sh,c}}$, the domain of its right-hand side being $\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^{s+1}_{\mathrm{sh,c}}$. For later use we also define

$$L^\varepsilon_{\mathrm{sh}} = (U^\varepsilon)^{-1} L^\varepsilon U^\varepsilon|_{\mathcal{X}_{\mathrm{sh}}}, \qquad L^\varepsilon_{\mathrm{c}} = (U^\varepsilon)^{-1} L^\varepsilon U^\varepsilon|_{\mathcal{X}^{t+1}_{\mathrm{c}}} \quad (t \geq 0)$$

as well as $R_{\mathrm{wh}} = R|_{\mathcal{X}_{\mathrm{wh}}}$, $R_{\mathrm{sh}} = R|_{\mathcal{X}_{\mathrm{sh}}}$, $R_{\mathrm{c}} = R|_{\mathcal{X}^t_{\mathrm{c}}}$. For all $t > \frac{1}{2}$ the nonlinearities in equations (2.10), (2.11) are analytic at the origin and satisfy the estimates

$$\|f^\varepsilon(z,q)\|_{\mathcal{X}_{\mathrm{wh}}} = \mathcal{O}(\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \|(z,q)\|_{\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^t_{\mathrm{sh,c}}} + \|(z,q)\|^3_{\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^t_{\mathrm{sh,c}}}), \tag{2.12}$$

$$\|g^\varepsilon_1(z,q)\|_t = \mathcal{O}(\|(z,q)\|_{\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^t_{\mathrm{sh,c}}}), \tag{2.13}$$

$$\|g^\varepsilon_2(z,q)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \|(z,q)\|_{\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^t_{\mathrm{sh,c}}}), \tag{2.14}$$

$$\|h^\varepsilon(z)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(\|z\|^2_{\mathcal{X}_{\mathrm{wh}}}). \tag{2.15}$$

Coordinates

We end this section by introducing suitable coordinates for the spaces $\mathcal{X}_{\mathrm{wh}}$ and $\mathcal{X}^s_{\mathrm{sh,c}}$ in equations (2.10), (2.11) which are spanned by the eigenvectors of L^0. Due to the symmetry of the spectrum of L^0 illustrated in Figure 2.2 we can decompose the spaces $E_{-m} \oplus E_m$, $m \in \mathbb{N}_0$, (and hence the spaces $\mathcal{X}_{\mathrm{wh}}$ and $\mathcal{X}^s_{\mathrm{sh,c}}$) into a direct sum of respectively four- and two-dimensional L^0-invariant subspaces. We first give explicit representations of the operator L^0 restricted to these subspaces.

$\boldsymbol{\lambda \neq \pm \mathrm{i}\omega}$ and $\mathbf{Im}\,\boldsymbol{\lambda \neq 0}$ if $\boldsymbol{m = 0}$: Suppose that (λ, u) is a mode m eigenpair of L^0 with $\lambda \neq \pm \mathrm{i}\omega$ and that $\operatorname{Im}\lambda \neq 0$ if $m = 0$. If $\operatorname{Re}\lambda \neq 0$, equations (2.5), (2.6) show that $\{u\mathrm{e}^{\mathrm{i}m\tau}, \overline{u}\mathrm{e}^{-\mathrm{i}m\tau}, T\overline{u}\mathrm{e}^{\mathrm{i}m\tau}, Tu\mathrm{e}^{-\mathrm{i}m\tau}\}$ is a basis for the L^0-invariant subspace associated with the eigenvalue quartet $\pm\lambda, \pm\overline{\lambda}$, and the matrices of L^0 and R with respect to this basis are respectively

$$\begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \overline{\lambda} & 0 & 0 \\ 0 & 0 & -\overline{\lambda} & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

If $\operatorname{Re}\lambda = 0$, Assumptions 2(i) imply that $\overline{u}$ and Tu are linearly dependent (see equations (2.5), (2.6)). By replacing u by $\mathrm{i}(\overline{u} - Tu)$ if necessary, we may assume that $\overline{u} = Tu$. Thus $\{u\mathrm{e}^{\mathrm{i}m\tau}, \overline{u}\mathrm{e}^{-\mathrm{i}m\tau}\}$ is a basis for the L^0-invariant subspace associated with the eigenvalue pair $\pm\lambda$ and the matrices of L^0 and R with respect to this basis are respectively

$$\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

$\boldsymbol{m = 0}$ and $\mathbf{Im}\,\boldsymbol{\lambda = 0}$: Suppose that (λ, u) is a mode 0 eigenpair with $\operatorname{Im}\lambda = 0$. Assumptions 2(i) force u and $\overline{u}$ to be linearly dependent, and by replacing u by $\mathrm{i}(u - \overline{u})$ if necessary we may assume that $u = \overline{u}$. Thus $\{u, Tu\}$ is a basis for the real L^0-invariant subspace associated with the eigenvalue pair $\pm\lambda$ and the matrices of L^0 and R with respect to this basis are respectively

$$\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

$\boldsymbol{\lambda = \pm i\omega}$: Consider the geometrically simple, algebraically double eigenvalues $\pm\mathrm{i}\omega$ in modes ± 1 (see Assumptions 2(ii)). There exist $u, v \in \mathbb{C}^n$ with

$$L^0 u\mathrm{e}^{\mathrm{i}\tau} = \mathrm{i}\omega u\mathrm{e}^{\mathrm{i}\tau}, \quad L^0 v\mathrm{e}^{\mathrm{i}\tau} = \mathrm{i}\omega v\mathrm{e}^{\mathrm{i}\tau} + u\mathrm{e}^{\mathrm{i}\tau} \tag{2.16}$$

and

$$L^0 \overline{u}\mathrm{e}^{-\mathrm{i}\tau} = -\mathrm{i}\omega\overline{u}\mathrm{e}^{-\mathrm{i}\tau}, \quad L^0 \overline{v}\mathrm{e}^{-\mathrm{i}\tau} = -\mathrm{i}\omega\overline{v}\mathrm{e}^{-\mathrm{i}\tau} + \overline{u}\mathrm{e}^{-\mathrm{i}\tau}. \tag{2.17}$$

By symmetry we also have that

$$L^0 Tu\mathrm{e}^{-\mathrm{i}\tau} = -\mathrm{i}\omega Tu\mathrm{e}^{-\mathrm{i}\tau}, \quad L^0 Tv\mathrm{e}^{-\mathrm{i}\tau} = -\mathrm{i}\omega Tv\mathrm{e}^{-\mathrm{i}\tau} - Tu\mathrm{e}^{-\mathrm{i}\tau}$$

and since the eigenvalue $-\mathrm{i}\omega$ is geometrically simple, we find Tu and $\overline{u}$ to be linearly dependent. By replacing u and v with respectively $\mathrm{i}(u - T\overline{u})$ and $\mathrm{i}(v + T\overline{v})$ if necessary, we may assume that $Tu = \overline{u}$ and $Tv = -\overline{v}$. Hence $\{u\mathrm{e}^{\mathrm{i}\tau}, v\mathrm{e}^{\mathrm{i}\tau}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}, \overline{v}\mathrm{e}^{-\mathrm{i}\tau}\}$ is a basis for the weakly hyperbolic subspace $\mathcal{X}_{\mathrm{wh}}$ and the matrices of L^0 and R with respect to this basis are respectively

$$\begin{pmatrix} \mathrm{i}\omega & 1 & 0 & 0 \\ 0 & \mathrm{i}\omega & 0 & 0 \\ 0 & 0 & -\mathrm{i}\omega & 1 \\ 0 & 0 & 0 & -\mathrm{i}\omega \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Assumptions 2(ii) state that for $\varepsilon > 0$ the operator L^ε has a pair of simple complex mode 1 eigenvalues $\mathrm{i}(\omega + \nu^\varepsilon) \pm \lambda^\varepsilon$ with the corresponding complex conjugate mode -1 eigenvalues $-\mathrm{i}(\omega + \nu^\varepsilon) \pm \lambda^\varepsilon$. Here λ^ε and ν^ε are (real-valued) analytic functions of ε. We make the following assumptions on the qualitative behaviour of λ^ε, ν^ε and the form of the mode ± 1 eigenvectors for $\varepsilon > 0$.

Assumptions 3. The functions λ^ε and ν^ε satisfy

$$\lambda^\varepsilon = \varepsilon + \mathcal{O}(\varepsilon^2), \quad \nu^\varepsilon = \mathcal{O}(\varepsilon).$$

For $\varepsilon > 0$ the vectors $(u + \lambda^\varepsilon v)\mathrm{e}^{\mathrm{i}\tau}$, $(u - \lambda^\varepsilon v)\mathrm{e}^{\mathrm{i}\tau}$ and $(\overline{u} + \lambda^\varepsilon\overline{v})\mathrm{e}^{-\mathrm{i}\tau}$, $(\overline{u} - \lambda^\varepsilon\overline{v})\mathrm{e}^{-\mathrm{i}\tau}$ are mode ± 1 eigenvectors of L^ε with corresponding eigenvalues $\mathrm{i}(\omega + \nu^\varepsilon) \pm \lambda^\varepsilon$ and $-\mathrm{i}(\omega + \nu^\varepsilon) \pm \lambda^\varepsilon$, that is

$$\begin{aligned} L^\varepsilon(u\mathrm{e}^{\mathrm{i}\tau} \pm \lambda^\varepsilon v\mathrm{e}^{\mathrm{i}\tau}) &= (\mathrm{i}(\omega + \nu^\varepsilon) \pm \lambda^\varepsilon)(u\mathrm{e}^{\mathrm{i}\tau} \pm \lambda^\varepsilon v\mathrm{e}^{\mathrm{i}\tau}), \\ L^\varepsilon(\overline{u}\mathrm{e}^{-\mathrm{i}\tau} \pm \lambda^\varepsilon\overline{v}\mathrm{e}^{-\mathrm{i}\tau}) &= (-\mathrm{i}(\omega + \nu^\varepsilon) \pm \lambda^\varepsilon)(\overline{u}\mathrm{e}^{-\mathrm{i}\tau} \pm \lambda^\varepsilon\overline{v}\mathrm{e}^{-\mathrm{i}\tau}). \end{aligned}$$

It follows that the matrix of the operator $L_{\mathrm{wh}}^{\varepsilon}$ for $\varepsilon > 0$ with respect to the ordered basis $\{u\mathrm{e}^{\mathrm{i}\tau}, v\mathrm{e}^{\mathrm{i}\tau}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}, \overline{v}\mathrm{e}^{-\mathrm{i}\tau}\}$ for $\mathcal{X}_{\mathrm{wh}}$ is given by

$$L_{\mathrm{wh}}^{\varepsilon} = \begin{pmatrix} \mathrm{i}(\omega+\nu^{\varepsilon}) & 1 & 0 & 0 \\ (\lambda^{\varepsilon})^2 & \mathrm{i}(\omega+\nu^{\varepsilon}) & 0 & 0 \\ 0 & 0 & -\mathrm{i}(\omega+\nu^{\varepsilon}) & 1 \\ 0 & 0 & (\lambda^{\varepsilon})^2 & -\mathrm{i}(\omega+\nu^{\varepsilon}) \end{pmatrix}.$$

We conclude from Assumptions 3 and

$$T\overline{(u+\lambda^{\varepsilon}v)} = (u-\lambda^{\varepsilon}v)$$

that

$$\begin{aligned} L^{\varepsilon}T(u+\lambda^{\varepsilon}v)\mathrm{e}^{-\mathrm{i}\tau} &= -(\mathrm{i}(\omega+\nu^{\varepsilon})+\lambda^{\varepsilon})T(u+\lambda^{\varepsilon}v)\mathrm{e}^{-\mathrm{i}\tau}, \\ L^{\varepsilon}T\overline{(u+\lambda^{\varepsilon}v)}\mathrm{e}^{\mathrm{i}\tau} &= (\mathrm{i}(\omega+\nu^{\varepsilon})-\lambda^{\varepsilon})T\overline{(u+\lambda^{\varepsilon}v)}\mathrm{e}^{\mathrm{i}\tau}, \end{aligned}$$

which verifies that Assumptions 3 are compatible with the symmetry shown in equations (2.5), (2.6).

We now introduce some notation for the individual eigenvalues and the corresponding eigenvectors of $L_{\mathrm{sh,c}}^{0}$. Define $r_{\pm 1} = n-2$ and $r_m = n$ for $m \in \mathbb{Z} \setminus \{\pm 1\}$, denote by $\lambda_{m,1}, \ldots, \lambda_{m,r_m}$ the simple mode m eigenvalues of $L_{\mathrm{sh,c}}^{0}$, and let $\{e_{m,1}\mathrm{e}^{\mathrm{i}m\tau}, \ldots, e_{m,r_m}\mathrm{e}^{\mathrm{i}m\tau}\}$ be a basis for the r_m-dimensional, $L_{\mathrm{sh,c}}^{0}$-invariant subspace associated to these eigenvalues with $|e_{m,r}| = 1$ for each $m \in \mathbb{Z}$, $r = 1, \ldots, r_m$. Denoting by σ_m the (even) number of strongly hyperbolic mode m eigenvalues (so that $\sigma_m = \sigma_{-m}$) we number the eigenvalues and eigenvectors so that $\lambda_{m,1}, \ldots, \lambda_{m,r_m-\sigma_m}$ are purely imaginary for $m \in \mathbb{Z}$ and $\overline{\lambda}_{m,r} = \lambda_{-m,r}$, $\overline{e}_{m,r} = e_{-m,r}$ for $m \in \mathbb{Z} \setminus \{0\}$. More precise book keeping for the mode 0 eigenvalues is necessary in order to formulate appropriate reality conditions for the coefficients in the expansions below. Let $r_0^{(1)}$ be the number of real mode 0 eigenvalues and $r_0^{(2)}$ be the number of eigenvalues with non-zero imaginary part (so that $r_0^{(1)}$ and $r_0^{(2)}$ are even). In keeping with the symmetry we denote these eigenvalues by

$$\lambda^{(1)}_{0,-\frac{1}{2}r_0^{(1)}}, \ldots, \lambda^{(1)}_{0,-1}, \lambda^{(1)}_{0,1}, \ldots, \lambda^{(1)}_{0,\frac{1}{2}r_0^{(1)}}, \qquad \lambda^{(2)}_{0,-\frac{1}{2}r_0^{(2)}}, \ldots, \lambda^{(2)}_{0,-1}, \lambda^{(2)}_{0,1}, \ldots, \lambda^{(2)}_{0,\frac{1}{2}r_0^{(2)}},$$

and number them such that

$$\lambda^{(1)}_{0,-r} = -\lambda^{(1)}_{0,r}, \quad r = 1, \ldots, \tfrac{1}{2}r_0^{(1)}$$

and

$$\lambda^{(2)}_{0,-r} = \overline{\lambda}^{(2)}_{0,r}, \quad r = 1, \ldots, \tfrac{1}{2}r_0^{(2)}.$$

In particular we have the decompositions

$$
\begin{aligned}
E_0 &= \bigoplus_{r=1}^{\frac{1}{2}r_0^{(1)}} E_{0,r}^{(1)} \oplus \bigoplus_{r=1}^{\frac{1}{2}r_0^{(2)}} E_{0,r}^{(2)},\\
E_{-1} \oplus E_1 &= \mathcal{X}_{\mathrm{wh}} \oplus \bigoplus_{r=1}^{r_1} E_{1,r},\\
E_{-m} \oplus E_m &= \bigoplus_{r=1}^{r_m} E_{m,r}, \quad m \geq 2,
\end{aligned}
$$

where

$$
\begin{aligned}
E_{0,r}^{(1)} &= \langle e_{0,r}^{(1)}, Te_{0,r}^{(1)} \rangle, \quad r = 1, \dots, \tfrac{1}{2}r_0^{(1)},\\
E_{0,r}^{(2)} &= \langle e_{0,r}^{(2)}, \overline{e}_{0,r}^{(2)} \rangle, \quad r = 1, \dots, \tfrac{1}{2}r_0^{(2)},
\end{aligned}
$$

and

$$
E_{m,r} = \langle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}, \overline{e}_{m,r}\mathrm{e}^{-\mathrm{i}m\tau} \rangle, \quad r = 1, \dots, r_m
$$

for $m \geq 1$.

Remark 2.7. The matrix of $L_{\mathrm{sh,c}}^0|_{E_{0,r}^{(1)}}$ with respect to the basis $\{e_{0,r}^{(1)}, Te_{0,r}^{(1)}\}$ is given by

$$
\begin{pmatrix} \lambda_{0,r}^{(1)} & 0 \\ 0 & -\lambda_{0,r}^{(1)} \end{pmatrix},
$$

the matrix of $L_{\mathrm{sh,c}}^0|_{E_{0,r}^{(2)}}$ with respect to the basis $\{e_{0,r}^{(2)}, \overline{e}_{0,r}^{(2)}\}$ is given by

$$
\begin{pmatrix} \lambda_{0,r}^{(2)} & 0 \\ 0 & \overline{\lambda}_{0,r}^{(2)} \end{pmatrix}
$$

and the matrix of $L_{\mathrm{sh,c}}^0|_{E_{m,r}}$, $m \geq 1$, with respect to the basis $\{e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}, \overline{e}_{m,r}\mathrm{e}^{-\mathrm{i}m\tau}\}$ is given by

$$
\begin{pmatrix} \lambda_{m,r} & 0 \\ 0 & \overline{\lambda}_{m,r} \end{pmatrix}.
$$

Real-valued functions z in $\mathcal{X}_{\mathrm{wh}}$ are written as

$$
z(\tau) = z_1 u\mathrm{e}^{\mathrm{i}\tau} + z_2 v\mathrm{e}^{\mathrm{i}\tau} + \overline{z}_1\overline{u}\mathrm{e}^{-\mathrm{i}\tau} + \overline{z}_2\overline{v}\mathrm{e}^{-\mathrm{i}\tau}, \quad z_1, z_2 \in \mathbb{C}. \tag{2.18}
$$

and real-valued functions q in $\mathcal{X}_{\mathrm{sh,c}}^t$ as

$$
\begin{aligned}
q(\tau) &= \sum_{\substack{m\in\mathbb{Z}\\ 1\leq r\leq r_m}} q_{m,r}e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\\
&= \sum_{\substack{-\frac{1}{2}r_0^{(1)}\leq r\leq\frac{1}{2}r_0^{(1)}\\ r\neq 0}} q_{0,r}^{(1)}e_{0,r}^{(1)} + \sum_{\substack{-\frac{1}{2}r_0^{(2)}\leq r\leq\frac{1}{2}r_0^{(2)}\\ r\neq 0}} q_{0,r}^{(2)}e_{0,r}^{(2)} + \sum_{\substack{m\in\mathbb{Z}\setminus\{0\}\\ 1\leq r\leq r_m}} q_{m,r}e_{m,r}\mathrm{e}^{\mathrm{i}m\tau},
\end{aligned}
$$

where the coefficients $q_{m,r} \in \mathbb{C}$ satisfy the reality conditions

$$\overline{q}_{m,r} = q_{-m,r}, \quad m \in \mathbb{Z} \setminus \{0\}, \tag{2.19}$$

as well as

$$q_{0,r}^{(1)} = \overline{q}_{0,r}^{(1)}, \quad r = \pm 1, \dots, \pm \tfrac{1}{2} r_0^{(1)}, \tag{2.20}$$
$$q_{0,-r}^{(2)} = \overline{q}_{0,r}^{(2)}, \quad r = 1, \dots, \tfrac{1}{2} r_0^{(2)}. \tag{2.21}$$

In Lemma 2.15 below we use the above coordinates particularly for the strongly hyperbolic part $Z = P_{\mathrm{sh}} q$ of q. Recall that the strongly hyperbolic subspace $\mathcal{X}_{\mathrm{sh}}$ of L^0 is finite-dimensional (see Proposition 2.3), that is

$$\mathcal{X}_{\mathrm{sh}} \subseteq E_0 \oplus (E_{-m_1} \oplus E_{m_1}) \oplus \cdots \oplus (E_{-m_d} \oplus E_{m_d})$$

for finitely many $m_1, \dots, m_d \in \mathbb{N}$ and in particular $\sigma_m = 0$ for $m \notin \{0, \pm m_1, \dots, \pm m_d\}$. For notational simplicity we define $m_0 = 0$ and therefore obtain

$$Z(\tau) = \sum_{\substack{m \in \{\pm m_0, \dots, \pm m_d\} \\ r_m - \sigma_m + 1 \le \sigma \le r_m}} Z_{m,\sigma} e_{m,\sigma} \mathrm{e}^{\mathrm{i} m \tau}$$

(due to the numbering of the eigenvalues), where we have written $Z_{m,\sigma}$ instead of $q_{m,\sigma}$ to denote the coefficients of $Z = P_{\mathrm{sh}} q$.

Next we consider the change of basis matrices Q_m, where

$$Q_1 = (e_{1,1} | \cdots | e_{1,n-2} | u | v), \quad Q_{-1} = (\overline{e}_{1,1} | \cdots | \overline{e}_{1,n-2} | \overline{u} | \overline{v})$$

and

$$Q_m = (e_{m,1} | \cdots | e_{m,n}) \quad (m \in \mathbb{Z} \setminus \{\pm 1\}),$$

from the canonical basis $\{e_1, \dots, e_n\}$ for $\mathbb{C}^n$ to the bases $\{e_{1,1}, \cdots, e_{1,n-2}, u, v\}$, $\{\overline{e}_{1,1}, \cdots, \overline{e}_{1,n-2}, \overline{u}, \overline{v}\}$ and $\{e_{m,1}, \dots, e_{m,n}\}$, respectively. We use the following auxiliary fact, which states that the inverse change of basis matrices Q_m^{-1} are bounded independently of $m \in \mathbb{Z}$, in order to avoid small divisor problems in the estimates for the normal-form transformations in Chapter 3 below.

Proposition 2.8. *The change of basis matrix Q_m satisfies*

$$\|Q_m^{-1}\| \lesssim 1$$

for all $m \in \mathbb{Z}$.

Proof. Suppose that $m \in \mathbb{Z} \setminus \{\pm 1\}$ and note that

$$Q_m^{-1} = \frac{1}{\det Q_m} Q_m^{\#}$$

where the entries of the adjunct matrix $Q_m^{\#}$ are given by

$$(Q_m^{\#})_{ij} = (-1)^{i+j} \det (e_{m,1}|\cdots|e_{m,i-1}|e_j|e_{m,i+1}|\cdots|e_{m,n}).$$

Recalling that, due to Proposition 2.2, the vectors $e_{m,r}$ depend continuously upon m and satisfy $\lim_{|m|\to\infty} e_{m,r} = v_r^0$, where $v_1^0,\ldots,v_n^0$ denote the normalised eigenvectors of the matrix $\mathrm{i}A^0(0)$, we find by continuity that

$$\lim_{|m|\to\infty} \det Q_m = \det P,$$

where

$$P = (v_1^0|\cdots|v_n^0).$$

Hence

$$|\det Q_m| \geq \frac{1}{2}|\det P|$$

for sufficiently large values of $|m|$. Similarly we find that

$$|(Q_m^{\#})_{ij}| \leq 2|\det (v_1^0|\cdots|v_{i-1}^0|e_j|v_{i+1}^0|\cdots|v_n^0)| \leq 2$$

for sufficiently large values of $|m|$. Every entry in the matrix Q_m^{-1} is therefore bounded and it follows that $\|Q_m^{-1}\| \lesssim 1$ for all $m \in \mathbb{Z} \setminus \{\pm 1\}$ und thus for all $m \in \mathbb{Z}$. ■

Consider the infinite-dimensional centre subspace $\mathcal{X}_\mathrm{c}^t$ of the operator $L^0\colon \mathcal{X}^{t+1} \to \mathcal{X}^t$. It is a direct consequence of Proposition 2.8 that

$$\{e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\}_{m\in\mathbb{Z}}^{r=1,\ldots,r_m-\sigma_m}$$

is a *Riesz basis* for $\mathcal{X}_\mathrm{c}^t$ (see Christensen [5, Definition 3.6.1]), since it is the image of the subset

$$\{e_r\mathrm{e}^{\mathrm{i}m\tau}\}_{m\in\mathbb{Z}}^{r=1,\ldots,r_m-\sigma_m}$$

of the orthonormal basis

$$\{e_r\mathrm{e}^{\mathrm{i}m\tau}\}_{m\in\mathbb{Z}}^{r=1,\ldots,n}$$

for $\mathcal{X}^t$ under the invertible bounded linear operator $Q\colon \mathcal{X}^t \to \mathcal{X}^t$ defined by

$$(Qu)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} (Q_m u_m)\mathrm{e}^{\mathrm{i}m\tau}.$$

Since by construction $e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}$ (for all $m \in \mathbb{Z}$, $r = 1,\ldots,r_m-\sigma_m$) are eigenvectors of L_c^ε, we conclude that the operator L_c^ε is a *Riesz spectral operator*, i.e. there exists a Riesz basis for $\mathcal{X}_\mathrm{c}^t$ consisting of (generalised) eigenvectors of L_c^ε.

Remark 2.9. We proceed by identifying $\mathcal{X}_\mathrm{wh}$ with $\mathbb{C}^4$ and substituting the *Ansatz* (2.18) into equations (2.10),(2.11). Note however that we can generalise these equations by allowing *any* analytic function $f^\varepsilon(z,q)$ in equation (2.10) which satisfies the estimate (2.12) and making the *Ansatz*

$$z(\tau) = z_1 u\mathrm{e}^{\mathrm{i}\tau} + z_2 v\mathrm{e}^{\mathrm{i}\tau} + \overline{z}_1\overline{u}\mathrm{e}^{-\mathrm{i}\tau} + \overline{z}_2\overline{v}\mathrm{e}^{-\mathrm{i}\tau} + \ell(z_1,z_2,\overline{z}_1,\overline{z}_2)(\tau),$$

where $\ell \in \mathcal{L}(\mathbb{C}^4, \mathcal{X}_\mathrm{wh})$, in equation (2.11). Our subsequent analysis applies to this generalisation, which is exploited in particular in our treatment of Klein-Gordon equations in Chapter 6.

2.2 Approximate pulse solutions

In general one cannot expect equations (2.10), (2.11) to admit genuine homoclinic solutions (z, q) with $(z(\xi), q(\xi)) \to (0, 0)$ as $|\xi| \to \infty$ since such solutions would arise as the intersection of *finite-dimensional* stable and unstable manifolds in an *infinite-dimensional* phase space. In this section we examine the special case $h^\varepsilon(z) = 0$ in which $\{q = 0\}$ is an invariant four-dimensional subspace and confirm the existence of a pair of homoclinic solutions $p^{\varepsilon\pm}$ in this subspace under a sign condition on the cubic, ε-independent part of $f^\varepsilon(z, 0)$. Our strategy for the general case is to use a sequence of normal-form transformations to make $h^\varepsilon(z)$ exponentially small in a suitable sense (see Chapter 3 below) and hence show that exponentially small perturbations of $p^{\varepsilon\pm}$ persist as solutions to equations (2.10), (2.11) over exponentially long intervals of values for ξ.

We perform the first of the normal-form transformations seperately since this transformation, which removes the quadratic, ε-independent part in the Maclaurin expansion of $h^\varepsilon(z)$, affects the cubic, ε-independent part of f^ε (see above). We say that a transformation $(z, q) \mapsto K(z, q)$ is *finite-dimensional* if it affects only finitely many modes, that is $P_m K \neq 0$ for only finitely many $m \in \mathbb{Z}$.

Lemma 2.10. *There exists a near-identity, finite dimensional change of coordinates which transforms the coupled system (2.10), (2.11) into*

$$z_\xi = L^\varepsilon_{\mathrm{wh}} z + \tilde{f}^\varepsilon(z, q), \tag{2.22}$$

$$q_\xi = L^\varepsilon_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}\Big(\tilde{g}^\varepsilon_1(z, q) q_\tau\Big) + \tilde{g}^\varepsilon_2(z, q) + \tilde{h}^\varepsilon(z), \tag{2.23}$$

and preserves the reversibility. The nonlinearities $\tilde{f}^\varepsilon$, $\tilde{g}^\varepsilon_1$ and $\tilde{g}^\varepsilon_2$, satisfy the same estimates as respectively f^ε, g^ε_1 and g^ε_2, while

$$\|\tilde{h}^\varepsilon(z)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(|z|^2 |(z, \varepsilon)|)$$

for $t > \frac{1}{2}$.

Proof. Let $h^0_2(z)$ denote the term in the Maclaurin expansion of $h^\varepsilon(z)$ which is homogeneous of degree two in z and does not depend upon ε, so that

$$\|h^\varepsilon(z) - h^0_2(z)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(|z|^2 |(z, \varepsilon)|).$$

The quadratic nonlinearity h^0_2 is a mapping from $\mathcal{X}_{\mathrm{wh}} \subseteq E_{-1} \oplus E_1$ to $E_0 \oplus (E_{-2} \oplus E_2)$ 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 $m \neq 0, \pm 2$. Substituting $q = \tilde{q} - \Gamma(z)$ into equations

(2.10), (2.11), we find that they are transformed into respectively (2.22), (2.23) with

$$\begin{aligned}
\tilde{f}^{\varepsilon}(z,\tilde{q}) &= f^{\varepsilon}(z,\tilde{q}-\Gamma(z)),\\
\tilde{g}_1^{\varepsilon}(z,\tilde{q}) &= g_1^{\varepsilon}(z,\tilde{q}-\Gamma(z)),\\
\tilde{g}_2(z,\tilde{q}) &= g_2^{\varepsilon}(z,\tilde{q}-\Gamma(z)) - g_2^{\varepsilon}(z,-\Gamma(z))\\
&\qquad - P_{\mathrm{sh,c}}\Big(g_1^{\varepsilon}(z,\tilde{q}-\Gamma(z)) - g_1^{\varepsilon}(z,-\Gamma(z))\Big)\partial_\tau\Gamma(z)\\
&\qquad + \mathrm{d}\Gamma[z]\Big(f^{\varepsilon}(z,\tilde{q}-\Gamma(z)) - f^{\varepsilon}(z,-\Gamma(z))\Big),
\end{aligned}$$

and

$$\begin{aligned}
\tilde{h}^{\varepsilon}(z) = &-L^{\varepsilon}_{\mathrm{sh,c}}\Gamma(z) + \mathrm{d}\Gamma[z](L^{\varepsilon}_{\mathrm{wh}}z) + \mathrm{d}\Gamma[z]\Big(f^{\varepsilon}(z,-\Gamma(z))\Big)\\
&- P_{\mathrm{sh,c}}\Big(g_1^{\varepsilon}(z,-\Gamma(z))\partial_\tau\Gamma(z)\Big) + g_2^{\varepsilon}(z,-\Gamma(z)) + h^{\varepsilon}(z).
\end{aligned}$$

Our objective of removing the ε-independent terms which are homogeneous of degree two in z from the Maclaurin expansion of $\tilde{h}^{\varepsilon}$ is achieved by choosing

$$\Gamma\colon \mathcal{X}_{\mathrm{wh}} \to E_0 \oplus (E_{-2}\oplus E_2)$$

to be a mapping which is homogeneous of degree two and satisfies the equation

$$(\mathcal{L}\Gamma)(z) = h_2^0(z) \tag{2.24}$$

for each $z \in \mathcal{X}_{\mathrm{wh}}$, where

$$(\mathcal{L}\Gamma)(z) = L^0_{\mathrm{sh,c}}\Gamma(z) - \mathrm{d}\Gamma[z](L^0_{\mathrm{wh}}z). \tag{2.25}$$

Note that a solution Γ to equation (2.25) commutes with the reverser R, so that the transformed system (2.22), (2.23) is reversible.

In order to solve equation (2.24), we write

$$\Gamma(z) = \sum_{i+j+k+l=2}\ \sum_{\substack{m\in\{-2,0,2\}\\ 1\le r\le r_m}} \Gamma_{m,ijkl,r} z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}$$

where the vectors $e_{m,r}$ are defined in Section 2.1 and $\Gamma_{m,ijkl,r}$ are complex numbers.

Let $\mathcal{P}^2$ denote the vector space of homogeneous polynomials of degree two in the variables z_1, z_2, $\overline{z}_1$ and $\overline{z}_2$ taking values in $E_0\oplus(E_{-2}\oplus E_2)$ and note that $\mathcal{P}^2$ decomposes into the direct sum

$$\bigoplus_{\substack{m\in\{-2,0,2\}\\ 1\le r\le r_m}} \mathcal{P}^2_{m,r},$$

in which the vector space $\mathcal{P}^2_{m,r}$ of homogeneous polynomials of degree two taking values in $\langle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle$ is invariant under $\mathcal{L}$. Equipping $\mathcal{P}^2_{m,r}$ with the basis

$$\mathcal{B}_{m,r} = \{z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r}\mathrm{e}^{\mathrm{i}m\tau} \colon i+j+k+l=2\}$$

and using the calculation

$$\begin{aligned}\mathcal{L}(z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r} \mathrm{e}^{\mathrm{i}m\tau}) = \lambda_{m,r} z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r} \mathrm{e}^{\mathrm{i}m\tau} - \mathrm{i}\omega(i+j-k-l) z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r} \mathrm{e}^{\mathrm{i}m\tau} \\ - i z_1^{i-1} z_2^{j+1} \overline{z}_1^k \overline{z}_2^l e_{m,r} \mathrm{e}^{\mathrm{i}m\tau} - k z_1^i z_2^j \overline{z}_1^{k-1} \overline{z}_2^{l+1} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau}\end{aligned}$$

to compute the matrix $\mathcal{A}_{m,r}$ of the operator $\mathcal{L}_{m,r} = \mathcal{L}|_{\mathcal{P}^2_{m,r}} : \mathcal{P}^2_{m,r} \to \mathcal{P}^2_{m,r}$ with respect to $\mathcal{B}_{m,r}$, we find that $\det \mathcal{A}_{m,r} = (\lambda_{m,r} - 2\mathrm{i}\omega)^3 \lambda_{m,r}^4 (\lambda_{m,r} + 2\mathrm{i}\omega)^3$. Since Assumptions 2(iii) guarantee that $\lambda_{m,r} \notin \{0, \pm 2\mathrm{i}\omega\}$, we conclude that $\mathcal{L}_{m,r}$ is invertible, so that equation (2.24) admits a unique solution for $\Gamma(z)$. ■

We now consider the dynamical system

$$z_\xi = L^\varepsilon_{\mathrm{wh}} z + \tilde{f}^\varepsilon(z, 0), \tag{2.26}$$

which is obtained from equations (2.22), (2.23) by neglecting $\tilde{h}^\varepsilon(z)$ and setting $q = 0$. Since the matrix of $R_{\mathrm{wh}} = R|_{\mathcal{X}_{\mathrm{wh}}}$ is given by

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

we know that equation (2.26) is invariant under the transformation

$$\xi \mapsto -\xi, \quad (z_1, z_2, \overline{z}_1, \overline{z}_2) \mapsto (\overline{z}_1, -\overline{z}_2, z_1, -z_2).$$

Moreover, the original system (2.1) is invariant under the translation $\tau \mapsto \tau + a$ for each $a \in \mathbb{R}$ and this translation invariance is inherited by equation (2.26). We therefore find that the system (2.26) has the $O(2)$-symmetry

$$(z_1, z_2, \overline{z}_1, \overline{z}_2) \mapsto (\mathrm{e}^{\mathrm{i}a} z_1, \mathrm{e}^{\mathrm{i}a} z_2, \mathrm{e}^{-\mathrm{i}a} \overline{z}_1, \mathrm{e}^{-\mathrm{i}a} \overline{z}_2).$$

Here we are interested in constructing reversible homoclinic solutions to (2.26), that is homoclinic solutions z satisfying $R_{\mathrm{wh}} z(0) = z(0)$.

Using the reversibility and the $O(2)$-symmetry we write equation (2.26) as

$$z_{1\xi} = \mathrm{i}(\omega + \nu^\varepsilon) z_1 + z_2 + \mathcal{R}^\varepsilon_1(z_1, z_2, \overline{z}_1, \overline{z}_2), \tag{2.27}$$
$$z_{2\xi} = (\lambda^\varepsilon)^2 z_1 + \mathrm{i}(\omega + \nu^\varepsilon) z_2 - \check{C} z_1 |z_1|^2 + \mathcal{R}^\varepsilon_2(z_1, z_2, \overline{z}_1, \overline{z}_2) \tag{2.28}$$

where the remainder terms $\mathcal{R}^\varepsilon_j$ satisfy

$$\overline{\mathcal{R}^\varepsilon_1(z_1, z_2, \overline{z}_1, \overline{z}_2)} = -\mathcal{R}^\varepsilon_1(\overline{z}_1, -\overline{z}_2, z_1, -z_2), \quad \overline{\mathcal{R}^\varepsilon_2(z_1, z_2, \overline{z}_1, \overline{z}_2)} = \mathcal{R}^\varepsilon_2(\overline{z}_1, -\overline{z}_2, z_1, -z_2)$$

as well as

$$\mathcal{R}^\varepsilon_j(\mathrm{e}^{\mathrm{i}a} z_1, \mathrm{e}^{\mathrm{i}a} z_2, \mathrm{e}^{-\mathrm{i}a} \overline{z}_1, \mathrm{e}^{-\mathrm{i}a} \overline{z}_2) = \mathrm{e}^{\mathrm{i}a} \mathcal{R}^\varepsilon_j(z_1, z_2, \overline{z}_1, \overline{z}_2), \quad a \in \mathbb{R}.$$

The coefficient $\check{C}$ can be computed explicitly via the formula

$$\check{C} = -P_{(2)}\Big(3f^{0;3,0}\big[\{u\mathrm{e}^{\mathrm{i}\tau}\}^{(2)}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}\big] + f^{0;1,1}\big[\overline{u}\mathrm{e}^{-\mathrm{i}\tau}, -\sum_{r=1}^{n}\Gamma_{2,2000,r}e_{m,r}\mathrm{e}^{2\mathrm{i}\tau}\big] + f^{0;1,1}\big[u\mathrm{e}^{\mathrm{i}\tau}, -\sum_{r=1}^{n}\Gamma_{0,1010,r}e_{0,r}\big]\Big),$$

where we have used the Maclaurin expansion

$$\begin{aligned}\tilde{f}^0(z,0) &= f^0(z,-\Gamma(z)) \\ &= \sum_{\substack{n_1+n_2\geq 2\\(n_1,n_2)\neq(2,0)}} f^{0;n_1,n_2}\Big[\Big\{z_1u\mathrm{e}^{\mathrm{i}\tau} + z_2v\mathrm{e}^{\mathrm{i}\tau} + \overline{z}_1\overline{u}\mathrm{e}^{-\mathrm{i}\tau} + \overline{z}_2\overline{v}\mathrm{e}^{-\mathrm{i}\tau}\Big\}^{(n_1)}, \\ &\qquad -\Big\{\sum_{i+j+k+l=2}\ \sum_{\substack{m\in\{-2,0,2\}\\1\leq r\leq r_m}}\Gamma_{m,ijkl,r}z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\Big\}^{(n_2)}\Big],\end{aligned}$$

in which

$$f^{0;n_1,n_2} = \frac{1}{n_1!n_2!}\mathrm{d}_{1,2}^{n_1,n_2}f^0[0,0],$$

and $P_{(2)}(z)$ denotes the second coordinate of the vector $z \in \mathcal{X}_{\mathrm{wh}}$ with respect to the ordered basis $\{u\mathrm{e}^{\mathrm{i}\tau}, v\mathrm{e}^{\mathrm{i}\tau}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}, \overline{v}\mathrm{e}^{-\mathrm{i}\tau}\}$ for $\mathcal{X}_{\mathrm{wh}}$.

Assumptions 4. The coefficient $\check{C}$ in equations (2.27), (2.28) is positive.

Next we use the transformation

$$z_1(\xi) = \mathrm{e}^{\mathrm{i}(\omega+\nu^\varepsilon)\xi}z_1'(\xi), \quad z_2(\xi) = \mathrm{e}^{\mathrm{i}(\omega+\nu^\varepsilon)\xi}z_2'(\xi)$$

so that equations (2.27), (2.28) become

$$z'_{1\xi} = z'_2 + \mathcal{R}_1^\varepsilon(z'_1, z'_2, \overline{z'_1}, \overline{z'_2}) \tag{2.29}$$

$$z'_{2\xi} = (\lambda^\varepsilon)^2 z'_1 - \check{C}z'_1|z'_1|^2 + \mathcal{R}_2^\varepsilon(z'_1, z'_2, \overline{z'_1}, \overline{z'_2}). \tag{2.30}$$

Changing to real coordinates by writing

$$z'_1 = x'_1 + \mathrm{i}x'_2, \quad z'_2 = y'_1 + \mathrm{i}y'_2$$

and working with the scaled variables

$$\check{\xi} = \lambda^\varepsilon\xi, \quad x'(\xi) = \lambda^\varepsilon\check{x}(\check{\xi}), \quad y'(\xi) = (\lambda^\varepsilon)^2\check{y}(\check{\xi}),$$

we arrive at the system

$$\check{x}_{1\check{\xi}} = \check{y}_1 + \mathcal{R}_{1,1}^\varepsilon(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2), \tag{2.31}$$

$$\check{y}_{1\check{\xi}} = \check{x}_1 - \check{C}\check{x}_1(\check{x}_1^2 + \check{x}_2^2) + \mathcal{R}_{1,2}^\varepsilon(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2), \tag{2.32}$$

$$\check{x}_{2\check{\xi}} = \check{y}_2 + \mathcal{R}_{2,1}^\varepsilon(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2), \tag{2.33}$$

$$\check{y}_{2\check{\xi}} = \check{x}_2 - \check{C}\check{x}_2(\check{x}_1^2 + \check{x}_2^2) + \mathcal{R}_{2,2}^\varepsilon(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2), \tag{2.34}$$

where the remainder terms satisfy $|\mathcal{R}^{\varepsilon}_{i,j}(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)| = \mathcal{O}(\varepsilon|(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)|^3)$. Moreover $\mathcal{R}^{\varepsilon}_{1,1}(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)$ and $\mathcal{R}^{\varepsilon}_{2,2}(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)$ are odd in $(\check{y}_1, \check{x}_2)$ while $\mathcal{R}^{\varepsilon}_{1,2}(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)$ and $\mathcal{R}^{\varepsilon}_{2,1}(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)$ are even in $(\check{y}_1, \check{x}_2)$.

The existence of reversible homoclinic solutions to equations (2.31)–(2.34) is established in Lemma 2.13 below using a method devised by Kirchgässner [17, Proposition 5.1]. We first state some necessary functional-analytic results involved in Kirchgässner's theory.

Proposition 2.11.

(i) The formula

$$L\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} f_1'' - f_1 \\ f_2'' - f_2 \end{pmatrix}$$

defines a bounded linear operator $C^2_{-\theta}(\mathbb{R})^2 \to C_{-\theta}(\mathbb{R})^2$ *and* $C^2_{-\theta,\mathrm{e}}(\mathbb{R}) \times C^2_{-\theta,\mathrm{o}}(\mathbb{R}) \to C_{-\theta,\mathrm{e}}(\mathbb{R}) \times C_{-\theta,\mathrm{o}}(\mathbb{R})$ *for each* $\theta \geq 0$.

(ii) For $0 \leq \eta < 1$ *the operator* $L\colon C^2_{-\eta}(\mathbb{R})^2 \to C_{-\eta}(\mathbb{R})^2$ *is invertible with bounded inverse given by*

$$(L^{-1}g)(\xi) = -\frac{1}{2}\int_{-\infty}^{\infty} \mathrm{e}^{-|\xi - r|} g(r)\,\mathrm{d}r,$$

where the integration is taken componentwise.

Proposition 2.12. *Suppose that* $g \in C_{-1}(\mathbb{R})$ *and* $0 \leq \eta < 1$. *The formula*

$$K_g f = L^{-1}\begin{pmatrix} -3\check{C}g^2 f_1 \\ -\check{C}g^2 f_2 \end{pmatrix}$$

defines a bounded linear operator $C_0(\mathbb{R})^2 \to C^2_{-\eta}(\mathbb{R})^2$ *and a compact operator* $C_{-\eta}(\mathbb{R})^2 \to C_{-\eta}(\mathbb{R})^2$.

Lemma 2.13. *Suppose that* $\eta \in [0, 1)$. *The system* (2.31)–(2.34) *has a pair* $\check{p}^{\varepsilon\pm}$ *of reversible (that is invariant with respect to the transformation*

$$\check{\xi} \mapsto -\check{\xi}, \quad (\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)^\top \mapsto (\check{x}_1, -\check{y}_1, -\check{x}_2, \check{y}_2)^\top)$$

homoclinic solutions which satisfy the estimates

$$|\check{p}^{\varepsilon\pm}(\check{\xi})|, |\check{p}^{\varepsilon\pm}_{\check{\xi}}(\check{\xi})| \lesssim \mathrm{e}^{-\eta|\check{\xi}|}.$$

Proof. First we note that for $\varepsilon = 0$, the system (2.31)–(2.34) has the explicit homoclinic solution

$$(\check{x}_1, \check{y}_1, \check{x}_2, \check{y}_2)^\top = (H, H_{\check{\xi}}, 0, 0)^\top,$$

where

$$H(\check{\xi}) = \sqrt{\frac{2}{\check{C}}}\,\mathrm{sech}(\check{\xi}).$$

In fact, we have the family

$$\Big\{(\check{x}_1,\check{y}_1,\check{x}_2,\check{y}_2)^\top : (\check{x}_1,\check{x}_2)^\top = R_a(H(\check{\xi}_0+\cdot),0)^\top,\\ (\check{y}_1,\check{y}_2)^\top = R_a(H_{\check{\xi}}(\check{\xi}_0+\cdot),0)^\top,\ a\in[0,2\pi),\ \check{\xi}_0\in\mathbb{R}\Big\}$$

of homoclinic solutions, where R_a denotes the 2×2 matrix representing a rotation through the angle a. Note that the solutions $\check{p}^+$ obtained from $a=0$, $\check{\xi}_0=0$ and $\check{p}^-$ obtained from $a=\pi$, $\check{\xi}_0=0$ are reversible.

Now we apply perturbation arguments to show that equations (2.31)–(2.34) also admit reversible homoclinic solutions when $\varepsilon>0$. Using the implicit function theorem, we can solve equations (2.31) and (2.33) for $\check{y}=\check{y}(\check{x},\check{x}_{\check{\xi}})$ and write

$$\begin{aligned}\check{y}_1 &= \check{x}_{1\check{\xi}} + s_1^\varepsilon(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}}),\\ \check{y}_2 &= \check{x}_{2\check{\xi}} + s_2^\varepsilon(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}}),\end{aligned}$$

where s_1^ε and s_2^ε are analytic at the origin, s_1^ε is odd in $\check{x}_{1\check{\xi}}$, $\check{x}_2$ and s_2^ε is even in $\check{x}_{1\check{\xi}}$, $\check{x}_2$ with

$$|s_j^\varepsilon(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}})| = \mathcal{O}(|(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}})|^2).$$

Hence we can rewrite equations (2.31)–(2.34) as the second order system

$$\check{x}_{1\check{\xi}\check{\xi}} = \check{x}_1 - \check{C}\check{x}_1(\check{x}_1^2+\check{x}_2^2) + S_1^\varepsilon(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}},\check{x}_{1\check{\xi}\check{\xi}},\check{x}_{2\check{\xi}\check{\xi}}) \tag{2.35}$$

$$\check{x}_{2\check{\xi}\check{\xi}} = \check{x}_2 - \check{C}\check{x}_2(\check{x}_1^2+\check{x}_2^2) + S_2^\varepsilon(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}},\check{x}_{1\check{\xi}\check{\xi}},\check{x}_{2\check{\xi}\check{\xi}}), \tag{2.36}$$

where S_1^ε and S_2^ε are analytic at the origin, S_1^ε is even in $\check{x}_{1\check{\xi}}$, $\check{x}_2$, $\check{x}_{2\check{\xi}\check{\xi}}$ and S_2^ε is odd in $\check{x}_{1\check{\xi}}$, $\check{x}_2$, $\check{x}_{2\check{\xi}\check{\xi}}$ with

$$|S_j^\varepsilon(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}},\check{x}_{1\check{\xi}\check{\xi}},\check{x}_{2\check{\xi}\check{\xi}})| = \mathcal{O}(|(\check{x}_1,\check{x}_2,\check{x}_{1\check{\xi}},\check{x}_{2\check{\xi}},\check{x}_{1\check{\xi}\check{\xi}},\check{x}_{2\check{\xi}\check{\xi}})|^2).$$

Now we write $\check{x}$ as a perturbation

$$\check{x}_1 = H + p_1, \quad \check{x}_2 = p_2$$

of the homoclinic solution $(H,0)^\top$ so that the system (2.35), (2.36) becomes

$$p_{1\check{\xi}\check{\xi}} - p_1 = -3\check{C}H^2p_1 + r_1^\varepsilon(p_1,p_2,p_{1\check{\xi}},p_{2\check{\xi}},p_{1\check{\xi}\check{\xi}},p_{2\check{\xi}\check{\xi}},H), \tag{2.37}$$

$$p_{2\check{\xi}\check{\xi}} - p_2 = -\check{C}H^2p_2 + r_2^\varepsilon(p_1,p_2,p_{1\check{\xi}},p_{2\check{\xi}},p_{1\check{\xi}\check{\xi}},p_{2\check{\xi}\check{\xi}},H), \tag{2.38}$$

with the obvious definitions of r_1^ε and r_2^ε. We study this system in the space $C^2_{-\eta}(\mathbb{R})^2$ and consider the nonlinearities r_1^ε and r_2^ε, with a slight abuse of notation, as analytic mappings $C^2_{-\eta}(\mathbb{R})^2 \to C_{-\eta}(\mathbb{R})^2$ with

$$\|r_j^\varepsilon(p_1,p_2)\|_{0,-\eta} = \mathcal{O}(\varepsilon) + \mathcal{O}(\|(p_1,p_2)\|^2_{2,-\eta}).$$

In terms of the operators L and K defined in Propositions 2.11 and 2.12, we can then write the system (2.37), (2.38) as

$$p = K_H p + L^{-1} r^\varepsilon(p), \tag{2.39}$$

where $p = (p_1, p_2)^\top$ and $r^\varepsilon = (r_1^\varepsilon, r_2^\varepsilon)^\top$. In order to solve equation (2.39), we show that the operator

$$I - K_H \colon C_{-\eta,\mathrm{e}}(\mathbb{R}) \times C_{-\eta,\mathrm{o}}(\mathbb{R}) \to C^2_{-\eta,\mathrm{e}}(\mathbb{R}) \times C^2_{-\eta,\mathrm{o}}(\mathbb{R})$$

is invertible.

As a first step we show that 1 is a geometrically double eigenvalue of

$$K_H \colon C_{-\eta}(\mathbb{R})^2 \to C_{-\eta}(\mathbb{R})^2,$$

noting that the eigenvalue problem

$$K_H p = p$$

is equivalent to the decoupled system

$$p_{1\check{\xi}\check{\xi}} = p_1 - 3\check{C}H^2 p_1, \tag{2.40}$$

$$p_{2\check{\xi}\check{\xi}} = p_2 - \check{C}H^2 p_2 \tag{2.41}$$

of ordinary differential equations. Let

$$\begin{aligned}
v_1(\check{\xi}) &= \operatorname{sech}(\check{\xi}) \tanh(\check{\xi}),\\
v_2(\check{\xi}) &= \operatorname{sech}(\check{\xi})(-3 + \cosh^2(\check{\xi}) + 3\check{\xi}\tanh(\check{\xi})),\\
w_1(\check{\xi}) &= \operatorname{sech}(\check{\xi}),\\
w_2(\check{\xi}) &= \operatorname{sech}(\check{\xi})(2\check{\xi} + \sinh(2\check{\xi})),
\end{aligned}$$

and note that $\{v_1, v_2\}$ is a fundamental solution set for equation (2.40) while $\{w_1, w_2\}$ is a fundamental solution set for equation (2.41). Since v_1, w_1 are bounded and v_2, w_2 are unbounded, we conclude that all bounded solutions of equation (2.40) are multiples of $v_1 = -H_{\check{\xi}}$ and all bounded solutions of equation (2.41) are multiples of $w_1 = (2/\check{C})^{-1/2}H$. The eigenspace of $K_H \colon C_{-\eta}(\mathbb{R})^2 \to C_{-\eta}(\mathbb{R})^2$ corresponding to the eigenvalue 1 is therefore

$$\left\langle \begin{pmatrix} H_{\check{\xi}} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ H \end{pmatrix} \right\rangle$$

and lies in $C_{-\eta,\mathrm{o}}(\mathbb{R}) \times C_{-\eta,\mathrm{e}}(\mathbb{R})$.

In particular 1 is not an eigenvalue of $K_H|_{C_{-\eta,\mathrm{e}}(\mathbb{R}) \times C_{-\eta,\mathrm{o}}(\mathbb{R})}$ and since K_H is a compact operator $C_{-\eta}(\mathbb{R})^2 \to C_{-\eta}(\mathbb{R})^2$ by Proposition 2.12, it follows that the spectrum of $K_H|_{C_{-\eta,\mathrm{e}}(\mathbb{R}) \times C_{-\eta,\mathrm{o}}(\mathbb{R})}$ consists only of eigenvalues, so that

$$I - K_H \colon C_{-\eta,\mathrm{e}}(\mathbb{R}) \times C_{-\eta,\mathrm{o}}(\mathbb{R}) \to C^2_{-\eta,\mathrm{e}}(\mathbb{R}) \times C^2_{-\eta,\mathrm{o}}(\mathbb{R})$$

is invertible. We can therefore solve equation (2.39) for sufficiently small values of $\varepsilon > 0$ using the implicit function theorem. The solution $p^\star$ satisfies $\|p^\star\|_{2,-\eta} = \mathcal{O}(\varepsilon)$. Returning to equations (2.31)–(2.34), we have found a solution

$$\check{p}^{\varepsilon+} = \begin{pmatrix} \check{p}_1^{\varepsilon+} \\ \check{q}_1^{\varepsilon+} \\ \check{p}_2^{\varepsilon+} \\ \check{q}_2^{\varepsilon+} \end{pmatrix} = \check{p}^+ + \begin{pmatrix} p_1^\star \\ p_{1\check{\xi}}^\star \\ p_2^\star \\ p_{2\check{\xi}}^\star \end{pmatrix} + \begin{pmatrix} 0 \\ s_1^\varepsilon(H + p_1^\star, p_2^\star, H_{\check{\xi}} + p_{1\check{\xi}}^\star, p_{2\check{\xi}}^\star) \\ 0 \\ s_2^\varepsilon(H + p_1^\star, p_2^\star, H_{\check{\xi}} + p_{1\check{\xi}}^\star, p_{2\check{\xi}}^\star) \end{pmatrix}.$$

This solution satisfies

$$|\check{p}^{\varepsilon+}(\check{\xi})|, |\check{p}_{\check{\xi}}^{\varepsilon+}(\check{\xi})| \lesssim \mathrm{e}^{-\eta|\check{\xi}|}.$$

The second homoclinic solution $\check{p}^{\varepsilon-}$ satisfying the above estimates is obtained analogously (from $\check{p}^-$). ■

Returning to the unscaled variables ξ, z_1 and z_2 we find from Lemma 2.13 that equation (2.26) has a pair of reversible homoclinic solutions of the form

$$p^{\varepsilon\pm}(\xi) = \begin{pmatrix} \mathrm{e}^{\mathrm{i}(\omega+\nu^\varepsilon)\xi}\lambda^\varepsilon(\check{p}_1^{\varepsilon\pm}(\lambda^\varepsilon\xi) + \mathrm{i}\check{p}_2^{\varepsilon\pm}(\lambda^\varepsilon\xi)) \\ \mathrm{e}^{\mathrm{i}(\omega+\nu^\varepsilon)\xi}(\lambda^\varepsilon)^2(\check{q}_1^{\varepsilon\pm}(\lambda^\varepsilon\xi) + \mathrm{i}\check{q}_2^{\varepsilon\pm}(\lambda^\varepsilon\xi)) \\ \mathrm{e}^{-\mathrm{i}(\omega+\nu^\varepsilon)\xi}\lambda^\varepsilon(\check{p}_1^{\varepsilon\pm}(\lambda^\varepsilon\xi) - \mathrm{i}\check{p}_2^{\varepsilon\pm}(\lambda^\varepsilon\xi)) \\ \mathrm{e}^{-\mathrm{i}(\omega+\nu^\varepsilon)\xi}(\lambda^\varepsilon)^2(\check{q}_1^{\varepsilon\pm}(\lambda^\varepsilon\xi) - \mathrm{i}\check{q}_2^{\varepsilon\pm}(\lambda^\varepsilon\xi)) \end{pmatrix}$$

which satisfy the estimate

$$|p^{\varepsilon\pm}(\xi)| \le c_\mathrm{h}\lambda^\varepsilon \mathrm{e}^{-\eta\lambda^\varepsilon|\xi|}, \quad \xi \in \mathbb{R},$$

for some $c_\mathrm{h} > 0$.

2.3 Existence theory – Foundations

In this section we apply some further near-identity transformations to equations (2.22), (2.23); the nonlinearities in the transformed system have a sufficient order of magnitude to ensure the convergence of a Kato iteration scheme in Chapters 4 and 5 below. We also summarise the results of the normal-form theory in Chapter 3 and present a reformulation of the hyperbolic part of the transformed system which is useful in proving the convergence of the iteration scheme.

Rescaling

A central requirement of the normal-form theory in Chapter 3 is that the parameter-independent part of the dynamical system (2.22) for z should be diagonalisable. Since (the matrix of) L_wh^0 on the right-hand side of equation (2.22) is evidently not diagonalisable, we use a change of parameter to 'replace' it by a diagonalisable matrix. We write

$$\begin{aligned}
\tilde{f}^\varepsilon(z,q) = \tilde{f}^\varepsilon(z_1,z_2,\overline{z}_1,\overline{z}_2,q) &= \begin{pmatrix} \tilde{f}_1^\varepsilon(z_1,z_2,\overline{z}_1,\overline{z}_2,q) \\ \tilde{f}_2^\varepsilon(z_1,z_2,\overline{z}_1,\overline{z}_2,q) \\ \overline{\tilde{f}_1^\varepsilon(z_1,z_2,\overline{z}_1,\overline{z}_2,q)} \\ \overline{\tilde{f}_2^\varepsilon(z_1,z_2,\overline{z}_1,\overline{z}_2,q)} \end{pmatrix}, \\
\tilde{g}_j^\varepsilon(z,q) &= \tilde{g}_j^\varepsilon(z_1,z_2,\overline{z}_1,\overline{z}_2,q), \quad j=1,2, \\
\tilde{h}^\varepsilon(z) &= \tilde{h}^\varepsilon(z_1,z_2,\overline{z}_1,\overline{z}_2),
\end{aligned}$$

introduce a new parameter μ by writing $\varepsilon = \mu^2$, and use the scaled variables

$$z_1' = (\lambda^\mu)^{-\frac{1}{2}} z_1, \quad z_2' = (\lambda^\mu)^{-\frac{3}{2}} z_2, \quad q' = (\lambda^\mu)^{-1} q,$$

where we have abused the notation to abbreviate $\lambda^\mu = \lambda^\varepsilon|_{\varepsilon=\mu^2}$ and $\nu^\mu = \nu^\varepsilon|_{\varepsilon=\mu^2}$ (see below). Omitting the primes for notational simplicity, equations (2.22), (2.23) are transformed into

$$z_\xi = L_{\mathrm{wh}}^\mu z + \check{f}_1^\mu(z) + \check{f}_2^\mu(z,q), \tag{2.42}$$

$$q_\xi = L_{\mathrm{sh,c}}^\mu q + P_{\mathrm{sh,c}}\Big(\check{g}_1^\mu(z,q) q_\tau\Big) + \check{g}_2^\mu(z,q) + \check{h}^\mu(z), \tag{2.43}$$

where we have (with a further abuse of notation) abbreviated $L_{\mathrm{sh,c}}^\varepsilon|_{\varepsilon=\mu^2}$ to $L_{\mathrm{sh,c}}^\mu$ and

$$\begin{aligned}
L_{\mathrm{wh}}^\mu &= \begin{pmatrix} \mathrm{i}(\omega+\nu^\mu) & \lambda^\mu & 0 & 0 \\ \lambda^\mu & \mathrm{i}(\omega+\nu^\mu) & 0 & 0 \\ 0 & 0 & -\mathrm{i}(\omega+\nu^\mu) & \lambda^\mu \\ 0 & 0 & \lambda^\mu & -\mathrm{i}(\omega+\nu^\mu) \end{pmatrix}, \\
\check{f}_1^\mu(z) &= \check{f}^\mu(z,0), \\
\check{f}_2^\mu(z,q) &= \check{f}^\mu(z,q) - \check{f}^\mu(z,0), \\
\check{f}^\mu(z,q) &= \begin{pmatrix} (\lambda^\mu)^{-\frac{1}{2}} \tilde{f}_1^\varepsilon\Big((\lambda^\mu)^{\frac{1}{2}} z_1, (\lambda^\mu)^{\frac{3}{2}} z_2, (\lambda^\mu)^{\frac{1}{2}} \overline{z}_1, (\lambda^\mu)^{\frac{3}{2}} \overline{z}_2, \lambda^\mu q\Big)\Big|_{\varepsilon=\mu^2} \\ (\lambda^\mu)^{-\frac{3}{2}} \tilde{f}_2^\varepsilon\Big((\lambda^\mu)^{\frac{1}{2}} z_1, (\lambda^\mu)^{\frac{3}{2}} z_2, (\lambda^\mu)^{\frac{1}{2}} \overline{z}_1, (\lambda^\mu)^{\frac{3}{2}} \overline{z}_2, \lambda^\mu q\Big)\Big|_{\varepsilon=\mu^2} \\ (\lambda^\mu)^{-\frac{1}{2}} \overline{\tilde{f}_1^\varepsilon\Big((\lambda^\mu)^{\frac{1}{2}} z_1, (\lambda^\mu)^{\frac{3}{2}} z_2, (\lambda^\mu)^{\frac{1}{2}} \overline{z}_1, (\lambda^\mu)^{\frac{3}{2}} \overline{z}_2, \lambda^\mu q\Big)}\Big|_{\varepsilon=\mu^2} \\ (\lambda^\mu)^{-\frac{3}{2}} \overline{\tilde{f}_2^\varepsilon((\lambda^\mu)^{\frac{1}{2}} z_1, (\lambda^\mu)^{\frac{3}{2}} z_2, (\lambda^\mu)^{\frac{1}{2}} \overline{z}_1, (\lambda^\mu)^{\frac{3}{2}} \overline{z}_2, \lambda^\mu q\Big)}\Big|_{\varepsilon=\mu^2} \end{pmatrix}, \\
\check{g}_1^\mu(z,q) &= \tilde{g}_1^\varepsilon\Big((\lambda^\mu)^{\frac{1}{2}} z_1, (\lambda^\mu)^{\frac{3}{2}} z_2, (\lambda^\mu)^{\frac{1}{2}} \overline{z}_1, (\lambda^\mu)^{\frac{3}{2}} \overline{z}_2, \lambda^\mu q\Big)\Big|_{\varepsilon=\mu^2}, \\
\check{g}_2^\mu(z,q) &= (\lambda^\mu)^{-1} \tilde{g}_2^\varepsilon\Big((\lambda^\mu)^{\frac{1}{2}} z_1, (\lambda^\mu)^{\frac{3}{2}} z_2, (\lambda^\mu)^{\frac{1}{2}} \overline{z}_1, (\lambda^\mu)^{\frac{3}{2}} \overline{z}_2, \lambda^\mu q\Big)\Big|_{\varepsilon=\mu^2}, \\
\check{h}^\mu(z) &= (\lambda^\mu)^{-1} \tilde{h}^\varepsilon\Big((\lambda^\mu)^{\frac{1}{2}} z_1, (\lambda^\mu)^{\frac{3}{2}} z_2, (\lambda^\mu)^{\frac{1}{2}} \overline{z}_1, (\lambda^\mu)^{\frac{3}{2}} \overline{z}_2\Big)\Big|_{\varepsilon=\mu^2}.
\end{aligned}$$

We use the analogous abbreviations L_{sh}^μ and L_{c}^μ for respectively $L_{\mathrm{sh}}^\varepsilon|_{\varepsilon=\mu^2}$ and $L_{\mathrm{c}}^\varepsilon|_{\varepsilon=\mu^2}$.

The transformed nonlinearities satisfy the estimates

$$\begin{aligned}
|\check{f}_1^\mu(z)| &= \mathcal{O}(|z|^3),\\
|\check{f}_2^\mu(z,q)| &= \mathcal{O}(\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}}\|(z,q)\|_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^t_{\mathrm{sh,c}}}),\\
\|\check{g}_1^\mu(z,q)\|_t &= \mathcal{O}(\mu\|(z,q)\|_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^t_{\mathrm{sh,c}}}),\\
\|\check{g}_2^\mu(z,q)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} &= \mathcal{O}(\mu|z|\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}}+\mu^2\|q\|^2_{\mathcal{X}^t_{\mathrm{sh,c}}}),\\
\|\check{h}^\mu(z)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} &= \mathcal{O}(\mu|z|^2|(z,\mu)|)
\end{aligned}$$

for $t>\frac{1}{2}$.

Preliminary transformations

Our existence theory for pulse solutions in Chapters 4 and 5 is based upon the construction of an iteration scheme, the convergence of which requires a sufficient power of μ in the corresponding estimates for the nonlinearities. The following transformations remove some lower-order terms contained in the Maclaurin expansions of the nonlinearities on the right-hand side of equations (2.42), (2.43) whose presence lead to difficulties in obtaining these estimates.

The first transformation affects the weakly hyperbolic part of the scaled system (2.42), (2.43) by removing terms whose order is less than four and which are linear in q from the Maclaurin expansion of $\check{f}_2^\mu(z,q)$ at the expense of modifying the higher-order terms in the Maclaurin expansions of $\check{f}_2^\mu(z,q)$ and the other nonlinearities.

Lemma 2.14. *There exists a near-identity, finite dimensional change of coordinates which transforms the coupled system* (2.42), (2.43) *into*

$$z_\xi = L^\mu_{\mathrm{wh}}z+\overline{f}_1^\mu(z)+\overline{f}_2^\mu(z,q), \tag{2.44}$$

$$q_\xi = L^\mu_{\mathrm{sh,c}}q+P_{\mathrm{sh,c}}\Big(\overline{g}_1^\mu(z,q)q_\tau\Big)+\overline{g}_2^\mu(z,q)+\overline{h}^\mu(z), \tag{2.45}$$

and preserves the reversibility. The nonlinearities $\overline{h}^\mu$, $\overline{g}_1^\mu$ and $\overline{g}_2^\mu$ satisfy the same estimates as respectively $\check{h}^\mu$, $\check{g}_1^\mu$ and $\check{g}_2^\mu$, while

$$|\overline{f}_1^\mu(z)-\check{f}_1^\mu(z)| = \mathcal{O}(|z|^3|(z,\mu)|), \tag{2.46}$$

$$|\overline{f}_2^\mu(z,q)| = \mathcal{O}(|z||(z,\mu)|^2\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}}+\|q\|^2_{\mathcal{X}^t_{\mathrm{sh,c}}}) \tag{2.47}$$

for $t>\frac{1}{2}$.

Proof. Since $\check{f}_2^\mu$ is an analytic function taking values in $\mathcal{X}_{\mathrm{wh}}$ with

$$|\check{f}_2^\mu(z,q)| = \mathcal{O}(\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}}\|(z,q)\|_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^t_{\mathrm{sh,c}}}),$$

we can write

$$\check{f}_2^\mu(z,q) = \sum_{j=1}^{2}\sum_{i=0}^{2-j}\mu^i\check{f}_2^{i;j,1}(z,q)+\mathcal{O}(|z||(z,\mu)|^2\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}}+\|q\|^2_{\mathcal{X}^t_{\mathrm{sh,c}}}), \tag{2.48}$$

where $\mu^i \check{f}_2^{i;j,k}(z,q)$ denotes the term in the Maclaurin expansion of $\check{f}_2^{\mu}(z,q)$ which is homogeneous of degree i, j and k in respectively μ, z and q.

To eliminate the term $\mu^i \check{f}_2^{i;j,1}(z,q)$ we use a finite-dimensional change of coordinates of the form

$$\tilde{z} = z + \mu^i d_{i,j}(z,q),$$

in which $d_{i,j} \colon \mathcal{X}_{\mathrm{wh}} \times \mathcal{X}_{\mathrm{sh,c}}^t \to \mathcal{X}_{\mathrm{wh}}$ is analytic at the origin for $t > \frac{1}{2}$, homogeneous of degree j in its first argument, and linear in its second. We write the inverse transformation as

$$z = \tilde{z} - \mu^i d_{i,j}(\tilde{z},q) + \mathcal{R}_1^{\mu}(\tilde{z},q)$$

where $\mathcal{R}_1^{\mu}$ is analytic at the origin with

$$|\mathcal{R}_1^{\mu}(\tilde{z},q)| = \mathcal{O}(\mu^i |\tilde{z}|^j \|q\|^2_{\mathcal{X}_{\mathrm{sh,c}}^t}).$$

Using equation (2.42), we find that the resulting equation for $\tilde{z}$ is

$$\tilde{z}_\xi = L_{\mathrm{wh}}^0 \tilde{z} + \check{f}_1^{\mu}(\tilde{z}) + \check{f}_2^{\mu}(\tilde{z},q) + \mu^i \mathrm{d}_1 d_{i,j}[\tilde{z},q](L_{\mathrm{wh}}^0 \tilde{z}) + \mu^i d_{i,j}(\tilde{z}, L_{\mathrm{sh,c}}^0 q) - \mu^i L_{\mathrm{wh}}^0 d_{i,j}(\tilde{z},q) + \mathcal{R}_2^{\mu}(\tilde{z},q),$$

where

$$|\mathcal{R}_2^{\mu}(\tilde{z},q)| = \mathcal{O}(|\tilde{z}||(\tilde{z},\mu)|^{i+j}\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t} + \|q\|^2_{\mathcal{X}_{\mathrm{sh,c}}^t} + |\tilde{z}|^3|(\tilde{z},\mu)|),$$

so that our objective is achieved by choosing $d_{i,j}$ so that

$$(\mathcal{L} d_{i,j})(\tilde{z},q) = -\check{f}_2^{i;j,1}(\tilde{z},q) \tag{2.49}$$

for each $(\tilde{z},q) \in \mathcal{X}_{\mathrm{wh}} \times \mathcal{X}_{\mathrm{sh,c}}^t$, where

$$(\mathcal{L} d_{i,j})(\tilde{z},q) = d_{i,j}(\tilde{z}, L_{\mathrm{sh,c}}^0 q) - L_{\mathrm{wh}}^0 d_{i,j}(\tilde{z},q) + \mathrm{d}_1 d_{i,j}[\tilde{z},q](L_{\mathrm{wh}}^0 \tilde{z}).$$

Note that a solution $d_{i,j}$ to equation (2.49) commutes with the reverser R, so that the transformed system (2.44), (2.45) is reversible.

In order to solve equation (2.49) we write

$$d_{i,j}(z,q) = \begin{pmatrix} d_{i,j,1}(z,q) \\ d_{i,j,2}(z,q) \\ \overline{d_{i,j,1}(z,q)} \\ \overline{d_{i,j,2}(z,q)} \end{pmatrix}, \quad \check{f}_2^{i;j,1}(z,q) = \begin{pmatrix} f_{i,j,1}(z,q) \\ f_{i,j,2}(z,q) \\ \overline{f_{i,j,1}(z,q)} \\ \overline{f_{i,j,2}(z,q)} \end{pmatrix},$$

where

$$d_{i,j,k}(z,q) = \sum_{|\boldsymbol{\alpha}|=j} \sum_{\substack{-j \le m \le j \\ 1 \le r \le r_{-m+1}}} d_{i,j,k}^{\boldsymbol{\alpha},m,r} z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} q_{-m+1,r}$$

and

$$f_{i,j,k}(z,q) = \sum_{|\boldsymbol{\alpha}|=j} \sum_{\substack{-j \le m \le j \\ 1 \le r \le r_{-m+1}}} f_{i,j,k}^{\boldsymbol{\alpha},m,r} z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} q_{-m+1,r},$$

for $k=1,2$, in which $d_{i,j,k}^{\alpha,m,r}$ and $f_{i,j,k}^{\alpha,m,r}$ are complex numbers and we have used the coordinates for q given in Section 2.1.

Noting that

$$\mathcal{L}\begin{pmatrix} d_1 z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} q_{-m+1,r} \\ d_2 z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} q_{-m+1,r} \\ \overline{d}_1 z_1^{\alpha_3} z_2^{\alpha_4} \overline{z}_1^{\alpha_1} \overline{z}_2^{\alpha_4} \overline{q}_{-m+1,r} \\ \overline{d}_2 z_1^{\alpha_3} z_2^{\alpha_4} \overline{z}_1^{\alpha_1} \overline{z}_2^{\alpha_4} \overline{q}_{-m+1,r} \end{pmatrix} = \begin{pmatrix} (\lambda_{-m+1,r} - \mathrm{i}\omega + m\mathrm{i}\omega) d_1 z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} q_{-m+1,r} \\ (\lambda_{-m+1,r} - \mathrm{i}\omega + m\mathrm{i}\omega) d_2 z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} q_{-m+1,r} \\ (\overline{\lambda}_{-m+1,r} + \mathrm{i}\omega - m\mathrm{i}\omega) \overline{d}_1 z_1^{\alpha_3} z_2^{\alpha_4} \overline{z}_1^{\alpha_1} \overline{z}_2^{\alpha_2} \overline{q}_{-m+1,r} \\ (\overline{\lambda}_{-m+1,r} + \mathrm{i}\omega - m\mathrm{i}\omega) \overline{d}_2 z_1^{\alpha_3} z_2^{\alpha_4} \overline{z}_1^{\alpha_1} \overline{z}_2^{\alpha_2} \overline{q}_{-m+1,r} \end{pmatrix}$$

for each $d_1, d_2 \in \mathbb{C}$, we find that equation (2.49) is equivalent to the equations

$$(\lambda_{-m+1,r} + (m-1)\mathrm{i}\omega) d_{i,j,k}^{\alpha,m,r} = -f_{i,j,k}^{\alpha,m,r}, \qquad k=1,2,$$

for the coefficients $d_{i,j,k}^{\alpha,m,r}$. Assumptions 2(iii) guarantee that $\lambda \neq l\mathrm{i}\omega$ for each $\lambda \in \sigma(L_{\mathrm{sh,c}}^0)$ and $l \in \mathbb{Z}$ and therefore these equations are always solvable.

We use the composition of the above changes of coordinates for each $j \in \{1,2\}$ and $i \in \{0,\dots,2-j\}$ in the order $(i,j)=(0,1)$, $(i,j)=(1,1)$, $(i,j)=(0,2)$ to establish the lemma. ∎

There are also terms in the central part whose presence leads to difficulties in the theory presented in Chapters 4 and 5, namely terms which are linear in μ, z and $Z = P_{\mathrm{sh}} q$.

Lemma 2.15. *There exists a near-identity, finite dimensional change of coordinates which transforms the coupled system (2.44), (2.45) into*

$$z_\xi = L_{\mathrm{wh}}^\mu z + \hat{f}_1^\mu(z) + \hat{f}_2^\mu(z,q), \tag{2.50}$$

$$q_\xi = L_{\mathrm{sh,c}}^\mu q + P_{\mathrm{sh,c}}\Big(\hat{g}_1^\mu(z,q) q_\tau\Big) + \hat{g}_2^\mu(z,q) + \hat{h}^\mu(z), \tag{2.51}$$

and preserves the reversibility. The nonlinearities $\hat{f}_2^\mu$, $\hat{g}_1^\mu$ and $\hat{h}^\mu$ satisfy the same estimates as respectively $\overline{f}_2^\mu$, $\overline{g}_1^\mu$ and $\overline{h}^\mu$ while

$$\|\hat{g}_2^\mu(z,q)\|_{\mathcal{X}_{\mathrm{sh,c}}^t} = \mathcal{O}(\mu|z|\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t} + \mu^2\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t}^2 + \mu\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t}^3)$$

for $t > \frac{1}{2}$. Moreover

$$\hat{f}_1^\mu(z) = \overline{f}_1^\mu(z), \quad P_{\mathrm{c}}\Big(\hat{g}_1^{1;1,0}(z,0)Z_\tau\Big) = 0, \quad P_{\mathrm{c}}\hat{g}_2^{1;1,1}(z,Z) = 0,$$

where $\mu^i \hat{g}_l^{i;j,k}(z,q)$ denotes the term in the Maclaurin expansion of $\hat{g}_l^\mu(z,q)$ which is homogeneous of degree i, j and k in respectively μ, z and q.

Proof. First note that if $b\colon \mathcal{X}_{\mathrm{wh}} \times \mathcal{X}_{\mathrm{sh}} \to \mathcal{X}_{\mathrm{c}}^t$ is bilinear, we have that $P_m b(z,Z) = 0$ for $m \notin \{-m_j \pm 1, m_j \pm 1 \colon j = 0,\dots,d\}$ (see the comments below (2.20), (2.21)). We therefore use a change of coordinates of the form

$$\tilde{q} = q + \mu\Lambda(z)Z,$$

in which $\Lambda\colon \mathcal{X}_{\mathrm{wh}} \to \mathcal{L}(\mathcal{X}_{\mathrm{sh}}, \mathcal{X}_{\mathrm{c}}^t)$ is linear for $t > \frac{1}{2}$, so that $P_m(\Lambda(z)Z) = 0$ if $m \notin \{-m_j \pm 1, m_j \pm 1\colon j = 0, \ldots, d\}$. In particular $\tilde{Z} = P_{\mathrm{sh}}\tilde{q} = P_{\mathrm{sh}}q = Z$ and using equation (2.45), we find that the resulting equation for $\tilde{q}$ is

$$\tilde{q}_\xi = L^\mu_{\mathrm{sh,c}}\tilde{q} + P_{\mathrm{sh,c}}\Big(\overline{g}^\mu_1(z,\tilde{q})\tilde{q}_\tau\Big) + \overline{g}^\mu_2(z,\tilde{q}) - \mu L^0_{\mathrm{c}}\Lambda(z)Z + \mu\Lambda(L^0_{\mathrm{wh}}z)Z + \mu\Lambda(z)L^0_{\mathrm{sh}}Z + \overline{h}^\mu(z) + \mathcal{R}(z,\tilde{q}),$$

where

$$\|\mathcal{R}(z,\tilde{q})\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(\mu|z|\|\tilde{q}\|_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu\|\tilde{q}\|^3_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu^2|z|^3|(z,\mu)|).$$

Our objective of removing terms which are linear in each of μ, z and Z in the centre part is achieved by choosing Λ so that

$$(\mathcal{L}\Lambda)(z)Z = G(z,Z), \tag{2.52}$$

where

$$(\mathcal{L}\Lambda)(z)Z = L^0_{\mathrm{c}}\Lambda(z)Z - \Lambda(z)L^0_{\mathrm{sh}}Z - \Lambda(L^0_{\mathrm{wh}}z)Z$$

and $G(z,Z) = P_{\mathrm{c}}\Big(\overline{g}^{1;1,0}_1(z,0)Z_\tau + \overline{g}^{1;1,1}_2(z,Z)\Big)$. Note that a solution Λ to equation (2.52) commutes with the reverser R, so that the transformed system (2.50), (2.51) is reversible.

The bilinear expressions $\Lambda(z)Z$ and $G(z,Z)$ in equation (2.52) are polynomials in the variables z and Z whose monomials are given by terms of the form

$$\Lambda^{m,r}_{\boldsymbol{\alpha},\beta} z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} Z_{\pm m_j,\sigma} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau}, \qquad G^{m,r}_{\boldsymbol{\alpha},\beta} z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} Z_{\pm m_j,\sigma} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau},$$

where $m \in \{-m_j \pm 1, m_j \pm 1\colon j = 0, \ldots, d\}$, $1 \leq r \leq r_m$, $\boldsymbol{\alpha} \in \mathbb{N}_0^4$ with $|\boldsymbol{\alpha}| = 1$, $\beta \in \{1, \ldots, \sigma_0 + 2(\sigma_{m_1} + \ldots + \sigma_{m_d})\}$ and $r_{m_j} - \sigma_{m_j} + 1 \leq \sigma \leq r_{m_j}$, in which we have used the coordinates for Z introduced below (2.20), (2.21) (recall that the eigenvalues are numbered such that $\lambda_{m_j, r_{m_j} - \sigma_{m_j} + 1}$ is the first strongly hyperbolic mode m_j eigenvalue).

Consider the vector space $\mathcal{Q}^2$ of bilinear polynomials in the variables z_1, z_2, $\overline{z}_1$, $\overline{z}_2$ and $\{Z_{\pm m_j,\sigma}\}$ taking values in $\mathcal{X}^{s+1}_{\mathrm{c}}$ and note that $\mathcal{Q}^2$ decomposes into the direct sum

$$\bigoplus_{\substack{0 \leq j \leq d \\ m \in \{-m_j \pm 1, m_j \pm 1\}}} \mathcal{Q}^2_{m,r},$$

in which the vector space $\mathcal{Q}^2_{m,r}$ of bilinear polynomials in the variables z_1, z_2, $\overline{z}_1$, $\overline{z}_2$ and $\{Z_{\pm m_j,\sigma}\}$ taking values in $\langle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle$ is invariant under $\mathcal{L}$.

In order to achieve cancellation of the terms which are linear in z_1 and $Z_{0,r_0-\sigma_0+1}$ in equation (2.52), we compute the corresponding monomial of $(\mathcal{L}\Lambda)(z)Z$ to be

$$(\underbrace{\mathrm{i}\mu_{m,r}}_{\in\,\sigma(L^0_{\mathrm{c}})} - \underbrace{\lambda_{0,r_0-\sigma_0+1}}_{\in\,\sigma(L^0_{\mathrm{sh}})} - \mathrm{i}\omega)\Lambda^{m,r}_{e_1,1} z_1 Z_{0,r_0-\sigma_0+1} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau},$$

in which $\boldsymbol{e}_1 = (1,0,0,0) \in \mathbb{N}_0^4$, so that it is necessary to choose $\Lambda^{m,r}_{\boldsymbol{e}_1,1}$ such that

$$(\mathrm{i}\mu_{m,r} - \lambda_{0,r_0-\sigma_0+1} - \mathrm{i}\omega)\Lambda^{m,r}_{\boldsymbol{e}_1,1} = G^{m,r}_{\boldsymbol{e}_1,1}.$$

Since $\mathrm{Re}(\lambda_{0,r_0-\sigma_0+1}) \neq 0$, we find that this choice is always possible. A similar argument applies to the remaining monomials so that $\mathcal{L}|_{\mathcal{Q}^2_{m,r}}$ is invertible and equation (2.52) can be solved for Λ. ■

Results of the normal-form theory

In Chapter 3 below we use a sequence of normal-form transformations which systematically remove terms in the Maclaurin expansion of the nonlinearity $\hat{h}^\mu$ in equation (2.51), so that the remaining terms become exponentially small in comparison with μ. Applying Lemma 2.13 to the transformed system yields homoclinic solutions $p^{\mu\pm}$ which can be used in a perturbation argument as approximations to genuine pulse solutions. Note that the normal-form transformations do not affect the leading-order terms in the scaled system (2.31)–(2.34) (see the comments above Lemma 2.10), so that the leading-order term in the asymptotic formula for $p^{\mu\pm}$ is also unaffected. The result of the theory in Chapter 3 is summarised in the following theorem.

Theorem 2.16. *There exist a constant $c^\star > 0$ and a near-identity, finite-dimensional change of coordinates which transforms the coupled system (2.50), (2.51) into*

$$z_\xi = L^\mu_{\mathrm{wh}} z + \tilde{f}^\mu_1(z) + \tilde{f}^\mu_2(z,q), \tag{2.53}$$

$$q_\xi = L^\mu_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}\Big(\tilde{g}^\mu_1(z,q)q_\tau\Big) + \tilde{g}^\mu_2(z,q) + \tilde{h}^\mu(z), \tag{2.54}$$

and preserves the reversibility. The transformed nonlinearities $\tilde{f}^\mu_1$, $\tilde{f}^\mu_2$ and $\tilde{g}^\mu_1$ satisfy the same estimates as respectively $\hat{f}^\mu_1$, $\hat{f}^\mu_2$ and $\hat{g}^\mu_1$ while

$$\|\tilde{g}^\mu_2(z,q)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} = \mathcal{O}(\mu|z|\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu^2\|q\|^2_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu\|q\|^3_{\mathcal{X}^t_{\mathrm{sh,c}}} + |z|^3\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}})$$

and

$$\|\tilde{h}^\mu(z)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \lesssim \mu^2 \mathrm{e}^{-c^\star/\mu}, \qquad \|\mathrm{d}\tilde{h}^\mu[z]\|_{\mathcal{L}(\mathcal{X}_{\mathrm{wh}},\mathcal{X}^t_{\mathrm{sh,c}})} \lesssim \mu \mathrm{e}^{-c^\star/\mu}$$

for $t > \frac{1}{2}$. Moreover

$$P_{\mathrm{c}}\Big(\tilde{g}^{1;1,0}_1(z,0)Z_\tau\Big) = P_{\mathrm{c}}\tilde{g}^{1;1,1}_2(z,Z) = 0$$

and

$$|\tilde{f}^\mu_1(z) - \hat{f}^\mu_1(z)| = \mathcal{O}(|z|^3|(z,\mu)|).$$

Here we gather some material which is used in the existence theory in Chapters 4 and 5 below. We begin with a reformulation of the coupled systems (2.53), (2.54) by writing z as a perturbation of the approximate pulses constructed in Section 2.2. We also split equation (2.54) into a finite-dimensional system for the strongly hyperbolic part $Z = P_{\mathrm{sh}}q$ and an infinite-dimensional system for the centre part $w = P_{\mathrm{c}}q$.

Applying Lemma 2.13 (with a brief return to unscaled variables) yields homoclinic solutions $p^{\mu\pm}$ to the equation

$$z_\xi = L^\mu_{\mathrm{wh}} z + \tilde{f}^\mu_1(z)$$

which satisfy the estimate

$$|p^{\mu\pm}(\xi)| \leq c_{\mathrm{h}} \mu \mathrm{e}^{-\eta\lambda^\mu|\xi|}, \quad \xi \in \mathbb{R}.$$

Writing

$$\tilde{f}^\mu_1(z) = \begin{pmatrix} \tilde{f}^\mu_{1,1}(z) \\ \tilde{f}^\mu_{1,2}(z) \\ \overline{\tilde{f}^\mu_{1,1}(z)} \\ \overline{\tilde{f}^\mu_{1,2}(z)} \end{pmatrix}$$

and using the formula

$$p^{\mu\pm}(\xi) = \sqrt{\lambda^\mu} \begin{pmatrix} \mathrm{e}^{\mathrm{i}(\omega+\nu^\mu)\xi}(\breve{p}^{\mu\pm}_1(\lambda^\mu\xi) + \mathrm{i}\breve{p}^{\mu\pm}_2(\lambda^\mu\xi)) \\ \mathrm{e}^{\mathrm{i}(\omega+\nu^\mu)\xi}(\breve{q}^{\mu\pm}_1(\lambda^\mu\xi) + \mathrm{i}\breve{q}^{\mu\pm}_2(\lambda^\mu\xi)) \\ \mathrm{e}^{-\mathrm{i}(\omega+\nu^\mu)\xi}(\breve{p}^{\mu\pm}_1(\lambda^\mu\xi) - \mathrm{i}\breve{p}^{\mu\pm}_2(\lambda^\mu\xi)) \\ \mathrm{e}^{-\mathrm{i}(\omega+\nu^\mu)\xi}(\breve{q}^{\mu\pm}_1(\lambda^\mu\xi) - \mathrm{i}\breve{q}^{\mu\pm}_2(\lambda^\mu\xi)) \end{pmatrix}$$

in which $\breve{p}^{\mu\pm}_j = \breve{p}^{\varepsilon\pm}_j|_{\varepsilon=\mu^2}$ (see Lemma 2.13), we conclude that

$$\sqrt{\lambda^\mu}\tilde{R}_{(\omega+\nu^\mu)\xi} \begin{pmatrix} \breve{p}^{\mu\pm}_1(\lambda^\mu\xi) \\ \breve{q}^{\mu\pm}_1(\lambda^\mu\xi) \\ \breve{p}^{\mu\pm}_2(\lambda^\mu\xi) \\ \breve{q}^{\mu\pm}_2(\lambda^\mu\xi) \end{pmatrix}$$

with

$$\tilde{R}_a = \begin{pmatrix} \cos a & 0 & -\sin a & 0 \\ 0 & \cos a & 0 & -\sin a \\ \sin a & 0 & \cos a & 0 \\ 0 & \sin a & 0 & \cos a \end{pmatrix}, \quad a \in \mathbb{R},$$

solves the equation

$$\begin{pmatrix} x_{1\xi} \\ y_{1\xi} \\ x_{2\xi} \\ y_{2\xi} \end{pmatrix} = \begin{pmatrix} -\omega x_2 \\ -\omega y_2 \\ \omega x_1 \\ \omega y_1 \end{pmatrix} + \begin{pmatrix} \mathrm{Re}\, \tilde{f}^\mu_{1,1}(z) \\ \mathrm{Re}\, \tilde{f}^\mu_{1,2}(z) \\ \mathrm{Im}\, \tilde{f}^\mu_{1,1}(z) \\ \mathrm{Im}\, \tilde{f}^\mu_{1,2}(z) \end{pmatrix}$$

where

$$z_1 = x_1 + \mathrm{i}x_2, \quad z_2 = y_1 + \mathrm{i}y_2.$$

We henceforth use real coordinates for real-valued functions in $\mathcal{X}_{\mathrm{wh}}$ and write

$$z = \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix}, \quad p^\mu = \sqrt{\lambda^\mu}\tilde{R}_{(\omega+\nu^\mu)\xi} \begin{pmatrix} \breve{p}^\mu_1(\lambda^\mu\xi) \\ \breve{q}^\mu_1(\lambda^\mu\xi) \\ \breve{p}^\mu_2(\lambda^\mu\xi) \\ \breve{q}^\mu_2(\lambda^\mu\xi) \end{pmatrix},$$

where $\check{p}_j^\mu$ and $\check{q}_j^\mu$ each denote either of the functions $\check{p}_j^{\mu\pm}$ and $\check{q}_j^{\mu\pm}$ respectively. Note that in real coordinates the matrix of L^μ_{wh} with respect to the basis

$$\{\mathrm{Re}(u\mathrm{e}^{\mathrm{i}\tau}), \mathrm{Re}(v\mathrm{e}^{\mathrm{i}\tau}), \mathrm{Im}(u\mathrm{e}^{\mathrm{i}\tau}), \mathrm{Im}(v\mathrm{e}^{\mathrm{i}\tau})\}$$

(see equations (2.16), (2.17)) is given by

$$\begin{pmatrix} 0 & \lambda^\mu & -(\omega+\nu^\mu) & 0 \\ \lambda^\mu & 0 & 0 & -(\omega+\nu^\mu) \\ \omega+\nu^\mu & 0 & 0 & \lambda^\mu \\ 0 & \omega+\nu^\mu & \lambda^\mu & 0 \end{pmatrix}.$$

In the existence theory in Chapters 4 and 5 we use a perturbation argument, writing $z = p^\mu + r$. Decoupling the system into equations for the strongly hyperbolic part $Z = P_{\mathrm{sh}}q$ and the centre part $w = P_{\mathrm{c}}q$ yields

$$Z_\xi = L^\mu_{\mathrm{sh}} Z + G_1^\mu(Z, p^\mu + r, w) + G_2^\mu(r), \tag{2.55}$$

$$r_\xi = L^\mu_{\mathrm{wh}} r + \mathcal{L}^\mu r + f_3^\mu(r) + f_4^\mu(Z, r, w), \tag{2.56}$$

$$w_\xi = L^\mu_{\mathrm{c}} w + P_{\mathrm{c}}\Big(\hat{g}_1^\mu(Z, p^\mu + r, w) w_\tau\Big) + \mu b(p^\mu + r, w) + \hat{g}_2^\mu(Z, p^\mu + r, w) + \hat{h}^\mu(r). \tag{2.57}$$

Here

$$\begin{aligned} G_1^\mu(Z,z,w) &= P_{\mathrm{sh}}\Big(\tilde{g}_1^\mu(z, Z+w)(Z+w)_\tau + \tilde{g}_2^\mu(z, Z+w)\Big), \\ G_2^\mu(r) &= P_{\mathrm{sh}}\tilde{h}^\mu(p^\mu + r), \\ \mathcal{L}^\mu &= \mathrm{d}\tilde{f}_1^\mu[p^\mu], \\ f_3^\mu(r) &= \tilde{f}_1^\mu(p^\mu + r) - \tilde{f}_1^\mu(p^\mu) - \mathrm{d}\tilde{f}_1^\mu[p^\mu](r), \\ f_4^\mu(Z, r, w) &= \tilde{f}_2^\mu(p^\mu + r, Z + w), \\ \hat{g}_1^\mu(Z,z,w) &= \tilde{g}_1^\mu(z, Z+w), \\ \hat{g}_2^\mu(Z,z,w) &= P_{\mathrm{c}}\Big(\tilde{g}_2^\mu(z, Z+w) - \mu\tilde{g}_2^{1;1,1}(z,w) + \tilde{g}_1^\mu(z, Z+w)Z_\tau\Big) \\ &= P_{\mathrm{c}}\Big(\tilde{g}_2^\mu(z, Z+w) - \mu\tilde{g}_2^{1;1,1}(z, Z+w) + \tilde{g}_1^\mu(z, Z+w)Z_\tau - \mu\tilde{g}_1^{1;1,0}(z,0)Z_\tau\Big), \\ b(z,w) &= P_{\mathrm{c}}\tilde{g}_2^{1;1,1}(z,w), \\ \hat{h}^\mu(r) &= P_{\mathrm{c}}\tilde{h}^\mu(p^\mu + r), \end{aligned}$$

where we have used the facts that

$$P_{\mathrm{c}}\tilde{g}_2^{1;1,1}(z,Z) = 0, \quad P_{\mathrm{c}}\Big(\tilde{g}_1^{1;1,0}(z,0)Z_\tau\Big) = 0.$$

The following estimates for the nonlinearities G_1^μ, $\hat{g}_1^\mu$ and $\hat{g}_2^\mu$ are a direct consequence of the estimates given in Theorem 2.16.

Proposition 2.17.

(i) The nonlinearity G_1^μ in equation (2.55) satisfies the estimates

$$|G_1^\mu(Z,z,w)| = \mathcal{O}(\mu|z|\|(Z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{c}}^t} + \mu\|(Z,w)\|^2_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{c}}^t} + |z|^3\|(Z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{c}}^t})$$

and

$$\|\mathrm{d}G_1^\mu[Z,z,w]\|_{\mathcal{L}(\mathcal{X}^t;\mathcal{X}_{\mathrm{sh}})} = \mathcal{O}(\mu\|(Z,z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}_{\mathrm{c}}^t} + |z|^2\|(Z,z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}_{\mathrm{c}}^t})$$

for $t > \frac{1}{2}$.

(ii) The nonlinearities $\hat{g}_1^\mu$ and $\hat{g}_2^\mu$ in equation (2.57) satisfy the estimates

$$\begin{aligned}\|\hat{g}_1^\mu(Z,z,w)\|_t &= \mathcal{O}(\mu\|(Z,z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}_{\mathrm{c}}^t}),\\ \|\mathrm{d}\hat{g}_1^\mu[Z,z,w]\|_{\mathcal{L}(\mathcal{X}^t;H^t_{\mathrm{per}})} &= \mathcal{O}(\mu)\end{aligned}$$

and

$$\begin{aligned}\|\hat{g}_2^\mu(Z,z,w)\|_{\mathcal{X}_{\mathrm{c}}^t} &= \mathcal{O}(|z||(z,\mu)|^2\|(Z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{c}}^t} + \mu\|(Z,w)\|^2_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{c}}^t}),\\ \|\mathrm{d}\hat{g}_2^\mu[Z,z,w]\|_{\mathcal{L}(\mathcal{X}^t;\mathcal{X}_{\mathrm{c}}^t)} &= \mathcal{O}(|(z,\mu)|^2\|(Z,z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}_{\mathrm{c}}^t} + \mu\|(Z,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{c}}^t})\end{aligned}$$

for $t > \frac{1}{2}$.

Our aim in Chapters 4 and 5 below is to construct exponentially small solutions (Z,r,w) to equations (2.55)–(2.57) in order to find solutions $(p^\mu + r, Z + w)$ to equations (2.53), (2.54) which are exponentially close to p^μ. The following result summarises the corresponding estimates for the relevant nonlinearities.

Proposition 2.18. *Suppose that $t > \frac{1}{2}$, $|r| \leq \mathrm{e}^{-c^\star/2\mu}$, $|Z| \leq \mathrm{e}^{-c^\star/2\mu}$ and $\|w\|_{\mathcal{X}_{\mathrm{c}}^{s+1}} \leq \mathrm{e}^{-c^\star/2\mu}$. The nonlinearities f_3^μ, f_4^μ, G_2^μ and $\hat{h}^\mu$ in equations (2.55)–(2.57) satisfy the estimates*

$$\begin{aligned}|f_3^\mu(r)| &\lesssim |r|^2, \quad \|\mathrm{d}f_3^\mu[r]\|_{\mathcal{L}(\mathcal{X}_{\mathrm{wh}};\mathcal{X}_{\mathrm{wh}})} \lesssim |r|\\ |f_4^\mu(Z,r,w)| &\lesssim \mu^3\|(Z,r,w)\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}_{\mathrm{c}}^t}, \quad \|\mathrm{d}f_4^\mu[Z,r,w]\|_{\mathcal{L}(\mathcal{X}^t;\mathcal{X}_{\mathrm{wh}})} \lesssim \mu^3,\\ |G_2^\mu(z)| &\lesssim \mu^2\mathrm{e}^{-c^\star/\mu}, \quad \|\mathrm{d}G_2^\mu[z]\|_{\mathcal{L}(\mathcal{X}_{\mathrm{wh}};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu\mathrm{e}^{-c^\star/\mu},\\ \|\hat{h}^\mu(z)\|_{\mathcal{X}_{\mathrm{c}}^t} &\lesssim \mu^2\mathrm{e}^{-c^\star/\mu}, \quad \|\mathrm{d}\hat{h}^\mu[z]\|_{\mathcal{L}(\mathcal{X}_{\mathrm{wh}};\mathcal{X}_{\mathrm{c}}^t)} \lesssim \mu\mathrm{e}^{-c^\star/\mu}.\end{aligned}$$

Reformulation of the hyperbolic part

The finite-dimensional strongly hyperbolic and weakly hyperbolic parts (2.55) and (2.56) are handled by reformulating them as integral equations

$$\begin{aligned}Z(\xi) = &\int_{-\infty}^{\xi} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{ss}}\Big(G_1^\mu(Z, p^\mu + \mathcal{E}_{\mathrm{wh}}r, \mathcal{E}_{\mathrm{c}}w)(t) + G_2^\mu(\mathcal{E}_{\mathrm{wh}}r)(t)\Big)\,\mathrm{d}t\\ &- \int_{\xi}^{\infty} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{su}}\Big(G_1^\mu(Z, p^\mu + \mathcal{E}_{\mathrm{wh}}r, \mathcal{E}_{\mathrm{c}}w)(t) + G_2^\mu(\mathcal{E}_{\mathrm{wh}}r)(t)\Big)\,\mathrm{d}t \qquad (2.58)\end{aligned}$$

and

$$r(\xi) = \sum_{i=1}^{2} \Big(\int_0^\xi \Big\langle \Big(f_3^\mu(r)(t) + f_4^\mu(Z,r,w)(t) \Big), (S_i^\mu)^*(t) \Big\rangle \,\mathrm{d}t \Big) S_i^\mu(\xi) \\ - \Big(\int_\xi^\infty \Big\langle \Big(f_3^\mu(r)(t) + f_4^\mu(Z,r,w)(t) \Big), (U_i^\mu)^*(t) \Big\rangle \,\mathrm{d}t \Big) U_i^\mu(\xi). \tag{2.59}$$

Here P_{ss} and P_{su} denote the projections onto the *stable subspace* of L_{sh}^μ spanned by the eigenvectors of L_{sh}^μ whose corresponding eigenvalues have negative real part, and the *unstable subspace* of L_{sh}^μ spanned by the eigenvectors of L_{sh}^μ whose corresponding eigenvalues have positive real part. Since there exist only finitely many eigenvalues of L^μ with non-zero real part due to Proposition 2.3, we find that there exist positive constants ω_1, ω_2 such that

$$\|\mathrm{e}^{\xi L_{\mathrm{sh}}^\mu} P_{\mathrm{ss}}\|_{\mathcal{L}(\mathcal{X}_{\mathrm{sh}})} \lesssim \mathrm{e}^{-\xi\omega_1}, \quad \|\mathrm{e}^{\xi L_{\mathrm{sh}}^\mu} P_{\mathrm{su}}\|_{\mathcal{L}(\mathcal{X}_{\mathrm{sh}})} \lesssim \mathrm{e}^{\xi\omega_2}, \qquad \xi \in \mathbb{R}. \tag{2.60}$$

The mappings $\mathcal{E}_{\mathrm{wh}}$ and $\mathcal{E}_{\mathrm{c}}$ are the extensions by reversibility defined by

$$(\mathcal{E}_{\mathrm{wh}} r)(\xi) = \begin{cases} r(\xi), & \xi \geq 0, \\ R_{\mathrm{wh}} r(-\xi), & \xi < 0, \end{cases} \quad (\mathcal{E}_{\mathrm{c}} w)(\xi) = \begin{cases} w(\xi), & \xi \geq 0, \\ R_{\mathrm{c}} w(-\xi), & \xi < 0. \end{cases}$$

A solution $r(\xi)$ to equation (2.59) which is defined for $\xi \geq 0$ and satisfies $R_{\mathrm{wh}} r(0) = r(0)$ automatically gives rise to a reversible solution $r(\xi)$ to equation (2.56) due to extension by reversibility. The reversibility of a solution Z to equation (2.58) defined for $\xi \in \mathbb{R}$ is verified *a posteriori* by observing that $\xi \mapsto R_{\mathrm{sh}} Z(-\xi)$ solves equation (2.58) whenever $\xi \mapsto Z(\xi)$ is a solution. By uniqueness we thus obtain a reversible solution to equation (2.55). Finally $\{S_1^\mu, S_2^\mu, U_1^\mu, U_2^\mu\}$ is a fundamental solution set for the linearised equation for r, a complete quantitative description of which is given by the following result. Its proof is deferred to Section 2.4.

Lemma 2.19. *The equation*

$$r_\xi = L_{\mathrm{wh}}^\mu r + \mathcal{L}^\mu r \tag{2.61}$$

has linearly independent solutions S_1^μ, S_2^μ, U_1^μ, U_2^μ on $[0,\infty)$ with

$$|S_1^\mu(\xi)|, |S_2^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \quad |U_1^\mu(\xi)|, |U_2^\mu(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \qquad \xi \in [0,\infty),$$

The dual basis $\{(S_1^\mu)^*(\xi), (S_2^\mu)^*(\xi), (U_1^\mu)^*(\xi), (U_2^\mu)^*(\xi)\}$ to $\{S_1^\mu(\xi), S_2^\mu(\xi), U_1^\mu(\xi), U_2^\mu(\xi)\}$ in $\mathbb{R}^4$ satisfies

$$|(S_1^\mu)^*(\xi)|, |(S_2^\mu)^*(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \quad |(U_1^\mu)^*(\xi)|, |(U_2^\mu)^*(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \qquad \xi \in [0,\infty).$$

The functions U_i^μ are reversible and the functions S_i^μ are anti-reversible, that is they satisfy

$$R_{\mathrm{wh}} U_i^\mu(0) = U_i^\mu(0), \quad R_{\mathrm{wh}} S_i^\mu(0) = -S_i^\mu(0), \qquad i = 1, 2.$$

Remark 2.20. In order to obtain satisfactory energy estimates to prove the convergence of an iteration scheme for equations (2.57)–(2.59) in the existence theory in Chapter 4 below, the domain of integration is restricted to respectively $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$ and $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$. We thus construct solutions to equations (2.53), (2.54) which are exponentially close to p^μ on the exponentially large interval $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$.

In the semilinear case (Chapter 5) we use a different iteration scheme to construct solutions to equations (2.53), (2.54) on the entire real line. For this purpose we need a similar result for the linearisation of the unperturbed equation (2.53), namely

$$z_\xi = L^\mu_{\mathrm{wh}} z + \tilde{f}^\mu_1(z) + \tilde{f}^\mu_2(Z, z, w),$$

which we reformulate as the integral equation

$$\begin{aligned} z(\xi) = \sum_{i=3}^{4} \Big(\int_{-\infty}^{\xi} \Big\langle \Big(\tilde{f}^\mu_1(z)(t) + \tilde{f}^\mu_2(Z, z, w)(t) \Big), (S^\mu_i)^*(t) \Big\rangle \,\mathrm{d}t \Big) S^\mu_i(\xi) \\ - \Big(\int_{\xi}^{\infty} \Big\langle \Big(\tilde{f}^\mu_1(z)(t) + \tilde{f}^\mu_2(Z, z, w)(t) \Big), (U^\mu_i)^*(t) \Big\rangle \,\mathrm{d}t \Big) U^\mu_i(\xi), \end{aligned} \tag{2.62}$$

where $\{S^\mu_3, S^\mu_4, U^\mu_3, U^\mu_4\}$ is the fundamental solution set for the linearised equation for z given by the following result.

Lemma 2.21. *The equation*

$$z_\xi = L^\mu_{\mathrm{wh}} z \tag{2.63}$$

has linearly independent solutions S^μ_3, S^μ_4, U^μ_3, U^μ_4 on $\mathbb{R}$ with

$$|S^\mu_3(\xi)|, |S^\mu_4(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \quad |U^\mu_3(\xi)|, |U^\mu_4(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \qquad \xi \in \mathbb{R},$$

The dual basis $\{(S^\mu_3)^*(\xi), (S^\mu_4)^*(\xi), (U^\mu_3)^*(\xi), (U^\mu_4)^*(\xi)\}$ to $\{S^\mu_3(\xi), S^\mu_4(\xi), U^\mu_3(\xi), U^\mu_4(\xi)\}$ in $\mathbb{R}^4$ satisfies

$$|(S^\mu_3)^*(\xi)|, |(S^\mu_4)^*(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \quad |(U^\mu_3)^*(\xi)|, |(U^\mu_4)^*(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \qquad \xi \in \mathbb{R}.$$

Proof. We use again the transformation

$$(x_1(\xi), y_1(\xi), x_2(\xi), y_2(\xi))^\top = \tilde{R}_{(\omega+\nu^\mu)\xi}(x_1'(\xi), y_1'(\xi), x_2'(\xi), y_2'(\xi))^\top,$$

so that equation (2.63) is transformed into the system

$$z_\xi = \lambda^\mu T_1(z),$$

where

$$T_1 \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ x_1 \\ y_2 \\ x_2 \end{pmatrix},$$

the fundamental solution set of which is given by

$$s_3^\mu(\xi) = \begin{pmatrix} -\mathrm{e}^{-\lambda^\mu \xi} \\ \mathrm{e}^{-\lambda^\mu \xi} \\ 0 \\ 0 \end{pmatrix}, \quad u_3^\mu(\xi) = \begin{pmatrix} \mathrm{e}^{\lambda^\mu \xi} \\ \mathrm{e}^{\lambda^\mu \xi} \\ 0 \\ 0 \end{pmatrix},$$

$$s_4^\mu(\xi) = \begin{pmatrix} 0 \\ 0 \\ -\mathrm{e}^{-\lambda^\mu \xi} \\ \mathrm{e}^{-\lambda^\mu \xi} \end{pmatrix}, \quad u_4^\mu(\xi) = \begin{pmatrix} 0 \\ 0 \\ \mathrm{e}^{\lambda^\mu \xi} \\ \mathrm{e}^{\lambda^\mu \xi} \end{pmatrix}$$

with

$$|s_3^\mu(\xi)|, |s_4^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \quad |u_3^\mu(\xi)|, |u_4^\mu(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \qquad \xi \in \mathbb{R}.$$

The dual basis is given by

$$(s_3^\mu)^*(\xi) = \frac{1}{2}\begin{pmatrix} -\mathrm{e}^{\lambda^\mu \xi} \\ \mathrm{e}^{\lambda^\mu \xi} \\ 0 \\ 0 \end{pmatrix}, \quad (u_3^\mu)^*(\xi) = \frac{1}{2}\begin{pmatrix} \mathrm{e}^{-\lambda^\mu \xi} \\ \mathrm{e}^{-\lambda^\mu \xi} \\ 0 \\ 0 \end{pmatrix},$$

$$(s_4^\mu)^*(\xi) = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ -\mathrm{e}^{\lambda^\mu \xi} \\ \mathrm{e}^{\lambda^\mu \xi} \end{pmatrix}, \quad (u_4^\mu)^*(\xi) = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ \mathrm{e}^{-\lambda^\mu \xi} \\ \mathrm{e}^{-\lambda^\mu \xi} \end{pmatrix}$$

and satisfies

$$|(s_3^\mu)^*(\xi)|, |(s_4^\mu)^*(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \quad |(u_3^\mu)^*(\xi)|, |(u_4^\mu)^*(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \qquad \xi \in \mathbb{R}.$$

The functions

$$S_i^\mu(\xi) = \tilde{R}_{(\omega+\nu^\mu)\xi} s_i^\mu(\xi), \quad U_i^\mu(\xi) = \tilde{R}_{(\omega+\nu^\mu)\xi} u_i^\mu(\xi), \qquad \xi \in \mathbb{R},$$

are thus solutions to equation (2.63) with the desired properties. ∎

2.4 Assorted proofs

Here we present the proofs of the technical results in Lemmata 2.4, 2.5 and 2.19.

Lemma 2.4. *Suppose that $T_m^\varepsilon \in \mathcal{L}(\mathbb{C}^d)$ depends analytically upon ε in a neighbourhood of the origin uniformly over $m \in \mathbb{Z}$. The formula*

$$(T^\varepsilon u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} (T_m^\varepsilon u_m) \mathrm{e}^{\mathrm{i}m\tau}$$

defines an operator $T^\varepsilon \in \mathcal{L}(\mathcal{X}^t)$ for all $t \geq 0$ which depends analytically upon ε in the same neighbourhood of the origin.

Proof. Due to the analyticity of T_m^ε we know that

$$T_m^\varepsilon = \sum_{k=0}^\infty M_k^m(\{\varepsilon\}^{(k)})$$

for $M_k^m \in \mathcal{L}_{\mathrm{sym}}^{(k)}(\mathbb{R}; \mathcal{L}(\mathbb{C}^d))$ with $M_0^m = T_m^0$ and there exist constants $C, r > 0$ which are independent of m such that

$$|M_k^m(\{\varepsilon\}^{(k)})v| \leq Cr^{-k}|\varepsilon|^k|v|$$

for each $v \in \mathbb{C}^d$.

We conclude that the symmetric, k-linear operator $\tilde{M}_k$ on $\mathbb{R}$ defined by the formula

$$\Big(\tilde{M}_k(\{\varepsilon\}^{(k)})u\Big)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \Big(M_k^m(\{\varepsilon\}^{(k)})u_m\Big)\mathrm{e}^{\mathrm{i}m\tau}$$

satisfies

$$\|\tilde{M}_k(\{\varepsilon\}^{(k)})u\|_t^2 = \sum_{m\in\mathbb{Z}} (1+m^2)^t |M_k^m(\{\varepsilon\}^{(k)})u_m|^2 \leq (Cr^{-k}|\varepsilon|^k)^2 \|u\|_t^2$$

and therefore is an element of $\mathcal{L}_{\mathrm{sym}}^{(k)}(\mathbb{R}; \mathcal{L}(\mathcal{X}^t))$. Moreover, for sufficiently small values of $|\varepsilon|$, we have that

$$\begin{aligned}
\sum_{m\in\mathbb{Z}} \sum_{k=0}^\infty \Big|\Big(M_k^m(\{\varepsilon\}^{(k)})u_m\Big)\mathrm{e}^{\mathrm{i}m\tau}\Big| &\leq \sum_{m\in\mathbb{Z}} \sum_{k=0}^\infty C\Big(\frac{|\varepsilon|}{r}\Big)^k |u_m| \\
&\lesssim \sum_{m\in\mathbb{Z}} |u_m| \\
&= \sum_{m\in\mathbb{Z}} (1+m^2)^{t/2}|u_m|(1+m^2)^{-t/2} \\
&\leq \Big(\sum_{m\in\mathbb{Z}} (1+m^2)^t|u_m|^2\Big)^{1/2} \Big(\sum_{m\in\mathbb{Z}} (1+m^2)^{-t}\Big)^{1/2} \\
&\lesssim \|u\|_t
\end{aligned}$$

for each $u \in \mathcal{X}^t$ and $\tau \in [0, 2\pi]$ so that the series on the left-hand side converges absolutely in $\mathbb{C}^d$, that is

$$\begin{aligned}
(T^\varepsilon u)(\tau) &= \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} (T_m^\varepsilon u_m)\mathrm{e}^{\mathrm{i}m\tau} \\
&= \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \sum_{k=0}^\infty \Big(M_k^m(\{\varepsilon\}^{(k)})u_m\Big)\mathrm{e}^{\mathrm{i}m\tau} \\
&= \frac{1}{\sqrt{2\pi}} \sum_{k=0}^\infty \sum_{m\in\mathbb{Z}} \Big(M_k^m(\{\varepsilon\}^{(k)})u_m\Big)\mathrm{e}^{\mathrm{i}m\tau} \\
&= \sum_{k=0}^\infty \Big(\tilde{M}_k(\{\varepsilon\}^{(k)})u\Big)(\tau).
\end{aligned}$$

The operator $T^\varepsilon \in \mathcal{L}(\mathcal{X}^t)$ thus depends analytically upon ε. ∎

Lemma 2.5. *Let X be a Banach space, $J \in \mathcal{L}(X)$ with the property that $J^2 = I$, and $P_1^\varepsilon, \dots, P_l^\varepsilon \in \mathcal{L}(X)$ be projections which depend analytically upon ε in a neighbourhood of the origin and satisfy*

$$P_i^\varepsilon J = J P_i^\varepsilon, \quad P_i^\varepsilon P_j^\varepsilon = \delta_{ij} P_i^\varepsilon, \qquad i, j = 1, \dots, l,$$

and

$$\sum_{i=1}^{l} P_i^\varepsilon = I.$$

There exists $U^\varepsilon \in \mathcal{L}(X)$ which is invertible such that

$$U^\varepsilon P_i^0 (U^\varepsilon)^{-1} = P_i^\varepsilon, \quad i = 1, \dots, l.$$

Furthermore, U^ε, $(U^\varepsilon)^{-1}$ depend analytically upon ε and $U^\varepsilon J = J U^\varepsilon$.

Proof. Define

$$Q^\varepsilon = -\sum_{i=1}^{l} P_i^\varepsilon (P_i^\varepsilon)',$$

where $(P^\varepsilon)'$ denotes differentiation of P^ε with respect to ε, and observe that the linear differential equations

$$\begin{aligned} (U^\varepsilon)' &= Q^\varepsilon U^\varepsilon, \\ U^0 &= I, \end{aligned}$$

and

$$\begin{aligned} (V^\varepsilon)' &= -V^\varepsilon Q^\varepsilon, \\ V^0 &= I \end{aligned}$$

have unique solutions $U^\varepsilon, V^\varepsilon \in \mathcal{L}(X)$ which depend analytically upon ε. Observe that the solution U^ε is constructed as the uniform limit of the sequence $\{U_n^\varepsilon\}_{n\in\mathbb{N}}$ in $\mathcal{L}(X)$ defined by

$$U_0^\varepsilon = I, \quad U_n^\varepsilon = I + \int_0^\varepsilon Q^s U_{n-1}^s \, \mathrm{d}s$$

(see Kato [15, p. 100]). By assumption the projections $P_1^\varepsilon, \dots, P_l^\varepsilon$ and hence the operator Q^ε commute with the mapping J. It follows by induction that $U_n^\varepsilon J = J U_n^\varepsilon$ and thus $U^\varepsilon J = J U^\varepsilon$.

The next step is to show that $V^\varepsilon = (U^\varepsilon)^{-1}$ (so that in particular U^ε is invertible and $(U^\varepsilon)^{-1}$ depends analytically upon ε). On the one hand we have that

$$(V^\varepsilon U^\varepsilon)' = (V^\varepsilon)' U^\varepsilon + V^\varepsilon (U^\varepsilon)' = -V^\varepsilon Q^\varepsilon U^\varepsilon + V^\varepsilon Q^\varepsilon U^\varepsilon = 0,$$

that is $V^\varepsilon U^\varepsilon$ is constant and hence $V^\varepsilon U^\varepsilon = V^0 U^0 = I$. On the other hand note that the linear differential equation

$$\begin{aligned} (W^\varepsilon)' &= [Q^\varepsilon, W^\varepsilon], \\ W^0 &= I, \end{aligned}$$

has a unique solution W^ε which depends analytically upon ε. However $W^\varepsilon = I$ and $W^\varepsilon = U^\varepsilon V^\varepsilon$ solve this equation (since $(U^\varepsilon V^\varepsilon)' = Q^\varepsilon U^\varepsilon V^\varepsilon - U^\varepsilon V^\varepsilon Q^\varepsilon = [Q^\varepsilon, U^\varepsilon V^\varepsilon]$), so that $U^\varepsilon V^\varepsilon = I$.

Finally we prove that $U^\varepsilon P_j^0 (U^\varepsilon)^{-1} = P_j^\varepsilon$ for each $j = 1, \dots, l$. To this end consider the linear differential equation

$$\begin{aligned}(Z^\varepsilon)' &= Q^\varepsilon Z^\varepsilon,\\ Z^0 &= P_j^0\end{aligned}$$

and note that $U^\varepsilon P_j^0$ solves this equation (since $(U^\varepsilon P_j^0)' = (U^\varepsilon)' P_j^0 = Q^\varepsilon U^\varepsilon P_j^0$ and $U^0 = I$). We show that $P_j^\varepsilon U^\varepsilon$ is also a solution to this equation. Differentiating the identity

$$P_i^\varepsilon P_j^\varepsilon = \delta_{ij} P_i^\varepsilon$$

yields

$$(P_i^\varepsilon)' P_j^\varepsilon + P_i^\varepsilon (P_j^\varepsilon)' = \delta_{ij} (P_i^\varepsilon)'$$

from which it follows by left multiplication with P_i^ε that

$$P_i^\varepsilon (P_i^\varepsilon)' P_j^\varepsilon + P_i^\varepsilon (P_j^\varepsilon)' = \delta_{ij} P_i^\varepsilon (P_i^\varepsilon)'.$$

Since $\delta_{ij} P_i^\varepsilon = P_j^\varepsilon P_i^\varepsilon$ we obtain

$$P_i^\varepsilon (P_i^\varepsilon)' P_j^\varepsilon + P_i^\varepsilon (P_j^\varepsilon)' = P_j^\varepsilon P_i^\varepsilon (P_i^\varepsilon)'$$

that is

$$[-P_i^\varepsilon (P_i^\varepsilon)', P_j^\varepsilon] = P_i^\varepsilon (P_j^\varepsilon)'$$

and summation over $i = 1, \dots, l$ yields

$$[Q^\varepsilon, P_j^\varepsilon] = (P_j^\varepsilon)'.$$

Hence

$$\begin{aligned}(P_j^\varepsilon U^\varepsilon)' &= (P_j^\varepsilon)' U^\varepsilon + P_j^\varepsilon (U^\varepsilon)'\\ &= Q^\varepsilon P_j^\varepsilon U^\varepsilon - P_j^\varepsilon Q^\varepsilon U^\varepsilon + P_j^\varepsilon Q^\varepsilon U^\varepsilon\\ &= Q^\varepsilon P_j^\varepsilon U^\varepsilon\end{aligned}$$

with $P_j^0 U^0 = P_j^0$, so that it follows by uniqueness of the solution that $P_j^\varepsilon U^\varepsilon = U^\varepsilon P_j^0$ for each $j = 1, \dots, l$. ■

Lemma 2.19. *The equation*

$$r_\xi = L_{\mathrm{wh}}^\mu r + \mathcal{L}^\mu r \tag{2.61}$$

has linearly independent solutions S_1^μ, S_2^μ, U_1^μ, U_2^μ on $[0, \infty)$ with

$$|S_1^\mu(\xi)|, |S_2^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \quad |U_1^\mu(\xi)|, |U_2^\mu(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \qquad \xi \in [0, \infty),$$

The dual basis $\{(S_1^\mu)^*(\xi), (S_2^\mu)^*(\xi), (U_1^\mu)^*(\xi), (U_2^\mu)^*(\xi)\}$ *to* $\{S_1^\mu(\xi), S_2^\mu(\xi), U_1^\mu(\xi), U_2^\mu(\xi)\}$ *in* $\mathbb{R}^4$ *satisfies*

$$|(S_1^\mu)^*(\xi)|, |(S_2^\mu)^*(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \quad |(U_1^\mu)^*(\xi)|, |(U_2^\mu)^*(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \qquad \xi \in [0, \infty).$$

The functions U_i^μ *are reversible and the functions* S_i^μ *are anti-reversible, that is they satisfy*

$$R_{\mathrm{wh}} U_i^\mu(0) = U_i^\mu(0), \quad R_{\mathrm{wh}} S_i^\mu(0) = -S_i^\mu(0), \qquad i = 1, 2.$$

Proof. We use the transformation

$$(x_1(\xi), y_1(\xi), x_2(\xi), y_2(\xi))^\top = \tilde{R}_{(\omega+\nu^\mu)\xi}(x_1'(\xi), y_1'(\xi), x_2'(\xi), y_2'(\xi))^\top$$

(see Section 2.2), so that the matrix of L_{wh}^0 in equation (2.61), which is given by

$$\begin{pmatrix} 0 & 0 & -\omega & 0 \\ 0 & 0 & 0 & -\omega \\ \omega & 0 & 0 & 0 \\ 0 & \omega & 0 & 0 \end{pmatrix},$$

becomes the zero matrix and $\tilde{f}_1^\mu$ is (with a slight abuse of notation) transformed into

$$\tilde{f}_1^\mu(z) = \begin{pmatrix} 0 \\ -\check{C}x_1(x_1^2 + x_2^2) \\ 0 \\ -\check{C}x_2(x_1^2 + x_2^2) \end{pmatrix} + \mathcal{O}(|z|^3 |(z, \mu)|),$$

and we have also used the fact that the preliminary transformations in Lemmata 2.14 and 2.15 as well as the normal-form transformation in Chapter 3 below only modify the higher order terms in the Maclaurin expansion of the nonlinearity $\tilde{f}^\varepsilon(z, 0)|_{\varepsilon=\mu^2}$ in equation (2.26).

Moreover after the above transformation (with a further slight abuse of notation) p^μ is given by

$$p^\mu(\xi) = \sqrt{\lambda^\mu} \begin{pmatrix} H(\lambda^\mu \xi) \\ \dot{H}(\lambda^\mu \xi) \\ 0 \\ 0 \end{pmatrix} + \mathcal{R}^\mu(\xi)$$

with

$$H(\check{\xi}) = \sqrt{\frac{2}{\check{C}}}\,\mathrm{sech}(\xi), \quad |\mathcal{R}^\mu(\xi)| \lesssim \mu^2 \mathrm{e}^{-\lambda^\mu \eta \xi},$$

so that that the transformed system reads

$$r_\xi = \lambda^\mu T_1(r) + T_2^\mu(r) + T_3^\mu(\xi) r \tag{2.64}$$

with

$$T_1 \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ x_1 \\ y_2 \\ x_2 \end{pmatrix}, \quad T_2^\mu \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -3\check{C}H^2(\lambda^\mu \xi)\lambda^\mu x_1 \\ 0 \\ -\check{C}H^2(\lambda^\mu \xi)\lambda^\mu x_2 \end{pmatrix}$$

and

$$|T_3^\mu(\xi)| \lesssim \mu^3 \mathrm{e}^{-\lambda^\mu \eta \xi}.$$

We first consider the system

$$r_\xi = \lambda^\mu T_1(r) + T_2^\mu(r),$$

which has the explicit solutions

$$s_1^\mu(\xi) = \begin{pmatrix} v_1(\lambda^\mu \xi) \\ v_1'(\lambda^\mu \xi) \\ 0 \\ 0 \end{pmatrix}, \quad u_1^\mu(\xi) = \begin{pmatrix} v_2(\lambda^\mu \xi) \\ v_2'(\lambda^\mu \xi) \\ 0 \\ 0 \end{pmatrix},$$

$$s_2^\mu(\xi) = \begin{pmatrix} 0 \\ 0 \\ w_1(\lambda^\mu \xi) \\ w_1'(\lambda^\mu \xi) \end{pmatrix}, \quad u_2^\mu(\xi) = \begin{pmatrix} 0 \\ 0 \\ w_2(\lambda^\mu \xi) \\ w_2'(\lambda^\mu \xi) \end{pmatrix},$$

where

$$\begin{aligned} v_1(\xi) &= \operatorname{sech}(\xi)\tanh(\xi), \\ v_2(\xi) &= \operatorname{sech}(\xi)(-3 + \cosh^2(\xi) + 3\xi\tanh(\xi)), \\ w_1(\xi) &= \operatorname{sech}(\xi), \\ w_2(\xi) &= \operatorname{sech}(\xi)(2\xi + \sinh(2\xi)), \end{aligned}$$

which satisfy

$$|s_1^\mu(\xi)|, |s_2^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \quad |u_1^\mu(\xi)|, |u_2^\mu(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \qquad \xi \in [0, \infty).$$

The dual basis $\{(s_1^\mu)^*(\xi), (s_2^\mu)^*(\xi), (u_1^\mu)^*(\xi), (u_2^\mu)^*(\xi)\}$ to $\{s_1^\mu(\xi), s_2^\mu(\xi), u_1^\mu(\xi), u_2^\mu(\xi)\}$ in $\mathbb{R}^4$ is given by

$$(s_1^\mu)^*(\xi) = \frac{1}{2}\begin{pmatrix} v_2'(\lambda^\mu \xi) \\ -v_2(\lambda^\mu \xi) \\ 0 \\ 0 \end{pmatrix}, \quad (u_1^\mu)^*(\xi) = \frac{1}{2}\begin{pmatrix} -v_1'(\lambda^\mu \xi) \\ v_1(\lambda^\mu \xi) \\ 0 \\ 0 \end{pmatrix},$$

$$(s_2^\mu)^*(\xi) = \frac{1}{4}\begin{pmatrix} 0 \\ 0 \\ w_2'(\lambda^\mu \xi) \\ -w_2(\lambda^\mu \xi) \end{pmatrix}, \quad (u_2^\mu)^*(\xi) = \frac{1}{4}\begin{pmatrix} 0 \\ 0 \\ -w_1'(\lambda^\mu \xi) \\ w_1(\lambda^\mu \xi) \end{pmatrix},$$

and satisfies

$$|(s_1^\mu)^*(\xi)|, |(s_2^\mu)^*(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}, \quad |(u_1^\mu)^*(\xi)|, |(u_2^\mu)^*(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}, \qquad \xi \in [0, \infty).$$

Next we turn to (2.64). Consider the integral equations

$$\begin{aligned} \tilde{s}_i^\mu(\xi) = s_i^\mu(\xi) &- \sum_{l=1}^{2} \int_\xi^\infty \langle T_3^\mu(t)\tilde{s}_i^\mu(t), (s_l^\mu)^*(t)\rangle \, \mathrm{d}t \; s_l^\mu(\xi) \\ &- \sum_{l=1}^{2} \int_\xi^\infty \langle T_3^\mu(t)\tilde{s}_i^\mu(t), (u_l^\mu)^*(t)\rangle \, \mathrm{d}t \; u_l^\mu(\xi), \quad i = 1, 2, \end{aligned} \tag{2.65}$$

$$\begin{aligned} \tilde{u}_i^\mu(\xi) = u_i^\mu(\xi) &- \sum_{l=1}^{2} \int_0^\xi \langle T_3^\mu(t)\tilde{u}_i^\mu(t), (s_l^\mu)^*(t)\rangle \, \mathrm{d}t \; s_l^\mu(\xi) \\ &- \sum_{l=1}^{2} \int_\xi^\infty \langle T_3^\mu(t)\tilde{u}_i^\mu(t), (u_l^\mu)^*(t)\rangle \, \mathrm{d}t \; u_l^\mu(\xi), \quad i = 1, 2, \end{aligned} \tag{2.66}$$

and note that any solution of equation (2.65) or (2.66) yields a solution of equation (2.64). Our task is therefore to find solutions to equations (2.65), (2.66) which satisfy the desired estimates.

To this end we denote the right-hand side of equation (2.65) by $\mathcal{F}_i(\tilde{s}_i^\mu)$ and consider

$$\|\tilde{s}_i^\mu\|_{0,-\lambda^\mu} = \sup_{\xi \geq 0} |\tilde{s}_i^\mu(\xi)| \mathrm{e}^{\lambda^\mu \xi}.$$

Using the estimates for s_i^μ and u_i^μ, we find that

$$\begin{aligned} |\mathcal{F}_i(\tilde{s}_i^\mu)(\xi) - s_i^\mu(\xi)| &\lesssim \sum_{l=1}^{2} \int_\xi^\infty |T_3^\mu(t)| \|\tilde{s}_i^\mu\|_{0,-\lambda^\mu} \mathrm{e}^{-\lambda^\mu t} \mathrm{e}^{\lambda^\mu t} \, \mathrm{d}t \; \mathrm{e}^{-\lambda^\mu \xi} \\ &\qquad + \sum_{l=1}^{2} \int_\xi^\infty |T_3^\mu(t)| \|\tilde{s}_i^\mu\|_{0,-\lambda^\mu} \mathrm{e}^{-\lambda^\mu t} \mathrm{e}^{-\lambda^\mu t} \, \mathrm{d}t \; \mathrm{e}^{\lambda^\mu \xi} \\ &\lesssim \int_\xi^\infty |T_3^\mu(t)| \, \mathrm{d}t \; \|\tilde{s}_i^\mu\|_{0,-\lambda^\mu} \mathrm{e}^{-\lambda^\mu \xi}, \end{aligned}$$

so that

$$\begin{aligned} \|\mathcal{F}_i(\tilde{s}_i^\mu) - s_i^\mu\|_{0,\lambda^\mu} &\lesssim \|\tilde{s}_i^\mu\|_{0,\lambda^\mu} \int_\xi^\infty |T_3^\mu(t)| \, \mathrm{d}t \\ &\lesssim \|\tilde{s}_i^\mu\|_{0,\lambda^\mu} \int_0^\infty \mu^3 \mathrm{e}^{-\lambda^\mu \eta t} \, \mathrm{d}t \\ &\lesssim \mu \|\tilde{s}_i^\mu\|_{0,\lambda^\mu}. \end{aligned}$$

It follows that $\|\mathcal{F}_i(\tilde{s}_i^\mu)\|_{0,-\lambda^\mu} < \infty$ for $\|\tilde{s}_i^\mu\|_{0,-\lambda^\mu} < \infty$ and

$$\|\mathcal{F}_i(\tilde{s}_{i(1)}^\mu) - \mathcal{F}_i(\tilde{s}_{i(2)}^\mu)\|_{0,-\lambda^\mu} = \|\mathcal{F}_j(\tilde{s}_{i(1)}^\mu - \tilde{s}_{i(2)}^\mu) - s_i^\mu\|_{0,-\lambda^\mu} \lesssim \mu \|\tilde{s}_{i(1)}^\mu - \tilde{s}_{i(2)}^\mu\|_{0,-\lambda^\mu}.$$

We conclude that for small enough values of $\mu \geq 0$, the mapping $\mathcal{F}_i$ is a contraction on

$$\overline{B}_1(s_i^\mu) \subseteq C_{-\lambda^\mu}([0,\infty);\mathbb{R}^4) = \{f \in C([0,\infty);\mathbb{R}^4) \colon \|f\|_{0,-\lambda^\mu} < \infty\}$$

and thus has a fixed point $\tilde{s}_i^\mu$ satisfying

$$|\tilde{s}_i(\xi) - s_i(\xi)|\mathrm{e}^{\lambda^\mu \xi} \leq 1, \quad \xi \in [0,\infty),$$

so that in particular $|\tilde{s}_i^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \xi}$. We also observe that

$$|\tilde{s}_i^\mu(\xi) - s_i^\mu(\xi)|\mathrm{e}^{\lambda^\mu \xi} \leq \int_\xi^\infty |T_3^\mu(t)| \,\mathrm{d}t\, \|\tilde{s}_i^\mu\|_{0,-\lambda^\mu} \to 0, \quad \xi \to \infty,$$

and conclude that

$$\lim_{\xi\to\infty} \tilde{s}_i^\mu(\xi)e^{\lambda^\mu\xi} = \lim_{\xi\to\infty} s_i^\mu(\xi)e^{\lambda^\mu\xi} = \begin{cases}(2,-2,0,0)^\top, & i=1,\\ (0,0,2,-2)^\top, & i=2.\end{cases} \tag{2.67}$$

In a similar fashion we construct a solution to equation (2.66) by denoting the right-hand side by $\mathcal{G}_i(\tilde{u}_i^\mu)$ and using the norm

$$\|\tilde{u}_i^\mu\|_{0,\lambda^\mu} = \sup_{\xi\geq 0} |\tilde{u}_i^\mu(\xi)|\mathrm{e}^{-\lambda^\mu\xi}.$$

Using the estimates for s_i^μ and u_i^μ, we find that

$$\begin{aligned}|\mathcal{G}_i(\tilde{u}_i^\mu)(\xi) - u_i^\mu(\xi)| &\lesssim \sum_{l=1}^2 \int_0^\xi |T_3^\mu(t)| \|\tilde{u}_i^\mu\|_{0,\lambda^\mu} \mathrm{e}^{\lambda^\mu t}\mathrm{e}^{\lambda^\mu t}\,\mathrm{d}t\, \mathrm{e}^{-\lambda^\mu\xi}\\ &\qquad + \sum_{l=1}^2 \int_\xi^\infty |T_3^\mu(t)| \|\tilde{u}_i^\mu\|_{0,\lambda^\mu} \mathrm{e}^{\lambda^\mu t}\mathrm{e}^{-\lambda^\mu t}\,\mathrm{d}t\, \mathrm{e}^{\lambda^\mu\xi}\\ &\lesssim \int_0^\infty |T_3^\mu(t)|\,\mathrm{d}t\, \|\tilde{u}_i^\mu\|_{0,\lambda^\mu}\mathrm{e}^{\lambda^\mu\xi},\end{aligned}$$

so that

$$\begin{aligned}\|\mathcal{G}_i(\tilde{u}_i^\mu) - u_i^\mu\|_{0,\lambda^\mu} &\lesssim \|\tilde{u}_i^\mu\|_{0,\lambda^\mu} \int_0^\infty |T_3^\mu(t)|\,\mathrm{d}t\\ &\lesssim \|\tilde{u}_i^\mu\|_{0,\lambda^\mu} \int_0^\infty \mu^3\mathrm{e}^{-\lambda^\mu\eta t}\,\mathrm{d}t\\ &\lesssim \mu\|\tilde{u}_i^\mu\|_{0,\lambda^\mu}.\end{aligned}$$

It follows that $\|\mathcal{G}_i(\tilde{u}_i^\mu)\|_{0,\lambda^\mu} < \infty$ for $\|\tilde{u}_i^\mu\|_{0,\lambda^\mu} < \infty$ and

$$\|\mathcal{G}_i(\tilde{u}_{i(1)}^\mu) - \mathcal{G}_i(\tilde{u}_{i(2)}^\mu)\|_{0,\lambda^\mu} = \|\mathcal{G}_i(\tilde{u}_{i(1)}^\mu - \tilde{u}_{i(2)}^\mu) - u_i^\mu\|_{0,\lambda^\mu} \lesssim \mu\|\tilde{u}_{i(1)}^\mu - \tilde{u}_{i(2)}^\mu\|_{0,\lambda^\mu}.$$

We conclude that for small enough values of $\mu \geq 0$, the mapping $\mathcal{G}_i$ is a contraction on

$$\overline{B}_1(u_i^\mu) \subseteq C_{\lambda^\mu}([0,\infty);\mathbb{R}^4) = \{f \in C([0,\infty);\mathbb{R}^4) \colon \|f\|_{0,\lambda^\mu} < \infty\}$$

and thus has a fixed point $\tilde{u}_i^\mu$ satisfying

$$|\tilde{u}_i^\mu(\xi) - u_i^\mu(\xi)|\mathrm{e}^{-\lambda^\mu \xi} \leq 1, \quad \xi \in [0,\infty),$$

so that in particular $|\tilde{u}_i^\mu(\xi)| \lesssim \mathrm{e}^{\lambda^\mu \xi}$. We also observe that

$$\begin{aligned}
|\tilde{u}_i^\mu(\xi) - u_i^\mu(\xi)|\mathrm{e}^{-\lambda^\mu \xi} &\lesssim \sum_{l=1}^{2} \int_0^\xi |T_3^\mu(t)| \|\tilde{u}_i^\mu\|_{0,\lambda^\mu} \mathrm{e}^{\lambda^\mu(t-2\xi)}\, \mathrm{d}t \\
&\qquad + \sum_{l=1}^{2} \int_\xi^\infty |T_3^\mu(t)| \|\tilde{u}_i^\mu\|_{0,\lambda^\mu} \mathrm{e}^{-\lambda^\mu t}\, \mathrm{d}t \\
&\lesssim \mathrm{e}^{-\lambda^\mu \xi} \int_0^\xi |T_3^\mu(t)|\, \mathrm{d}t + \int_\xi^\infty |T_3^\mu(t)|\, \mathrm{d}t \\
&\to 0, \quad \xi \to \infty,
\end{aligned}$$

and conclude that

$$\lim_{\xi\to\infty} \tilde{u}_i^\mu(\xi) e^{-\lambda^\mu \xi} = \lim_{\xi\to\infty} u_i^\mu(\xi) e^{-\lambda^\mu \xi} = \begin{cases} (\frac{1}{2}, \frac{1}{2}, 0, 0)^\top, & i = 1, \\ (0,0,1,1)^\top, & i = 2. \end{cases} \tag{2.68}$$

The next step is to compute the dual basis $\{(\tilde{s}_1^\mu)^*(\xi), (\tilde{s}_2^\mu)^*(\xi), (\tilde{u}_1^\mu)^*(\xi), (\tilde{u}_2^\mu)^*(\xi)\}$ in $\mathbb{R}^4$. To this end we define the matrices

$$M(\xi) = \left(\tilde{s}_1^\mu(\xi)\middle|\tilde{s}_2^\mu(\xi)\middle|\tilde{u}_1^\mu(\xi)\middle|\tilde{u}_2^\mu(\xi)\right), \quad M^*(\xi) = \left((\tilde{s}_1^\mu)^*(\xi)\middle|(\tilde{s}_2^\mu)^*(\xi)\middle|(\tilde{u}_1^\mu)^*(\xi)\middle|(\tilde{u}_2^\mu)^*(\xi)\right)$$

and note that $M^* = (M^\top)^{-1}$ by definition of the dual basis. In order to compute M^*, we first note that the Wronskian $W = \det M$ is constant since $W_\xi = (\operatorname{tr} \mathcal{L}^\mu)W$ and the diagonal entries of $\mathcal{L}^\mu$ are zero because $R_{\mathrm{wh}}\mathcal{L}^\mu = -\mathcal{L}^\mu R_{\mathrm{wh}}$. Using the multilinearity and continuity of the determinant together with equations (2.67), (2.68) we compute

$$\begin{aligned}
\det M &= \lim_{\xi\to\infty} \det M \\
&= \lim_{\xi\to\infty} \det\left(\tilde{s}_1^\mu(\xi)\mathrm{e}^{\lambda^\mu\xi}\middle|\tilde{s}_2^\mu(\xi)\mathrm{e}^{\lambda^\mu\xi}\middle|\tilde{u}_1^\mu(\xi)\mathrm{e}^{-\lambda^\mu\xi}\middle|\tilde{u}_2^\mu(\xi)\mathrm{e}^{-\lambda^\mu\xi}\right) \\
&= \lim_{\xi\to\infty} \det\left(s_1^\mu(\xi)\mathrm{e}^{\lambda^\mu\xi}\middle|s_2^\mu(\xi)\mathrm{e}^{\lambda^\mu\xi}\middle|u_1^\mu(\xi)\mathrm{e}^{-\lambda^\mu\xi}\middle|u_2^\mu(\xi)\mathrm{e}^{-\lambda^\mu\xi}\right) \\
&= \lim_{\xi\to\infty} \underbrace{\det\left(s_1^\mu(\xi)\middle|s_2^\mu(\xi)\middle|u_1^\mu(\xi)\middle|u_2^\mu(\xi)\right)}_{=8} \\
&= 8.
\end{aligned}$$

In terms of the adjunct matrix $M^\#$ the entries of M^* can be computed via

$$(M^*)^\top = \frac{1}{\det M} M^\#$$

with

$$M^\#_{ij} = (-1)^{i+j} \det(c_1^M|\cdots|c_{i-1}^M|e_j|c_{i+1}^M|\cdots|c_4^M),$$

in which c_j^M denotes the j-th column of M and $\{e_1, \dots, e_4\}$ is the usual basis for $\mathbb{R}^4$. We thus obtain the explicit formulae

$$(\tilde{s}_1^\mu)^*(\xi) = \frac{1}{8}\begin{pmatrix} \det(e_1|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \\ -\det(e_2|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \\ \det(e_3|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \\ -\det(e_4|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \end{pmatrix}, \quad (\tilde{u}_1^\mu)^*(\xi) = \frac{1}{8}\begin{pmatrix} \det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|e_1|\tilde{u}_2^\mu(\xi)) \\ -\det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|e_2|\tilde{u}_2^\mu(\xi)) \\ \det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|e_3|\tilde{u}_2^\mu(\xi)) \\ -\det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|e_4|\tilde{u}_2^\mu(\xi)) \end{pmatrix}$$

$$(\tilde{s}_2^\mu)^*(\xi) = \frac{1}{8}\begin{pmatrix} -\det(\tilde{s}_1^\mu(\xi)|e_1|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \\ \det(\tilde{s}_1^\mu(\xi)|e_2|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \\ -\det(\tilde{s}_1^\mu(\xi)|e_3|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \\ \det(\tilde{s}_1^\mu(\xi)|e_4|\tilde{u}_1^\mu(\xi)|\tilde{u}_2^\mu(\xi)) \end{pmatrix}, \quad (\tilde{u}_2^\mu)^*(\xi) = \frac{1}{8}\begin{pmatrix} -\det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|e_1) \\ \det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|e_2) \\ -\det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|e_3) \\ \det(\tilde{s}_1^\mu(\xi)|\tilde{s}_2^\mu(\xi)|\tilde{u}_1^\mu(\xi)|e_4) \end{pmatrix}$$

and again using the multilinearity of the determinant, we find that

$$\begin{aligned} |(\tilde{s}_1^\mu)^*(\xi)| &\lesssim |\tilde{s}_2^\mu(\xi)||\tilde{u}_1^\mu(\xi)||\tilde{u}_2^\mu(\xi)| \lesssim \mathrm{e}^{\lambda^\mu\xi}, \\ |(\tilde{s}_2^\mu)^*(\xi)| &\lesssim |\tilde{s}_1^\mu(\xi)||\tilde{u}_1^\mu(\xi)||\tilde{u}_2^\mu(\xi)| \lesssim \mathrm{e}^{\lambda^\mu\xi}, \\ |(\tilde{u}_1^\mu)^*(\xi)| &\lesssim |\tilde{s}_1^\mu(\xi)||\tilde{s}_2^\mu(\xi)||\tilde{u}_2^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu\xi}, \\ |(\tilde{u}_2^\mu)^*(\xi)| &\lesssim |\tilde{s}_1^\mu(\xi)||\tilde{s}_2^\mu(\xi)||\tilde{u}_1^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu\xi}, \end{aligned}$$

for $\xi \in [0, \infty)$.

It remains to demonstrate the reversibility of $\tilde{u}_i^\mu$ and the anti-reversibility of $\tilde{s}_i^\mu$. It follows from equation (2.66) that $\tilde{u}_1^\mu(0), \tilde{u}_2^\mu(0) \in \langle u_1^\mu(0), u_2^\mu(0)\rangle$, and since the functions u_1^μ and u_2^μ are reversible, we conclude that $\tilde{u}_1^\mu$ and $\tilde{u}_2^\mu$ are reversible. Moreover, recall that equation (2.26), written in the new variables as

$$z_\xi = \tilde{F}_1^\mu(z),$$

has the family

$$\{\tilde{R}_a p^\mu(\xi_0 + \cdot)\colon a \in [0, 2\pi),\ \xi_0 \in \mathbb{R}\}$$

of homoclinic solutions. Hence the functions

$$\tilde{\tilde{s}}_1 = \frac{\mathrm{d}}{\mathrm{d}\xi_0}\tilde{R}_a p^\mu(\xi_0 + \cdot)\Big|_{a,\xi_0=0} = (p^\mu)', \quad \tilde{\tilde{s}}_2 = \frac{\mathrm{d}}{\mathrm{d}a}\tilde{R}_a p^\mu(\xi_0 + \cdot)\Big|_{a,\xi_0=0} = \tilde{R}_{\pi/2} p^\mu$$

are linearly independent, anti-reversible, homoclinic solutions of equation (2.64), and the following argument shows that the set $\{\tilde{\tilde{s}}_1^\mu, \tilde{\tilde{s}}_2^\mu, \tilde{u}_1^\mu, \tilde{u}_2^\mu\}$ is a fundamental solution set to equation (2.64). Suppose that

$$\alpha_1\tilde{\tilde{s}}_1^\mu(\xi) + \alpha_2\tilde{\tilde{s}}_2^\mu(\xi) = -\beta_1\tilde{u}_1^\mu(\xi) - \beta_2\tilde{u}_2^\mu(\xi), \quad \xi \in [0, \infty),$$

for some scalars α_j and β_j. The function on the left-hand side is anti-reversible, while the function on the right-hand side is reversible, so that

$$\alpha_1\tilde{\tilde{s}}_1^\mu(\xi) + \alpha_2\tilde{\tilde{s}}_2^\mu(\xi) = 0, \quad \beta_1\tilde{u}_1^\mu(\xi) + \beta_2\tilde{u}_2^\mu(\xi) = 0, \qquad \xi \in [0, \infty),$$

and thus $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$, since $\tilde{\tilde{s}}_1^\mu, \tilde{\tilde{s}}_2^\mu$ as well as $\tilde{u}_1^\mu, \tilde{u}_2^\mu$ are linearly independent. By construction $\{\tilde{s}_1^\mu, \tilde{s}_2^\mu, \tilde{u}_1^\mu, \tilde{u}_2^\mu\}$ is also a fundamental solution set to equation (2.61), so that each of the bounded solutions $\tilde{s}_1^\mu$ and $\tilde{s}_2^\mu$ is a linear combination of $\tilde{\tilde{s}}_1^\mu$ and $\tilde{\tilde{s}}_2^\mu$. In particular $\tilde{s}_1^\mu$ and $\tilde{s}_2^\mu$ are anti-reversible (note that we cannot simply take $\tilde{s}_1^\mu = \tilde{\tilde{s}}_1^\mu$, $\tilde{s}_2^\mu = \tilde{\tilde{s}}_2^\mu$ since we *a priori* know only that $|\tilde{\tilde{s}}_1^\mu(\xi)|, |\tilde{\tilde{s}}_2^\mu(\xi)| \lesssim \mathrm{e}^{-\lambda^\mu \eta \xi}$ for $\xi \in [0, \infty)$).

The solutions to equation (2.61) with the desired properties are now obtained by returning to the original coordinates and defining

$$S_i^\mu(\xi) = \tilde{R}_{(\omega+\nu^\mu)\xi} \tilde{s}_i^\mu(\xi), \quad U_i^\mu(\xi) = \tilde{R}_{(\omega+\nu^\mu)\xi} \tilde{u}_i^\mu(\xi), \qquad \xi \in [0, \infty).$$

■

3 Normal-form theory

In this chapter we consider equations (2.50), (2.51) derived in Lemma 2.15, namely

$$z_\xi = L_{\mathrm{wh}}^\mu z + f_1^\mu(z) + f_2^\mu(z,q), \tag{3.1}$$

$$q_\xi = L_{\mathrm{sh,c}}^\mu q + P_{\mathrm{sh,c}}\Big(g_1^\mu(z,q)q_\tau\Big) + g_2^\mu(z,q) + h^\mu(z), \tag{3.2}$$

where the hats have been dropped for notational simplicity. Recall that

$$|f_1^\mu(z)| = \mathcal{O}(|z|^3), \tag{3.3}$$

$$|f_2^\mu(z,q)| = \mathcal{O}(|z||(z,\mu)|^2\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t} + \|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t}^2), \tag{3.4}$$

$$\|g_1^\mu(z,q)\|_t = \mathcal{O}(\mu\|(z,q)\|_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}_{\mathrm{sh,c}}^t}), \tag{3.5}$$

$$\|g_2^\mu(z,q)\|_{\mathcal{X}_{\mathrm{sh,c}}^t} = \mathcal{O}(\mu|z|\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t} + \mu^2\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t}^2 + \mu\|q\|_{\mathcal{X}_{\mathrm{sh,c}}^t}^3), \tag{3.6}$$

$$\|h^\mu(z)\|_{\mathcal{X}_{\mathrm{sh,c}}^t} = \mathcal{O}(\mu|z|^2|(z,\mu)|) \tag{3.7}$$

for $t > \frac{1}{2}$ and

$$P_{\mathrm{c}}\Big(g_1^{1;1,0}(z,0)Z_\tau\Big) = 0, \quad P_{\mathrm{c}}g_2^{1;1,1}(z,Z) = 0.$$

We use the following strategy devised by Groves & Schneider [10]: we perform a sequence of changes of variable which systematically remove terms from the Maclaurin series of $h^\mu(z)$ up to a specific order p while preserving the overall structure of equations (3.1), (3.2). We show that it is possible to make an optimal choice for p so that the remaining part of $h^\mu(z)$ is exponentially small in comparison with μ and hence prove Theorem 2.16. The analysis is based upon a theory for finite-dimensional dynamical systems given by Iooss and Lombardi [14]. We refer to several of their combinatorial results here.

3.1 Construction of the normal-form transformation

In this section we show that for arbitrary order $k \geq 2$ there exists a near-identity change of coordinates which transforms equations (3.1), (3.2) into equations of the same form with the additional property that the Maclaurin series of the transformed nonlinearity $\tilde{h}^\mu(z)$ does not contain any terms of order less than or equal k.

We begin by introducing notation for Banach-space valued polynomials of five complex variables and stating some facts which are useful in the following analysis.

Definition 3.1. Suppose that X is a complex Banach space and let $k \in \mathbb{N}_0$.

(i) We call $\Psi^k \colon \mathbb{C}^5 \to X$ a *homogeneous polynomial of degree k* if

$$\Psi^k(x) = \sum_{|\boldsymbol{\alpha}|=k} v_{\boldsymbol{\alpha}} x_1^{\alpha_1} \cdot \ldots \cdot x_5^{\alpha_5}$$

for some $v_{\boldsymbol{\alpha}} \in X$, $\boldsymbol{\alpha} \in \mathbb{N}_0^5$. We denote the vector space of all homogeneous polynomials of degree k from $\mathbb{C}^5$ to X by $\mathcal{P}^k(\mathbb{C}^5; X)$.

(ii) We call $\Psi \colon \mathbb{C}^5 \to X$ a *polynomial of degree k* if

$$\Psi = \sum_{j=0}^{k} \Psi^j,$$

where $\Psi^j \in \mathcal{P}^j(\mathbb{C}^5; X)$ and $\Psi^k \neq 0$.

Remark 3.2. Suppose that X is a complex Banach space and let $k \in \mathbb{N}_0$.

(i) Observe that $\Psi^k \in \mathcal{P}^k(\mathbb{C}^5; X)$ if and only if there exists $A^k \in \mathcal{L}^{(k)}_{\mathrm{sym}}(\mathbb{C}^5; X)$ satisfying $\Psi^k(x) = A^k(\{x\}^{(k)})$ for each $x \in \mathbb{C}^5$. According to Definition 3.1, if $\Psi^k \in \mathcal{P}^k(\mathbb{C}^5; X)$ is given by

$$\Psi^k(x) = \sum_{|\boldsymbol{\alpha}|=k} v_{\boldsymbol{\alpha}} x_1^{\alpha_1} \cdot \ldots \cdot x_5^{\alpha_5},$$

then $v_{\boldsymbol{\alpha}}$ and A^k are related by the formula

$$v_{\boldsymbol{\alpha}} = \binom{k}{\boldsymbol{\alpha}} A^k\Big(\{e_1\}^{(\alpha_1)}, \ldots, \{e_5\}^{(\alpha_5)}\Big),$$

where $\{e_1, \ldots, e_5\}$ denotes the usual basis for $\mathbb{C}^5$.

(ii) The expressions

$$|\Psi^k|_2 = \sqrt{\sum_{|\boldsymbol{\alpha}|=k} \|v_{\boldsymbol{\alpha}}\|_X^2 \alpha_1! \cdot \ldots \cdot \alpha_5!},$$

$$|\Psi^k|_{2,k} = \frac{1}{\sqrt{k!}} |\Phi|_2$$

and

$$|\Psi^k|_{0,k} = \sup_{x \in \mathbb{R}^5} \frac{\|\Phi^k(x)\|_X}{|x|^k}$$

define norms on $\mathcal{P}^k(\mathbb{C}^5; X)$.

(iii) Suppose that $k > 0$ and $\Psi^k \in \mathcal{P}^k(\mathbb{C}^5; X)$ is given by

$$\Psi^k(x) = \sum_{|\boldsymbol{\alpha}|=k} v_{\boldsymbol{\alpha}} x_1^{\alpha_1} \cdot \ldots \cdot x_5^{\alpha_5} = A^k(\{x\}^{(k)}).$$

It follows that

$$\begin{aligned} \mathrm{d}\Psi^k[x](\tilde{x}) &= kA^k(\{x\}^{(k-1)}, \tilde{x}) \\ &= \sum_{|\boldsymbol{\alpha}|=k-1} \Big(\sum_{i=1}^{5} (\alpha_i + 1) v_{\boldsymbol{\alpha}+\boldsymbol{e}_i} \tilde{x}_i\Big) x_1^{\alpha_1} \cdot \ldots \cdot x_5^{\alpha_5} \end{aligned}$$

for $\tilde{x} \in \mathbb{C}^5$, where $\boldsymbol{e}_i \in \mathbb{N}_0^5$ is the multi-index whose entries are 1 at position i and 0 elsewhere. In particular the derivative $\mathrm{d}\Psi^k$ is a polynomial in $\mathcal{P}^{k-1}(\mathbb{C}^5; \mathcal{L}(\mathbb{C}^5; X))$.

(iv) For our specific use we consider polynomials taking values in either $\mathcal{X}^t_{\mathrm{sh,c}}$ for $t > \frac{1}{2}$, or $\mathbb{C}^5$. To this end we abbreviate $\mathcal{P}^k_t = \mathcal{P}^k(\mathbb{C}^5; \mathcal{X}^t_{\mathrm{sh,c}})$ and $\mathcal{Q}^k = \mathcal{P}^k(\mathbb{C}^5; \mathbb{C}^5)$. The norms on $\mathcal{P}^k_t$ will be written as respectively $|\cdot|^t_2$, $|\cdot|^t_{2,k}$ and $|\cdot|^t_{0,k}$.

The relation between the different norms defined in Remark 3.2(ii) are given in the following two lemmata, the proofs of which are similar to those of Iooss and Lombardi [14, Lemmata 2.10 and 2.11].

Lemma 3.3. *Suppose that X is a complex Banach space and let $k \in \mathbb{N}_0$.*

(i) The estimates

$$|\Psi^k|_{0,k} \le |\Psi^k|_{2,k} \le \sqrt{5}k^2 |\Psi^k|_{0,k}$$

hold for each $\Psi^k \in \mathcal{P}^k(\mathbb{C}^5; X)$.

(ii) Suppose that $p \in \mathbb{N}_0$, $k \neq 0$ and $\Psi^k \in \mathcal{P}^k(\mathbb{C}^5; X)$, $N^p \in \mathcal{Q}^p$. The operator

$$x \mapsto \mathrm{d}\Psi^k[x](N^p(x))$$

lies in $\mathcal{P}^l(\mathbb{C}^5; X)$ with $l = k + p - 1$ and satisfies the estimate

$$|\mathrm{d}\Psi^k[\cdot](N^p)|_{2,l} \le \sqrt{k^2 + 4k}\, |\Psi^k|_{2,k} |N^p|_{2,p}.$$

Lemma 3.4. *Suppose that $q \in \mathbb{N}$, $i \in \{0, \ldots, q\}$, $p_1, \ldots, p_{q-i} \in \mathbb{N}$ and that $R_q \in \mathcal{L}^{(q)}(\mathbb{C}^{5i} \times (\mathcal{X}^t_{\mathrm{sh,c}})^{q-i}; \mathcal{X}^t_{\mathrm{sh,c}})$. For each choice of $\Psi^{p_j} \in \mathcal{P}^j_t$, $j \in \{1, \ldots, q-i\}$, the operator*

$$x \mapsto R_q[\{x\}^{(i)}, \Psi^{p_1}(x), \ldots, \Psi^{p_{q-i}}(x)]$$

lies in $\mathcal{P}^l_t$ with $l = i + p_1 + \ldots + p_{q-i}$ and satisfies the estimate

$$|R_q[\{\cdot\}^{(i)}, \Psi^{p_1}, \ldots, \Psi^{p_{q-i}}]|^t_{2,l} \le \|R_q\|_{\mathcal{L}^{(q)}} \sqrt{5}^i |\Psi^{p_1}|^t_{2,p_1} \cdot \ldots \cdot |\Psi^{p_{q-i}}|^t_{2,p_{q-i}}.$$

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

Another useful fact is the following proposition concerning the relation between the norm of a homogeneous polynomial and the norm of its derivative.

Proposition 3.5. *Suppose that X is a complex Banach space and let $k \in \mathbb{N}$. The estimate*

$$|\mathrm{d}\Psi^k|_{2,k-1} \leq \sqrt{5}k|\Psi^k|_{2,k},$$

holds for each $\Psi^k \in \mathcal{P}^k(\mathbb{C}^5; X)$, where $|\cdot|_{2,k-1}$ denotes the corresponding norm for $\mathcal{P}^{k-1}(\mathbb{C}^5; \mathcal{L}(\mathbb{C}^5; X))$.

Proof. Suppose that

$$\Psi^k(x) = \sum_{|\boldsymbol{\alpha}|=k} v_{\boldsymbol{\alpha}} x_1^{\alpha_1} \cdot \ldots \cdot x_5^{\alpha_5}, \quad v_{\boldsymbol{\alpha}} \in X.$$

According to Remark 3.2(iii), one can write

$$\mathrm{d}\Psi^k[x] = \sum_{|\boldsymbol{\alpha}|=k-1} w_{\boldsymbol{\alpha}} x_1^{\alpha_1} \cdot \ldots \cdot x_5^{\alpha_5}$$

with

$$w_{\boldsymbol{\alpha}}(\tilde{x}) = \sum_{i=1}^{5} (\alpha_i + 1) v_{\boldsymbol{\alpha}+\boldsymbol{e}_i} \tilde{x}_i,$$

so that

$$\begin{aligned}
|\mathrm{d}\Psi^k|_{2,k-1}^2 &= \frac{1}{(k-1)!} \sum_{|\boldsymbol{\alpha}|=k-1} \|w_{\boldsymbol{\alpha}}\|_{\mathcal{L}(\mathbb{C}^5;X)}^2 \alpha_1! \cdot \ldots \cdot \alpha_5! \\
&= \frac{k}{k!} \sum_{|\boldsymbol{\alpha}|=k-1} \sup_{|\tilde{x}|=1} \left\| \sum_{i=1}^{5} (\alpha_i+1) v_{\boldsymbol{\alpha}+\boldsymbol{e}_i} \tilde{x}_i \right\|_X^2 \alpha_1! \cdot \ldots \cdot \alpha_5! \\
&\leq \frac{k}{k!} \sum_{|\boldsymbol{\alpha}|=k-1} \left(\sum_{i=1}^{5} (\alpha_i+1) \|v_{\boldsymbol{\alpha}+\boldsymbol{e}_i}\|_X \right)^2 \alpha_1! \cdot \ldots \cdot \alpha_5! \\
&\leq \frac{5k}{k!} \sum_{|\boldsymbol{\alpha}|=k-1} \sum_{i=1}^{5} (\alpha_i+1)^2 \|v_{\boldsymbol{\alpha}+\boldsymbol{e}_i}\|_X^2 \alpha_1! \cdot \ldots \cdot \alpha_5! \\
&\leq \frac{5k^2}{k!} \sum_{|\boldsymbol{\alpha}|=k-1} \sum_{i=1}^{5} \|v_{\boldsymbol{\alpha}+\boldsymbol{e}_i}\|_X^2 \alpha_1! \cdot \ldots \cdot (\alpha_i+1)! \cdot \ldots \cdot \alpha_5! \\
&= \frac{5k^2}{k!} \sum_{|\boldsymbol{\alpha}|=k} \|v_{\boldsymbol{\alpha}}\|_X^2 \alpha_1! \cdot \ldots \cdot \alpha_5! \\
&= 5k^2 |\Psi^k|_{2,k}^2.
\end{aligned}$$

∎

By writing $y = (z, \mu)$ and attaching the additional equation $\partial_\xi \mu = 0$ to equation (3.1), we reformulate equations (3.1), (3.2) as

$$\begin{aligned} y_\xi &= \tilde{L}^\mu_{\mathrm{wh}} y + f_1(y) + f_2(y, q), \\ q_\xi &= L^\mu_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}\Big(g_1(y, q) q_\tau\Big) + g_2(y, q) + h(y), \end{aligned}$$

where

$$\begin{aligned} \tilde{L}^\mu_{\mathrm{wh}} &= \begin{pmatrix} L^\mu_{\mathrm{wh}} & 0 \\ 0 & 0 \end{pmatrix}, \\ f_1(y) &= \begin{pmatrix} f^\mu_1(z) \\ 0 \end{pmatrix}, \quad f_2(y, q) = \begin{pmatrix} f^\mu_2(z, q) \\ 0 \end{pmatrix}, \\ g_j(y, q) &= g^\mu_j(z, q), \quad j = 1, 2, \\ h(y) &= h^\mu(z). \end{aligned}$$

To achieve our objective of removing terms from the Maclaurin expansion of $h(y)$ up to some specific order p, we use a change of variable of the form

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

where $\Phi \colon \mathbb{C}^5 \to \mathcal{X}^{t+1}_{\mathrm{sh,c}}$ is a polynomial of degree p for $t > \frac{1}{2}$, that is

$$\Phi(y) = \sum_{k=2}^{p} \Phi^k(y), \quad \Phi^k \in \mathcal{P}^k_{t+1},$$

which commutes with the reverser R (we slightly abuse the notation writing (Rz, μ) as Ry). We arrive at the equations

$$y_\xi = \tilde{L}^\mu_{\mathrm{wh}} y + \tilde{f}_1(y) + \tilde{f}_2(y, \tilde{q}), \tag{3.8}$$

$$\tilde{q}_\xi = L^\mu_{\mathrm{sh,c}} \tilde{q} + P_{\mathrm{sh,c}}\Big(\tilde{g}_1(y, \tilde{q}) \tilde{q}_\tau\Big) + \tilde{g}_2(y, \tilde{q}) + \tilde{h}(y) \tag{3.9}$$

with

$$\begin{aligned} \tilde{f}_1(y) &= f_1(y) + f_2(y, -\Phi(y)), \\ \tilde{f}_2(y, \tilde{q}) &= f_2(y, \tilde{q} - \Phi(y)) - f_2(y, -\Phi(y)), \\ \tilde{g}_1(y, \tilde{q}) &= g_1(y, \tilde{q} - \Phi(y)), \\ \tilde{g}_2(y, \tilde{q}) &= g_2(y, \tilde{q} - \Phi(y)) - g_2(y, -\Phi(y)) \\ &\qquad - P_{\mathrm{sh,c}}\Big((g_1(y, \tilde{q} - \Phi(y)) - g_1(y, -\Phi(y)))\partial_\tau \Phi(y)\Big) \\ &\qquad + \mathrm{d}\Phi[y]\Big(f(y, \tilde{q} - \Phi(y)) - f(y, -\Phi(y))\Big), \\ \tilde{h}(y) &= -L^0_{\mathrm{sh,c}} \Phi(y) + \mathrm{d}\Phi[y](\tilde{L}^0_{\mathrm{wh}} y) + N(y) \end{aligned}$$

in which

$$f(y, q) = (\tilde{L}^\mu_{\mathrm{wh}} - \tilde{L}^0_{\mathrm{wh}}) y + f_1(y) + f_2(y, q)$$

and

$$N(y) = -\,(L^{\mu}_{\mathrm{sh,c}} - L^{0}_{\mathrm{sh,c}})\Phi(y) + \mathrm{d}\Phi[y]\Big(f(y, -\Phi(y))\Big) \\ + h(y) + g_2(y, -\Phi(y)) - P_{\mathrm{sh,c}}\Big(g_1(y, -\Phi(y))\partial_\tau \Phi(y)\Big).$$

To eliminate the k-th order term $\tilde{h}^k(y)$ in the Maclaurin expansion of $\tilde{h}$, we choose Φ^k in such a way that

$$L^0_{\mathrm{sh,c}}\Phi^k(y) - \mathrm{d}\Phi^k[y](\tilde{L}^0_{\mathrm{wh}}y) = N^k(y), \quad k = 2, \dots, p. \tag{3.10}$$

By (3.7), we know that $\tilde{h}^2 = \tilde{h}^3 = 0$ and hence $\Phi^2 = \Phi^3 = 0$. However we include these terms in the expansions for notational simplicity in the combinatorics below. Note that a solution Φ^k to equation (3.10) commutes with the reverser R.

In order to prove that equation (3.10) has a solution for every $k = 2, \dots, p$ we consider its left-hand side as an operator $\mathcal{L}_k \colon \mathcal{P}^k_{t+1} \to \mathcal{P}^k_t$ and show that $\mathcal{L}_k$ is invertible with an inverse $\mathcal{L}_k^{-1}$ which is bounded independently of k. In the construction of the estimate for $\|\mathcal{L}_k^{-1}\|_{\mathcal{L}(\mathcal{P}^k_t;\mathcal{P}^k_{t+1})}$ we change between the canonical basis $\{e_1, \dots, e_n\}$ and respectively the bases $\{e_{1,1}, \cdots, e_{1,n-2}, u, v\}$, $\{\overline{e}_{1,1}, \cdots, \overline{e}_{1,n-2}, \overline{u}, \overline{v}\}$, $\{e_{m,1}, \dots, e_{m,n}\}$ for $\mathbb{C}^n$ (see Section 2.1) and use the fact that the change of basis matrices Q_m^{-1}, where

$$Q_1 = (e_{1,1}|\cdots|e_{1,n-2}|u|v), \quad Q_{-1} = (\overline{e}_{1,1}|\cdots|\overline{e}_{1,n-2}|\overline{u}|\overline{v})$$

and

$$Q_m = (e_{m,1}|\cdots|e_{m,n}) \quad (m \in \mathbb{Z} \setminus \{\pm 1\}),$$

are bounded independently of $m \in \mathbb{Z}$ (see Proposition 2.8).

Proposition 3.6. *The operator $\mathcal{L}_k \colon \mathcal{P}^k_{t+1} \to \mathcal{P}^k_t$ defined by*

$$(\mathcal{L}_k\Phi^k)(y) = L^0_{\mathrm{sh,c}}\Phi^k(y) - \mathrm{d}\Phi^k[y](\tilde{L}^0_{\mathrm{wh}}y)$$

is invertible and the operator norm

$$\gamma := \|\mathcal{L}_k^{-1}\|_{\mathcal{L}(\mathcal{P}^k_t;\mathcal{P}^k_{t+1})} = \sup_{|\Psi^k|^t_2 = 1} |\mathcal{L}_k^{-1}\Psi^k|^{t+1}_2$$

is independent of k.

Proof. First note that polynomials $\Psi^k \in \mathcal{P}^k_t$ and $\Phi^k \in \mathcal{P}^k_{t+1}$ take values in the finite dimensional subspace

$$E_0 \oplus (E_{-1} \oplus E_1) \oplus \cdots \oplus (E_{-k} \oplus E_k)$$

of respectively $\mathcal{X}^t$ and $\mathcal{X}^{t+1}$. Writing

$$\Psi^k(y) = \sum_{|\boldsymbol{\alpha}|=k} \sum_{\substack{-k \le m \le k \\ 1 \le r \le r_m}} \Psi_{m,\boldsymbol{\alpha},r} z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} \mu^{\alpha_5} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau},$$

$$\Phi^k(y) = \sum_{|\boldsymbol{\alpha}|=k} \sum_{\substack{-k \le m \le k \\ 1 \le r \le r_m}} \Phi_{m,\boldsymbol{\alpha},r} z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} \mu^{\alpha_5} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau}$$

for some complex numbers $\Phi_{m,\boldsymbol{\alpha},r}$ and $\Psi_{m,\boldsymbol{\alpha},r}$, we therefore find that the equation

$$\mathcal{L}_k\Phi^k = \Psi^k$$

is equivalent to

$$(\lambda_{m,r} + mi\omega)\Phi_{m,\boldsymbol{\alpha},r} = \Psi_{m,\boldsymbol{\alpha},r},$$

and since Assumptions 2(iii) guarantee that $\lambda_{m,r} \neq mi\omega$ for each $m \in \mathbb{Z}$ we conclude that $\mathcal{L}_k$ is invertible with inverse

$$(\mathcal{L}_k^{-1}\Psi^k)(y) = \sum_{|\boldsymbol{\alpha}|=k} \sum_{\substack{-k\le m\le k\\ 1\le r\le r_m}} \frac{\Psi_{m,\boldsymbol{\alpha},r}}{\lambda_{m,r} + mi\omega} z_1^{\alpha_1} z_2^{\alpha_2} \overline{z}_1^{\alpha_3} \overline{z}_2^{\alpha_4} \mu^{\alpha_5} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau}.$$

Next we use the estimate for $\|Q_m^{-1}\|$ in Proposition 2.8 and the fact that the vectors $e_{m,r}$ are normalised to compute

$$\begin{aligned}
&(|\mathcal{L}_k^{-1}\Psi^k|_2^{t+1})^2\\
&= \sum_{|\boldsymbol{\alpha}|=k} \sum_{m=-k}^{k} (1+m^2)^{t+1} \left| \sum_{r=1}^{r_m} \frac{\Psi_{m,\boldsymbol{\alpha},r}}{\lambda_{m,r} + mi\omega} e_{m,r} \right|^2 \alpha_1! \cdot \ldots \cdot \alpha_5!\\
&\le \sum_{|\boldsymbol{\alpha}|=k} \sum_{m=-k}^{k} (1+m^2)^{t} \max_{r=1,\ldots,r_m} \frac{1+m^2}{|\lambda_{m,r} + mi\omega|^2} \sum_{r=1}^{r_m} |\Psi_{m,\boldsymbol{\alpha},r}|^2 \alpha_1! \cdot \ldots \cdot \alpha_5!\\
&= \sum_{|\boldsymbol{\alpha}|=k} \sum_{m=-k}^{k} (1+m^2)^{t} \max_{r=1,\ldots,r_m} \frac{1+m^2}{|\lambda_{m,r} + mi\omega|^2} \left| \sum_{r=1}^{r_m} \Psi_{m,\boldsymbol{\alpha},r} e_r \right|^2 \alpha_1! \cdot \ldots \cdot \alpha_5!\\
&= \sum_{|\boldsymbol{\alpha}|=k} \sum_{m=-k}^{k} (1+m^2)^{t} \max_{r=1,\ldots,r_m} \frac{1+m^2}{|\lambda_{m,r} + mi\omega|^2} \left| Q_m^{-1} \sum_{r=1}^{r_m} \Psi_{m,\boldsymbol{\alpha},r} e_{m,r} \right|^2 \alpha_1! \cdot \ldots \cdot \alpha_5!\\
&\lesssim \sum_{|\boldsymbol{\alpha}|=k} \sum_{m=-k}^{k} (1+m^2)^{t} \max_{r=1,\ldots,r_m} \frac{1+m^2}{|\lambda_{m,r} + mi\omega|^2} \left| \sum_{r=1}^{r_m} \Psi_{m,\boldsymbol{\alpha},r} e_{m,r} \right|^2 \alpha_1! \cdot \ldots \cdot \alpha_5!.
\end{aligned}$$

We use the same notation as in the proof of Proposition 2.3 and recall that the eigenvalues $\lambda_1^0(m), \ldots, \lambda_n^0(m)$ of the matrix $\mathrm{i}\tilde{A}^0(0) + \frac{1}{m}\partial\tilde{B}^0[0]$ are purely imaginary for sufficiently large $|m|$ with

$$\lim_{|m|\to\infty} \lambda_r^0(m) = \mathrm{i}\mu_r^0, \quad r = 1, \ldots, n,$$

where $\mu_1^0, \ldots, \mu_n^0$ are the distinct, real eigenvalues of the matrix $\tilde{A}^0(0)$. Since

$$\lambda_{m,r} = m\lambda_r^0(m),$$

we conclude that

$$\lim_{|m|\to\infty} \frac{1}{m}(\lambda_{m,r} + mi\omega) = \mathrm{i}(\mu_r^0 + \omega),$$

where the right-hand side is non-zero since $\mu_r^0 \neq -\omega$ due to Assumptions 2(iv). We conclude that

$$|\lambda_{m,r} + mi\omega| \gtrsim m,$$

so that

$$C = \sup_{m\in\mathbb{Z}} \max_{1\le r\le r_m} \frac{1+m^2}{|\lambda_{m,r}+mi\omega|^2}$$

is finite and $\|\mathcal{L}_k^{-1}\|_{\mathcal{L}(\mathcal{P}_t^k;\mathcal{P}_{t+1}^k)} \lesssim C$ independently of k. ∎

3.2 Estimates for the normal-form transformation

Our aim is to determine an optimal order p for our transformation such that the remaining nonlinear terms

$$\sum_{m\ge p+1} \tilde{h}^m(y)$$

in the Maclaurin expansion of $\tilde{h}$ become exponentially small with respect to μ. We begin by estimating the size of Φ, that is we derive estimates for $|\Phi^m|_{2,m}^{t+1}$ and $\|\Phi^m(y)\|_{\mathcal{X}_{\mathrm{sh,c}}^{t+1}}$. Here we make extensive use of the relations between the different norms given in Lemmata 3.3 and 3.4.

By equation (3.10), we have that

$$\Phi^m = \mathcal{L}_m^{-1}N^m, \quad m=2,\dots,p.$$

Defining $\phi_m = |\Phi^m|_{2,m}^{t+1}$ and $\nu_m = |N^m|_{2,m}^{t}$, we use Proposition 3.6 to conclude that

$$\phi_m \le \gamma\nu_m,$$

and we proceed by estimating ν_m for $m\in\{2,\dots,p\}$.

We use the notation

$$\begin{aligned} f^m[y_1,\dots,y_m] &= \frac{1}{m!}\mathrm{d}^m f[0](y_1,\dots,y_m), \\ f^{m_1,m_2}[y_1,\dots,y_{m_1},w_1,\dots,w_{m_2}] &= \frac{1}{m_1!m_2!}\mathrm{d}^{m_1,m_2} f[0](y_1,\dots,y_{m_1},w_1,\dots,w_{m_2}), \end{aligned}$$

where $m, m_1, m_2 \in \mathbb{N}_0$ for analytic functions of one and two variables and

$$\begin{aligned} \Phi_{\boldsymbol{p}} &= (\Phi^{p_1},\dots,\Phi^{p_q}), \\ \phi_{\boldsymbol{p}} &= \prod_{j=1}^{q} |\Phi^{p_j}|_{2,p_j}^{t+1} \end{aligned}$$

for $\boldsymbol{p}\in\mathbb{N}_2^q$, where $\mathbb{N}_2 := \mathbb{N}\setminus\{1\}$. We also write

$$(U^\mu)^{-1}\tilde{A}^\mu(0)U^\mu = \sum_{i=0}^{\infty} A_i\mu^i, \quad (U^\mu)^{-1}\partial\tilde{B}^\mu(0)U^\mu = \sum_{i=0}^{\infty} B_i\mu^i,$$

in which $A_i, B_i \in \mathcal{L}(\mathcal{X}^t)$ are operator-valued coefficients with $A_0 = \tilde{A}^0(0)$ and $B_0 = \partial \tilde{B}^0(0)$, so that

$$(L^\mu_{\mathrm{sh,c}} - L^0_{\mathrm{sh,c}})\Phi(y) = \sum_{i=1}^\infty A_i \mu^i \partial_\tau \Phi(y) + \sum_{i=1}^\infty B_i \mu^i \Phi(y).$$

It follows that

$$\begin{aligned} N^m(y) &= h^m[\{y\}^{(m)}] \\ &\quad + \sum_{2\le k\le m-1} \mathrm{d}\Phi^k[y]\Bigg(\sum_{2\le q\le m-k+1} \sum_{i=0}^q \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i} \\ |\boldsymbol{p}|=m-k+1-i}} f^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)]\Bigg) \\ &\quad - \sum_{1\le q\le m-1} \sum_{i=0}^q \sum_{\substack{(\boldsymbol{p},r)\in\mathbb{N}_2^{q-i+1} \\ |\boldsymbol{p}|=m-i-r}} P_{\mathrm{sh,c}}\Big(g_1^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)]\partial_\tau \Phi^r(y)\Big) \\ &\quad + \sum_{2\le q\le m} \sum_{i=0}^q \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i} \\ |\boldsymbol{p}|=m-i}} g_2^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)] \\ &\quad - \sum_{2\le k\le m-1} A_{m-k}\mu^{m-k}\partial_\tau\Phi^k(y) - \sum_{2\le k\le m-1} B_{m-k}\mu^{m-k}\Phi^k(y) \end{aligned} \tag{3.11}$$

for $m \ge 2$.

Because the functions $f\colon \mathbb{C}^5 \times \mathcal{X}^t_{\mathrm{sh,c}} \to \mathbb{C}^5$, $g_1\colon \mathbb{C}^5 \times \mathcal{X}^t_{\mathrm{sh,c}} \to H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^{n\times n})$, $g_2\colon \mathbb{C}^5 \times \mathcal{X}^t_{\mathrm{sh,c}} \to \mathcal{X}^t_{\mathrm{sh,c}}$ and $h\colon \mathbb{C}^5 \to \mathcal{X}^t_{\mathrm{sh,c}}$ are analytic at the origin for $t > \frac{1}{2}$, there exist universal constants $a > 1$ and $\rho < 1$ such that

$$\begin{aligned} &|f^{m_1,m_2}[y_1,\dots,y_{m_1},q_1,\dots,q_{m_2}]| \\ &\qquad \le \frac{a}{\rho^{m_1+m_2}}|y_1|\cdot\ldots\cdot|y_{m_1}|\cdot\|q_1\|_{\mathcal{X}^t_{\mathrm{sh,c}}}\cdot\ldots\cdot\|q_{m_2}\|_{\mathcal{X}^t_{\mathrm{sh,c}}}, \\ &\Big\|P_{\mathrm{sh,c}}\Big(g_1^{m_1,m_2}[y_1,\dots,y_{m_1},q_1,\dots,q_{m_2}]\partial_\tau q_{m_2+1}\Big)\Big\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \\ &\qquad \le \frac{a}{4\rho^{m_1+m_2}}|y_1|\cdot\ldots\cdot|y_{m_1}|\cdot\|q_1\|_{\mathcal{X}^t_{\mathrm{sh,c}}}\cdot\ldots\cdot\|q_{m_2}\|_{\mathcal{X}^t_{\mathrm{sh,c}}}\cdot\|q_{m_2+1}\|_{\mathcal{X}^{t+1}_{\mathrm{sh,c}}}, \\ &\|g_2^{m_1,m_2}[y_1,\dots,y_{m_1},q_1,\dots,q_{m_2}]\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \\ &\qquad \le \frac{a}{4\rho^{m_1+m_2}}|y_1|\cdot\ldots\cdot|y_{m_1}|\cdot\|q_1\|_{\mathcal{X}^t_{\mathrm{sh,c}}}\cdot\ldots\cdot\|q_{m_2}\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \end{aligned}$$

and

$$\|h^{m_1}[y_1,\dots y_{m_1}]\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \le \frac{a}{\rho^{m_1}}|y_1|\cdot\ldots\cdot|y_{m_1}|$$

for each $m_1, m_2 \in \mathbb{N}_0$. We furthermore choose a and ρ such that

$$\|A_m\|_{\mathcal{L}(\mathcal{X}^t)} \le \frac{a}{4\rho^m}, \quad \|B_m\|_{\mathcal{L}(\mathcal{X}^t)} \le \frac{a}{4\rho^m}$$

for each $m \in \mathbb{N}_0$. Using these estimates together with Lemmata 3.3 and 3.4, we find from equation (3.11) that

$$\begin{aligned}\phi_m \leq \gamma\Bigg(&\frac{a}{\rho^m}\sqrt{5}^m + \sum_{2\leq k\leq m-1}\sqrt{k^2+4k}\phi_k \sum_{2\leq q\leq m-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-k+1-i}}\frac{a}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}\\&+\frac{1}{4}\sum_{1\leq q\leq m-1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\|\boldsymbol{p}|=m-i}}\frac{a}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}} + \frac{1}{4}\sum_{2\leq q\leq m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-i}}\frac{a}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}\\&+\frac{1}{2}\sum_{2\leq k\leq m-1}\frac{a}{\rho^{m-k}}\sqrt{5}^{m-k}\phi_k\Bigg).\end{aligned}\tag{3.12}$$

Observe that

$$\begin{aligned}\sum_{2\leq q\leq m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-i}}\frac{a}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}} &\geq \sum_{2\leq q\leq m}\sum_{\substack{p_1\in\mathbb{N}_2\\p_1=m-(q-1)}}\frac{a}{\rho^q}\sqrt{5}^{q-1}\phi_{p_1}\\&=\sum_{2\leq q\leq m-1}\frac{a}{\rho^q}\sqrt{5}^{q-1}\phi_{m-(q-1)}\\&\geq\sum_{2\leq q\leq m-1}\frac{a}{\rho^{q-1}}\sqrt{5}^{q-1}\phi_{m-(q-1)}\\&=\sum_{2\leq k\leq m-1}\frac{a}{\rho^{m-k}}\sqrt{5}^{m-k}\phi_k,\end{aligned}$$

in which we have written $k = m-(q-1)$ in the last step, and

$$\begin{aligned}\sum_{1\leq q\leq m-1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\|\boldsymbol{p}|=m-i}}\frac{a}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}} &= \sum_{2\leq q\leq m}\sum_{i=0}^{q-1}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-i}}\frac{a}{\rho^{q-1}}\sqrt{5}^i\phi_{\boldsymbol{p}}\\&\leq\sum_{2\leq q\leq m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-i}}\frac{a}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}.\end{aligned}$$

Using these combinatorial results, we find from inequality (3.12) that

$$\begin{aligned}\phi_m \leq{}& \frac{a\gamma}{\rho^m}\sqrt{5}^m + \sum_{2\leq k\leq m-1}\sqrt{k^2+4k}\sum_{2\leq q\leq m-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-k+1-i}}\frac{a\gamma}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}\\&+\sum_{2\leq q\leq m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-i}}\frac{a\gamma}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}.\end{aligned}\tag{3.13}$$

Our aim is to obtain a non-recursive estimate for ϕ_m, $m \geq 2$, and to this end we consider the sequence $\{\beta_l\}_{l\in\mathbb{N}}$ defined recursively by the formulae

$$\begin{aligned}
\beta_1 &= 1,\\
\beta_m &= \left(\frac{\rho}{a\gamma}\right)^{m-2} + \sum_{2\leq q\leq m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-i}}\left(\frac{\rho}{a\gamma}\right)^{q-2}\beta_{\boldsymbol{p}}\\
&\qquad + 5\sum_{2\leq k\leq m-1}k\beta_k\sum_{2\leq q\leq m-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\|\boldsymbol{p}|=m-k+1-i}}\left(\frac{\rho}{a\gamma}\right)^{q-2}\beta_{\boldsymbol{p}}
\end{aligned}$$

for $m \geq 2$, where we have used the notation

$$\beta_{\boldsymbol{p}} = \prod_{j=1}^{q-i}\beta_{p_j}.$$

The first step is to establish an estimate of the form

$$\phi_m \leq c_m\beta_m, \quad m \geq 2,$$

where $\{c_m\}_{m\in\mathbb{N}}$ is a non-recursively defined sequence of positive real numbers.

Proposition 3.7. *The estimate*

$$\phi_m \leq \sqrt{5}\left(\frac{\sqrt{5}a\gamma}{\rho^2}\right)^{m-1}\beta_m$$

holds for each $m \geq 2$.

Proof. We establish this result by mathematical induction. If $m = 2$ we have that $\beta_2 = 2$ and inequality (3.13) shows that

$$\phi_2 \leq 2\frac{a\gamma}{\rho^2}\sqrt{5}^2 = \sqrt{5}\left(\frac{\sqrt{5}a\gamma}{\rho^2}\right)\beta_2.$$

Now suppose that $m \geq 3$ and the result holds for all $j < m$. Using inequality (3.13) we find that

$$
\begin{aligned}
\phi_m \le{}& \frac{a\gamma}{\rho^m}\sqrt{5}^m + \sum_{2\le q\le m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=m-i}} \frac{a\gamma}{\rho^q}\sqrt{5}^q\Big(\frac{\sqrt{5}a\gamma}{\rho^2}\Big)^{m-q}\beta_{\boldsymbol{p}}\\
&+\sum_{2\le k\le m-1}\sqrt{k^2+4k}\sqrt{5}\Big(\frac{\sqrt{5}a\gamma}{\rho^2}\Big)^{k-1}\beta_k\\
&\qquad\times\sum_{2\le q\le m-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=m-k+1-i}} \frac{a\gamma}{\rho^q}\sqrt{5}^q\Big(\frac{\sqrt{5}a\gamma}{\rho^2}\Big)^{m-k+1-q}\beta_{\boldsymbol{p}}\\
={}& \frac{a\gamma}{\rho^m}\sqrt{5}^m + \sum_{2\le q\le m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=m-i}} \frac{a\gamma}{\rho^q}\sqrt{5}^m\Big(\frac{a\gamma}{\rho^2}\Big)^{m-q}\beta_{\boldsymbol{p}}\\
&+\sum_{2\le k\le m-1}\sqrt{k^2+4k}\sqrt{5}\beta_k\sum_{2\le q\le m-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=m-k+1-i}} \frac{a\gamma}{\rho^q}\sqrt{5}^m\Big(\frac{a\gamma}{\rho^2}\Big)^{m-q}\beta_{\boldsymbol{p}}
\end{aligned}
$$

and since $\sqrt{k^2+4k}\sqrt{5}\le 5k$ and

$$
\begin{aligned}
\frac{1}{\sqrt{5}}\Big(\frac{\rho^2}{\sqrt{5}a\gamma}\Big)^{m-1}\frac{a\gamma}{\rho^m}\sqrt{5}^m &= \Big(\frac{\rho}{a\gamma}\Big)^{m-2},\\
\frac{1}{\sqrt{5}}\Big(\frac{\rho^2}{\sqrt{5}a\gamma}\Big)^{m-1}\frac{a\gamma}{\rho^q}\sqrt{5}^m\Big(\frac{a\gamma}{\rho^2}\Big)^{m-q} &= \Big(\frac{\rho}{a\gamma}\Big)^{q-2},
\end{aligned}
$$

we conclude that

$$
\begin{aligned}
\phi_m \le{}& \sqrt{5}\Big(\frac{\sqrt{5}a\gamma}{\rho^2}\Big)^{m-1}\Bigg(\Big(\frac{\rho}{a\gamma}\Big)^{m-2} + \sum_{2\le q\le m}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=m-i}}\Big(\frac{\rho}{a\gamma}\Big)^{q-2}\beta_{\boldsymbol{p}}\\
&\qquad+5\sum_{2\le k\le m-1}k\beta_k\sum_{2\le q\le m-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=m-k+1-i}}\Big(\frac{\rho}{a\gamma}\Big)^{q-2}\beta_{\boldsymbol{p}}\Bigg)\\
={}& \sqrt{5}\Big(\frac{\sqrt{5}a\gamma}{\rho^2}\Big)^{m-1}\beta_m.
\end{aligned}
$$

■

The second step is to show that the sequence $\{\beta_l\}_{l\in\mathbb{N}}$ is bounded by a non-recursively defined sequence $\{\alpha_l\}_{l\in\mathbb{N}}$ and to use Proposition 3.7 to obtain the desired non-recursive estimate for ϕ_m, $m\ge 2$.

Proposition 3.8. *Consider the sequence $\{\alpha_m\}_{m\in\mathbb{N}}$ defined by the formulae*

$$\alpha_1 = 1,$$
$$\alpha_m = \theta^{m-2}(m-2)!, \quad m \geq 2,$$

where $\theta \geq 1$ is a constant satisfying $\rho/a\gamma\theta \leq 1/5$ and

$$\frac{21}{40} + \frac{129}{2\theta} \leq 1.$$

The estimate

$$\beta_m \leq 2^m \alpha_m$$

holds for each $m \geq 1$.

Proof. This result is also established using mathematical induction. Note that

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

so that the statement is true for $m = 1$ and $m = 2$. Now suppose that $m \geq 3$ and that $\beta_j \leq 2^j\alpha_j$ for all $j < m$. Since $\beta_1 = 1$, we find that

$$\sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=m-i}} \beta_{\boldsymbol{p}} = \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=m}} \beta_{\boldsymbol{p}},$$

so that we can rewrite the recursive formula for $\beta_m,\ m \geq 2$, as

$$\begin{aligned}\beta_m =&\left(\frac{\rho}{a\gamma}\right)^{m-2} + \sum_{2\leq q\leq m} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=m}} \left(\frac{\rho}{a\gamma}\right)^{q-2} \beta_{\boldsymbol{p}}\\ &+ 5\sum_{2\leq k\leq m-1} k\beta_k \sum_{2\leq q\leq m-k+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=m-k+1}} \left(\frac{\rho}{a\gamma}\right)^{q-2} \beta_{\boldsymbol{p}}.\end{aligned}$$

It follows by the induction hypothesis that

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

where

$$\begin{aligned}\Delta_m^1 &= \sum_{3\leq q\leq m} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=m}} \left(\frac{\rho}{a\gamma}\right)^{q-2} \alpha_{\boldsymbol{p}},\\ \Delta_m^2 &= 10\sum_{2\leq k\leq m-1} k\alpha_k \sum_{3\leq q\leq m-k+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=m-k+1}} \left(\frac{\rho}{a\gamma}\right)^{q-2} \alpha_{\boldsymbol{p}},\\ \Delta_m^3 &= \sum_{1\leq k\leq m-1} \alpha_k\alpha_{m-k},\\ \Delta_m^4 &= 10\sum_{2\leq k\leq m-1} k\alpha_k \sum_{1\leq j\leq m-k} \alpha_j\alpha_{m-k+1-j}\end{aligned}$$

and we have used the notation

$$\alpha_{\boldsymbol{p}} = \prod_{j=1}^{q} \alpha_{p_j}.$$

In order to estimate the quantities $\Delta_m^1, \dots, \Delta_m^4$, we make use of the inequalities

$$\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q \\ |\boldsymbol{p}|=m}} \alpha_{\boldsymbol{p}} \leq \frac{2}{\theta^{q-2}}\alpha_m, \qquad 3 \leq q \leq m, \tag{3.14}$$

$$\sum_{2\leq k\leq m-1} \frac{k(k-2)!(m-k-1)!}{(m-2)!} \leq \frac{5}{2}, \qquad m \geq 3, \tag{3.15}$$

$$\sum_{2\leq k\leq m-2} \frac{(k-2)!(m-k-2)!}{(m-2)!} \leq \frac{1}{2}, \qquad m \geq 4, \tag{3.16}$$

which where established by Iooss and Lombardi [14, Lemma 2.13, steps 2.1, 2.2 and 4]. Inequality (3.14) shows that

$$\begin{aligned}
\Delta_m^1 &= \sum_{3\leq q\leq m} \left(\frac{\rho}{a\gamma}\right)^{q-2} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q \\ |\boldsymbol{p}|=m}} \alpha_{\boldsymbol{p}} \\
&\leq 2 \sum_{3\leq q\leq m} \left(\frac{\rho}{a\gamma\theta}\right)^{q-2} \alpha_m \\
&= 2 \sum_{1\leq q\leq m-2} \left(\frac{\rho}{a\gamma\theta}\right)^{q} \alpha_m \\
&\leq 2\left(\frac{1}{1-\frac{\rho}{a\gamma\theta}} - 1\right)\alpha_m \\
&\leq \frac{\alpha_m}{2},
\end{aligned}$$

since $\rho/a\gamma\theta \leq 1/5$. Similarly, it follows from inequalities (3.14), (3.15) that

$$\begin{aligned}
\Delta_m^2 &= 10 \sum_{2\leq k\leq m-1} k\alpha_k \sum_{3\leq q\leq m-k+1} \left(\frac{\rho}{a\gamma}\right)^{q-2} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q \\ |\boldsymbol{p}|=m-k+1}} \alpha_{\boldsymbol{p}} \\
&\leq 5 \sum_{2\leq k\leq m-1} k\alpha_k\alpha_{m-k+1} \\
&= 5 \sum_{2\leq k\leq m-1} k\theta^{m-3}(k-2)!(m-k-1)! \\
&= \frac{5}{\theta}\alpha_m \sum_{2\leq k\leq m-1} \frac{k(k-2)!(m-k-1)!}{(m-2)!} \\
&\leq \frac{25}{2\theta}\alpha_m.
\end{aligned}$$

Next we show that

$$\Delta_m^3 = \sum_{1\le k\le m-1} \alpha_k \alpha_{m-k} \le \frac{2}{\theta}\alpha_m, \quad m \ge 3.$$

If $m = 3$, we have that

$$\sum_{1\le k\le 2} \alpha_k \alpha_{3-k} = 2\alpha_2 = 2 = \frac{2}{\theta}\theta = \frac{2}{\theta}\alpha_3$$

and in the case $m \ge 4$ we can use induction to show that

$$\frac{2}{\theta(m-2)} + \frac{1}{2\theta^2} \le \frac{2}{\theta},$$

which together with inequality (3.16) yields

$$\begin{aligned}
\sum_{1\le k\le m-1} \alpha_k \alpha_{m-k} &= 2\alpha_1\alpha_{m-1} + \sum_{2\le k\le m-2} \alpha_k \alpha_{m-k} \\
&= 2\theta^{m-3}(m-3)! + \theta^{m-4} \sum_{2\le k\le m-2} (k-2)!(m-k-2)! \\
&= \alpha_m \left(\frac{2}{\theta(m-2)} + \frac{1}{\theta^2} \sum_{2\le k\le m-2} \frac{(k-2)!(m-k-2)!}{(m-2)!} \right) \\
&\le \alpha_m \left(\frac{2}{\theta(m-2)} + \frac{1}{2\theta^2} \right) \\
&\le \frac{2}{\theta}\alpha_m,
\end{aligned}$$

proving the estimate for Δ_m^3. Turning to Δ_m^4, we use the estimate for Δ_m^3, $m \ge 3$, together with inequality (3.15) to conclude that

$$
\begin{aligned}
\Delta_m^4 &= 10 \sum_{2\le k\le m-1} k\alpha_k \sum_{1\le j\le m-k} \alpha_j \alpha_{m-k+1-j} \\
&= 10(m-1)\alpha_{m-1}\alpha_1^2 + 10 \sum_{2\le k\le m-2} k\alpha_k \sum_{1\le j\le m-k} \alpha_j \alpha_{m-k+1-j} \\
&= 10(m-1)\alpha_{m-1} + 10 \sum_{2\le k\le m-2} k\alpha_k \Delta^3_{m-k+1} \\
&\le 10(m-1)\alpha_{m-1} + \frac{20}{\theta} \sum_{2\le k\le m-2} k\alpha_k \alpha_{m-k+1} \\
&= 10(m-1)\theta^{m-3}(m-3)! + 20\theta^{m-4} \sum_{2\le k\le m-2} k(k-2)!(m-k-1)! \\
&= \alpha_m \biggl(\frac{10(m-1)}{\theta(m-2)} + \frac{20}{\theta^2} \sum_{2\le k\le m-2} \frac{k(k-2)!(m-k-1)!}{(m-2)!} \biggr) \\
&\le \alpha_m \biggl(\frac{20}{\theta}\frac{m-1}{m-2} + \frac{20}{\theta} \sum_{2\le k\le m-2} \frac{k(k-2)!(m-k-1)!}{(m-2)!} \biggr) \\
&= \alpha_m \frac{20}{\theta} \sum_{2\le k\le m-1} \frac{k(k-2)!(m-k-1)!}{(m-2)!} \\
&\le \frac{50}{\theta}\alpha_m.
\end{aligned}
$$

Finally note that

$$
\Bigl(\frac{\rho}{a\gamma}\Bigr)^{m-2} = \Bigl(\frac{\rho}{a\gamma\theta}\Bigr)^{m-2}\theta^{m-2} \le \frac{1}{5}\theta^{m-2} \le \frac{2^m}{40}\alpha_m,
$$

since $5 \cdot 2^m \ge 40$ for every $m \ge 3$. Altogether we have that

$$
\begin{aligned}
\beta_m &\le \Bigl(\frac{1}{2} + \frac{25}{2\theta} + \frac{2}{\theta} + \frac{50}{\theta} + \frac{1}{40}\Bigr)2^m\alpha_m \\
&= \Bigl(\frac{21}{40} + \frac{129}{2\theta}\Bigr)2^m\alpha_m \\
&\le 2^m\alpha_m.
\end{aligned}
$$

■

For the remainder of this thesis, we fix $\theta \ge 1$ to be a constant satisfying the assumptions of Proposition 3.8. It follows from Propositions 3.7 and 3.8 that

$$
\begin{aligned}
\phi_m &\le \sqrt{5}\Bigl(\frac{\sqrt{5}a\gamma}{\rho^2}\Bigr)^{m-1}\beta_m \\
&\le \sqrt{5}\Bigl(\frac{\sqrt{5}a\gamma}{\rho^2}\Bigr)^{m-1}2^m\alpha_m \\
&= \sqrt{5}\Bigl(\frac{\sqrt{5}a\gamma}{\rho^2}\Bigr)^{m-1}2^m\theta^{m-2}(m-2)! \\
&= \frac{20a\gamma}{\rho^2}\Bigl(\frac{2\sqrt{5}a\gamma\theta}{\rho^2}\Bigr)^{m-2}(m-2)! \qquad (3.17)
\end{aligned}
$$

for each $m \geq 2$. Using this estimate together with a mutual constraint on the order p of the normal form and the maximum size δ of $|y|$, we can establish estimates for $\|\Phi(y)\|_{\mathcal{X}_{\mathrm{sh,c}}^{t+1}}$ and $\|\mathrm{d}\Phi[y]\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}_{\mathrm{sh,c}}^{t+1})}$ from which the estimate for $\|\tilde{h}(y)\|_{\mathcal{X}_{\mathrm{sh,c}}^{t}}$ is derived.

Proposition 3.9. *Let $p \geq 2$ be a natural number and $\delta > 0$ such that*

$$\delta p \leq \frac{\rho^2}{4\sqrt{5}a\gamma\theta}.$$

The estimates

$$\begin{aligned}
\Big\| \sum_{2\leq k\leq p} \Phi^k(y) \Big\|_{\mathcal{X}_{\mathrm{sh,c}}^{t+1}} &\leq \frac{\sqrt{5}\delta}{\theta}, \\
\Big\| \sum_{2\leq k\leq p} \mathrm{d}\Phi^k[y] \Big\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}_{\mathrm{sh,c}}^{t+1})} &\leq \frac{10}{\theta}, \\
\Big\| \sum_{2\leq k\leq p} \mathrm{d}^2\Phi^k[y] \Big\|_{\mathcal{L}_{\mathrm{sym}}^{(2)}(\mathbb{C}^5;\mathcal{X}_{\mathrm{sh,c}}^{t+1})} &\leq \frac{10\sqrt{5}}{\theta\delta}
\end{aligned}$$

hold for each $y \in \overline{B}_\delta(0) \subseteq \mathbb{C}^5$.

Proof. Take $y \in \overline{B}_\delta(0)$. Using Lemma 3.3(i) together with inequality (3.17), we find that

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

and since $(k-2)!/p^{k-1} \leq 1/p$ for $2 \leq k \leq p$, we conclude that

$$\Big\| \sum_{2\leq k\leq p} \Phi^k(y) \Big\|_{\mathcal{X}_{\mathrm{sh,c}}^{t+1}} \leq \frac{2\sqrt{5}\delta}{\theta p} \sum_{2\leq k\leq p} \left(\frac{1}{2}\right)^{k-1} \leq \frac{\sqrt{5}\delta}{\theta}.$$

For the second estimate we additionally use Proposition 3.5 to obtain

$$
\begin{aligned}
\Big\| \sum_{2\le k\le p} \mathrm{d}\Phi^k[y] \Big\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^{t+1}_{\mathrm{sh,c}})} &\le \sum_{2\le k\le p} |\mathrm{d}\Phi^k[y]|^{t+1}_{0,k-1} |y|^{k-1} \\
&\le \sum_{2\le k\le p} |\mathrm{d}\Phi^k[y]|^{t+1}_{2,k-1} |y|^{k-1} \\
&\le \sum_{2\le k\le p} \sqrt{5} k \phi_k \delta^{k-1} \\
&\le \sum_{2\le k\le p} \frac{20\sqrt{5}a\gamma\delta}{\rho^2} \left(\frac{2\sqrt{5}a\gamma\theta\delta}{\rho^2}\right)^{k-2} k(k-2)! \\
&= \frac{10}{\theta} \sum_{2\le k\le p} \left(\frac{2\sqrt{5}a\gamma\theta\delta}{\rho^2}\right)^{k-1} k(k-2)! \\
&\le \frac{10}{\theta} \sum_{2\le k\le p} \left(\frac{1}{2p}\right)^{k-1} k(k-2)!,
\end{aligned}
$$

and since $k(k-2)!/p^{k-1} \le 1$ for $2 \le k \le p$, we conclude that

$$
\Big\| \sum_{2\le k\le p} \mathrm{d}\Phi^k[y] \Big\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^{t+1}_{\mathrm{sh,c}})} \le \frac{10}{\theta} \sum_{2\le k\le p} \left(\frac{1}{2}\right)^{k-1} \le \frac{10}{\theta}.
$$

The estimate for the second derivative is proved in a similar manner by estimating

$$
\|\mathrm{d}^2\Phi^k[y]\|_{\mathcal{L}^{(2)}_{\mathrm{sym}}(\mathbb{C}^5;\mathcal{X}^{t+1}_{\mathrm{sh,c}})} \le 5k(k-1)\phi_k\delta^{k-2}
$$

and $k!/p^{k-1} \le 1$ for $2 \le k \le p$. ∎

3.3 Exponentially small estimate

In this section we make an appropriate choice for the order p of the normal form so that $\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}}$ is exponentially small with respect to μ. Since by construction $\tilde{h}^m = 0$ for $m \leq p$, we find that

$$\begin{aligned}
\tilde{h}(y) = {} & \sum_{p+1\leq q} h^q[\{y\}^{(q)}] - \sum_{p\leq q} A_q\mu^q \sum_{2\leq k\leq p} \partial_\tau \Phi^k(y) - \sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}} A_q\mu^q \partial_\tau \Phi^r(y) \\
& - \sum_{p\leq q} B_q\mu^q \sum_{2\leq k\leq p} \Phi^k - \sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}} B_q\mu^q \Phi^r(y) \\
& - \sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{(\boldsymbol{p},r)\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i-r}} P_{\mathrm{sh,c}}\Big(g_1^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)]\partial_\tau\Phi^r(y)\Big) \\
& - \sum_{p\leq q}\sum_{i=0}^{q} P_{\mathrm{sh,c}}\left(g_1^{i,q-i}\left[\{y\}^{(i)}, -\Big\{\sum_{2\leq k\leq p}\Phi^k(y)\Big\}^{(q-i)}\right]\sum_{2\leq k\leq p}\partial_\tau\Phi^k(y)\right) \\
& + \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} g_2^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)] \\
& + \sum_{p+1\leq q}\sum_{i=0}^{q} g_2^{i,q-i}\left[\{y\}^{(i)}, -\Big\{\sum_{2\leq k\leq p}\Phi^k(y)\Big\}^{(q-i)}\right] \\
& + \sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}} \mathrm{d}\Phi^k[y]\Bigg(\sum_{2\leq q\leq l-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}} f^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)]\Bigg) \\
& + \Big(\sum_{k=2}^{p}\mathrm{d}\Phi^k[y]\Big)\Bigg(\sum_{\substack{2\leq q\leq p\\ l\geq p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} f^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)] \\
& \qquad + \sum_{p+1\leq q}\sum_{i=0}^{q} f^{i,q-i}\left[\{y\}^{(i)}, -\Big\{\sum_{2\leq k\leq p}\Phi^k(y)\Big\}^{(q-i)}\right]\Bigg),
\end{aligned} \tag{3.18}$$

in which we have set $\Phi^{p_j} = 0$ for $p_j \geq p+1$ for notational simplicity. We estimate each term in the above formula using Proposition 3.9 and Lemmata 3.3 and 3.4 as well as the analyticity of the involved nonlinearities.

Suppose that $\delta > 0$ and $p \geq 2$ satisfy the constraint given in Proposition 3.9 and note that

$$\|\Phi^k(y)\|_{\mathcal{X}^{t+1}_{\mathrm{sh,c}}} \leq |\Phi^k|^{t+1}_{0,k}|y|^k \leq \phi_k|y|^k \leq \phi_k\delta^k$$

for $|y| \leq \delta$ by definition of the relevant norms and Lemmata 3.3 and 3.4. It follows from equation (3.18) that

$$
\begin{aligned}
\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \leq{} & \sum_{p+1\leq q} \frac{a}{\rho^q}(\sqrt{5}\delta)^q + \frac{1}{2}\sum_{p\leq q}\frac{a}{\rho^q}(\sqrt{5}\delta)^q\frac{\sqrt{5}\delta}{\theta} + \frac{1}{2}\sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}} \frac{a}{\rho^q}\sqrt{5}^q\delta^{q+r}\phi_r \\
& + \frac{1}{4}\sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}} \frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}} \\
& + \frac{1}{4}\sum_{p\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{\sqrt{5}\delta}{\theta} \\
& + \left(\frac{1}{4}+\frac{10}{\theta}\right)\sum_{\substack{2\leq q\leq p\\ l\geq p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}} \\
& + \left(\frac{1}{4}+\frac{10}{\theta}\right)\sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} \\
& + \sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}}\sqrt{5}k\phi_k\sum_{2\leq q\leq l-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}.
\end{aligned} \tag{3.19}
$$

Next we simplify the expression on the right-hand side of inequality (3.19). First note that

$$
\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} \geq \frac{a}{\rho^q}(\sqrt{5}\delta)^q
$$

and therefore

$$
\begin{aligned}
& \sum_{p+1\leq q}\frac{a}{\rho^q}(\sqrt{5}\delta)^q + \frac{1}{2}\sum_{p\leq q}\frac{a}{\rho^q}(\sqrt{5}\delta)^q\frac{\sqrt{5}\delta}{\theta} + \frac{1}{4}\sum_{p\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{\sqrt{5}\delta}{\theta} \\
& \qquad + \left(\frac{1}{4}+\frac{10}{\theta}\right)\sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} \\
& \leq \left(\frac{5}{4}+\frac{10}{\theta}\right)\sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} + \frac{3}{4}\sum_{p\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{\sqrt{5}\delta}{\theta} \\
& \leq \left(\frac{5}{4}+\frac{10}{\theta}\right)\sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} + \frac{3}{4}\sum_{p+1\leq q}\sum_{i=0}^{q-1}\frac{a}{\rho^{q-1}}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} \\
& \leq \left(2+\frac{10}{\theta}\right)\sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} \\
& = \left(2+\frac{10}{\theta}\right)\sum_{p+1\leq q}\frac{a}{\rho^q}(\sqrt{5}\delta)^q\sum_{i=0}^{q}\left(\frac{1}{\theta}\right)^{q-i} \\
& \leq \left(2+\frac{10}{\theta}\right)\sum_{p+1\leq q}\frac{a}{\rho^q}(\sqrt{5}\delta)^q(q+1).
\end{aligned}
$$

Moreover

$$\sum_{\substack{1\le q\le p-1\\ l\ge p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}\ge\sum_{\substack{1\le q\le p-1\\ l\ge p+1}}\sum_{\substack{r\in\mathbb{N}_2\\ r=l-q}}\frac{a\delta^l}{\rho^q}\sqrt{5}^q\phi_r$$
$$=\sum_{\substack{1\le q\le p-1\\ p+1-q\le r\le p}}\frac{a}{\rho^q}\delta^{q+r}\sqrt{5}^q\phi_r$$

and defining $\phi_1=\sqrt{5}$ yields

$$\sum_{\substack{1\le q\le p-1\\ l\ge p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}=\sum_{\substack{2\le q\le p\\ l\ge p+1}}\sum_{i=0}^{q-1}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^{q-1}}\sqrt{5}^i\phi_{\boldsymbol{p}}$$
$$\le\sum_{\substack{2\le q\le p\\ l\ge p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}$$
$$=\sum_{\substack{2\le q\le p\\ l\ge p+1}}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}}\frac{a\delta^l}{\rho^q}\phi_{\boldsymbol{p}}.$$

Hence

$$\frac{1}{2}\sum_{\substack{1\le q\le p-1\\ p+1-q\le r\le p}}\frac{a}{\rho^q}\sqrt{5}^q\delta^{q+r}\phi_r+\frac{1}{4}\sum_{\substack{1\le q\le p-1\\ l\ge p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}$$
$$+\left(\frac{1}{4}+\frac{10}{\theta}\right)\sum_{\substack{2\le q\le p\\ l\ge p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}$$
$$\le\left(1+\frac{10}{\theta}\right)\sum_{\substack{2\le q\le p\\ l\ge p+1}}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}}\frac{a\delta^l}{\rho^q}\phi_{\boldsymbol{p}}.$$

Finally we find that

$$\sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}}\sqrt{5}k\phi_k\sum_{2\le q\le l-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}$$
$$=\sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}}\sqrt{5}k\phi_k\sum_{2\le q\le l-k+1}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}}\frac{a\delta^l}{\rho^q}\phi_{\boldsymbol{p}}$$

and altogether we conclude from inequality (3.19) that

$$\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}}\le\left(2+\frac{10}{\theta}\right)\Delta_1^p+\left(1+\frac{10}{\theta}\right)\Delta_2^p+\Delta_3^p,\tag{3.20}$$

where

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

$$\Delta_2^p = \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} \frac{a\delta^l}{\rho^q}\phi_{\boldsymbol{p}},$$

$$\Delta_3^p = \sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}} \sqrt{5}k\phi_k \sum_{2\leq q\leq l-k+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} \frac{a\delta^l}{\rho^q}\phi_{\boldsymbol{p}}$$

and we have defined $\phi_j = 0$ if $j \geq p+1$.

We now construct an estimate for $\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}}$ of a non-recursive nature. To this end we need the following combinatorial results. Throughout the following theory we define $(-1)! = 0! = 1$ for notational simplicity.

Lemma 3.10. *Suppose that $p \geq 2$ is a natural number.*

(i) The inequality

$$\sum_{l=p+1}^{pq} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} \left(\frac{1}{p}\right)^l (p_1-2)!\cdot\ldots\cdot(p_q-2)! \leq \left(\sum_{j=1}^{p}\left(\frac{1}{p}\right)^j (j-2)!\right)^q$$

holds for each $q \in \{2,\ldots,p\}$.

(ii) The inequality

$$\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} (p_1-2)!\cdot\ldots\cdot(p_q-2)! \leq 2(l-k-1)!$$

holds for each $k \in \{2,\ldots,p\}$, $l \in \{p+1,\ldots,p+k-1\}$ and $q \in \{2,\ldots,l-k+1\}$.

Proof.

(i) Using the multinomial theorem, we find that

$$\begin{aligned}\left(\sum_{j=1}^{p}\left(\frac{1}{p}\right)^{j}(j-2)!\right)^{q} &= \sum_{\substack{0\le k_1,\dots,k_p\le q\\ k_1+\ldots+k_p=q}} \frac{q}{k_1!\cdot\ldots\cdot k_p!}\left(\frac{1}{p}\right)^{k_1+2k_2+\ldots+pk_p}\\ &\qquad\times(-1)!^{k_1}\cdot\ldots\cdot(p-2)!^{k_p}\\ &= \sum_{l=q}^{pq}\sum_{\substack{0\le k_1,\dots,k_p\le q\\ k_1+\ldots+k_p=q\\ k_1+2k_2+\ldots+pk_p=l}} \underbrace{\frac{q}{k_1!\cdot\ldots\cdot k_p!}}_{\ge 1}\left(\frac{1}{p}\right)^{l}(-1)!^{k_1}\cdot\ldots\cdot(p-2)!^{k_p}\\ &\ge \sum_{l=p+1}^{pq}\sum_{\substack{0\le k_1,\dots,k_p\le q\\ k_1+\ldots+k_p=q\\ k_1+2k_2+\ldots+pk_p=l}} \left(\frac{1}{p}\right)^{l}(-1)!^{k_1}\cdot\ldots\cdot(p-2)!^{k_p}\\ &\ge \sum_{l=p+1}^{pq}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}}\left(\frac{1}{p}\right)^{l}(p_1-2)!\cdot\ldots\cdot(p_q-2)!,\end{aligned}$$

since each term in the sum

$$\sum_{l=p+1}^{pq}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}}\left(\frac{1}{p}\right)^{l}(p_1-2)!\cdot\ldots\cdot(p_q-2)!$$

is included in the sum

$$\sum_{l=p+1}^{pq}\sum_{\substack{0\le k_1,\dots,k_p\le q\\ k_1+\ldots+k_p=q\\ k_1+2k_2+\ldots+pk_p=l}} \left(\frac{1}{p}\right)^{l}(-1)!^{k_1}\cdot\ldots\cdot(p-2)!^{k_p}$$

(take $p_1,\dots,p_q\in\mathbb{N}$ with $p_1+\ldots+p_q=l$ and set $k_{p_i}=1$ for $i=1,\dots,q$ and $k_j=0$ for $j\in\{1,\dots,p\}\setminus\{p_1,\dots,p_q\}$. It follows that

$$\left(\frac{1}{p}\right)^{l}(p_1-2)!\cdot\ldots\cdot(p_q-2)! = \left(\frac{1}{p}\right)^{l}(-1)!^{k_1}\cdot\ldots\cdot(p-2)!^{k_p}$$

with $k_1+\ldots+k_p=q$ and $k_1+2k_2+\ldots+pk_p=p_1+\ldots+p_q=l$).

(ii) Here we distinguish two cases. If $q=l-k+1$, then

$$\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}}(p_1-2)!\cdot\ldots\cdot(p_q-2)! = (-1)!^q = 1\le 2(l-k-1)!.$$

Now suppose that $2\le q<l-k+1$ and let $\boldsymbol{p}\in\mathbb{N}^q$ with $|\boldsymbol{p}|=l-k+1$, so that there exists $i\in\{1,\dots,q\}$ with $p_i\ge 2$. It follows that

$$\prod_{j=1}^{q}(p_j-2)!\le(p_i-2)!\prod_{\substack{j=1\\ j\ne i}}^{q}(p_j-1)!\le(l-k+1-(q-1)-2)! = (l-k-q)!,$$

since $n_1!n_2! \leq (n_1+n_2)!$. Using the fact that the number of compositions of $l-k+1$ into q parts is given by $\binom{l-k}{q-1}$ (see Iooss and Lombardi [14, Lemma A.2]), we conclude that

$$\begin{aligned}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\|\boldsymbol{p}|=l-k+1}} (p_1-2)!\cdot\ldots\cdot(p_q-2)! &\leq \binom{l-k}{q-1}(l-k-q)!\\ &= (l-k-1)!\frac{l-k}{(l-k-q+1)(q-1)!}\\ &= (l-k-1)!\Big(\frac{1}{(q-1)!}+\frac{q-1}{(l-k-q+1)(q-1)!}\Big)\\ &\leq 2(l-k-1)!.\end{aligned}$$

■

To obtain the non-recursive estimate for $\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}}$, we proceed as in Section 3.2 and impose a mutual constraint on p and δ which is slightly stricter than that given in Proposition 3.9.

Proposition 3.11. *Suppose that $\delta>0$ and $p\geq 2$ satisfy*

$$\delta p \leq \frac{\rho^2}{2\sqrt{5}\mathrm{e}a\gamma\theta}, \qquad r := \frac{\rho}{a\gamma\theta} < \frac{1}{5}.$$

The estimate

$$\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \leq \frac{11}{5}a\Big((C\delta)^{p+1}p! + \frac{1}{\mathrm{e}^{p+1}p^2}\Big),$$

holds for each $y \in \overline{B}_\delta(0) \subseteq \mathbb{C}^5$, where

$$C = \frac{2\sqrt{5}a\gamma\theta}{\rho^2}.$$

Proof. We proceed by estimating the quantities Δ_1^p, Δ_2^p and Δ_3^p appearing in inequality (3.20). First note that the assumption

$$\delta \leq \frac{1}{p}\frac{\rho^2}{2\sqrt{5}\mathrm{e}a\gamma\theta}$$

implies

$$\frac{2\sqrt{5}\delta}{\rho} \leq \frac{r}{p\mathrm{e}} \leq \frac{r}{2p} \leq \frac{1}{10p} \leq \frac{1}{20}$$

and since $q+1 \leq 2^q$ for $q \geq 2$, we obtain

$$\begin{aligned}
\Delta_1^p &= \sum_{p+1\leq q} \frac{a}{\rho^q}(\sqrt{5}\delta)^q(q+1) \\
&\leq a \sum_{p+1\leq q} \left(\frac{2\sqrt{5}\delta}{\rho}\right)^q \\
&= a\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1} \sum_{q=0}^{\infty} \left(\frac{2\sqrt{5}\delta}{\rho}\right)^q \\
&\leq a\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1} \sum_{q=0}^{\infty} \left(\frac{1}{20}\right)^q \\
&= a\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1} \frac{20}{19}.
\end{aligned}$$

To estimate the terms Δ_2^p and Δ_3^p we notice that the inequality

$$\phi_m \leq \frac{20a\gamma}{\rho^2}\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{m-2}\alpha_m, \quad m \in \{2,\dots,p\},$$

derived from Propositions 3.7 and 3.8 continues to hold for $m=1$ and $m \geq p+1$, since we have defined $\phi_1 = \sqrt{5}$, so that

$$\frac{20a\gamma}{\rho^2}\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{-1}\alpha_1 = 2\sqrt{5} \geq \sqrt{5} = \phi_1,$$

and $\phi_j = 0$ for $j \geq p+1$. Furthermore, since $\theta \geq 1$, we have that

$$\theta^{m-2} \leq \theta^{m-1}, \quad m \geq 2,$$

so that

$$\alpha_m \leq \theta^{m-1}(m-2)!, \quad m \geq 2,$$

and this inequality also continues to hold for $m=1$ since $\alpha_1 = 1$.

We therefore find that

$$
\begin{aligned}
\Delta_2^p &= \sum_{2\le q\le p} \sum_{l\ge p+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} \frac{a\delta^l}{\rho^q} \phi_{\boldsymbol{p}} \\
&= \sum_{2\le q\le p} \sum_{l=p+1}^{pq} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} \frac{a\delta^l}{\rho^q} \phi_{\boldsymbol{p}} \\
&\le \sum_{2\le q\le p} \sum_{l=p+1}^{pq} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} \frac{a\delta^l}{\rho^q} \left(\frac{20a\gamma}{\rho^2}\right)^q \left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{l-2q} \alpha_{\boldsymbol{p}} \\
&\le \sum_{2\le q\le p} \sum_{l=p+1}^{pq} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} \frac{a\delta^l}{\rho^q} (2\sqrt{5})^q \left(\frac{2\sqrt{5}a\gamma\theta}{\rho^2}\right)^{l-q} (p_1-2)!\cdot\ldots\cdot(p_q-2)! \\
&= a \sum_{2\le q\le p} r^q \sum_{l=p+1}^{pq} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} (C\delta)^l (p_1-2)!\cdot\ldots\cdot(p_q-2)!.
\end{aligned}
$$

By assumption we have that $\delta p \le 1/\mathrm{e}C$, that is $C\delta \le 1/\mathrm{e}p$. Using Lemma 3.10(i) we conclude that

$$
\begin{aligned}
\Delta_2^p &\le \frac{a}{\mathrm{e}^{p+1}} \sum_{2\le q\le p} r^q \sum_{l=p+1}^{pq} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} \left(\frac{1}{p}\right)^l (p_1-2)!\cdot\ldots\cdot(p_q-2)! \\
&\le \frac{a}{\mathrm{e}^{p+1}} \sum_{2\le q\le p} r^q \left(\sum_{j=1}^{p} \left(\frac{1}{p}\right)^j (j-2)!\right)^q \\
&= \frac{a}{\mathrm{e}^{p+1}} \sum_{2\le q\le p} r^q \left(\frac{1}{p} + \frac{1}{p^2} \underbrace{\sum_{j=2}^{p} \frac{(j-2)!}{p^{j-2}}}_{\le 1}\right)^q \\
&\le \frac{a}{\mathrm{e}^{p+1}} \sum_{2\le q\le p} r^q \left(\frac{1}{p} + \frac{p-1}{p^2}\right)^q \\
&\le \frac{a}{\mathrm{e}^{p+1}} \sum_{2\le q\le p} \left(\frac{2r}{p}\right)^q \\
&= \frac{a}{\mathrm{e}^{p+1}} \frac{4r^2}{p^2} \sum_{0\le q\le p-2} \left(\frac{2r}{p}\right)^q \\
&\le \frac{a}{p^2\mathrm{e}^{p+1}} \frac{4r^2}{1-r} \\
&\le \frac{a}{5p^2\mathrm{e}^{p+1}}
\end{aligned}
$$

where we have used the facts that $r \le 1/5$ and $2/p \le 1$.

Turning to Δ_3^p, we use the estimates for ϕ_m and α_m, $m \in \mathbb{N}$, to obtain

$$\begin{aligned}
\Delta_3^p &= \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \sqrt{5}k\phi_k \sum_{2\le q\le l-k+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} \frac{a\delta^l}{\rho^q}\phi_{\boldsymbol{p}} \\
&\le \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \sqrt{5}k\frac{20a\gamma}{\rho^2}\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{k-2}\alpha_k \\
&\qquad\qquad \times \sum_{2\le q\le l-k+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} \frac{a\delta^l}{\rho^q}\left(\frac{20a\gamma}{\rho^2}\right)^q\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{l-k+1-2q}\alpha_{\boldsymbol{p}} \\
&= \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \frac{20\sqrt{5}}{\gamma}\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{l-1} k\alpha_k \sum_{2\le q\le l-k+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} \left(\frac{\rho}{a\gamma}\right)^{q-2}\alpha_{\boldsymbol{p}}\delta^l \\
&\le \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \frac{20\sqrt{5}}{\gamma}\delta^l\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{l-1}\theta^{l-2}k(k-2)! \\
&\qquad\qquad \times \sum_{2\le q\le l-k+1} r^{q-2} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} (p_1-2)!\cdot\ldots\cdot(p_q-2)!
\end{aligned}$$

and using Lemma 3.10(ii) and the assumption $r \le 1/5$, we obtain

$$\begin{aligned}
\Delta_3^p &\le \frac{40\sqrt{5}}{\gamma} \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \delta^l\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{l-1}\theta^{l-2}k(k-2)!(l-k-1)! \sum_{0\le q\le l-k-1} r^q \\
&\le \frac{50\sqrt{5}}{\gamma} \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \delta^l\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{l-1}\theta^{l-2}k(k-2)!(l-k-1)! \\
&= \frac{25\rho^2}{a\gamma^2\theta^2} \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} (C\delta)^l k(k-2)!(l-k-1)! \\
&= \frac{25\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1} \sum_{2\le k\le p} k(k-2)! \sum_{p+1\le l\le p+k-1} (C\delta)^{l-p-1}(l-k-1)!.
\end{aligned}$$

By assumption we have that $C\delta \le 1/ep \le 1/2p$ and since $p+1 \le l \le p+k-1$ implies that $p-k \le l-k-1 \le p-2 \le p$, we find that

$$\begin{aligned}
(C\delta)^{l-p-1}(l-k-1)! &\le \left(\frac{1}{2}\right)^{l-p-1}\frac{(l-k-1)!}{p^{l-p-1}} \\
&= \left(\frac{1}{2}\right)^{l-p-1}(p-k)!\prod_{j=0}^{l-p-2}\underbrace{\frac{l-k-1-j}{p}}_{\le 1} \\
&\le \left(\frac{1}{2}\right)^{l-p-1}(p-k)!.
\end{aligned}$$

Moreover

$$\begin{aligned}\binom{p}{k}(k-1) &= \frac{p!}{k(k-2)!(p-k)!}\\ &= (p-1)\frac{p}{k}\prod_{j=0}^{k-3}\frac{p-2-j}{k-2-j}\\ &\geq p-1,\end{aligned}$$

so that

$$\begin{aligned}\Delta_3^p &\leq \frac{25\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1}\sum_{2\leq k\leq p}k(k-2)!(p-k)!\sum_{p+1\leq l\leq p+k-1}\left(\frac{1}{2}\right)^{l-p-1}\\ &\leq \frac{50\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1}\sum_{2\leq k\leq p}k(k-2)!(p-k)!\\ &= \frac{50\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1}p!\sum_{2\leq k\leq p}\binom{p}{k}^{-1}\frac{1}{k-1}\\ &\leq \frac{50\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1}p!\sum_{2\leq k\leq p}\frac{1}{p-1}\\ &= \frac{50\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1}p!.\end{aligned}$$

Using inequality (3.20), we conclude from the estimates for Δ_1^p, Δ_2^p and Δ_3^p that

$$\begin{aligned}\|\tilde{h}(y)\|_{\mathcal{X}^{t+1}_{\mathrm{sh,c}}} &\leq \left(2+\frac{10}{\theta}\right)a\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1}\frac{20}{19}+\left(1+\frac{10}{\theta}\right)\frac{a}{5p^2\mathrm{e}^{p+1}}+\frac{50\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1}p!\\ &\leq \frac{240}{19}a\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1}+\frac{11a}{5p^2\mathrm{e}^{p+1}}+\frac{50\rho^2}{a\gamma^2\theta^2}(C\delta)^{p+1}p!\\ &= \frac{240}{19}a(rC\delta)^{p+1}+\frac{11a}{5p^2\mathrm{e}^{p+1}}+50ar^2(C\delta)^{p+1}p!\\ &\leq \frac{240}{19}a\left(\frac{1}{5}\right)^3(C\delta)^{p+1}+\frac{11a}{5p^2\mathrm{e}^{p+1}}+2a(C\delta)^{p+1}p!\\ &= \frac{11}{5}a\left(\left(\underbrace{\frac{1200}{209p!}\left(\frac{1}{5}\right)^3}_{\leq 1/11}+\frac{10}{11}\right)(C\delta)^{p+1}p!+\frac{1}{\mathrm{e}^{p+1}p^2}\right)\\ &\leq \frac{11}{5}a\left((C\delta)^{p+1}p!+\frac{1}{\mathrm{e}^{p+1}p^2}\right).\end{aligned}$$

■

The final step in the proof that $\|\tilde{h}(y)\|_{\mathcal{X}^{t}_{\mathrm{sh,c}}}$ is exponentially small with respect to y is to determine an optimal order p up to which the normal form should be constructed. For this purpose we need the following proposition, which was proved by Iooss and Lombardi [14, Lemma 2.19].

Proposition 3.12. *Let $\nu > 0$. The function $f_\nu \colon \mathbb{N} \to \mathbb{R}$ defined by*

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

satisfies

$$f_\nu\left(\left\lfloor \frac{1}{\nu \mathrm{e}} \right\rfloor\right) \leq M \sqrt{\frac{\nu}{\mathrm{e}}} \mathrm{e}^{-2/\nu \mathrm{e}},$$

where $\lfloor x \rfloor$ denotes the integer part of the real number x and

$$M = \sup_{p \in \mathbb{N}} \frac{\mathrm{e}^2 p!}{p^{p+1/2} \mathrm{e}^{-p}}.$$

Now choose $\delta > 0$ and define $p_{\mathrm{opt}} \in [2, (\mathrm{e}C\delta)^{-1}]$ by

$$p_{\mathrm{opt}} = \lfloor (\mathrm{e}C\delta)^{-1} \rfloor,$$

so that p_{opt} and δ satisfy the mutual constraint given in Proposition 3.11 and

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

Using Proposition 3.12 with $\nu = C\delta$ we conclude from the estimate given in Proposition 3.11 that

$$\begin{aligned}
\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} &\leq \frac{11}{5} a \left(f_{C\delta}(p_{\mathrm{opt}}) + \frac{1}{\mathrm{e}^{p_{\mathrm{opt}}+1} p_{\mathrm{opt}}^2} \right) \\
&\leq \frac{11}{5} a \left(M \sqrt{\frac{C\delta}{\mathrm{e}}} \mathrm{e}^{-2/\mathrm{e}C\delta} + (2\mathrm{e}C\delta)^2 \mathrm{e}^{-1/\mathrm{e}C\delta} \right) \\
&= \frac{11}{5} a (\mathrm{e}C\delta)^2 \mathrm{e}^{-1/\mathrm{e}C\delta} \Bigg(\frac{M}{\mathrm{e}} \underbrace{(\mathrm{e}C\delta)^{-3/2} \mathrm{e}^{-1/\mathrm{e}C\delta}}_{\leq 1} + 4 \Bigg) \\
&\leq \frac{11}{5} a (\mathrm{e}C\delta)^2 \mathrm{e}^{-1/\mathrm{e}C\delta} \left(\frac{M}{\mathrm{e}} + 4 \right) \\
&= \frac{11}{5} a C^2 (M\mathrm{e} + 4\mathrm{e}^2) \delta^2 \mathrm{e}^{-1/\mathrm{e}C\delta}.
\end{aligned} \tag{3.21}$$

In Chapters 4 and 5 we use a perturbation argument and write $z(\xi) = p^\mu(\xi) + r(\xi)$ to construct a generalised pulse solution to the transformed equations (3.8), (3.9), where p^μ is the rescaling of either of the homoclinic solutions $p^{\varepsilon\pm}$ constructed in Section 2.2. Recall that p^μ satisfies the estimate

$$|p^\mu(\xi)| \leq c_{\mathrm{h}} \mu \mathrm{e}^{-\eta\lambda^\mu \xi} \leq c_{\mathrm{h}} \mu$$

for each $\xi \in [0,\infty)$. Therefore, to guarantee that $|y| = |(z,\mu)| \le \delta$, we define $\delta = (2c_\mathrm{h}+1)\mu_0$ and restrict the perturbation r to $\{|r| \le c_\mathrm{h}\mu_0\}$, that is (z,μ) to

$$\{|z| \le 2c_\mathrm{h}\mu_0,\ 0 \le \mu \le \mu_0\}.$$

Defining $c^\star = \left(\mathrm{e}C(2c_\mathrm{h}+1)\right)^{-1}$, we thus find from inequality (3.21) that

$$\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} \lesssim \mu^2 \mathrm{e}^{-c^\star/\mu},$$

showing that $\tilde{h}(y) = \tilde{h}^\mu(z)$ is exponentially small with respect to μ for $p = p_\mathrm{opt}$.

3.4 Further estimates

In this section we construct a corresponding estimate for the derivative $\mathrm{d}\tilde{h}[y]$ of $\tilde{h}(y)$ with respect to y. In this context we use the notation

$$\Phi_{\boldsymbol{p}}^{(\boldsymbol{l})}(y;\tilde{y}) = \left(\mathrm{d}^{l_1}\Phi^{p_1}[y](\tilde{y}),\ldots,\mathrm{d}^{l_q}\Phi^{p_q}[y](\tilde{y})\right)$$

for $q \in \mathbb{N}$, $\boldsymbol{p} \in \mathbb{N}_2^q$ and $\boldsymbol{l} \in \mathbb{N}_0^q$.

Remark 3.13. It follows from Proposition 3.5 that

$$\phi_{\boldsymbol{p}}^{(\boldsymbol{l})} = \prod_{j=1}^{q} |\mathrm{d}^{l_j}\Phi^{p_j}|_{2,p_j-l_j} \le \sqrt{5}^{|\boldsymbol{l}|} \prod_{j=1}^{q} \frac{p_j!}{(p_j-l_j)!}\phi_{\boldsymbol{p}}.$$

Now we compute an expression for $\mathrm{d}\tilde{h}[y]$ by differentiating equation (3.18). The result is

$$
\begin{aligned}
&\mathrm{d}\tilde{h}[y](\tilde{y}) \\
&= \sum_{p+1\le q} q h^q[\{y\}^{(q-1)},\tilde{y}] - \sum_{p\le q} qA_q\mu^{q-1}\tilde{\mu}\sum_{2\le k\le p}\partial_\tau\Phi^k(y) - \sum_{p\le q}A_q\mu^q\sum_{2\le k\le p}\partial_\tau\mathrm{d}\Phi^k[y](\tilde{y}) \\
&\quad - \sum_{\substack{1\le q\le p-1\\ p+1-q\le r\le p}} qA_q\mu^{q-1}\tilde{\mu}\partial_\tau\Phi^r(y) - \sum_{\substack{1\le q\le p-1\\ p+1-q\le r\le p}} A_q\mu^q\partial_\tau\mathrm{d}\Phi^r[y](\tilde{y}) \\
&\quad - \sum_{p\le q} qB_q\mu^{q-1}\tilde{\mu}\sum_{2\le k\le p}\Phi^k(y) - \sum_{p\le q}B_q\mu^q\sum_{2\le k\le p}\mathrm{d}\Phi^k[y](\tilde{y}) \\
&\quad - \sum_{\substack{1\le q\le p-1\\ p+1-q\le r\le p}} qB_q\mu^{q-1}\tilde{\mu}\Phi^r(y) - \sum_{\substack{1\le q\le p-1\\ p+1-q\le r\le p}} B_q\mu^q\mathrm{d}\Phi^r[y](\tilde{y}) \\
&\quad - \sum_{\substack{1\le q\le p-1\\ l\ge p+1}}\sum_{i=0}^{q}\sum_{\substack{(\boldsymbol{p},r)\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i-r}} P_{\mathrm{sh,c}}\Bigg(i g_1^{i,q-i}[\{y\}^{(i-1)},\tilde{y},-\Phi_{\boldsymbol{p}}(y)]\partial_\tau\Phi^r(y) \\
&\qquad\qquad\qquad + \sum_{\substack{\boldsymbol{l}\in\mathbb{N}_0^{q-i}\\ |\boldsymbol{l}|=1}} g_1^{i,q-i}[\{y\}^{(i)},-\Phi_{\boldsymbol{p}}^{(\boldsymbol{l})}(y;\tilde{y})]\partial_\tau\Phi^r(y) \\
&\qquad\qquad\qquad + g_1^{i,q-i}[\{y\}^{(i)},-\Phi_{\boldsymbol{p}}(y)]\partial_\tau\mathrm{d}\Phi^r[y](\tilde{y})\Bigg) \\
&\quad - \sum_{p\le q}\sum_{i=0}^{q} P_{\mathrm{sh,c}}\Bigg(i g_1^{i,q-i}\Big[\{y\}^{(i-1)},\tilde{y},-\Big\{\sum_{2\le k\le p}\Phi^k(y)\Big\}^{(q-i)}\Big]\sum_{2\le k\le p}\partial_\tau\Phi^k(y) \\
&\qquad + (q-i)g_1^{i,q-i}\Big[\{y\}^{(i)},-\Big\{\sum_{2\le k\le p}\Phi^k(y)\Big\}^{q-i-1}, \\
&\qquad\qquad - \sum_{2\le k\le p}\mathrm{d}\Phi^k[y](\tilde{y})\Big]\sum_{2\le k\le p}\partial_\tau\Phi^k(y) \\
&\qquad + g_1^{i,q-i}\Big[\{y\}^{(i)},-\Big\{\sum_{2\le k\le p}\Phi^k(y)\Big\}^{(q-i)}\Big]\sum_{2\le k\le p}\partial_\tau\mathrm{d}\Phi^k[y](\tilde{y})\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{\substack{2\le q\le p\\ l\ge p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} i g_2^{i,q-i}[\{y\}^{(i-1)}, \tilde{y}, -\Phi_{\boldsymbol{p}}(y)] + \sum_{\substack{\boldsymbol{l}\in\mathbb{N}_0^{q-i}\\ |\boldsymbol{l}|=1}} g_2^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}^{(\boldsymbol{l})}(y;\tilde{y})]\\
&+ \sum_{p+1\le q} \sum_{i=0}^{q} i g_2^{i,q-i}\Big[\{y\}^{(i-1)}, \tilde{y}, -\Big\{\sum_{2\le k\le p} \Phi^k(y)\Big\}^{(q-i)}\Big]\\
&\qquad\qquad + (q-i) g_2^{i,q-i}\Big[\{y\}^{(i)}, -\Big\{\sum_{2\le k\le p} \Phi^k(y)\Big\}^{q-i-1}, -\sum_{2\le k\le p} \mathrm{d}\Phi^k[y](\tilde{y})\Big]\\
&+ \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \mathrm{d}^2\Phi^k[y]\Big(\sum_{2\le q\le l-k+1} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}} f^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)], \tilde{y}\Big)\\
&+ \sum_{\substack{2\le k\le p\\ p+1\le l\le p+k-1}} \mathrm{d}\Phi^k[y]\Big(\sum_{2\le q\le l-k+1} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}} i f^{i,q-i}[\{y\}^{(i-1)}, \tilde{y}, -\Phi_{\boldsymbol{p}}(y)]\\
&\qquad\qquad\qquad\qquad + \sum_{\substack{\boldsymbol{l}\in\mathbb{N}_0^{q-i}\\ |\boldsymbol{l}|=1}} f^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}^{(\boldsymbol{l})}(y;\tilde{y})]\Big)\\
&+ \Big(\sum_{k=2}^{p} \mathrm{d}^2\Phi^k[y]\Big)\Big(\sum_{\substack{2\le q\le p\\ l\ge p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} f^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}(y)]\\
&\qquad\qquad + \sum_{p+1\le q} \sum_{i=0}^{q} f^{i,q-i}\Big[\{y\}^{(i)}, -\Big\{\sum_{2\le k\le p} \Phi^k(y)\Big\}^{(q-i)}\Big], \tilde{y}\Big)\\
&+ \Big(\sum_{k=2}^{p} \mathrm{d}\Phi^k[y]\Big)\Big(\sum_{\substack{2\le q\le p\\ l\ge p+1}} \sum_{i=0}^{q} \Big(\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} i f^{i,q-i}[\{y\}^{(i-1)}, \tilde{y}, -\Phi_{\boldsymbol{p}}(y)]\\
&\qquad\qquad + \sum_{\substack{\boldsymbol{l}\in\mathbb{N}_0^{q-i}\\ |\boldsymbol{l}|=1}} f^{i,q-i}[\{y\}^{(i)}, -\Phi_{\boldsymbol{p}}^{(\boldsymbol{l})}(y;\tilde{y})]\Big)\\
&\qquad\qquad + \sum_{p+1\le q} \sum_{i=0}^{q} i f^{i,q-i}\Big[\{y\}^{(i-1)}, \tilde{y}, -\Big\{\sum_{2\le k\le p} \Phi^k(y)\Big\}^{(q-i)}\Big]\\
&\qquad\qquad\qquad + (q-i) f^{i,q-i}\Big[\{y\}^{(i)}, -\Big\{\sum_{2\le k\le p} \Phi^k(y)\Big\}^{(q-i-1)},\\
&\qquad\qquad\qquad\qquad - \sum_{2\le k\le p} \mathrm{d}\Phi^k[y](\tilde{y})\Big]\Big). \qquad (3.22)
\end{aligned}
$$

Repeating the arguments employed for the estimate for $\|\tilde{h}(y)\|_{\mathcal{X}^t_{\mathrm{sh,c}}}$ in Section 3.3 we find, using additionally Remark 3.13 and the estimate for the second derivative of $\Phi(y)$ given in Proposition 3.9, that

$$
\begin{aligned}
&\|\mathrm{d}\tilde{h}[y]\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^t_{\mathrm{sh,c}})}\\
&\leq \sum_{p+1\leq q} q\frac{a}{\rho^q}\sqrt{5}^q\delta^{q-1} + \frac{1}{2}\sum_{p\leq q} q\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{q-1}\frac{\sqrt{5}\delta}{\theta} + \frac{1}{2}\sum_{p\leq q}\frac{a}{\rho^q}(\sqrt{5}\delta)^q\frac{10}{\theta}\\
&\quad + \frac{1}{2}\sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}} q\frac{a}{\rho^q}\sqrt{5}^{q+1}\phi_r\delta^{q-1+r} + \frac{1}{2}\sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}}\frac{a}{\rho^q}\sqrt{5}^{q+1}r\phi_r\delta^{q-1+r}\\
&\quad + \frac{1}{4}\sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&\quad + \frac{1}{4}\sum_{p\leq q}\sum_{i=0}^{q}\Bigg(i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{\sqrt{5}\delta}{\theta}\\
&\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\frac{\sqrt{5}\delta}{\theta}\\
&\qquad\qquad + \frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{10}{\theta}\Bigg)\\
&\quad + \frac{1}{4}\sum_{\substack{2\leq q\leq p\\ l\geq p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&\quad + \frac{1}{4}\sum_{p+1\leq q}\sum_{i=0}^{q} i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i} + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\\
&\quad + \sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}} 5k(k-1)\phi_k\sum_{2\leq q\leq l-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}}\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}}\\
&\quad + \sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}}\sqrt{5}k\phi_k\sum_{2\leq q\leq l-k+1}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}}(l-k+1)\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&\quad + \frac{10\sqrt{5}}{\theta\delta}\Bigg(\sum_{\substack{2\leq q\leq p\\ l\geq p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}}\frac{a\delta^l}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}} + \sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\Bigg)\\
&\quad + \frac{10}{\theta}\Bigg(\sum_{\substack{2\leq q\leq p\\ l\geq p+1}}\sum_{i=0}^{q}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&\qquad\quad + \sum_{p+1\leq q}\sum_{i=0}^{q}\Bigg(i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\\
&\qquad\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\Bigg)\Bigg)
\end{aligned}
\tag{3.23}
$$

where we still use the definition $\Phi^j = 0$ for $j \geq p+1$.

Next we proceed as in the estimate for $\tilde{h}(y)$ and simplify the right-hand side of (3.23), recalling that we have also defined $\phi_1 = \sqrt{5}$. We have that

$$\begin{aligned}
&\sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}} \Bigg(5k(k-1)\phi_k \sum_{2\leq q\leq l-k+1} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}} \frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^i\phi_{\boldsymbol{p}} \\
&\qquad\qquad + \sqrt{5}k\phi_k \sum_{2\leq q\leq l-k+1} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-k+1-i}} (l-k+1)\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\Bigg) \\
&\quad = \sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}} 5k\phi_k \sum_{2\leq q\leq l-k+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} l\frac{a\delta^{l-1}}{\rho^q}\phi_{\boldsymbol{p}}
\end{aligned}$$

and

$$\begin{aligned}
\sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}} &\geq \sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{q+1}\phi_{l-q} \\
&= \sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}} (q+r)\frac{a}{\rho^q}\sqrt{5}^{q+1}\phi_r\delta^{q-1+r}.
\end{aligned}$$

Hence

$$
\begin{aligned}
&\frac{1}{2} \sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}} q\frac{a}{\rho^q}\sqrt{5}^{q+1}\phi_r\delta^{q-1+r} + \frac{1}{2} \sum_{\substack{1\leq q\leq p-1\\ p+1-q\leq r\leq p}} \frac{a}{\rho^q}\sqrt{5}^{q+1} r\phi_r\delta^{q-1+r}\\
&\quad + \frac{1}{4} \sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}} + \frac{1}{4} \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&\quad + \frac{10\sqrt{5}}{\theta\delta} \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} \frac{a\delta^{l}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}} + \frac{10}{\theta} \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&\leq \frac{3}{4} \sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i+1}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}} + \left(\frac{1}{4}+\frac{10}{\theta}\right) \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} l\frac{a\delta^{l-1}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&\qquad + \frac{10\sqrt{5}}{\theta\delta} \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{i=0}^{q} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}_2^{q-i}\\ |\boldsymbol{p}|=l-i}} \frac{a\delta^{l}}{\rho^q}\sqrt{5}^{i+1}\phi_{\boldsymbol{p}}\\
&= \frac{3\sqrt{5}}{4} \sum_{\substack{1\leq q\leq p-1\\ l\geq p+1}} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^{q+1}\\ |\boldsymbol{p}|=l}} l\frac{a\delta^{l-1}}{\rho^q}\phi_{\boldsymbol{p}} + \left(\frac{\sqrt{5}}{4}+\frac{10\sqrt{5}}{\theta}\right) \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^{q}\\ |\boldsymbol{p}|=l}} l\frac{a\delta^{l-1}}{\rho^q}\phi_{\boldsymbol{p}}\\
&\qquad\qquad + \frac{50}{\theta} \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^{q}\\ |\boldsymbol{p}|=l}} \frac{a\delta^{l-1}}{\rho^q}\phi_{\boldsymbol{p}}\\
&\leq \left(\sqrt{5}+\frac{10\sqrt{5}+50}{\theta}\right) \sum_{\substack{2\leq q\leq p\\ l\geq p+1}} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^{q}\\ |\boldsymbol{p}|=l}} l\frac{a\delta^{l-1}}{\rho^q}\phi_{\boldsymbol{p}}.
\end{aligned}
$$

Finally we find that

$$
\begin{aligned}
&\sum_{p+1\le q} q\frac{a}{\rho^q}\sqrt{5}^q\delta^{q-1} + \frac{1}{2}\sum_{p\le q} q\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{q-1}\frac{\sqrt{5}\delta}{\theta} + \frac{1}{2}\sum_{p\le q}\frac{a}{\rho^q}(\sqrt{5}\delta)^q\frac{10}{\theta}\\
&\quad + \frac{1}{4}\sum_{p\le q}\sum_{i=0}^{q}\Bigg(i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{\sqrt{5}\delta}{\theta}\\
&\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\frac{\sqrt{5}\delta}{\theta}\\
&\qquad\qquad + \frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{10}{\theta}\Bigg)\\
&\quad + \left(\frac{1}{4}+\frac{10}{\theta}\right)\sum_{p+1\le q}\sum_{i=0}^{q}\Bigg(i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\\
&\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\Bigg)\\
&\quad + \frac{10\sqrt{5}}{\theta\delta}\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}\\
&\le \frac{3}{4}\sum_{p\le q}\sum_{i=0}^{q}\Bigg(i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{\sqrt{5}\delta}{\theta}\\
&\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\frac{\sqrt{5}\delta}{\theta}\\
&\qquad\qquad + \frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\frac{10}{\theta}\Bigg)\\
&\quad + \left(1+\frac{1}{4}+\frac{10}{\theta}\right)\sum_{p+1\le q}\sum_{i=0}^{q}\Bigg(i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\\
&\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\Bigg)\\
&\quad + \frac{10\sqrt{5}}{\theta\delta}\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}\\
&= \frac{3}{4}\sum_{p+1\le q}\sum_{i=0}^{q-1}\Bigg(i\frac{a}{\rho^{q-1}}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\\
&\qquad\qquad + (q-i)\frac{a}{\rho^{q-1}}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\Bigg)\\
&\quad + \left(1+\frac{1}{4}+\frac{10}{\theta}\right)\sum_{p+1\le q}\sum_{i=0}^{q} i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}\\
&\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^i\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i-1}\frac{10}{\theta}\\
&\quad + \frac{10\sqrt{5}}{\theta}\sum_{p+1\le q}\sum_{i=0}^{q}\frac{a}{\rho^q}\sqrt{5}^i\delta^{i-1}\left(\frac{\sqrt{5}\delta}{\theta}\right)^{q-i}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \Big(2+\frac{10}{\theta}\Big)\sum_{p+1\leq q}\sum_{i=0}^{q}\bigg(i\frac{a}{\rho^q}\sqrt{5}(\sqrt{5}\delta)^{i-1}\Big(\frac{\sqrt{5}\delta}{\theta}\Big)^{q-i}\\
&\qquad\qquad\qquad + (q-i)\frac{a}{\rho^q}(\sqrt{5}\delta)^{i}\Big(\frac{\sqrt{5}\delta}{\theta}\Big)^{q-i-1}\frac{10}{\theta}\bigg)\\
&\quad+\frac{10\sqrt{5}}{\theta}\sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}\sqrt{5}^i\delta^{i-1}\Big(\frac{\sqrt{5}\delta}{\theta}\Big)^{q-i}\\
&\leq \Big(4\sqrt{5}+\frac{20\sqrt{5}}{\theta}\Big)\sum_{p+1\leq q}\sum_{i=0}^{q} q\frac{a}{\rho^q}\sqrt{5}^q\delta^{q-1}\Big(\frac{1}{\theta}\Big)^{q-i}+\frac{10\sqrt{5}}{\theta}\sum_{p+1\leq q}\sum_{i=0}^{q}\frac{a}{\rho^q}\sqrt{5}^q\delta^{q-1}\Big(\frac{1}{\theta}\Big)^{q-i}\\
&\leq \Big(4\sqrt{5}+\frac{30\sqrt{5}}{\theta}\Big)\sum_{p+1\leq q} q(q+1)\frac{a}{\rho^q}\sqrt{5}^q\delta^{q-1}.
\end{aligned}
$$

Using the previous estimates, we find from inequality (3.23) that

$$
\|\mathrm{d}\tilde{h}[y]\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^t_{\mathrm{sh,c}})} \leq \Big(4\sqrt{5}+\frac{30\sqrt{5}}{\theta}\Big)\hat{\Delta}_1^p+\Big(\sqrt{5}+\frac{10\sqrt{5}+50}{\theta}\Big)\hat{\Delta}_2^p+\sqrt{5}\hat{\Delta}_3^p, \tag{3.24}
$$

where

$$
\begin{aligned}
\hat{\Delta}_1^p &= \sum_{p+1\leq q} q(q+1)\frac{a}{\rho^q}\sqrt{5}^q\delta^{q-1},\\
\hat{\Delta}_2^p &= \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l}} l\frac{a\delta^{l-1}}{\rho^q}\phi_{\boldsymbol{p}},\\
\hat{\Delta}_3^p &= \sum_{\substack{2\leq k\leq p\\ p+1\leq l\leq p+k-1}}\sqrt{5}k\phi_k\sum_{2\leq q\leq l-k+1}\sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q\\ |\boldsymbol{p}|=l-k+1}} l\frac{a\delta^{l-1}}{\rho^q}\phi_{\boldsymbol{p}}.
\end{aligned}
$$

Using arguments similar to those in the proof of Proposition 3.11 we now establish a non-recursive estimate for $\mathrm{d}\tilde{h}$ involving the order p of the transformation. Here we impose a mutual constraint on δ and p which is stronger than that given in Proposition 3.11 and make use of the fact that $\Phi_2=\Phi_3=0$.

Proposition 3.14. *Suppose that $\delta>0$ and $p\geq 4$ satisfy*

$$
\delta p \leq \frac{\rho^2}{4\sqrt{5}\mathrm{e}a\gamma\theta},\qquad r:=\frac{\rho}{a\gamma\theta}<\frac{1}{5}.
$$

The estimate

$$
\|\mathrm{d}\tilde{h}[y]\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^t_{\mathrm{sh,c}})} \leq \frac{15}{\delta}a\Big((C\delta)^{p+1}p!+\frac{1}{\mathrm{e}^{p+1}p^2}\Big),
$$

holds for each $y\in\overline{B}_\delta(0)\subseteq\mathbb{C}^5$, where

$$
C=\frac{4\sqrt{5}a\gamma\theta}{\rho^2}.
$$

Proof. We proceed by estimating the quantities $\hat{\Delta}_1^p$, $\hat{\Delta}_2^p$ and $\hat{\Delta}_3^p$ appearing in inequality (3.24). First note that the assumption

$$\delta \leq \frac{1}{p}\frac{\rho^2}{4\sqrt{5}a\gamma\theta}$$

implies that

$$\frac{2\sqrt{5}\delta}{\rho} \leq \frac{r}{2\mathrm{e}p} \leq \frac{r}{4p} \leq \frac{1}{20p} \leq \frac{1}{40}$$

(where we have estimated $p \geq 2$) and since $(q+1)q \leq 2^q$ for $q \geq 5$, we obtain

$$\begin{aligned}
\hat{\Delta}_1^p &= \sum_{p+1\leq q} \frac{a}{\rho^q} q(q+1)\sqrt{5}^q \delta^{q-1} \\
&\leq \frac{a}{\delta} \sum_{p+1\leq q} \left(\frac{2\sqrt{5}\delta}{\rho}\right)^q \\
&\leq \frac{40a}{39\delta}\left(\frac{2\sqrt{5}\delta}{\rho}\right)^{p+1}.
\end{aligned}$$

Turning to $\hat{\Delta}_2^p$ and $\hat{\Delta}_3^p$, we estimate $l \leq 2^l$, $l \in \mathbb{N}$, and proceed as in the proof of Proposition 3.11 to obtain

$$\begin{aligned}
\hat{\Delta}_2^p &\leq \frac{a}{\delta} \sum_{2\leq q\leq p} \sum_{l\geq p+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q \\ |\boldsymbol{p}|=l}} \left(\frac{4\sqrt{5}a\gamma\theta\delta}{\rho^2}\right)^l \left(\frac{\rho}{a\gamma\theta}\right)^q (p_1-2)! \cdot \ldots \cdot (p_q-2)! \\
&= \frac{a}{\delta} \sum_{2\leq q\leq p} r^q \sum_{l\geq p+1} \sum_{\substack{\boldsymbol{p}\in\mathbb{N}^q \\ |\boldsymbol{p}|=l}} (C\delta)^l (p_1-2)! \cdot \ldots \cdot (p_q-2)! \\
&\leq \frac{a}{\delta}\frac{4r^2}{\mathrm{e}^{p+1}p^2}\frac{1}{1-r}
\end{aligned}$$

and

$$\begin{aligned}
\hat{\Delta}_3^p &\leq \frac{50\sqrt{5}}{\gamma} \sum_{\substack{2\leq k\leq p \\ p+1\leq l\leq p+k-1}} l\delta^{l-1}\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^{l-1} \theta^{l-2} k(k-2)!(l-k-1)! \\
&= \frac{25ar^2}{\delta} \sum_{\substack{2\leq k\leq p \\ p+1\leq l\leq p+k-1}} l\left(\frac{2\sqrt{5}a\gamma}{\rho^2}\right)^l k(k-2)!(l-k-1)! \\
&\leq \frac{25ar^2}{\delta} \sum_{\substack{2\leq k\leq p \\ p+1\leq l\leq p+k-1}} (C\delta)^l k(k-2)!(l-k-1)! \\
&\leq \frac{50ar^2}{\delta}(C\delta)^{p+1}p!.
\end{aligned}$$

It thus follows from inequality (3.24) that

$$\begin{aligned}
\|\mathrm{d}\tilde{h}[y]\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^t_{\mathrm{sh,c}})} &\leq \Big(4\sqrt{5}+\frac{30\sqrt{5}}{\theta}\Big)\frac{40a}{39\delta}\Big(\frac{2\sqrt{5}\delta}{\rho}\Big)^{p+1} \\
&\qquad + \Big(\sqrt{5}+\frac{10\sqrt{5}+50}{\theta}\Big)\frac{a}{\delta}\frac{4r^2}{\mathrm{e}^{p+1}p^2}\frac{1}{1-r}+\frac{50\sqrt{5}ar^2}{\delta}(C\delta)^{p+1}p! \\
&\leq \frac{3080}{39}\frac{a}{\delta}\Big(\frac{2\sqrt{5}\delta}{\rho}\Big)^{p+1}+\frac{300r^2}{\mathrm{e}^{p+1}p^2}\frac{1}{1-r}\frac{a}{\delta}+50\sqrt{5}r^2\frac{a}{\delta}(C\delta)^{p+1}p! \\
&= \frac{a}{\delta}\Big(\frac{3080}{39}\Big(\frac{rC\delta}{2}\Big)^{p+1}+\frac{300r^2}{\mathrm{e}^{p+1}p^2(1-r)}+50\sqrt{5}r^2(C\delta)^{p+1}p!\Big) \\
&\leq \frac{a}{\delta}\Big(\frac{3080}{39}\Big(\frac{1}{10}\Big)^3(C\delta)^{p+1}+15\frac{1}{\mathrm{e}^{p+1}p^2}+2\sqrt{5}(C\delta)^{p+1}p!\Big) \\
&\leq \frac{15}{\delta}a\Big(\Big(\underbrace{\frac{46200}{39p!}\Big(\frac{1}{10}\Big)^3+\frac{2\sqrt{5}}{15}}_{\leq 2/3}\Big)(C\delta)^{p+1}p!+\frac{1}{\mathrm{e}^{p+1}p^2}\Big) \\
&\leq \frac{15}{\delta}a\Big((C\delta)^{p+1}p!+\frac{1}{\mathrm{e}^{p+1}p^2}\Big).
\end{aligned}$$

■

Using Proposition 3.12 and arguing as before, we find from Proposition 3.14 that

$$\|\mathrm{d}\tilde{h}[y]\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^t_{\mathrm{sh,c}})} \leq 15aC^2(M\mathrm{e}+4\mathrm{e}^2)\mathrm{e}^{-1/\mathrm{e}C\delta}, \quad y\in\overline{B}_\delta(0),$$

for $p=p_{\mathrm{opt}}$ and with $\delta=(2c_{\mathrm{h}}+1)\mu_0$ and $c^\star=\Big(\mathrm{e}C(2c_{\mathrm{h}}+1)\Big)^{-1}$. Restricting $y=(z,\mu)$ to

$$\{|z|\leq 2c_{\mathrm{h}}\mu_0, 0\leq\mu\leq\mu_0\},$$

we conclude that

$$\|\mathrm{d}\tilde{h}[y]\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^t_{\mathrm{sh,c}})} \lesssim \mu\mathrm{e}^{-c^\star/\mu}.$$

We conclude this chapter by stating estimates for the transformed nonlinearities $\tilde{F}_1$, $\tilde{F}_2$, $\tilde{g}_1$ and $\tilde{g}_2$ appearing in equations (3.8), (3.9). Here we make use of the fact that $p\geq 4$ with $\Phi^2=\Phi^3=0$ and the original estimate (3.7) for h^μ so that

$$\begin{aligned}
\Big\|\sum_{2\leq k\leq p}\Phi^k(y)\Big\|_{\mathcal{X}^{t+1}_{\mathrm{sh,c}}} &= \mathcal{O}(|z|^2|y|^2), \\
\Big\|\sum_{2\leq k\leq p}\mathrm{d}\Phi^k[y]\Big\|_{\mathcal{L}(\mathbb{C}^5;\mathcal{X}^{t+1}_{\mathrm{sh,c}})} &= \mathcal{O}(|y|^3).
\end{aligned}$$

We return to the original notation by writing

$$\begin{aligned}
\tilde{f}_1^\mu(z) &= \tilde{f}_1(y), \quad \tilde{f}_2^\mu(z,q)=\tilde{f}_2(y,q), \\
\tilde{g}_j^\mu(z,q) &= \tilde{g}_j(y,q), \quad j=1,2.
\end{aligned}$$

The definitions of $\tilde{f}_1$, $\tilde{f}_2$, $\tilde{g}_1$ and $\tilde{g}_2$, the estimates (3.3)–(3.6) and the facts that

$$\tilde{g}_1^{1;1,0}(z,0) = g_1^{1;1,0}(z,0), \quad \tilde{g}_2^{1;1,1}(z,Z) = g_1^{1;1,1}(z,Z)$$

yield the following result which is included in the statement of Theorem 2.16.

Proposition 3.15. *The transformed nonlinearities $\tilde{f}_1$, $\tilde{f}_2$, $\tilde{g}_1$ and $\tilde{g}_2$ satisfy the estimates*

$$\begin{aligned}
|\tilde{f}_1^\mu(z)| &= \mathcal{O}(\mu^2|z| + |z|^3),\\
|\tilde{f}_2^\mu(z,q)| &= \mathcal{O}(|z||(z,\mu)|^2\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}} + \|q\|^2_{\mathcal{X}^t_{\mathrm{sh,c}}}),\\
\|\tilde{g}_1^\mu(z,q)\|_t &= \mathcal{O}(\mu\|(z,q)\|_{\mathcal{X}_{\mathrm{wh}}\times\mathcal{X}^t_{\mathrm{sh,c}}}),\\
\|\tilde{g}_2^\mu(z,q)\|_{\mathcal{X}^t_{\mathrm{sh,c}}} &= \mathcal{O}(\mu|z|\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu^2\|q\|^2_{\mathcal{X}^t_{\mathrm{sh,c}}} + \mu\|q\|^3_{\mathcal{X}^t_{\mathrm{sh,c}}} + |z|^3\|q\|_{\mathcal{X}^t_{\mathrm{sh,c}}})
\end{aligned}$$

for $t > \frac{1}{2}$, with

$$P_{\mathrm{c}}\Big(\tilde{g}_1^{1;1,0}(z,0)Z_\tau\Big) = P_{\mathrm{c}}\tilde{g}_2^{1;1,1}(z,Z) = 0$$

and

$$|\tilde{f}_1^\mu(z) - f_1^\mu(z)| = \mathcal{O}(|z|^3|(z,\mu)|).$$

4 Existence theory

In this chapter we construct reversible solutions to the transformed equations (2.53), (2.54) whose pointwise distance to the approximate pulse p^μ does not exceed $\mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$ (see Figure 4.1). To this end we construct exponentially small reversible solutions to equations (2.55)–(2.57), meaning reversible solutions which are bounded by $\mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$. Here we use the reformulations (2.58), (2.59) of equations (2.55), (2.56) as integral equations.

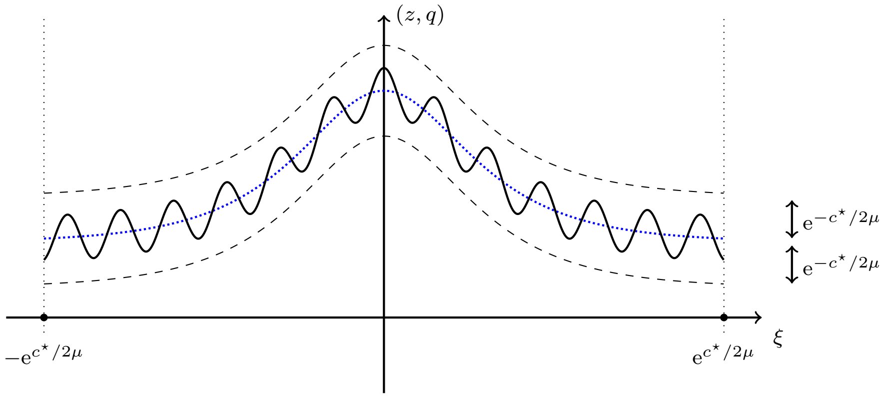

Figure 4.1: A generalised pulse solution to equations (2.53), (2.54) lies in an exponentially thin tubular neighbourhood of $(p^\mu, 0)$ on the exponentially large interval $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$.

Consider the following iteration scheme for equations (2.57)–(2.59), the convergence of which we prove in Section 4.2 below. Suppose that $Z_{(0)}$, $r_{(0)}$ and $w_{(0)}$ are appropriately chosen initial values (see below) and let

$$\begin{aligned} Z_{(j+1)}(\xi) = &\int_{-\mathrm{e}^{-c^\star/2\mu}}^{\xi} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{ss}}^\mu \Big(G_1^\mu(Z_{(j)}, p^\mu + \mathcal{E}_{\mathrm{wh}} r_{(j)}, \mathcal{E}_{\mathrm{c}} w_{(j)})(t) + G_2^\mu(\mathcal{E}_{\mathrm{wh}} r_{(j)})(t) \Big)\,\mathrm{d}t \\ &- \int_{\xi}^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{su}}^\mu \Big(G_1^\mu(Z_{(j)}, p^\mu + \mathcal{E}_{\mathrm{wh}} r_{(j)}, \mathcal{E}_{\mathrm{c}} w_{(j)})(t) + G_2^\mu(\mathcal{E}_{\mathrm{wh}} r_{(j)})(t) \Big)\,\mathrm{d}t, \end{aligned}$$

$$\begin{aligned} r_{(j+1)}(\xi) = \sum_{i=1}^{2} &\Big(\int_0^\xi \Big\langle \Big(f_3^\mu(r_{(j)})(t) + f_4^\mu(Z_{(j)}, r_{(j)}, w_{(j)})(t) \Big), S_i^*(t) \Big\rangle\,\mathrm{d}t \Big) S_i(\xi) \\ &- \Big(\int_\xi^\infty \Big\langle \Big(f_3^\mu(r_{(j)})(t) + f_4^\mu(Z_{(j)}, r_{(j)}, w_{(j)})(t) \Big), U_i^*(t) \Big\rangle\,\mathrm{d}t \Big) U_i(\xi), \end{aligned}$$

and $w_{(j+1)}$ be the unique solution of the initial value problem

$$\begin{aligned}\partial_\xi w_{(j+1)} &= L_{\mathrm{c}}^\mu w_{(j)} + P_{\mathrm{c}}\Big(\hat{g}_1^\mu(Z_{(j)}, p^\mu + r_{(j)}, w_{(j)})\partial_\tau w_{(j+1)}\Big) + \mu b(p^\mu + r_{(j)}, w_{(j)}) \\ &\qquad + \hat{g}_2^\mu(Z_{(j)}, p^\mu + r_{(j)}, w_{(j)}) + \hat{h}^\mu(r_{(j)}), \\ w_{(j+1)}(0) &= w_{(0)}(0)\end{aligned}$$

for fixed $j \in \mathbb{N}_0$. This iteration scheme yields a family of exponentially small solutions to equations (2.55)–(2.57) parameterised by the initial values $Z_{(0)}$, $r_{(0)}$ and $w_{(0)}$ on a *local centre-stable manifold* $W_{\mathrm{loc}}^{\mathrm{cs}}$ (see Definition 4.27); in this chapter and in Chapter 5 below we make use of familiar dynamical systems terminology regarding invariant manifolds (see Kelley [16]). An exponentially small solution to the reformulated system (2.55)–(2.57) corresponds to a solution of equations (2.53), (2.54) which is exponentially close to p^μ (see above).

The above definition of the iterate $w_{(j+1)}$ requires the solvability of an associated linear Cauchy problem. The relevant theory is given in Section 4.1 and uses standard methods for symmetrisable hyperbolic systems as presented by Benzoni-Gavage & Serre [1, Chapter 2], who consider the class of equations

$$\partial_\xi u + A(\xi,\tau)\partial_\tau u = B(\xi,\tau)u + f(\xi,\tau)$$

with coefficients A and B which depend smoothly on ξ and τ. It is necessary to extend their methods by treating coefficients with values in finite-order Sobolev spaces and replacing the function B by a general linear operator acting on the function u.

4.1 The linear Cauchy problem

In this section we study the linear initial-value problem

$$Lu = f, \tag{4.1}$$
$$u(0) = h \tag{4.2}$$

for fixed $f \in C\Big([0,T]; H_{\mathrm{per}}^t(0,2\pi;\mathbb{R}^n)\Big)$, $h \in H_{\mathrm{per}}^{t+1}(0,2\pi;\mathbb{R}^n)$, and

$$L = \partial_\xi + A\partial_\tau + \mathcal{L},$$

where A, $\mathcal{L}$ are respectively matrix- and operator-valued functions (see below). Here t is an integer with $t > \frac{5}{2}$ and ∂_ξ, ∂_τ denote respectively the partial derivatives with respect to the 'temporal' variable $\xi \in [0,T]$ and the periodic 'spatial' variable $\tau \in [0,2\pi]$. Our aim is to construct for arbitrary $T > 0$ a unique solution

$$u \in C\Big([0,T]; H_{\mathrm{per}}^t(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H_{\mathrm{per}}^{t-1}(0,2\pi;\mathbb{R}^n)\Big)$$

to the system (4.1), (4.2) in the case

$$A \in C\Big([0,T]; H_{\mathrm{per}}^t(0,2\pi;\mathbb{R}^{n\times n})\Big)$$

and

$$\mathcal{L} \in C\Big([0,T]; \mathcal{L}\big(H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)\Big)$$

for all $p \in \mathbb{N}$ under the following assumptions on A.

Assumptions 4.1. The matrix-valued function A is *symmetrisable*, that is there exists a symmetric, matrix-valued function

$$S \in C\big([0,T]; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^{n\times n})\big) \cap C^1\big([0,T]; H^{t-1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^{n\times n})\big)$$

such that SA is symmetric and S is uniformly positive definite, that is

$$\beta|v|^2 \leq \langle S(\xi,\tau)v, v\rangle \leq \frac{1}{\beta}|v|^2, \quad v \in \mathbb{R}^n \setminus \{0\},$$

uniformly over $(\xi,\tau) \in [0,T] \times [0,2\pi]$ for some constant $\beta > 0$.

Remark 4.2.

(i) Let $p \in \mathbb{N}$. The H^p-adjoint operator-valued function $\mathcal{L}^*_p$ defined via the duality pairing

$$\langle \mathcal{L}(\xi)u, v\rangle_{H^p,H^{-p}} = \langle u, \mathcal{L}^*_p(\xi)v\rangle_{H^p,H^{-p}}$$

satisfies

$$\mathcal{L}^*_p \in C\Big([0,T]; \mathcal{L}\big(H^{-p}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)\Big),$$

since $\|\mathcal{L}^*_p(\xi)\|_{\mathcal{L}(H^{-p}_{\mathrm{per}})} \leq \|\mathcal{L}(\xi)\|_{\mathcal{L}(H^p_{\mathrm{per}})}$ for each $\xi \in [0,T]$.

(ii) In the setting of Benzoni-Gavage & Serre [1, Chapter 2] we have that

$$\mathcal{L}(\xi)u = B(\xi)u$$

for a matrix-valued function

$$B \in C\big([0,T]; C^\infty_{\mathrm{per}}([0,2\pi];\mathbb{R}^{n\times n})\big).$$

It follows that

$$\mathcal{L} \in C\Big([0,T]; \mathcal{L}\big(H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)\Big)$$

for each $p \in \mathbb{N}$. The theory presented below also applies to the weaker case

$$B \in C\big([0,T]; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^{n\times n})\big)$$

if we consider an appropriate regularisation of B (see Lemma 4.11 and Remark 4.15).

Preliminary estimates

We gather some technical results and *a priori* estimates which are used in the following theory. We first establish an L^2-product estimate for the derivatives of two functions. For its proof we need the following special case of the Gagliardo-Nirenberg inequality for bounded domains (see Friedman [7, Part 1, Theorem 10.1]).

Proposition 4.3. *Let $p \in \mathbb{N} \setminus \{1\}$ and $g \in H^p_{\mathrm{per}}(0, 2\pi; \mathbb{R})$. The estimate*

$$\|\partial_\tau^j g\|_{L^{2p/j}} \lesssim \|g\|_\infty^{1-j/p} \|g\|_p^{j/p}$$

holds for each $j \in \mathbb{N}$ with $j < p$.

Corollary 4.4. *Let $p \in \mathbb{N}$ and $g_1, g_2 \in H^p_{\mathrm{per}}(0, 2\pi; \mathbb{R})$. The estimate*

$$\|\partial_\tau^j g_1 \partial_\tau^k g_2\|_0 \lesssim \|g_1\|_\infty \|g_2\|_p + \|g_1\|_p \|g_2\|_\infty$$

holds for each $j, k \in \mathbb{N}_0$ with $j + k = p$.

Proof. We suppose that j and k are both non-zero since otherwise the result follows from Hölder's inequality.

Using Hölder's inequality, Proposition 4.3 and Young's inequality we find that

$$\begin{aligned}
\|\partial_\tau^j g_1 \partial_\tau^k g_2\|_0 &\leq \|\partial_\tau^j g_1\|_{L^{2p/j}} \|\partial_\tau^k g_2\|_{L^{2p/k}} \\
&\lesssim \|g_1\|_\infty^{1-j/p} \|g_1\|_t^{j/p} \|g_2\|_\infty^{1-k/p} \|g_2\|_t^{k/p} \\
&= \big(\|g_1\|_\infty \|g_2\|_p\big)^{k/p} \big(\|g_1\|_p \|g_2\|_\infty\big)^{j/p} \\
&\lesssim \|g_1\|_\infty \|g_2\|_p + \|g_1\|_p \|g_2\|_\infty.
\end{aligned}$$

■

For later purposes we rather consider products of matrix- and vector-valued functions. In the proof of the following Moser-type inequality we apply the above estimates componentwise.

Lemma 4.5 (Moser estimate)**.** *The estimate*

$$\|\partial_\tau^p (Bg) - B\partial_\tau^p g\|_0 \lesssim \|\partial_\tau B\|_\infty \|g\|_{p-1} + \|\partial_\tau B\|_{p-1} \|g\|_\infty$$

holds for each $p \in \mathbb{N} \setminus \{1\}$ and

$$B \in H^p(0, 2\pi; \mathbb{R}^{n\times n}), \quad g \in H^p(0, 2\pi; \mathbb{R}^n).$$

Proof. It follows from the formula

$$\partial_\tau^p (Bg) = \sum_{k+l=p} \binom{p}{l} \partial_\tau^l B \partial_\tau^k g$$

that

$$\partial_\tau^p (Bg) - B\partial_\tau^p g = \sum_{\substack{l+k=p \\ l\neq 0}} \binom{p}{l} \partial_\tau^l B \partial_\tau^k g = \sum_{j+k=p-1} \binom{p}{j+1} \partial_\tau^j \partial_\tau B \partial_\tau^k g.$$

Applying Corollary 4.4 (with p replaced by $p-1$) to each term in the sum on the right hand side, we conclude that

$$\begin{aligned}\|\partial_\tau^p(Bg) - B\partial_\tau^p g\|_0 &\le \sum_{j+k=p-1} \binom{p}{j+1} \|\partial_\tau^j \partial_\tau B \partial_\tau^k g\|_0 \\ &\lesssim \|\partial_\tau B\|_\infty \|g\|_{p-1} + \|\partial_\tau B\|_{p-1}\|g\|_\infty.\end{aligned}$$

■

In order to guarantee that products of ξ-dependent matrix- and vector-valued functions lie in the correct function spaces we need the following result for bilinear expressions of respectively continuous and square-integrable functions.

Proposition 4.6. *Let X, Y and Z be Banach spaces and M be a bounded, bilinear mapping $X \times Y \to Z$. If $\phi \in C([0,T];X)$ and $\psi \in L^2(0,T;Y)$, then $M(\phi(\cdot),\psi(\cdot)) \in L^2(0,T;Z)$.*

Proof. Observe that

$$\int_0^T \|M(\phi(\xi),\psi(\xi))\|_Z^2 \,\mathrm{d}\xi \lesssim \int_0^T \|\phi(\xi)\|_X^2 \|\psi(\xi)\|_Y^2 \,\mathrm{d}\xi \le \|\phi\|_{C([0,T];X)}^2 \|\psi\|_{L^2(0,T;Y)}^2,$$

from which it follows that $M(\phi(\cdot),\psi(\cdot)) \in L^2(0,T;Z)$. ■

We also need the following product estimate for products of functions in $H^{-p}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$ and $H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$.

Proposition 4.7. *The formula*

$$(C,v) \mapsto Cv$$

defines a bounded, bilinear mapping $H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^{n\times n}) \times H^{-p}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n) \to H^{-p}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$ for all $p \in \mathbb{N}$.

Proof. The estimate

$$\begin{aligned}\|Cv\|_{-p} &= \sup_{\|\phi\|_p=1} |\langle Cv,\phi\rangle_{H^{-p},H^p}| \\ &= \sup_{\|\phi\|_p=1} |\langle v, C^\top\phi\rangle_{H^{-p},H^p}| \\ &\le \sup_{\|\phi\|_p=1} \|v\|_{-p}\|C^\top\phi\|_p\end{aligned}$$

implies that

$$\|Cv\|_{-p} \lesssim \|C\|_p \|v\|_{-p}.$$

■

Finally, we need the following version of Gronwall's inequality (see Sogge [19, Lemma 3.3]).

Proposition 4.8 (Gronwall's inequality)**.** *Let v, a and b be non-negative, continuous functions $[0,T] \to \mathbb{R}$, where a is monotonically increasing. Suppose further that*

$$v(\xi) \le a(\xi) + \int_0^\xi b(r)v(r)\,\mathrm{d}r$$

for all $\xi \in [0,T]$. The estimate

$$v(\xi) \le a(\xi) \exp\Big(\int_0^\xi b(r)\,\mathrm{d}r\Big)$$

holds for all $\xi \in [0,T]$.

The next result is an *a priori* estimate for a solution to the Cauchy problem (4.1), (4.2).

Lemma 4.9. *Let $p \in \{t-1, t\}$, $\xi_0 \in [0,T]$ and*

$$u \in C\Big([0,T]; H^{p+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

The estimate

$$\|u(\xi)\|_p^2 \lesssim \|u(\xi_0)\|_p^2 + \int_{\xi_0}^\xi \|(Lu)(r)\|_p^2\,\mathrm{d}r$$

holds for all $\xi \in [\xi_0, T]$.

Proof. Suppose that $k \in \mathbb{N}_0$ with $k \le p$. We note that by definition

$$\partial_\xi u(\xi) = Lu(\xi) - A(\xi)\partial_\tau u(\xi) - \mathcal{L}(\xi)u(\xi)$$

and compute

$$\begin{aligned}
&\partial_\xi \langle S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle \\
&\quad = \langle \partial_\xi S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle + 2\langle S(\xi)\partial_\tau^k \partial_\xi u(\xi), \partial_\tau^k u(\xi)\rangle \\
&\quad = \langle \partial_\xi S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle + 2\langle S(\xi)\partial_\tau^k (Lu(\xi)), \partial_\tau^k u(\xi)\rangle \\
&\qquad - 2\langle S(\xi)\partial_\tau^k (A(\xi)\partial_\tau u(\xi)) - S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi), \partial_\tau^k u(\xi)\rangle \\
&\qquad - 2\langle S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi), \partial_\tau^k u(\xi)\rangle - 2\langle S(\xi)\partial_\tau^k (\mathcal{L}(\xi)u(\xi)), \partial_\tau^k u(\xi)\rangle.
\end{aligned}$$

Integrating with respect to τ over one period yields

$$\begin{aligned}
&\partial_\xi \int_0^{2\pi} \langle S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\quad = 2\int_0^{2\pi} \langle S(\xi)\partial_\tau^k (Lu)(\xi), \partial_\tau^k u(\xi)\rangle + \int_0^{2\pi} \langle \partial_\xi S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\qquad - 2\int_0^{2\pi} \langle S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\qquad - 2\int_0^{2\pi} \langle S(\xi)\partial_\tau^k (A(\xi)\partial_\tau u(\xi)) - S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\qquad - 2\int_0^{2\pi} \langle S(\xi)\partial_\tau^k (\mathcal{L}(\xi)u(\xi)), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau.
\end{aligned}$$

Since the matrix SA is symmetric, integration by parts yields

$$\begin{aligned}
&\int_0^{2\pi} \langle S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\quad = \int_0^{2\pi} \langle \partial_\tau^{k+1} u(\xi), S(\xi)A(\xi)\partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\quad = -\int_0^{2\pi} \Big\langle \partial_\tau^k u(\xi), \partial_\tau\Big(S(\xi)A(\xi)\partial_\tau^k u(\xi)\Big)\Big\rangle \,\mathrm{d}\tau \\
&\quad = -\int_0^{2\pi} \Big\langle \partial_\tau^k u(\xi), \Big(\partial_\tau S(\xi)A(\xi) + S(\xi)\partial_\tau A(\xi)\Big)\partial_\tau^k u(\xi)\Big\rangle \,\mathrm{d}\tau \\
&\qquad - \int_0^{2\pi} \langle \partial_\tau^k u(\xi), S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi)\rangle \,\mathrm{d}\tau,
\end{aligned}$$

so that

$$\begin{aligned}
&2\int_0^{2\pi} \langle S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\quad = -\int_0^{2\pi} \Big\langle \partial_\tau^k u(\xi), \Big(\partial_\tau S(\xi)A(\xi) + S(\xi)\partial_\tau A(\xi)\Big)\partial_\tau^k u(\xi)\Big\rangle \,\mathrm{d}\tau
\end{aligned}$$

and hence

$$\begin{aligned}
&\partial_\xi \int_0^{2\pi} \langle S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \\
&\quad \leq 2\int_0^{2\pi} |\langle S(\xi)\partial_\tau^k (Lu)(\xi), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau + \int_0^{2\pi} |\langle \partial_\xi S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau \\
&\qquad + \int_0^{2\pi} \Big|\Big\langle \Big(\partial_\tau S(\xi)A(\xi) + S(\xi)\partial_\tau A(\xi)\Big)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\Big\rangle\Big| \,\mathrm{d}\tau \\
&\qquad + 2\int_0^{2\pi} |\langle S(\xi)\partial_\tau^k (A(\xi)\partial_\tau u(\xi)) - S(\xi)A(\xi)\partial_\tau^{k+1} u(\xi), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau \\
&\qquad + 2\int_0^{2\pi} |\langle S(\xi)\partial_\tau^k (\mathcal{L}(\xi)u(\xi)), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau.
\end{aligned}$$

An application of the Cauchy-Schwarz and Young's inequalities shows that

$$\begin{aligned}2\int_0^{2\pi} |\langle S(\xi)\partial_\tau^k(Lu)(\xi), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau &\le \frac{2}{\beta}\int_0^{2\pi} |\partial_\tau^k(Lu)(\xi)||\partial_\tau^k u(\xi)| \,\mathrm{d}\tau\\ &\le \frac{1}{\beta}\|\partial_\tau^k(Lu)(\xi)\|_0^2 + \frac{1}{\beta}\|\partial_\tau^k u(\xi)\|_0^2\\ &\le \frac{1}{\beta}\|(Lu)(\xi)\|_p^2 + \frac{1}{\beta}\|u(\xi)\|_p^2\end{aligned}$$

as well as

$$\begin{aligned}2\int_0^{2\pi} |\langle S(\xi)\partial_\tau^k(\mathcal{L}(\xi)u(\xi)), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau &\le \frac{1}{\beta}\|\mathcal{L}(\xi)u(\xi)\|_p^2 + \frac{1}{\beta}\|u(\xi)\|_p^2\\ &\le \frac{1}{\beta}\Big(1 + \|\mathcal{L}(\xi)\|_{\mathcal{L}(H^p_{\mathrm{per}})}^2\Big)\|u(\xi)\|_p^2.\end{aligned}$$

We additionally use the Moser estimate (Lemma 4.5) to obtain

$$\begin{aligned}&\int_0^{2\pi} |\langle S(\xi)\partial_\tau^k(A(\xi)\partial_\tau u(\xi)) - S(\xi)A(\xi)\partial_\tau^{k+1}u(\xi), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau\\ &\quad\le \frac{1}{\beta}\|\partial_\tau^k(A(\xi)\partial_\tau u(\xi)) - A(\xi)\partial_\tau^{k+1}u(\xi)\|_0^2 + \frac{1}{\beta}\|\partial_\tau^k u(\xi)\|_0^2\\ &\quad\le \frac{1}{\beta}\Big(\|\partial_\tau u(\xi)\|_\infty\|\partial_\tau A(\xi)\|_{k-1} + \|\partial_\tau u\|_{k-1}\|\partial_\tau A(\xi)\|_\infty\Big)^2 + \frac{1}{\beta}\|u(\xi)\|_p^2\\ &\quad\le \frac{1}{\beta}\Big(\|\partial_\tau u(\xi)\|_{p-1}\|A(\xi)\|_p + \|u\|_p\|\partial_\tau A(\xi)\|_{p-1}\Big)^2 + \frac{1}{\beta}\|u(\xi)\|_p^2\\ &\quad\le \frac{1}{\beta}\Big(1 + 4\|A(\xi)\|_p^2\Big)\|u(\xi)\|_p^2\end{aligned}$$

(note that $H^{p-1}_{\mathrm{per}}(0, 2\pi; \mathbb{R}) \hookrightarrow L^\infty_{\mathrm{per}}(0, 2\pi; \mathbb{R})$, since $p - 1 \ge t - 2 > \frac{1}{2}$). Moreover

$$\int_0^{2\pi} |\langle \partial_\xi S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle| \,\mathrm{d}\tau \le \|\partial_\xi S(\xi)\|_\infty\|u(\xi)\|_p^2$$

and

$$\begin{aligned}&\int_0^{2\pi} \Big|\Big\langle\Big(\partial_\tau S(\xi)A(\xi) + S(\xi)\partial_\tau A(\xi)\Big)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\Big\rangle\Big| \,\mathrm{d}\tau\\ &\quad\le \|\partial_\tau S(\xi)A(\xi) + S(\xi)\partial_\tau A(\xi)\|_\infty\|u(\xi)\|_p^2.\end{aligned}$$

Altogether we conclude that

$$\begin{aligned}\partial_\xi\int_0^{2\pi} \langle S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau &\le \frac{1}{\beta}\|(Lu)(\xi)\|_p^2 + \frac{1}{\beta}\|u(\xi)\|_p^2 + \|\partial_\xi S(\xi)\|_\infty\|u(\xi)\|_p^2\\ &\quad + \|\partial_\tau S(\xi)A(\xi) + S(\xi)\partial_\tau A(\xi)\|_\infty\|u(\xi)\|_p^2\\ &\quad + \frac{2}{\beta}\Big(1 + 4\|A(\xi)\|_p^2\Big)\|u(\xi)\|_p^2\\ &\quad + \frac{1}{\beta}\Big(1 + \|\mathcal{L}(\xi)\|_{\mathcal{L}(H^p_{\mathrm{per}})}^2\Big)\|u(\xi)\|_p^2\\ &= \frac{1}{\beta}\|(Lu)(\xi)\|_p^2 + c\|u(\xi)\|_p^2,\end{aligned}$$

where

$$c = \frac{1}{\beta} + \|\partial_\xi S\|_{C([0,T];L^\infty_{\text{per}})} + \|(\partial_\tau S)A + S(\partial_\tau A)\|_{C([0,T];L^\infty_{\text{per}})} \\ + \frac{2}{\beta}\Big(1 + 4\|A\|^2_{C([0,T];H^p_{\text{per}})}\Big) + \frac{1}{\beta}\Big(1 + \|\mathcal{L}\|^2_{C([0,T];\mathcal{L}(H^p_{\text{per}}))}\Big),$$

and by integrating over (ξ_0, ξ) we obtain

$$\int_0^{2\pi} \langle S(\xi)\partial_\tau^k u(\xi), \partial_\tau^k u(\xi)\rangle \,\mathrm{d}\tau \le \int_0^{2\pi} \langle S(\xi_0)\partial_\tau^k u(\xi_0), \partial_\tau^k u(\xi_0)\rangle \,\mathrm{d}\tau \\ + \int_{\xi_0}^{\xi} \left(\frac{1}{\beta}\|(Lu)(r)\|_p^2 + c\|u(r)\|_p^2\right) \mathrm{d}r$$

from which it follows that

$$\|\partial_\tau^k u(\xi)\|_0^2 \le \frac{1}{\beta^2}\|u(\xi_0)\|_p^2 + \int_{\xi_0}^{\xi} \left(\frac{1}{\beta^2}\|(Lu)(r)\|_p^2 + \frac{c}{\beta}\|u(r)\|_p^2\right) \mathrm{d}r.$$

Summation over $k \le p$ yields

$$\|u(\xi)\|_p^2 \le (p+1)\left(\frac{1}{\beta^2}\|u(\xi_0)\|_p^2 + \int_{\xi_0}^{\xi} \left(\frac{1}{\beta^2}\|(Lu)(r)\|_p^2 + \frac{c}{\beta}\|u(r)\|_p^2\right) \mathrm{d}r\right)$$

and Gronwall's inequality with

$$a(\xi) = (p+1)\frac{1}{\beta^2}\left(\|u(\xi_0)\|_p^2 + \int_{\xi_0}^{\xi} \|(Lu)(r)\|_p^2 \,\mathrm{d}r\right), \quad v(\xi) = \|u(\xi)\|_p^2, \quad b(\xi) = (p+1)\frac{c}{\beta}$$

(note that a is continuous due to Proposition 4.6) finally implies that

$$\|u(\xi)\|_p^2 \le (p+1)\frac{1}{\beta^2}\mathrm{e}^{(p+1)\frac{cT}{\beta}}\left(\|u(\xi_0)\|_p^2 + \int_{\xi_0}^{\xi} \|(Lu)(r)\|_p^2 \,\mathrm{d}r\right).$$

■

Corollary 4.10. *Let $p \in \{t-1, t\}$ and*

$$u \in C\Big([0,T]; H^{p+1}_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^p_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

The estimate

$$\|u(\xi)\|_p^2 \lesssim \|u(T)\|_p^2 + \int_0^T \|(Lu)(r)\|_p^2 \,\mathrm{d}r$$

holds for all $\xi \in [0,T]$.

Proof. The arguments in the proof of Lemma 4.9 remain true if we replace the operator L by the 'reversed' operator $\tilde{L} = \partial_\xi - A\partial_\tau - \mathcal{L}$. Lemma 4.9 with $\xi_0 = 0$ implies that

$$\|u(T-\xi)\|_p^2 \lesssim \|u(T)\|_p^2 + \int_0^T \|(\tilde{L}u)(T-r)\|_p^2 \,\mathrm{d}r$$

for all $\xi \in [0,T]$. We conclude the result by substituting $\tau = T - r$. ■

Regularisation

In order to solve equation (4.1) we first study the sequence of regularised initial value problems

$$L^{(N)}u_N = f_{(N)} \tag{4.3}$$
$$u_N(0) = 0 \tag{4.4}$$

where

$$L^{(N)} = \partial_\xi + (S^{(N)})^{-1}M^{(N)}\partial_\tau + \mathcal{L} \tag{4.5}$$

and

$$(S^{(N)})^{-1}M^{(N)} \in C\Big([0,T]; C^\infty_{\text{per}}([0,2\pi];\mathbb{R}^{n\times n})\Big), \quad f_{(N)} \in C\Big([0,T]; C^\infty_{\text{per}}([0,2\pi];\mathbb{R}^{n})\Big)$$

have the property that $\|A-(S^{(N)})^{-1}M^{(N)}\|_{C([0,T];H^t_{\text{per}})}, \|f-f_{(N)}\|_{C([0,T];H^t_{\text{per}})} \to 0$ as $N\to\infty$. The functions $S^{(N)}$, $M^{(N)}$ and $f^{(N)}$ are given by the truncated Fourier series of respectively S, SA and f. The following result shows that the convergence of these Fourier series is uniform with respect to ξ. The invertibility of $S^{(N)}$ is established in Lemma 4.13 and Corollary 4.14 below.

Lemma 4.11. *Suppose that $g \in C\Big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R})\Big)$ is given by the Fourier series expansion*

$$g(\xi,\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}} g_m(\xi)\mathrm{e}^{\mathrm{i}m\tau}.$$

The sequence $\{g_{(N)}\}_{N\in\mathbb{N}}$ defined by

$$g_{(N)}(\xi,\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m=-N}^{N} g_m(\xi)\mathrm{e}^{\mathrm{i}m\tau}$$

has the properties

(i) $g_{(N)}(\xi) \in C^\infty_{\text{per}}([0,2\pi];\mathbb{R})$ *for each* $\xi\in[0,T]$,

(ii) $g_{(N)} \in C\Big([0,T]; H^p_{\text{per}}(0,2\pi;\mathbb{R})\Big)$ *for each* $p\in\mathbb{N}_0$,

(iii) $\lim_{N\to\infty}\|g_{(N)}-g\|_{C([0,T];H^t_{\text{per}})} = 0$,

(iv) $\|g_{(N)}\|_{C([0,T];H^t_{\text{per}})} \lesssim 1$ *uniformly over* N.

Proof.

(i) Since $g_{(N)}(\xi)$ is a trigonometric polynomial, it follows that

$$g_{(N)}(\xi) \in C^\infty_{\text{per}}([0,2\pi];\mathbb{R})$$

for each $\xi\in[0,T]$.

(ii) For $p \in \mathbb{N}_0$ and $\xi_1, \xi_2 \in [0,T]$ we find that

$$\begin{aligned}\|g_{(N)}(\xi_1) - g_{(N)}(\xi_2)\|_p^2 &= \sum_{m=-N}^{N} (1+m^2)^p |g_m(\xi_1) - g_m(\xi_2)|^2 \\ &\leq c_N \sum_{m=-N}^{N} (1+m^2)^t |g_m(\xi_1) - g_m(\xi_2)|^2 \\ &\leq c_N \|g(\xi_1) - g(\xi_2)\|_t^2,\end{aligned}$$

in which $c_N = \max_{m=-N,\dots,N}(1+m^2)^{p-t}$, that is $g_{(N)} \in C\Big([0,T]; H^p_{\mathrm{per}}(0,2\pi;\mathbb{R})\Big)$.

(iii) First note that

$$\|g_{(N)}(\xi) - g(\xi)\|_t^2 = \sum_{|m|>N} (1+m^2)^t |g_m(\xi)|^2 \to 0$$

as $N \to \infty$ for each $\xi \in [0,T]$, since $g(\xi) \in H^t_{\mathrm{per}}(0,2\pi;\mathbb{R})$, that is

$$\|g(\xi)\|_t^2 = \sum_{m\in\mathbb{Z}} (1+m^2)^t |g_m(\xi)|^2 < \infty.$$

To prove the uniform convergence of $\{g_{(N)}\}_{N\in\mathbb{N}}$ we observe that

$$g_{(N)} - g \in C\Big([0,T]; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R})\Big),$$

so that

$$\begin{aligned}\|g_{(N)} - g\|_{C([0,T];H^t_{\mathrm{per}})} &= \sup_{\xi\in[0,T]} \|g_{(N)}(\xi) - g(\xi)\|_t \\ &= \max_{\xi\in[0,T]} \|g_{(N)}(\xi) - g(\xi)\|_t \\ &= \|g_{(N)}(\xi_N) - g(\xi_N)\|_t\end{aligned}$$

for some $\xi_N \in [0,T]$. Since $[0,T]$ is compact there exists $\xi^* \in [0,T]$ such that $\xi_N \to \xi^*$ as $N \to \infty$ up to subsequences. It follows that

$$\lim_{N\to\infty} \|g_{(N)} - g\|_{C([0,T];H^t_{\mathrm{per}})} = \lim_{N\to\infty} \|g_{(N)}(\xi_N) - g(\xi_N)\|_t$$

and

$$\|g_{(N)}(\xi_N) - g(\xi_N)\|_t \leq \|g_{(N)}(\xi_N) - g_{(N)}(\xi^*)\|_t + \|g_{(N)}(\xi^*) - g(\xi^*)\|_t + \|g(\xi^*) - g(\xi_N)\|_t.$$

We observe that

$$\|g_{(N)}(\xi_N) - g_{(N)}(\xi^*)\|_t^2 = \sum_{m=-N}^{N} (1+m^2)^t |g_m(\xi_N) - g_m(\xi^*)|^2 \leq \|g(\xi_N) - g(\xi^*)\|_t$$

and since g is continuous, we conclude that

$$\|g_{(N)}(\xi_N) - g_{(N)}(\xi^*)\|_t, \|g(\xi^*) - g(\xi_N)\|_t \to 0$$

as $N \to \infty$. Moreover, the pointwise convergence of $\{g_{(N)}\}_{N\in\mathbb{N}}$ implies that

$$\|g_{(N)}(\xi^*) - g(\xi^*)\|_t \to 0$$

as $N \to \infty$. Altogether it follows that

$$\lim_{N\to\infty} \|g_{(N)} - g\|_{C([0,T];H^t_{\mathrm{per}})} = 0$$

(up to subsequences).

(iv) We have that

$$\begin{aligned}
\|g_{(N)}\|_{C([0,T];H^t_{\mathrm{per}})} &= \sup_{\xi\in[0,T]} \|g_{(N)}(\xi)\|_t \\
&= \sup_{\xi\in[0,T]} \sum_{m=-N}^{N} (1+m^2)^t |g_m(\xi)|^2 \\
&\leq \sup_{\xi\in[0,T]} \|g(\xi)\|_t \\
&= \|g\|_{C([0,T];H^t_{\mathrm{per}})},
\end{aligned}$$

so that $\|g_{(N)}\|_{C([0,T];H^t_{\mathrm{per}})} \lesssim 1$ uniformly over N.

■

Corollary 4.12. *Suppose that $j \in \mathbb{N}_0$ and $g \in C^j\Big([0,T]; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R})\Big)$ is given by the Fourier series expansion*

$$g(\xi,\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} g_m(\xi)\mathrm{e}^{\mathrm{i}m\tau}.$$

The sequence $\{g_{(N)}\}_{N\in\mathbb{N}}$ defined by

$$g_{(N)}(\xi,\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m=-N}^{N} g_m(\xi)\mathrm{e}^{\mathrm{i}m\tau}$$

has the properties

(i) $g_{(N)}(\xi) \in C^\infty_{\mathrm{per}}([0,2\pi];\mathbb{R})$ *for each* $\xi \in [0,T]$,

(ii) $g_{(N)} \in C^j\Big([0,T]; H^p_{\mathrm{per}}(0,2\pi;\mathbb{R})\Big)$ *for each* $p \in \mathbb{N}_0$,

(iii) $\lim_{N\to\infty} \|g_{(N)} - g\|_{C([0,T];H^t_{\mathrm{per}})} = 0$,

(iv) $\|g_{(N)}\|_{C^j([0,T];H^t_{\mathrm{per}})} \lesssim 1$ *uniformly over* N.

Proof. Proposition 1.20 shows that the truncated Fourier series of $\partial_\xi^l g$ is given by

$$\partial_\xi^l g_{(N)}(\xi,\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m=-N}^{N} g_m^{(l)}(\xi) \mathrm{e}^{\mathrm{i}m\tau}.$$

The result therefore follows by applying the arguments in the proof of Lemma 4.11 to $\partial_\xi^l g$, $l = 0, \dots, j$. ■

Write

$$L = \partial_\xi + A\partial_\tau + \mathcal{L} = \partial_\xi + S^{-1}SA\partial_\tau + \mathcal{L}$$

and use Corollary 4.12 componentwise to find families

$$\{S^{(N)}\}_{N\in\mathbb{N}} \subset C^1\Big([0,T]; C^\infty_{\text{per}}([0,2\pi];\mathbb{R}^{n\times n})\Big),\ \{M^{(N)}\}_{N\in\mathbb{N}} \subset C\Big([0,T]; C^\infty_{\text{per}}([0,2\pi];\mathbb{R}^{n\times n})\Big)$$

of smooth (with respect to the spatial variable τ), symmetric matrix-valued functions such that

$$\|S - S^{(N)}\|_{C([0,T];H^t_{\text{per}})}, \|S - S^{(N)}\|_{C^1([0,T];H^{t-1}_{\text{per}})} \to 0, \quad \|SA - M^{(N)}\|_{C([0,T];H^t_{\text{per}})} \to 0$$

as $N \to \infty$. In particular the mapping

$$(\xi, u) \mapsto S^{(N)}(\xi)u$$

lies in $C\Big([0,T]; \mathcal{L}\Big(H^p_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big)\Big)$ for arbitrary $p \in \mathbb{N}$.

In order to define the symmetrisable regularised operator $L^{(N)}$ in (4.5) with the desired properties we prove that the regularisation $S^{(N)}$ of S is uniformly positive-definite and therefore invertible.

Lemma 4.13. *There exists a constant $\tilde\beta > 0$ which does not depend upon N such that*

$$\tilde\beta |v|^2 \leq \langle S^{(N)}(\xi,\tau)v, v\rangle \leq \frac{1}{\tilde\beta}|v|^2, \quad v \in \mathbb{R}^n \setminus \{0\}. \tag{4.6}$$

Proof. Write

$$\langle S^{(N)}(\xi,\tau)v,v\rangle = \langle S(\xi,\tau)v,v\rangle + \langle (S^{(N)}(\xi,\tau) - S(\xi,\tau))v, v\rangle,$$

so that

$$\beta|v|^2 + \langle (S^{(N)}(\xi,\tau) - S(\xi,\tau))v,v\rangle \leq \langle S^{(N)}(\xi,\tau)v,v\rangle \leq \frac{1}{\beta}|v|^2 + \langle (S^{(N)}(\xi,\tau) - S(\xi,\tau))v,v\rangle.$$

Denoting $S(\xi,\tau) = (s_{ij}(\xi,\tau))_{i,j=1}^n$ and $S^{(N)}(\xi,\tau) = (s_{ij}^{(N)}(\xi,\tau))_{i,j=1}^n$ we recall that

$$\|s_{ij}^{(N)} - s_{ij}\|_{C^1([0,T];H^t_{\text{per}})} \to 0, \quad i,j = 1,\dots,n,$$

as $N \to \infty$.

Since
$$|s_{ij}^{(N)}(\xi,\tau) - s_{ij}(\xi,\tau)| \le \|s_{ij}^{(N)} - s_{ij}\|_{C^1([0,T];H^t_{\mathrm{per}})},$$
we conclude that
$$|s_{ij}^{(N)}(\xi,\tau) - s_{ij}(\xi,\tau)| \to 0$$
as $N \to \infty$ uniformly over $(\xi,\tau) \in [0,T] \times [0,2\pi]$ for each $i,j = 1,\dots,n$ and hence

$$|(S^{(N)}(\xi,\tau) - S(\xi,\tau))v| \le \frac{\beta}{2}|v|$$

for sufficiently large values of N. It thus follows that

$$\frac{\beta}{2}|v|^2 \le \langle S^{(N)}(\xi,\tau)v, v\rangle \le \Big(\frac{1}{\beta} + \frac{\beta}{2}\Big)|v|^2, \quad v \in \mathbb{R}^n \setminus \{0\},$$

uniformly over $(\xi,\tau) \in [0,T] \times [0,2\pi]$. ■

Corollary 4.14. *Suppose that $\tilde{\beta} > 0$ is the constant given in Lemma 4.13. The matrix-valued function $S^{(N)}$ is invertible and satisfies*

$$\tilde{\beta}|v|^2 \le \langle (S^{(N)})^{-1}(\xi,\tau)v, v\rangle \le \frac{1}{\tilde{\beta}}|v|^2, \quad v \in \mathbb{R}^n \setminus \{0\}.$$

Furthermore the matrix-valued function $(S^{(N)})^{-1}$ satisfies

$$(S^{(N)})^{-1} \in C^1\Big([0,T]; C^\infty_{\mathrm{per}}([0,2\pi];\mathbb{R}^{n\times n})\Big).$$

Proof. It is a direct consequence of inequality (4.6) that S^N is invertible with

$$\tilde{\beta}|v|^2 \le \langle (S^{(N)})^{-1}(\xi,\tau)v, v\rangle \le \frac{1}{\tilde{\beta}}|v|^2, \quad v \in \mathbb{R}^n \setminus \{0\}.$$

Note that

$$(S^{(N)})^{-1} = \frac{1}{\det S^{(N)}}(S^{(N)})^{\#} \tag{4.7}$$

with the adjunct matrix $(S^{(N)})^{\#}$ of $S^{(N)}$ and we have that $\det S^{(N)}(\xi,\tau) \ge (\frac{\tilde{\beta}}{2})^n$ for all $(\xi,\tau) \in [0,T] \times [0,2\pi]$ due to inequality (4.6). The denominator $\det S^{(N)}$ on the right-hand side of (4.7) is therefore bounded (away from zero) uniformly over $(\xi,\tau) \in [0,T] \times [0,2\pi]$ and we conclude that

$$(S^{(N)})^{-1} \in C^1\Big([0,T]; C^\infty_{\mathrm{per}}([0,2\pi];\mathbb{R}^{n\times n})\Big).$$

■

We now consider the regularised operator

$$L^{(N)} = \partial_\xi + (S^{(N)})^{-1} M^{(N)} \partial_\tau + \mathcal{L},$$

and, noting that $(S^{(N)})^{-1} M^{(N)}, M^{(N)}(S^{(N)})^{-1} \in C\Big([0,T]; C^\infty_{\mathrm{per}}([0,2\pi]; \mathbb{R}^{n\times n})\Big)$ due to Corollary 4.14, the family of adjoint operators

$$L_p^{(N)*} = -\partial_\xi - M^{(N)}(S^{(N)})^{-1}\partial_\tau - \partial_\tau\Big(M^{(N)}(S^{(N)})^{-1}\Big) + \mathcal{L}_p^*, \quad p \in \mathbb{N}$$

(see Remark 4.2(i)). For notational simplicity below we abbreviate $L^{(N)} = \partial_\xi + P^{(N)}$, so that $L_p^{(N)*} = -\partial_\xi + P_p^{(N)*}$, with the obvious definitions of $P^{(N)}$ and $P_p^{(N)*}$, and also write $R^{(N)} = (S^{(N)})^{-1} M^{(N)}$, so that $(R^{(N)})^\top = M^{(N)}(S^{(N)})^{-1}$.

Remark 4.15.

(i) The matrices $M^{(N)}$ as well as $S^{(N)}$ and thus $(S^{(N)})^{-1}$ are symmetric so that the regularised operators $L^{(N)}$ and $L_p^{(N)*}$ are symmetrisable with symmetrisers $S^{(N)}$ and $(S^{(N)})^{-1}$, respectively.

(ii) Suppose that $\mathcal{L}(\xi)u = B(\xi)u$ with a matrix-valued function

$$B \in C\Big([0,T]; H^t_{\mathrm{per}}(0,2\pi; \mathbb{R}^{n\times n})\Big)$$

(see Remark 4.2(ii)) and consider the regularisation $B^{(N)}$ of B in the sense of Corollary 4.12. The regularised operator $\mathcal{L}^{(N)}$ defined by $\mathcal{L}^{(N)}(\xi)u = B^{(N)}(\xi)u$ satisfies

$$\mathcal{L}^{(N)} \in C\Big([0,T]; \mathcal{L}\Big(H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)\Big)$$

for arbitrary $p \in \mathbb{N}$ and applying the arguments below with $\mathcal{L}$ replaced by $\mathcal{L}^{(N)}$ yields the existence and uniqueness result for Cauchy problems of the type

$$\partial_\xi u + A(\xi,\tau)\partial_\tau u = B(\xi,\tau)u + f(\xi,\tau)$$

with non-smooth coefficients.

We use the symmetrisability of the adjoint operator $L_p^{(N)*}$ together with the fact that $(S^{(N)})^{-1} \in C\Big([0,T]; C^\infty_{\mathrm{per}}([0,2\pi]; \mathbb{R}^{n\times n})\Big)$ to prove an *a priori* estimate for $L_p^{(N)*}$. The proof requires the following commutator identity.

Proposition 4.16. *Let $p \in \mathbb{N}$, suppose that*

$$v \in H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n), \quad B \in C\Big([0,T]; C^\infty_{\mathrm{per}}([0,2\pi]; \mathbb{R}^{n\times n})\Big)$$

and consider the commutator

$$[(I-\partial_\tau^2)^p, B\partial_\tau] = (I-\partial_\tau^2)^p B\partial_\tau - B(I-\partial_\tau^2)^p\partial_\tau.$$

For each $k \in \mathbb{N}_0$ with $k \leq 2p$ there exists a function

$$C_k \in C\Big([0,T]; C^\infty_{\text{per}}([0,2\pi];\mathbb{R}^{n\times n})\Big)$$

such that

$$[(I-\partial_\tau^2)^p, B\partial_\tau]v = \sum_{k=0}^{2p} C_k \partial_\tau^k v.$$

In particular $[(I-\partial_\tau^2)^p, B\partial_\tau] \in C\Big([0,T]; \mathcal{L}\Big(H^p(0,2\pi;\mathbb{R}^n); H^{-p}(0,2\pi;\mathbb{R}^n)\Big)\Big)$.

Proof. We prove the statement by induction on p. For $p=1$ we have that

$$\begin{aligned}
[(I-\partial_\tau^2)^p, B\partial_\tau]v &= (I-\partial_\tau^2)B\partial_\tau v - B(I-\partial_\tau^2)\partial_\tau v \\
&= B\partial_\tau^3 v - \partial_\tau^2(B\partial_\tau v) \\
&= -\partial_\tau^2 B\partial_\tau v - 2\partial_\tau B\partial_\tau^2 v.
\end{aligned}$$

Now suppose that the statement is true for fixed $p \in \mathbb{N}$. We compute

$$[(I-\partial_\tau^2)^{p+1}, B\partial_\tau]v = (I-\partial_\tau^2)[(I-\partial_\tau^2)^p, B\partial_\tau]v + [(I-\partial_\tau^2), B\partial_\tau](I-\partial_\tau^2)^p v$$

and

$$\begin{aligned}
(I-\partial_\tau^2)[(I-\partial_\tau^2)^p, B\partial_\tau]v &= (I-\partial_\tau^2)\sum_{k=0}^{2p} C_k\partial_\tau^k v \\
&= \sum_{k=0}^{2p}\Big(C_k\partial_\tau^k v - \partial_\tau^2 C_k\partial_\tau^k v - 2\partial_\tau C_k\partial_\tau^{k+1} v - C_k\partial_\tau^{k+2} v\Big),
\end{aligned}$$

as well as

$$\begin{aligned}
&[(I-\partial_\tau^2), B\partial_\tau](I-\partial_\tau^2)^p v \\
&\quad= [(I-\partial_\tau^2), B\partial_\tau]\sum_{j=0}^{p}\binom{p}{j}(-1)^{p-j}\partial_\tau^{2(p-j)} v \\
&\quad= -\sum_{j=0}^{p}\binom{p}{j}(-1)^{p-j}\Big(\partial_\tau^2 B\partial_\tau^{2(p-j)+1} v + 2\partial_\tau B\partial_\tau^{2(p-j)+2} v\Big).
\end{aligned}$$

■

Lemma 4.17. *Suppose that $p \geq t-1$, $\xi_0 \in [0,T]$ and let*

$$u \in C\Big([0,T]; H^{-p+1}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{-p}(0,2\pi;\mathbb{R}^n)\Big).$$

The estimate

$$\|u(\xi)\|_{-p}^2 \leq c_N\left(\|u(\xi_0)\|_{-p}^2 + \int_{\xi_0}^{\xi} \|(L_p^{(N)*}u)(r)\|_{-p}^2 \,\mathrm{d}r\right)$$

holds for all $\xi \in [\xi_0, T]$.

Proof. We use the identity

$$(I-\partial_\tau^2)^{p/2}\phi = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}}(1+m^2)^{p/2}\phi_m \mathrm{e}^{\mathrm{i}m\tau}, \quad \phi(\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}}\phi_m \mathrm{e}^{\mathrm{i}m\tau}$$

for $\phi\in H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$ together with the fact that $\phi\mapsto\|(I-\partial_\tau^2)^{p/2}\phi\|_0$ defines the usual norm for $H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$. Consider the function $v(\xi)=(I-\partial_\tau^2)^{-p}u(\xi)$ and note that

$$\|v(\xi)\|_p = \|(I-\partial_\tau^2)^{p/2}v(\xi)\|_0 = \|(I-\partial_\tau^2)^{-p/2}u(\xi)\|_0 = \|u(\xi)\|_{-p},$$

as well as $\|v(\xi)\|_{p+1}=\|u(\xi)\|_{-p+1}$, so that

$$v\in C\Big([0,T];H^{p+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)\cap C^1\Big([0,T];H^p(0,2\pi;\mathbb{R}^n)\Big)$$

and Lemma 4.9 (for all $p\geq t-1$ since $L^{(N)}$ has smooth coefficients) implies that

$$\|v(\xi)\|_p^2 \leq c_N\Big(\|v(\xi_0)\|_p^2 + \int_{\xi_0}^{\xi}\|(L^{(N)}v)(r)\|_p^2\,\mathrm{d}r\Big)$$

for all $\xi\in[\xi_0,T]$. It follows that

$$\begin{aligned}
&\|u(\xi)\|_{-p}^2\\
&\quad\leq c_N\Big(\|u(\xi_0)\|_{-p}^2 + \int_{\xi_0}^{\xi}\Big\|(I-\partial_\tau^2)^{p/2}\Big(\partial_\xi v(\xi)+R^{(N)}(r)\partial_\tau v(r)\Big)\Big\|_0^2 + \|\mathcal{L}(r)v(r)\|_p^2\,\mathrm{d}r\Big)\\
&\quad\lesssim c_N\Big(\|u(\xi_0)\|_{-p}^2\\
&\qquad\qquad + \int_{\xi_0}^{\xi}\Big\|(I-\partial_\tau^2)^{-p/2}(I-\partial_\tau^2)^{p}\Big(\partial_\xi v(\xi)+R^{(N)}(r)\partial_\tau v(r)\Big)\Big\|_0^2 + \|v(r)\|_p^2\,\mathrm{d}r\Big)\\
&\quad= c_N\Big(\|u(\xi_0)\|_{-p}^2 + \int_{\xi_0}^{\xi}\Big\|(I-\partial_\tau^2)^{-p/2}\Big(\partial_\xi(I-\partial_\tau^2)^p v(r)+R^{(N)}(r)\partial_\tau(I-\partial_\tau^2)^p v(r)\Big)\Big\|_0^2\\
&\qquad\qquad + \Big\|(I-\partial_\tau^2)^{-p/2}([(I-\partial_\tau^2)^p,R^{(N)}\partial_\tau]v)(r)\Big\|_0^2\,\mathrm{d}r\Big)
\end{aligned}\tag{4.8}$$

since the operator $(I-\partial_\tau^2)^p$ commutes with the partial derivative ∂_ξ.

Next we use Proposition 4.16 to find that

$$\begin{aligned}
\Big\|(I-\partial_\tau^2)^{-p/2}([(I-\partial_\tau^2)^p,R^{(N)}\partial_\tau]v)(r)\Big\|_0^2 &= \|([(I-\partial_\tau^2)^p,R^{(N)}\partial_\tau]v)(r)\|_{-p}^2\\
&\leq c_N\|v(r)\|_p^2\\
&= c_N\|u(r)\|_{-p}^2.
\end{aligned}$$

We further estimate

$$\begin{aligned}
&\Big\|(I-\partial_\tau^2)^{-p/2}\Big(\partial_\xi(I-\partial_\tau^2)^p v(r)+R^{(N)}(r)\partial_\tau(I-\partial_\tau^2)^p v(r)\Big)\Big\|_0^2\\
&\quad= \|\partial_\xi u(r)+R^{(N)}(r)\partial_\tau u(r)\|_{-p}^2\\
&\quad= \|(L_p^{(N)*}u)(r)+\partial_\tau R^{(N)}(r)u(r)-\mathcal{L}_p^*(r)u(r)\|_{-p}^2\\
&\quad\leq \|(L_p^{(N)*}u)(r)\|_{-p}^2+\|\partial_\tau R^{(N)}(r)u(r)\|_{-p}^2+\|\mathcal{L}_p^*(r)u(r)\|_{-p}^2\\
&\quad\leq \|(L_p^{(N)*}u)(r)\|_{-p}^2 + c_N\|u(r)\|_{-p}^2,
\end{aligned}$$

in which the last line follows from the fact that $R^{(N)}(r) \in C^\infty_{\mathrm{per}}([0,2\pi];\mathbb{R}^{n\times n})$ and Proposition 4.7.

Inserting the two previous estimates into (4.8), we conclude that

$$\|u(\xi)\|^2_{-p} \le c_N\Big(\|u(\xi_0)\|^2_{-p} + \int_{\xi_0}^{\xi} \|(L_p^{(N)*}u)(r)\|^2_{-p} + c_N\|u(r)\|^2_{-p}\,\mathrm{d}r\Big).$$

Gronwall's inequality finally yields

$$\|u(\xi)\|^2_{-p} \le c_N \mathrm{e}^{c_N T}\Big(\|u(\xi_0)\|^2_{-p} + \int_{\xi_0}^{\xi} \|(L_p^{(N)*}u)(r)\|^2_{-p}\,\mathrm{d}r\Big)$$

for all $\xi \in [\xi_0, T]$. ∎

Corollary 4.18. *Suppose that $p \ge t-1$ and let*

$$u \in C\Big([0,T]; H^{-p+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{-p}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

The estimate

$$\|u(\xi)\|^2_{-p} \le c_N\Big(\|u(T)\|^2_{-p} + \int_0^T \|(L_p^{(N)*}u)(r)\|^2_{-p}\,\mathrm{d}r\Big)$$

holds for all $\xi \in [0,T]$.

The following theorem is the existence and uniqueness result for the initial value problem (4.3), (4.4). Here we construct distributional solutions to equation (4.3) using the set of test functions

$$\mathcal{E} = \Big\{\phi \in C^\infty\Big([0,T]; C^\infty_{\mathrm{per}}([0,2\pi];\mathbb{R}^n)\Big) : \phi(T) = 0\Big\}.$$

Theorem 4.19. *The regularised initial value problem (4.3), (4.4) posesses a unique solution*

$$u_N \in C\Big([0,T]; H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{t}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

Proof. We start by proving that equation (4.3) admits a distributional solution

$$u_N \in L^2\Big(0,T; H^{t+2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

To this end we consider the adjoint operator $L^{(N)*}_{t+2}$ on the set $\mathcal{E}$. The first step is to show that $L^{(N)*}_{t+2}$ maps $\mathcal{E}$ into $L^2\Big(0,T;H^{-t-2}(0,2\pi;\mathbb{R}^n)\Big)$. Suppose that $\phi \in \mathcal{E}$ with $\psi = L^{(N)*}_{t+2}\phi \in L^{(N)*}_{t+2}\mathcal{E}$. By definition we conclude that

$$\phi \in C^j\Big([0,T]; H^p_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big), \qquad j,p \in \mathbb{N}_0.$$

It follows that $\phi \in C^1\Big([0,T]; H^{-t-2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$, since $(1+m^2)^p \ge (1+m^2)^{-t-2}$ for each $p \in \mathbb{N}_0$. Proposition 4.7 implies that $\psi \in C\Big([0,T]; H^{-t-2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$ and thus $\psi \in L^2\Big(0,T; H^{-t-2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$, that is $(L^{(N)})^*\mathcal{E} \subset L^2\Big(0,T;H^{-t-2}(0,2\pi;\mathbb{R}^n)\Big)$.

Next we prove that the operator

$$L_{t+2}^{(N)*}\colon \mathcal{E} \to L_{t+2}^{(N)*}\mathcal{E} \subset L^2\Big(0,T; H_{\mathrm{per}}^{-t-2}(0,2\pi;\mathbb{R}^n)\Big)$$

is injective. Let $v_1, v_2 \in \mathcal{E}$ with $L_{t+2}^{(N)*}v_1 = L_{t+2}^{(N)*}v_2$. Corollary 4.18 for $p = t+2$ implies that

$$\begin{aligned}
&\|v_1(\xi) - v_2(\xi)\|_{-t-2}^2 \\
&\quad \lesssim \underbrace{\|v_1(T) - v_2(T)\|_{-t-2}^2}_{=0} + \int_0^T \underbrace{\|(L_{t+2}^{(N)*}v_1)(r) - (L_{t+2}^{(N)*}v_2)(r)\|_{-t-2}^2}_{=0} \,\mathrm{d}r \\
&\quad = 0,
\end{aligned}$$

that is $v_1 = v_2$.

Consider the linear functional $l\colon L_{t+2}^{(N)*}\mathcal{E} \to \mathbb{C}$ defined by

$$l(\psi) = \int_0^T \Big\langle \Big((L_{t+2}^{(N)*})^{-1}\psi\Big)(\xi), f_{(N)}(\xi)\Big\rangle_{H^{-t-2},H^{t+2}} \,\mathrm{d}\xi$$

(note that $((L_{t+2}^{(N)*})^{-1}\psi)(\xi) \in C_{\mathrm{per}}^{\infty}([0,2\pi];\mathbb{R}^n) \subset H_{\mathrm{per}}^{-t-2}(0,2\pi;\mathbb{R}^n))$. Using the Cauchy-Schwarz inequality we conclude that

$$\begin{aligned}
|l(\psi)| &\le \int_0^T \Big\|\Big((L_{t+2}^{(N)*})^{-1}\psi\Big)(\xi)\Big\|_{-t-2} \|f_{(N)}(\xi)\|_{t+2} \,\mathrm{d}\xi \\
&= \left(\left(\int_0^T \Big\|\Big((L_{t+2}^{(N)*})^{-1}\psi\Big)(\xi)\Big\|_{-t-2} \|f_{(N)}(\xi)\|_{t+2} \,\mathrm{d}\xi\right)^2\right)^{\frac{1}{2}} \\
&\le \left(\int_0^T \Big\|\Big((L_{t+2}^{(N)*})^{-1}\psi\Big)(\xi)\Big\|_{-t-2}^2 \,\mathrm{d}\xi \int_0^T \|f_{(N)}(\xi)\|_{t+2}^2 \,\mathrm{d}\xi\right)^{\frac{1}{2}} \\
&= \|f_{(N)}\|_{L^2(0,T;H_{\mathrm{per}}^{t+2})} \left(\int_0^T \Big\|\Big((L_{t+2}^{(N)*})^{-1}\psi\Big)(\xi)\Big\|_{-t-2}^2 \,\mathrm{d}\xi\right)^{\frac{1}{2}}.
\end{aligned} \tag{4.9}$$

Applying Corollary 4.18 with $p = t+2$ to a function

$$v \in C\Big([0,T]; H_{\mathrm{per}}^{-t-1}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H_{\mathrm{per}}^{-t-2}(0,2\pi;\mathbb{R}^n)\Big)$$

yields

$$\|v(\xi)\|_{-t-2}^2 \le c_N\Big(\|v(T)\|_{-t-2}^2 + \int_0^T \|(L_{t+2}^{(N)*}v)(r)\|_{-t-2}^2 \,\mathrm{d}r\Big).$$

and using this inequality with $v = (L_{t+2}^{(N)*})^{-1}\psi$ (noting that $((L_{t+2}^{(N)*})^{-1}\psi)(T) = 0$) we conclude from the estimate (4.9) that

$$\begin{aligned}
|l(\psi)| &\le c_N \|f_{(N)}\|_{L^2(0,T;H_{\mathrm{per}}^{t+2})} \left(\int_0^T \int_0^T \Big\|\Big(L_{t+2}^{(N)*}(L_{t+2}^{(N)*})^{-1}\psi\Big)(r)\Big\|_{-t-2}^2 \,\mathrm{d}r\mathrm{d}\xi\right)^{\frac{1}{2}} \\
&\le c_N\sqrt{T}\|f_{(N)}\|_{L^2(0,T;H_{\mathrm{per}}^{t+2})}\|\psi\|_{L^2(0,T;H_{\mathrm{per}}^{-t-2})}.
\end{aligned}$$

It follows from the Hahn-Banach theorem that l extends continuously to a linear form on $L^2\big(0,T;H^{-t-2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ and the Riesz representation theorem implies the existence of a unique function $u_N \in L^2\big(0,T;H^{t+2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ such that

$$l(\psi) = \int_0^T \langle \psi(\xi), u_N(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi$$

for each $\psi \in L^{(N)*}_{t+2}\mathcal{E}$. It follows that

$$\int_0^T \Big\langle (L^{(N)*}_{t+2}\phi)(\xi), u_N(\xi)\Big\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi = \int_0^T \langle \phi(\xi), f_{(N)}(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi \tag{4.10}$$

for each $\phi \in \mathcal{E}$, that is u_N is a distributional solution of the equation

$$L^{(N)}u_N = f_{(N)}.$$

Observe that

$$\partial_\xi u_N = -R^{(N)}\partial_\tau u_N - \mathcal{L}u_N$$

with $\partial_\tau u_N \in L^2\big(0,T;H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$, since $u_N \in L^2\big(0,T;H^{t+2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$. The bilinear mappings

$$(R^{(N)},\partial_\tau u_N) \mapsto R^{(N)}\partial_\tau u_N, \quad (\mathcal{L},u_N) \mapsto \mathcal{L}u_N$$

thus lie in $L^2\big(0,T;H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ due to Proposition 4.6. It follows that $\partial_\xi u_N \in L^2\big(0,T;H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$, and therefore $u_N \in H^1\big(0,T;H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$. Hence $u_N \in C\big([0,T];H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ with $\partial_\xi u_N \in C\big([0,T];H^{t}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ by Proposition 4.6 (see above). Finally, it follows from equation (4.10) that

$$\begin{aligned}\int_0^T \langle P^{(N)*}_{t+2}(\xi)\phi(\xi), u_N(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi = &\int_0^T \langle \partial_\xi\phi(\xi), u_N(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi\\ &+\int_0^T \langle \phi(\xi), f_{(N)}(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi\end{aligned}$$

and integration by parts yields

$$\begin{aligned}\int_0^T \langle \phi(\xi), P^{(N)}(\xi)u_N(\xi) - f_{(N)}(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi = &-\langle \phi(0), u_N(0)\rangle_{H^{-t-2},H^{t+2}}\\ &-\int_0^T \langle \phi(\xi), \partial_\xi u_N(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi,\end{aligned}$$

so that

$$0 = \int_0^T \langle \phi(\xi), (L^{(N)}u_N)(\xi) - f_{(N)}(\xi)\rangle_{H^{-t-2},H^{t+2}}\,\mathrm{d}\xi = -\langle \phi(0), u_N(0)\rangle_{H^{-t-2},H^{t+2}}$$

for each $\phi \in \mathcal{E}$, that is $u_N(0) = 0$.

It remains to prove the uniqueness of the solution. Let u_N^1 and u_N^2 be solutions of the initial value problem (4.3), (4.4). It follows that

$$u_N^1 - u_N^2 \in C\Big([0,T]; H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^t(0,2\pi;\mathbb{R}^n)\Big)$$

and Lemma 4.9 with $\xi_0 = 0$ implies

$$\begin{aligned}
\|u_N^1(\xi) - u_N^2(\xi)\|_t^2 &\lesssim \| \underbrace{u_N^1(0) - u_N^2(0)}_{=0} \|_t^2 + \int_0^\xi \Big\| \Big(L^{(N)}(u_N^1 - u_N^2) \Big)(r) \Big\|_t^2 \,\mathrm{d}r \\
&= \int_0^\xi \| (L^{(N)} u_N^1)(r) - (L^{(N)} u_N^2)(r) \|_t^2 \,\mathrm{d}r \\
&= \int_0^\xi \| f_{(N)}(r) - f_{(N)}(r) \|_t^2 \,\mathrm{d}r \\
&= 0,
\end{aligned}$$

so that $u_N^1 = u_N^2$. ∎

Solving the Cauchy problem

We show that the sequence $\{u_N\}_{N\in\mathbb{N}}$ of solutions to the regularised initial value problem (4.3), (4.4) converges to a distributional solution of the initial value problem (4.1), (4.2) the regularity of which we verify *a posteriori.*

We start with the case $h = 0$. Our aim is to show that $\{u_N\}_{N\in\mathbb{N}}$ converges to

$$u \in C\Big([0,T]; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{t-1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

We first establish the weak convergence of $\{u_N\}_{N\in\mathbb{N}}$. In the following theorem the result is formulated in terms of the adjoint operator

$$L_t^* u = -\partial_\xi u - \partial_\tau A u - A \partial_\tau u + \mathcal{L}_t^* u.$$

Theorem 4.20. *The sequence $\{u_N(\xi)\}_{N\in\mathbb{N}}$ converges weakly in $H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$, uniformly over $\xi \in [0,T]$, to a function*

$$u \in L^2\Big(0,T; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$$

such that

$$\int_0^T \langle (L_t^*\phi)(\xi), u(\xi) \rangle_{H^{-t},H^t} \,\mathrm{d}\xi = \int_0^T \langle \phi(\xi), f(\xi) \rangle_{H^{-t},H^t} \,\mathrm{d}\xi$$

for each $\phi \in \mathcal{E}$. In particular u is a distributional solution of the equation

$$Lu = f.$$

Proof. The *a priori* estimate for the H^t-norm in Lemma 4.9 yields

$$\|u_N(\xi)\|_t^2 \lesssim \int_0^\xi \|(L^{(N)}u_N)(r)\|_t^2 \,\mathrm{d}r \lesssim \|f\|^2_{C([0,T];H^t_{\mathrm{per}})}.$$

The sequence $\{u_N\}_{N\in\mathbb{N}}$ is hence bounded in $C\Big([0,T];H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$. The *a priori* estimate for the H^{t-1}-norm in Lemma 4.9 further yields

$$\begin{aligned}
&\|u_N(\xi)-u_M(\xi)\|_{t-1}^2\\
&\quad\lesssim \int_0^\xi \|(L(u_N-u_M))(r)\|_{t-1}^2\,\mathrm{d}r\\
&\quad\le \int_0^\xi \|((L-L^{(N)})u_N)(r)+((L^{(M)}-L)u_M)(r)+f_{(N)}(r)-f_{(M)}(r)\|_{t-1}^2\,\mathrm{d}r. \qquad (4.11)
\end{aligned}$$

By assumption $t-1>\frac{3}{2}$, so that

$$\begin{aligned}
\|((L-L^{(N)})u_N)(r)\|_{t-1}^2 &= \Big\|\Big((S^{(N)})^{-1}(r)M^{(N)}(r)-A(r)\Big)\partial_\tau u_N(r)\Big\|_{t-1}^2\\
&\le \|(S^{(N)})^{-1}(r)M^{(N)}(r)-A(r)\|_{t-1}^2\|\partial_\tau u_N(r)\|_{t-1}^2\\
&\le \|(S^{(N)})^{-1}(r)M^{(N)}(r)-A(r)\|_{H^t}^2\|u_N(r)\|_t^2\\
&\lesssim \|(S^{(N)})^{-1}M^{(N)}-A\|^2_{C([0,T];H^t_{\mathrm{per}})}.
\end{aligned}$$

It follows that

$$\|(L-L^{(N)})u_N\|_{C([0,T];H^{t-1}_{\mathrm{per}})}\to 0$$

as $N\to\infty$ and since also

$$\|f_{(N)}-f_{(M)}\|_{C([0,T];H^{t-1})}\to 0$$

as $M,N\to\infty$, we conclude from inequality (4.11) that $\{u_N\}_{N\in\mathbb{N}}$ is a Cauchy sequence in $C\Big([0,T];H^{t-1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$ and thus has a limit $u\in C\Big([0,T];H^{t-1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$.

Since $\|u_N(\xi)\|_t\lesssim 1$ uniformly over $N\in\mathbb{N}$ and $\xi\in[0,T]$, there exists a function $\zeta(\xi)\in H^t(0,2\pi;\mathbb{R}^n)$ such that $u_N(\xi)\rightharpoonup\zeta(\xi)$ in $H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$ (up to subsequences) and $\|\zeta(\xi)\|_t\lesssim 1$. It follows that $u_N(\xi)\rightharpoonup\zeta(\xi)$ in $H^{t-1}(0,2\pi;\mathbb{R}^n)$ and since we already know that $u_N\to u$ in $C\Big([0,T];H^{t-1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$ we conclude that $u(\xi)=\zeta(\xi)$, that is $u(\xi)\in H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$ with $\|u(\xi)\|_t\lesssim 1$, for all $\xi\in[0,T]$.

Applying the same argument to the sequence $\{u(\xi_k)\}_{k\in\mathbb{N}}$, where $\xi_k\to\xi\in[0,T]$, yields $\|u(\xi_k)\|_t\lesssim 1$ and $u(\xi_k)\rightharpoonup u(\xi)$ (up to subsequences). It follows that $u\in C\Big([0,T];H^t_{\mathrm{per,w}}(0,2\pi;\mathbb{R}^n)\Big)$, where $H^t_{\mathrm{per,w}}(0,2\pi;\mathbb{R}^n)$ denotes the space $H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$ equipped with the weak topology.

Finally, $\|u_N(\xi)\|_t\lesssim 1$ uniformly over $\xi\in[0,T]$ implies $\sup_{\xi\in[0,T]}\|u(\xi)\|_t^2<\infty$, that is $u\in L^2\Big(0,T;H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$. For $\phi\in\mathcal{E}$ we conclude using the Cauchy-Schwarz inequality and Proposition 4.7 that

$$|\langle(L_t^{(N)*}\phi)(\xi),u_N(\xi)\rangle_{H^{-t},H^t}|\le\|(L_t^{(N)*}\phi)(\xi)\|_{-t}\|u_N(\xi)\|_t\lesssim\|\phi\|_{C([0,T];H^p_{\mathrm{per}})}$$

and

$$|\langle \phi(\xi), f_{(N)}(\xi)\rangle_{H^{-t},H^t}| \le \|\phi\|_{C([0,T];H^p_{\mathrm{per}})}\|f\|_{C([0,T];H^t_{\mathrm{per}})}$$

for any $p \in \mathbb{N}$. The dominated convergence theorem allows us to pass to the limit $N \to \infty$ in the equation

$$\int_0^T \langle (L_t^{(N)*}\phi)(\xi), u_N(\xi)\rangle_{H^{-t},H^t}\,\mathrm{d}\xi = \int_0^T \langle \phi(\xi), f_{(N)}(\xi)\rangle_{H^{-t},H^t}\,\mathrm{d}\xi$$

for all $\phi \in \mathcal{E}$ (see equation (4.10)), so that u is a distributional solution to the equation $Lu = f$. ∎

The next step is to show that the square-integrable distributional solution u found in Theorem 4.20 is continuous. To this end we consider the family of operators

$$(T_k u)(\tau) = \left(\left(I - \frac{1}{k}\partial_\tau^2\right)^{-1} u\right)(\tau) = \frac{1}{\sqrt{2\pi}}\sum_{m\in\mathbb{Z}}\left(1+\frac{m^2}{k}\right)^{-1} u_m \mathrm{e}^{\mathrm{i}m\tau}, \quad k \in \mathbb{N},$$

on the space $L^2\big(0,T;H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ (where $(Tu)(\xi,\tau) = (Tu(\xi))(\tau)$). The continuity of any distributional solution u to equation (4.1) (see Corollary 4.23 below) is deduced from the following convergence result for the sequence $\{T_k u\}_{k\in\mathbb{N}}$.

Lemma 4.21. *Let $u \in L^2\big(0,T;H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ be a distributional solution to the equation $Lu = f$. The functions $T_k u$ have the properties*

(i) $T_k u \in C\big([0,T];H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big) \cap C^1\big([0,T];H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ *for each* $k \in \mathbb{N}$,

(ii) $\|T_k u - u\|_{L^2(0,T;H^t_{\mathrm{per}})} \to 0$ *as* $k \to \infty$,

(iii) $\|LT_k u - f\|_{L^2(0,T;H^t_{\mathrm{per}})} \to 0$ *as* $k \to \infty$.

Proof.

(i) Using Proposition 4.6, we conclude from $u \in L^2\big(0,T;H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ and $Lu = f$ in the distributional sense, that

$$u \in C\big([0,T];H^{t-1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big) \cap C^1\big([0,T];H^{t-2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$$

(see the argument below equation (4.10) in the proof of Theorem 4.19). It follows that

$$\begin{aligned}\|T_k u(\xi_1) - T_k u(\xi_2)\|_{t+1}^2 &= \sum_{m\in\mathbb{Z}} (1+m^2)^{t+1}\left(1+\frac{m^2}{k}\right)^{-2}|u_m(\xi_1) - u_m(\xi_2)|^2\\ &= k^2\sum_{m\in\mathbb{Z}}\frac{(1+m^2)^2}{(k+m^2)^2}(1+m^2)^{t-1}|u_m(\xi_1)-u_m(\xi_2)|^2\\ &\le k^2\|u(\xi_1)-u(\xi_2)\|_{t-1}^2,\end{aligned}$$

for all $\xi_1, \xi_2 \in [0,T]$, that is $T_k u \in C\Big([0,T]; H^{t+1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$. We differentiate $T_k u$ with respect to ξ (note that T_k^{-1} and thus T_k commute with ∂_ξ) and use a similar calculation with $T_k u$ replaced by $\partial_\xi(T_k u)$ to conclude that $\partial_\xi(T_k u) \in C\Big([0,T]; H^{t}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$.

(ii) First we calculate

$$\begin{aligned}
\|T_k u - u\|^2_{L^2(0,T;H^t_{\mathrm{per}})} &= \int_0^T \|T_k u(\xi) - u(\xi)\|_t^2 \,\mathrm{d}\xi \\
&= \int_0^T \sum_{m\in\mathbb{Z}} (1+m^2)^t \left(\left(1 + \frac{m^2}{k}\right)^{-1} - 1\right)^2 |u_m(\xi)|^2 \,\mathrm{d}\xi \\
&= \int_0^T \sum_{m\in\mathbb{Z}} \frac{m^4}{(k+m^2)^2} (1+m^2)^t |u_m(\xi)|^2 \,\mathrm{d}\xi.
\end{aligned}$$

Because

$$\frac{m^4}{(k+m^2)^2}(1+m^2)^t |u_m(\xi)|^2 \le (1+m^2)^t |u_m(\xi)|^2$$

and $u(\xi) \in H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)$ for all $\xi \in [0,T]$, we conclude that

$$\lim_{k\to\infty} \sum_{m\in\mathbb{Z}} \frac{m^4}{(k+m^2)^2}(1+m^2)^t |u_m(\xi)|^2 = 0$$

for all $\xi \in [0,T]$. Moreover, since $\|u(\cdot)\|^2_{H^t} \in L^1(0,T)$, the dominated convergence theorem implies that

$$\lim_{k\to\infty} \|T_k u - u\|^2_{L^2(0,T;H^t_{\mathrm{per}})} = \int_0^T \lim_{k\to\infty} \sum_{m\in\mathbb{Z}} \frac{m^4}{(k+m^2)^2}(1+m^2)^t |u_m(\xi)|^2 \,\mathrm{d}\xi = 0.$$

(iii) For the proof of the third property we consider the commutator

$$[T_k, L]u = T_k L u - L T_k u = T_k f - L T_k u.$$

The calculation in the proof of part (ii) with the functions u and $T_k u$ replaced by respectively f and $T_k f$ shows that

$$T_k f \to f$$

as $k \to \infty$ in $L^2\Big(0,T; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$. In order to prove that $LT_k u \to f$ as $k \to \infty$ in $L^2\Big(0,T; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$, it thus suffices to show that

$$LT_k u - T_k f \to 0$$

as $k \to \infty$ in $L^2\Big(0,T; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$, that is

$$\lim_{k\to\infty} \|[T_k, L]u\|_{L^2(0,T;H^t_{\mathrm{per}})} = 0.$$

The operator T_k commutes with the partial derivative ∂_ξ, so that

$$[T_k, L]u = T_k(A\partial_\tau u) - A\partial_\tau T_k u + T_k\mathcal{L}u - \mathcal{L}T_k u.$$

Next we compute

$$\begin{aligned}
T_k^{-1}(A\partial_\tau T_k u) &= \Big(I - \frac{1}{k}\partial_\tau^2\Big)(A\partial_\tau T_k u) \\
&= A\partial_\tau T_k u - \frac{1}{k}\partial_\tau^2(A\partial_\tau T_k u) \\
&= A\partial_\tau T_k u - \frac{1}{k}(\partial_\tau^2 A\partial_\tau T_k u + 2\partial_\tau A\partial_\tau^2 T_k u + A\partial_\tau^3 T_k u) \\
&= AT_k^{-1}(\partial_\tau T_k u) - \frac{1}{k}(\partial_\tau^2 A\partial_\tau T_k u + 2\partial_\tau A\partial_\tau^2 T_k u).
\end{aligned} \tag{4.12}$$

Considering the second order differential operator F_1 defined by

$$F_1(\xi)v = \partial_\tau^2 A(\xi)\partial_\tau v + 2\partial_\tau A(\xi)\partial_\tau^2 v,$$

we find, since by assumption $t - 2 > \frac{1}{2}$, that

$$\|F_1(\xi)T_k u(\xi) - F_1(\xi)u(\xi)\|_{H^{t-2}} \lesssim \|A(\xi)\|_{H^t}\|T_k u(\xi) - u(\xi)\|_{H^t}.$$

By part (ii) it follows that $F_1 T_k u \to F_1 u$ as $k \to \infty$ in $L^2\Big(0, T; H^{t-2}_{\mathrm{per}}(0, 2\pi; \mathbb{R}^n)\Big)$. Using (4.12) we conclude that

$$\begin{aligned}
&\|T_k(A\partial_\tau u) - A\partial_\tau T_k u\|^2_{L^2(0,T;H^t_{\mathrm{per}})} \\
&= \int_0^T \Big\|T_k\Big(A(\xi)\partial_\tau u(\xi)\Big) - A(\xi)\partial_\tau T_k u(\xi)\Big\|_t^2 \,\mathrm{d}\xi \\
&= \int_0^T \Big\|T_k\Big(A(\xi)\partial_\tau T_k^{-1}T_k u(\xi)\Big) - A(\xi)\partial_\tau T_k u(\xi)\Big\|_t^2 \,\mathrm{d}\xi \\
&= \int_0^T \Big\|T_k\Big(A(\xi)T_k^{-1}\partial_\tau T_k u(\xi)\Big) - A(\xi)\partial_\tau T_k u(\xi)\Big\|_t^2 \,\mathrm{d}\xi \\
&= \int_0^T \Big\|T_k\Big(T_k^{-1}\Big(A(\xi)\partial_\tau T_k u(\xi)\Big) + \frac{1}{k}F_1(\xi)T_k u(\xi)\Big) - A(\xi)\partial_\tau T_k u(\xi)\Big\|_t^2 \,\mathrm{d}\xi \\
&= \int_0^T \Big\|T_k\Big(\frac{1}{k}F_1(\xi)T_k u(\xi)\Big)\Big\|_t^2 \,\mathrm{d}\xi \\
&= \int_0^T \sum_{m\in\mathbb{Z}} (1+m^2)^t\Big(1 + \frac{m^2}{k}\Big)^{-2}\frac{1}{k^2}\Big|[F_1(\xi)T_k u(\xi)]_m\Big|^2 \,\mathrm{d}\xi \\
&= \int_0^T \sum_{m\in\mathbb{Z}} \frac{(1+m^2)^2}{(k+m^2)^2}(1+m^2)^{t-2}\Big|[F_1(\xi)T_k u(\xi)]_m\Big|^2 \,\mathrm{d}\xi,
\end{aligned}$$

in which we have used the notation $[\cdot]_m$ to denote the m-th Fourier component of a function.

Since

$$\lim_{k\to\infty} \|F_1 T_k u - F_1 u\|_{L^2(0,T;H^{t-2}_{\mathrm{per}})} = 0$$

(see above), that is $\|F_1(\cdot)T_k u(\cdot)\|_{t-2}^2 \to \|F_1(\cdot)u(\cdot)\|_{t-2}^2$ as $k \to \infty$ in $L^1(0,T)$, Theorem 4.9 of Brezis [3] implies the existence of a subsequence of $\{T_k\}_{k\in\mathbb{N}}$ (which we still write as $\{T_k\}_{k\in\mathbb{N}}$) and a function $h_1 \in L^1(0,T)$ such that

$$\lim_{k\to\infty} \|F_1(\xi)T_k u(\xi)\|_{t-2}^2 = \|F_1(\xi)u(\xi)\|_{t-2}^2$$

and $\|F_1(\xi)T_k u(\xi)\|_{t-2}^2 \leq h_1(\xi)$ for almost every $\xi \in (0,T)$. For each such ξ it follows that

$$\Big\{(1+m^2)^{t-2}\Big|[F_1(\xi)T_k u(\xi)]_m\Big|^2\Big\}_{m\in\mathbb{Z}} \to \Big\{(1+m^2)^{t-2}\Big|[F_1(\xi)u(\xi)]_m\Big|^2\Big\}_{m\in\mathbb{Z}}$$

in $\ell^1(\mathbb{Z})$ and again Theorem 4.9 of Brezis [3] implies that there exist a further subsequence of $\{T_k\}_{k\in\mathbb{N}}$ (which we still write as $\{T_k\}_{k\in\mathbb{N}}$) and a sequence $h_2 \in \ell^1(\mathbb{Z})$ such that

$$\lim_{k\to\infty} (1+m^2)^{t-2}\Big|[F_1(\xi)T_k u(\xi)]_m\Big|^2 = (1+m^2)^{s-2}\Big|[F_1(\xi)u(\xi)]_m\Big|^2$$

and

$$(1+m^2)^{t-2}\Big|[F_1(\xi)T_k u(\xi)]_m\Big|^2 \leq h_2(m)$$

for almost every $m \in \mathbb{Z}$. The dominated convergence theorem therefore yields

$$\begin{aligned}
&\lim_{k\to\infty} \Big\|T_k(A\partial_\tau u) - A\partial_\tau T_k u\Big\|^2_{L^2(0,T;H^t_{\mathrm{per}})} \\
&\quad = \int_0^T \lim_{k\to\infty} \sum_{m\in\mathbb{Z}} \frac{(1+m^2)^2}{(k+m^2)^2}(1+m^2)^{t-2}\Big|[F_1(\xi)T_k u(\xi)]_m\Big|^2 \,\mathrm{d}\xi \\
&\quad = \int_0^T \sum_{m\in\mathbb{Z}} \lim_{k\to\infty} \frac{(1+m^2)^2}{(k+m^2)^2}(1+m^2)^{t-2}\Big|[F_1(\xi)T_k u(\xi)]_m\Big|^2 \,\mathrm{d}\xi \\
&\quad = 0
\end{aligned}$$

up to subsequences.

Finally note that

$$T_k^{-1}(\mathcal{L}T_k u) = \mathcal{L}T_k^{-1}(T_k u) - \frac{1}{k}[\partial_\tau^2, \mathcal{L}]T_k u$$

and

$$\|[\partial_\tau^2, \mathcal{L}(\xi)]v\|_{t-2} \lesssim \|v\|_t$$

uniformly over $\xi \in [0,T]$, since

$$\mathcal{L} \in C\Big([0,T]; \mathcal{L}\Big(H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)\Big) \cap C\Big([0,T]; \mathcal{L}\Big(H^{t-2}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)\Big)$$

due to the assumption $t-2>\frac{1}{2}$. The above calculation with $A(\xi)\partial_\tau$ and $F_1(\xi)$ replaced by $\mathcal{L}(\xi)$ and

$$F_2(\xi)=[\partial_\tau^2,\mathcal{L}(\xi)]$$

therefore shows that

$$\lim_{k\to\infty}\|T_k\mathcal{L}u-\mathcal{L}T_ku\|^2_{L^2(0,T;H^t)}=0.$$

■

Corollary 4.22. *Let $u\in L^2\big(0,T;H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ be a distributional solution to the equation $Lu=f$. The sequence $\{T_ku\}_{k\in\mathbb{N}}$ converges in $C\big([0,T];H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$.*

Proof. Take $k_1,k_2>0$ with $k_1\neq k_2$. Applying Lemma 4.9 with $\xi_0=0$ to the function $T_{k_1}u-T_{k_2}u$ yields

$$\|T_{k_1}u(\xi)-T_{k_2}u(\xi)\|_t^2\lesssim\|\underbrace{T_{k_1}u(0)-T_{k_2}u(0)}_{=0}\|_t+\int_0^\xi\|(LT_{k_1}u)(r)-(LT_{k_2}u)(r)\|_t^2\,\mathrm{d}r$$

and taking the supremum over $[0,T]$ we find that

$$\|T_{k_1}u-T_{k_2}u\|^2_{C([0,T];H^t_{\mathrm{per}})}\lesssim\|LT_{k_1}u-LT_{k_2}u\|^2_{L^2(0,T;H^t_{\mathrm{per}})}. \tag{4.13}$$

Due to Lemma 4.21 there exists a subsequence of $\{T_k\}_{k\in\mathbb{N}}$ (which we still write as $\{T_k\}_{k\in\mathbb{N}}$) such that

$$\|LT_ku-f\|_{L^2(0,T;H^t_{\mathrm{per}})}\to 0$$

as $k\to\infty$. The sequence $\{LT_ku\}_{k\in\mathbb{N}}$ is thus a Cauchy sequence and hence convergent in $L^2\big(0,T;H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$. Inequality (4.13) therefore implies that the sequence $\{T_ku\}_{k\in\mathbb{N}}$ is a Cauchy sequence and hence convergent in $C\big([0,T];H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$. ■

Corollary 4.23. *Let $u\in L^2\big(0,T;H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ be a distributional solution to the equation $Lu=f$. It follows that $u\in C\big([0,T];H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$.*

Proof. Due to Corollary 4.22 there exists $w\in C\big([0,T];H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$ such that

$$\|T_ku-w\|_{C([0,T];H^t_{\mathrm{per}})}\to 0$$

as $k\to\infty$, from which it follows that

$$\|T_ku-w\|_{L^2(0,T;H^t_{\mathrm{per}})}\to 0$$

as $k\to\infty$. Since

$$\|T_ku-u\|_{L^2(0,T;H^t_{\mathrm{per}})}\to 0$$

as $k\to\infty$ due to Lemma 4.21(ii), the uniqueness of limits implies that $u=w$, that is $u\in C\big([0,T];H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\big)$. ■

We now solve the initial value problem (4.1), (4.2) starting with the case $h = 0$.

Theorem 4.24. *For each $f \in C\big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\big)$ the initial value problem*

$$\begin{aligned} Lu &= f, \\ u(0) &= 0 \end{aligned}$$

posesses a solution

$$u \in C\Big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{t-1}_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

Proof. The existence of a distributional solution $u \in L^2\big(0,T; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\big)$ to the equation $Lu = f$ with

$$\int_0^T \langle (L_t^*\phi)(\xi), u(\xi)\rangle_{H^{-t},H^t} \, \mathrm{d}\xi = \int_0^T \langle \phi(\xi), f(\xi)\rangle_{H^{-t},H^t} \, \mathrm{d}\xi$$

for each $\phi \in \mathcal{E}$ was shown in Theorem 4.20. Corollary 4.23 therefore implies that

$$u \in C\Big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big)$$

and using Proposition 4.6, we conclude that

$$u \in C\Big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{t-1}_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big)$$

(see the argument below equation (4.10) in the proof of Theorem 4.19) with $u(0) = 0$. ■

The general result follows from the special case $h = 0$.

Corollary 4.25. *For each $f \in C\big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\big)$ and each $h \in H^{t+1}_{\text{per}}(0,2\pi;\mathbb{R}^n)$ the initial value problem (4.1), (4.2) posesses a unique solution*

$$u \in C\Big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{t-1}_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big).$$

Proof. Define

$$\tilde{f} = f - Lh = f - A\partial_\tau h - \mathcal{L}h$$

and note that $\tilde{f} \in C\big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\big)$. Due to Theorem 4.24 there exists a function

$$\tilde{u} \in C\Big([0,T]; H^t_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{t-1}_{\text{per}}(0,2\pi;\mathbb{R}^n)\Big)$$

such that respectively

$$\begin{aligned} L\tilde{u} &= \tilde{f}, \\ \tilde{u}(0) &= 0 \end{aligned}$$

and

$$\begin{aligned} L(\tilde{u} + h) &= f, \\ \tilde{u}(0) &= 0, \end{aligned}$$

that is $u = \tilde{u} + h$ solves the initial value problem (4.1), (4.2).

It remains to prove the uniqueness of the solution. Let u_1 and u_2 be solutions to the initial value problem (4.1), (4.2). We have that

$$u_1 - u_2 \in C\Big([0,T]; H^t_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big) \cap C^1\Big([0,T]; H^{t-1}_{\mathrm{per}}(0,2\pi;\mathbb{R}^n)\Big)$$

and Lemma 4.9 with $\xi_0 = 0$ implies

$$\begin{aligned} \|u_1(\xi) - u_2(\xi)\|^2_{t-1} &\lesssim \|\underbrace{u_1(0) - u_2(0)}_{=0}\|^2_{t-1} + \int_0^\xi \Big\|\Big(L(u_1 - u_2)\Big)(r)\Big\|^2_{t-1} \,\mathrm{d}r \\ &= \int_0^\xi \|(Lu_1)(r) - (Lu_2)(r)\|^2_{t-1} \,\mathrm{d}r \\ &= \int_0^\xi \|f(r) - f(r)\|^2_{t-1} \,\mathrm{d}r \\ &= 0, \end{aligned}$$

that is $u_1 = u_2$. ■

4.2 Solution of the nonlinear problem

In this section we construct exponentially small solutions to equations (2.55)–(2.57) on the exponentially large interval $[-\mathrm{e}^{-c^\star/2\mu}, \mathrm{e}^{-c^\star/2\mu}]$. In particular we prove the following theorem.

Theorem 4.26. *Equations (2.55)–(2.57) admit reversible solutions* $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ *in*

$$C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{sh}}) \times C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{wh}}) \times C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}^{s+1}_{\mathrm{c}}),$$

parameterised by $w_0 \in \mathcal{X}^{s+2}_{\mathrm{c}}$ *with* $\|w_0\|_{\mathcal{X}^{s+2}_{\mathrm{c}}} \leq \mu^2 \mathrm{e}^{-c^\star/2\mu}$ *and* $R_{\mathrm{c}} w_0 = w_0$*, such that*

$$\|Z^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})}, \|r^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})}, \|w^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1}_{\mathrm{c}})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

These solutions satisfy

$$\begin{aligned} \|Z^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} &\lesssim \mu^2 \mathrm{e}^{-c^\star/2\mu}, \\ \|r^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})} &\lesssim \mu \mathrm{e}^{-c^\star/2\mu}, \\ \|w^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1}_{\mathrm{c}})} &\lesssim \mu^\delta \mathrm{e}^{-c^\star/2\mu} \end{aligned}$$

for arbitrary but fixed $\delta \in [0, \frac{1}{2})$.

The iteration scheme

In order to prove Theorem 4.26 we use the reformulations (2.58), (2.59) of equations (2.55), (2.56) as integral equations and employ the following iteration scheme. Choose $w_0 \in \mathcal{X}_\mathrm{c}^{s+2}$ with $\|w_0\|_{\mathcal{X}_\mathrm{c}^{s+2}} \leq \mu^2 \mathrm{e}^{-c^\star/2\mu}$, $R_\mathrm{c} w_0 = w_0$, let $j \in \mathbb{N}_0$ and suppose that

$$Z_{(j)} \in C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{sh}), \quad r_{(j)} \in C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{wh}), \quad w_{(j)} \in C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{c}^{s+2})$$

satisfy

$$\|Z_{(j)}\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{sh})} + \|r_{(j)}\|_{C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{wh})} + \|w_{(j)}\|_{C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{c}^{s+1})} \lesssim \mathrm{e}^{-c^\star/2\mu}$$

with $R_\mathrm{wh} r_{(j)}(0) = r_{(j)}(0)$ and $R_\mathrm{c} w_{(j)}(0) = w_{(j)}(0)$. We define

$$Z_{(j+1)} \in C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{sh})$$

by

$$\begin{aligned} Z_{(j+1)}(\xi) = &\int_{-\mathrm{e}^{c^\star/2\mu}}^{\xi} \mathrm{e}^{(\xi-t)L_\mathrm{sh}^\mu} P_\mathrm{ss}(G_{1(j)}^\mu(t) + G_{2(j)}^\mu(t))\, \mathrm{d}t \\ &- \int_{\xi}^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{(\xi-t)L_\mathrm{sh}^\mu} P_\mathrm{su}(G_{1(j)}^\mu(t) + G_{2(j)}^\mu(t))\, \mathrm{d}t, \end{aligned} \tag{4.14}$$

where the projections P_ss, P_su are defined in Section 2.3 (see the comments below equations (2.58), (2.59)). Similarly

$$r_{(j+1)} \in C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{wh})$$

is defined by

$$\begin{aligned} r_{(j+1)}(\xi) = \sum_{i=1}^{2} \Bigg(&\int_0^{\xi} \langle (f_{3(j)}^\mu(t) + f_{4(j)}^\mu(t)), (S_i^\mu)^*(t) \rangle \, \mathrm{d}t \Bigg) S_i^\mu(\xi) \\ &- \left(\int_{\xi}^{\mathrm{e}^{c^\star/2\mu}} \langle (f_{3(j)}^\mu(t) + f_{4(j)}^\mu(t)), (U_i^\mu)^*(t) \rangle \, \mathrm{d}t \right) U_i^\mu(\xi), \end{aligned} \tag{4.15}$$

where $\{S_1^\mu, S_2^\mu, U_1^\mu, U_2^\mu\}$ is the fundamental system constructed in Lemma 2.19. Finally, let

$$w_{(j+1)} \in C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{c}^{s+2})$$

be the unique solution of the initial value problem

$$\partial_\xi w_{(j+1)} = L_\mathrm{c}^\mu w_{(j)} + P_\mathrm{c}\Big(\hat{g}_{1(j)}^\mu \partial_\tau w_{(j+1)}\Big) + \mu b(z_{(j)}, w_{(j+1)}) + \hat{g}_{2(j)}^\mu + \hat{h}_{(j)}^\mu, \tag{4.16}$$

$$w_{(j+1)}(0) = w_0 \tag{4.17}$$

(see Lemma 4.28 below). Here we have abbreviated $z_{(j)} = p^\mu + r_{(j)}$ and $G_{1(j)}^\mu = G_1^\mu(Z_{(j)}, p^\mu + \mathcal{E}_\mathrm{wh} r_{(j)}, \mathcal{E}_\mathrm{c} w_{(j)})$ with similar abbreviations for the remaining nonlinearities. Note that $R_\mathrm{wh} r_{(j+1)}(0) = r_{(j+1)}(0)$, since the functions U_1^μ, U_2^μ are reversible, and $R_\mathrm{c} w_{(j+1)}(0) = w_{(j+1)}(0)$ by construction. In Lemmata 4.30 and 4.31 below we prove the convergence of the iteration scheme (4.14)–(4.16) and show that its limit $(Z^\star, r^\star, w^\star)$ lies in

$$\begin{aligned}&\Big\{(Z,r,w) \in C\Big([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{sh}\Big) \times C\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{wh}\Big) \times C\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_\mathrm{c}^{s+1}\Big) \colon \\ &\quad \|Z\|_{C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{sh})},\ \|r\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{wh})},\ \|w\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{c}^{s+1})} \leq \mathrm{e}^{-c^\star/2\mu}, \\ &\quad R_\mathrm{wh} r(0) = r(0),\ R_\mathrm{c} w(0) = w(0)\Big\}.\end{aligned}$$

We verify *a posteriori* that $Z^\star(\xi) = R_\mathrm{sh} Z^\star(-\xi)$; extending $r^\star$ and $w^\star$ by reversibility to $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$ (see also the comments below inequalities (2.60)) thus establishes Theorem 4.26.

Definition 4.27. The set

$$W_\mathrm{loc}^\mathrm{cs} = \Big\{\Big(Z_{w_0}^\star(0), r_{w_0}^\star(0), w_{w_0}^\star(0)\Big) \colon w_0 \in \mathcal{X}_\mathrm{c}^{s+2},\ R_\mathrm{c} w_0 = w_0,\ \|w_0\|_{s+2} \leq \mu^2 \mathrm{e}^{-c^\star/2\mu}\Big\}$$

is called a *local centre-stable manifold for solutions to equations* (2.55)–(2.57) *at* $\xi = 0$.

The first step is evidently to show that the iterate $w_{(j+1)}$ in equation (4.16) is well-defined. To this end we reformulate the initial value problem (4.16), (4.17) as a Cauchy problem of the type (4.1), (4.2) and apply the existence and uniqueness result in Corollary 4.25. We start by writing equation (4.16) as

$$\partial_\xi w_{(j+1)} = L_\mathrm{c}^\mu w_{(j+1)} + \hat{g}_{1(j)}^\mu \partial_\tau w_{(j+1)} - P_\mathrm{h}\Big(\hat{g}_{1(j)}^\mu \partial_\tau w_{(j+1)}\Big) + \mu b(z_{(j)}, w_{(j+1)}) + \hat{g}_{2(j)}^\mu + \hat{h}_{(j)}^\mu \tag{4.18}$$

and note that $\hat{g}_{1(j)}^\mu$ is of the form

$$(U^\mu)^{-1} \hat{S}^\mu(0)^{1/2}\Big(\hat{A}_{(j)}^\mu - \hat{A}^\mu(0)\Big)\hat{S}^\mu(0)^{-1/2} U^\mu. \tag{4.19}$$

Here $U^\mu = U^\varepsilon|_{\varepsilon=\mu^2} \colon \mathcal{X}^t \to \mathcal{X}^t$ $(t \geq 0)$ is the Kato similarity transformation introduced in Section 2.1 (see the comments below equations (2.8), (2.9)); the symbols $\hat{A}$, $\hat{S}$ denote the matrix-valued functions which are obtained by applying the rescaling in Section 2.3 (see equations (2.42), (2.43)) and the changes of variables in Lemmata 2.10, 2.14 and 2.15 and Theorem 2.16 to the functions

$$(z,q) \mapsto \tilde{A}^\varepsilon(z + U^\varepsilon q), \quad (z,q) \mapsto S^\varepsilon(z + U^\varepsilon q).$$

In particular we have that

$$\hat{A}^\mu(0) = \tilde{A}^\varepsilon(0)|_{\varepsilon=\mu^2} = \tilde{A}^\mu(0), \quad \hat{S}^\mu(0) = S^\varepsilon(0)|_{\varepsilon=\mu^2}$$

in the notation of Section 2.3.

The similarity transformation U^μ in the formula (4.19) above does not necessarily preserve the symmetrisability of the matrix-valued function

$$\hat{S}^\mu(0)^{1/2}\Big(\hat{A}^\mu_{(j)} - \hat{A}^\mu(0)\Big)\hat{S}^\mu(0)^{-1/2}.$$

In order to obtain a symmetrisable system in the sense of Assumptions 4.1 we therefore apply the invertible operator $U^\mu\colon \mathcal{X}^{s+1} \to \mathcal{X}^{s+1}$ to equation (4.18) and the initial condition (4.17). The result is the initial value problem

$$\begin{aligned}
\partial_\xi U^\mu w_{(j+1)} &= L^\mu U^\mu w_{(j+1)} + (\hat{S}^\mu_0)^{\frac{1}{2}}\hat{A}^\mu_{1(j)}(\hat{S}^\mu_0)^{-\frac{1}{2}}\partial_\tau U^\mu w_{(j+1)} - U^\mu P_{\mathrm{h}}\Big(\hat{g}^\mu_{1(j)}\partial_\tau w_{(j+1)}\Big)\\
&\qquad + \mu U^\mu b(z_{(j)}, w_{(j+1)}) + U^\mu \hat{g}^\mu_{2(j)} + U^\mu \hat{h}^\mu_{(j)},\\
&= (\hat{S}^\mu_0)^{\frac{1}{2}}\hat{A}^\mu_{(j)}(\hat{S}^\mu_0)^{-\frac{1}{2}}\partial_\tau U^\mu w_{(j+1)} + (\hat{S}^\mu_0)^{\frac{1}{2}}\partial B^\mu[0](\hat{S}^\mu_0)^{-\frac{1}{2}}U^\mu w_{(j+1)}\\
&\qquad - U^\mu P_{\mathrm{h}}\Big(\hat{g}^\mu_{1(j)}\partial_\tau w_{(j+1)}\Big) + \mu U^\mu b(z_{(j)}, w_{(j+1)}) + U^\mu \hat{g}^\mu_{2(j)} + U^\mu \hat{h}^\mu_{(j)}, \qquad (4.20)\\
U^\mu w_{(j+1)}(0) &= U^\mu w_0, \qquad (4.21)
\end{aligned}$$

where we recall that $L^\mu_{\mathrm{c}} = (U^\mu)^{-1}L^\mu U^\mu|_{\mathcal{X}^{s+1}_{\mathrm{c}}}$. Here we have abbreviated $\hat{A}^\mu_0 = \hat{A}^\mu(0)$ and $\hat{S}^\mu_0 = \hat{S}^\mu(0)$ as well as

$$\hat{A}^\mu_1(Z,z,w) = \hat{A}^\mu(Z,z,w) - \hat{A}^\mu(0), \quad \hat{S}^\mu_1(Z,z,w) = \hat{S}^\mu(Z,z,w) - \hat{S}^\mu(0).$$

In particular the matrix

$$\hat{S}^\mu(Z,z,w)\hat{A}^\mu(Z,z,w) = \hat{S}^\mu_0\hat{A}^\mu_0 + \hat{S}^\mu_0\hat{A}^\mu_1(Z,z,w) + \hat{S}^\mu_1(Z,z,w)\hat{A}^\mu(Z,z,w), \qquad (4.22)$$

and thus the matrix

$$\hat{S}^\mu_0\hat{A}^\mu_1(Z,z,w) + \hat{S}^\mu_1(Z,z,w)\hat{A}^\mu(Z,z,w), \qquad (4.23)$$

is symmetric.

The existence and uniqueness result for the initial value problem (4.20), (4.21) is given by the following lemma. The corresponding result for the initial value problem (4.16), (4.17) follows from the fact that the operator $U^\mu \in \mathcal{L}(\mathcal{X}^{s+2})$ is invertible.

Lemma 4.28. *The initial value problem (4.20), (4.21) posesses a unique solution*

$$U^\mu w_{(j+1)} \in C\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}^{s+2}\Big) \cap C^1\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}^{s+1}\Big).$$

Proof. The initial value problem (4.20), (4.21) is a linear Cauchy problem of the type (4.1), (4.2) for the function $U^\mu w_{(j+1)} \in \mathcal{X}^{s+2}$, where $s+2 > \frac{5}{2}$. The matrix $(\hat{S}^\mu_0)^{\frac{1}{2}}\hat{A}^\mu_{(j)}(\hat{S}^\mu_0)^{-\frac{1}{2}}$ is symmetrisable with symmetriser $(\hat{S}^\mu_0)^{-\frac{1}{2}}\hat{S}^\mu_{(j)}(\hat{S}^\mu_0)^{-\frac{1}{2}}$ which is uniformly positive definite due to Assumptions 1(iii). The mapping $\xi \mapsto \mathcal{L}(\xi)$ with

$$\mathcal{L}(\xi)\phi = (\hat{S}^\mu_0)^{\frac{1}{2}}\partial B^\mu[0](\hat{S}^\mu_0)^{-\frac{1}{2}}\phi - U^\mu P_{\mathrm{h}}\Big(\hat{g}^\mu_{1(j)}(\xi)(U^\mu)^{-1}\partial_\tau\phi\Big) + \mu U^\mu b(z_{(j)}(\xi), (U^\mu)^{-1}\phi)$$

lies in $C\Big([0, e^{c^\star/2\mu}]; \mathcal{L}(\mathcal{X}^t)\Big)$ for all $t > \frac{1}{2}$, since $U^\mu \in \mathcal{L}(\mathcal{X}^t)$ is invertible for each t, the projection P_h has finite-dimensional range and $z_{(j)}(\xi)$ lies in the finite-dimensional space $\mathcal{X}_\mathrm{wh}$ and thus in $\mathcal{X}^t$ for arbitrary t; finally the nonlinearity $\tilde{g}_2^\mu$ is an analytic mapping $\mathcal{X}^t \to \mathcal{X}_\mathrm{c}^t \subset \mathcal{X}^t$ for all $t > \frac{1}{2}$ so that $b(z_{(j)}(\xi), (U^\mu)^{-1}\phi) \in \mathcal{X}^t$ whenever $\phi \in \mathcal{X}^t$. Corollary 4.25 therefore implies that there is a unique solution

$$U^\mu w_{(j+1)} \in C\Big([0, e^{c^\star/2\mu}]; \mathcal{X}^{s+2}\Big) \cap C^1\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}^{s+1}\Big)$$

to the linear Cauchy problem (4.20), (4.21). ■

Reformulation of the centre part

The proof of satisfactory energy estimates below requires an appropriate further reformulation of equation (4.20). Since our aim is to establish an estimate for the norm $\|w_{(j)}\|_{s+1}$ we first differentiate (4.20) k-times with respect to τ (where $k = 0, \ldots, s+1$) to obtain

$$\begin{aligned}\partial_\xi \partial_\tau^k U^\mu w_{(j+1)} = {} & L^\mu \partial_\tau^k U^\mu w_{(j+1)} + \partial_\tau^k \Big((\hat{S}_0^\mu)^{\frac{1}{2}} \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau U^\mu w_{(j+1)}\Big) \\ & - \partial_\tau^k U^\mu P_\mathrm{h}\Big(\hat{g}_{1(j)}^\mu \partial_\tau w_{(j+1)}\Big) + \mu \partial_\tau^k U^\mu b(z_{(j)}, w_{(j+1)}) \\ & + \partial_\tau^k U^\mu \hat{g}_{2(j)}^\mu + \partial_\tau^k U^\mu \hat{h}_{(j)}^\mu. \end{aligned} \tag{4.24}$$

The term $\partial_\tau^k \Big((\hat{S}_0^\mu)^{\frac{1}{2}} \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau U^\mu w_{(j+1)}\Big)$ includes derivatives of order $(k+1)$ and can be written as

$$\begin{aligned}&\partial_\tau^k \Big((\hat{S}_0^\mu)^{\frac{1}{2}} \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau U^\mu w_{(j+1)}\Big) \\ &\quad = \partial_\tau^k \Big((\hat{S}_0^\mu)^{\frac{1}{2}} \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau U^\mu w_{(j+1)}\Big) - (\hat{S}_0^\mu)^{\frac{1}{2}} \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau^{k+1} U^\mu w_{(j+1)} \\ &\qquad + (\hat{S}_0^\mu)^{\frac{1}{2}} \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau^{k+1} U^\mu w_{(j+1)}, \end{aligned} \tag{4.25}$$

in which the final term is symmetrisable via the near-identity symmetrisation $I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}}$ (the difference of the first two terms on the right-hand side of equation (4.25) is handled below using the Moser estimate given in Lemma 4.5).

The next step is therefore to multiply equation (4.24) by the invertible matrix

$I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}}$ to obtain

$$\begin{aligned}
&\Big(I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}}\Big) \partial_\xi \partial_\tau^k U^\mu w_{(j+1)} \\
&\quad = L^\mu \partial_\tau^k U^\mu w_{(j+1)} + \partial_\tau^k \Big((\hat{S}_0^\mu)^{\frac{1}{2}} \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau U^\mu w_{(j+1)} \Big) \\
&\qquad\quad + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu \hat{A}_0^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau^{k+1} U^\mu w_{(j+1)} + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu \partial_\tau^k \Big(\hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau U^\mu w_{(j+1)} \Big) \\
&\qquad\quad + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu \partial B^\mu[0] (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau^k U^\mu w_{(j+1)} \\
&\qquad\quad + \Big(I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}}\Big) \Big(- \partial_\tau^k U^\mu P_{\mathrm{h}} \Big(\hat{g}_{1(j)}^\mu \partial_\tau w_{(j+1)} \Big) + \mu \partial_\tau^k U^\mu b(z_{(j)}, w_{(j+1)}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \partial_\tau^k U^\mu \hat{g}_{2(j)}^\mu + \partial_\tau^k U^\mu \hat{h}_{(j)}^\mu \Big) \\
&\quad = L^\mu \partial_\tau^k U^\mu w_{(j+1)} + \underline{\hat{g}}_{1(j)}^\mu \partial_\tau^{k+1} U^\mu w_{(j+1)} \\
&\qquad\quad + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{(j)}^\mu \Big(\partial_\tau^k \Big(\hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau U^\mu w_{(j+1)} \Big) - \hat{A}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau^{k+1} U^\mu w_{(j+1)} \Big) \\
&\qquad\quad + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu \partial B^\mu[0] (\hat{S}_0^\mu)^{-\frac{1}{2}} \partial_\tau^k U^\mu w_{(j+1)} \\
&\qquad\quad + \Big(I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu (\hat{S}_0^\mu)^{-\frac{1}{2}}\Big) \Big(- \partial_\tau^k U^\mu P_{\mathrm{h}} \Big(\hat{g}_{1(j)}^\mu \partial_\tau w_{(j+1)} \Big) + \mu \partial_\tau^k U^\mu b(z_{(j)}, w_{(j+1)}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \partial_\tau^k U^\mu \hat{g}_{2(j)}^\mu + \partial_\tau^k U^\mu \hat{h}_{(j)}^\mu \Big),
\end{aligned} \tag{4.26}$$

where $\underline{\hat{g}}_1^\mu$ is the symmetric matrix-valued function

$$\underline{\hat{g}}_1^\mu(Z,z,w) = (\hat{S}_0^\mu)^{-\frac{1}{2}} \Big(\hat{S}_0^\mu \hat{A}_1^\mu(Z,z,w) + \hat{S}_1^\mu(Z,z,w) \hat{A}^\mu(Z,z,w) \Big) (\hat{S}_0^\mu)^{-\frac{1}{2}}$$

(the term $\partial_\tau^k U^\mu P_{\mathrm{h}}\Big(\hat{g}_{1(j)}^\mu \partial_\tau w_{(j+1)}\Big)$ also includes derivatives of order $(k+1)$ and is handled below using the fact that P_{h} is smoothing). Note that $\hat{A}_1^\mu$, $\hat{S}_1^\mu$ and $\underline{\hat{g}}_1^\mu$ satisfy the estimates

$$\|\hat{A}_1^\mu(Z,z,w)\|_{H_{\mathrm{per}}^{s+1}}, \|\hat{S}_1^\mu(Z,z,w)\|_{H_{\mathrm{per}}^{s+1}}, \|\underline{\hat{g}}_1^\mu(Z,z,w)\|_{H_{\mathrm{per}}^{s+1}} = \mathcal{O}(\mu \|(Z,z,w)\|_{\mathcal{X}_{\mathrm{sh}} \times \mathcal{X}_{\mathrm{wh}} \times \mathcal{X}_{\mathrm{c}}^{s+1}}), \tag{4.27}$$

$$\|\mathrm{d}\hat{A}_1^\mu[Z,z,w]\|_{\mathcal{L}(\mathcal{X}^{s+1}; H_{\mathrm{per}}^{s+1})}, \|\mathrm{d}\hat{S}_1^\mu[Z,z,w]\|_{\mathcal{L}(\mathcal{X}^{s+1}; H_{\mathrm{per}}^{s+1})}, \|\mathrm{d}\underline{\hat{g}}_1^\mu[Z,z,w]\|_{\mathcal{L}(\mathcal{X}^{s+1}; H_{\mathrm{per}}^{s+1})} = \mathcal{O}(\mu). \tag{4.28}$$

Finally we diagonalise the operator L^μ in equation (4.26) via an appropriate change of bases in the mode m subspaces E_m. Since a diagonal operator with purely imaginary eigenvalues is skew-Hermitian, this procedure eliminates the term

$$\Big\langle L^\mu \partial_\tau^k U^\mu w_{(j+1)}(\xi), \partial_\tau^k U^\mu w_{(j+1)}(\xi) \Big\rangle_{L^2},$$

which is not dominated by $\mu^\delta \|w_{(j+1)}(\xi)\|_{s+1}^2$ for any $\delta > 0$, in our subsequent analysis. The proof of the following result is presented in Section 4.3.

Lemma 4.29. *There exists an invertible bounded linear operator $Q^\mu \colon \mathcal{X}^t \to \mathcal{X}^t$ (for all $t > \frac{1}{2}$) such that*

$$((Q^\mu)^{-1} L^\mu Q^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} (D^\mu_m u_m) \mathrm{e}^{\mathrm{i}m\tau},$$

where D^μ_m are diagonal matrices depending analytically upon μ in a neighbourhood of the origin which is independent of $m \in \mathbb{Z}$. Moreover

$$Q^\mu = T^\mu + R^\mu_1, \quad (Q^\mu)^{-1} = (T^\mu)^{-1} + R^\mu_2,$$

where $T^\mu \colon \mathcal{X}^t \to \mathcal{X}^t$ is unitary and $R^\mu_1, R^\mu_2 \colon \mathcal{X}^t \to \mathcal{X}^{t+1}$ are invertible bounded linear operators. The operators T^μ and R^μ_1, R^μ_2 depend analytically upon μ.

With respect to the bases used in the proof of Lemma 4.29 (see Section 4.3) the m-th Fourier components $P_m w$, $P_m U^\mu w$ of the functions $w \in \mathcal{X}^t_\mathrm{c}$ and $U^\mu w \in \mathcal{X}^t$ have a coordinate vector representation of the form

$$(w_{m,1}, \ldots, w_{m,r_m-\sigma_m}, 0, \ldots, 0)^\top. \tag{4.29}$$

The coordinate vector (4.29) has $\sigma_m + 2\delta_{m,\pm 1}$ zero entries, where we recall that σ_m denotes the number of strongly hyperbolic mode m eigenvalues of the operator L^0, and the (appropriately numbered) mode m eigenvalues $\lambda^\mu_{m,1}, \ldots, \lambda^\mu_{m,r_m-\sigma_m}$ are purely imaginary. The infinite-dimensional change of coordinates $U^\mu w = Q^\mu W$ (applied to the sequence $\{w_{(j)}\}_{j\in\mathbb{N}_0}$) therefore transforms equation (4.26) into

$$\begin{aligned}
&\Big(I + (Q^\mu)^{-1}(\hat{S}^\mu_0)^{-\frac{1}{2}} \hat{S}^\mu_{1(j)} (\hat{S}^\mu_0)^{-\frac{1}{2}} Q^\mu\Big) \partial_\xi \partial^k_\tau W_{(j+1)} \\
&\quad = (Q^\mu)^{-1} L^\mu Q^\mu \partial^k_\tau W_{(j+1)} + (T^\mu)^* \underline{\hat{g}}^\mu_{1(j)} T^\mu \partial^{k+1}_\tau W_{(j+1)} \\
&\qquad + (T^\mu)^* \underline{\hat{g}}^\mu_{1(j)} R^\mu_1 \partial^{k+1}_\tau W_{(j+1)} + R^\mu_2 \underline{\hat{g}}^\mu_{1(j)} Q^\mu \partial^{k+1}_\tau W_{(j+1)} \\
&\qquad + (Q^\mu)^{-1}(\hat{S}^\mu_0)^{-\frac{1}{2}} \hat{S}^\mu_{(j)} \Big(\partial^k_\tau \Big(\hat{A}^\mu_{1(j)} (\hat{S}^\mu_0)^{-\frac{1}{2}} Q^\mu \partial_\tau W_{(j+1)} \Big) \\
&\qquad\qquad - \hat{A}^\mu_{1(j)} (\hat{S}^\mu_0)^{-\frac{1}{2}} Q^\mu \partial^{k+1}_\tau W_{(j+1)} \Big) \\
&\qquad + (Q^\mu)^{-1}(\hat{S}^\mu_0)^{-\frac{1}{2}} \hat{S}^\mu_{1(j)} \partial B^\mu[0] (\hat{S}^\mu_0)^{-\frac{1}{2}} Q^\mu \partial^k_\tau W_{(j+1)} \\
&\qquad + (Q^\mu)^{-1} \Big(I + (\hat{S}^\mu_0)^{-\frac{1}{2}} \hat{S}^\mu_{1(j)} (\hat{S}^\mu_0)^{-\frac{1}{2}} \Big) \Big(- \partial^k_\tau U^\mu P_\mathrm{h} \Big(\hat{g}^\mu_{1(j)} (U^\mu)^{-1} Q^\mu \partial_\tau W_{(j+1)} \Big) \\
&\qquad\qquad + \mu \partial^k_\tau U^\mu b(z_{(j)}, (U^\mu)^{-1} Q^\mu W_{(j+1)}) \\
&\qquad\qquad + \partial^k_\tau U^\mu \hat{g}^\mu_{2(j)} + \partial^k_\tau U^\mu \hat{h}^\mu_{(j)} \Big),
\end{aligned} \tag{4.30}$$

where the subscript (j) now indicates evaluation at $(Z_{(j)}, p^\mu + r_{(j)}, (U^\mu)^{-1} Q^\mu W_{(j)})$ and the operator $(Q^\mu)^{-1} L^\mu Q^\mu \in \mathcal{L}(\mathcal{X}^t)$ is diagonal in the sense that

$$(Q^\mu)^{-1} L^\mu Q^\mu W = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}} \Big(\mathrm{diag}(\lambda^\mu_{m,1}, \ldots, \lambda^\mu_{m,r_m-\sigma_m}, 0, \ldots, 0) W_m \Big) \mathrm{e}^{\mathrm{i}m\tau}$$

in which $W_m = (w_{m,1}, \dots, w_{m,r_m-\sigma_m}, 0, \dots, 0)$ (see Lemma 4.29 and the coordinate representation (4.29)).

Estimates for the iteration scheme

To prove that the sequence $\{(Z_{(j)}, r_{(j)}, W_{(j)})\}_{j\in\mathbb{N}_0}$ converges in

$$C\Big([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{sh}}\Big) \times C\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{wh}}\Big) \times C\Big([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}^{s+1}\Big)$$

to a limit

$$(Z^\star, r^\star, W^\star) \in \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0),$$

we show that it lies in $\overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0)$ for appropriately chosen initial values. The result is summarised in the following lemma. For notational simplicity we use the abbreviations

$$\|(Z,r,W)\| = \|Z\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} + \|r\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})} + \|W\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1})},$$

and

$$\|(Z,W)\| = \|Z\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} + \|W\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1})}.$$

Lemma 4.30. *Suppose that $Z_{(0)} \in C([-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{sh}})$, $r_{(0)} \in C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}_{\mathrm{wh}})$ and $W_{(0)} \in C([0, \mathrm{e}^{c^\star/2\mu}]; \mathcal{X}^{s+1})$ satisfy*

$$\begin{aligned}
\|Z_{(0)}\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} &\leq \mathrm{e}^{-c^\star/2\mu},\\
\|r_{(0)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})} &\leq \mathrm{e}^{-c^\star/2\mu},\\
\|W_{(0)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1})} &\leq \mathrm{e}^{-c^\star/2\mu}
\end{aligned}$$

and $\|W_0\|_{s+1} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}$. The estimates

$$\begin{aligned}
\sup_{j\in\mathbb{N}_0} \|Z_{(j)}\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} &\leq \mathrm{e}^{-c^\star/2\mu},\\
\sup_{j\in\mathbb{N}_0} \|r_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})} &\leq \mathrm{e}^{-c^\star/2\mu}
\end{aligned}$$

and

$$\sup_{j\in\mathbb{N}_0} \|W_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1})} \leq \mathrm{e}^{-c^\star/2\mu}$$

hold for all sufficiently small values of μ.

In order to prove Lemma 4.30 we work inductively and demonstrate that

$$\|Z_{(j+1)}\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} \lesssim \mu^2 \mathrm{e}^{-c^\star/2\mu}, \tag{4.31}$$

$$\|r_{(j+1)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}, \tag{4.32}$$

$$\|W_{(j+1)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1})} \lesssim \sqrt{\mu}|\log\mu|\mathrm{e}^{-c^\star/2\mu} \tag{4.33}$$

whenever

$$\|Z_{(j)}\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} \le \mathrm{e}^{-c^\star/2\mu}, \tag{4.34}$$

$$\|r_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})} \le \mathrm{e}^{-c^\star/2\mu}, \tag{4.35}$$

$$\|W_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1})} \le \mathrm{e}^{-c^\star/2\mu} \tag{4.36}$$

for any $j \in \mathbb{N}_0$.

We begin with the result (4.31) for $Z_{(j+1)}$ on $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$. Using the estimates for G_1, G_2 in Propositions 2.17 and 2.18, we find that

$$\begin{aligned}
|G^\mu_{1(j)}(\xi)| &= |G^\mu_1(Z_{(j)}(\xi),(p^\mu+\mathcal{E}_{\mathrm{wh}}r_{(j)})(\xi),(\mathcal{E}_{\mathrm{c}}(U^\mu)^{-1}Q^\mu W_{(j)})(\xi))|\\
&\lesssim \mu|p^\mu+\mathcal{E}_{\mathrm{wh}}r_{(j)}|\|(Z_{(j)},\mathcal{E}_{\mathrm{c}}(U^\mu)^{-1}Q^\mu W_{(j)})\| + \mu\|(Z_{(j)},\mathcal{E}_{\mathrm{c}}(U^\mu)^{-1}Q^\mu W_{(j)})\|^2\\
&\qquad + |p^\mu+\mathcal{E}_{\mathrm{wh}}r_{(j)}|^3\|(Z_{(j)},\mathcal{E}_{\mathrm{c}}(U^\mu)^{-1}Q^\mu W_{(j)})\|\\
&\lesssim \mu\mathrm{e}^{-c^\star/\mu} + \mu^2\mathrm{e}^{-c^\star/2\mu}\\
&\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}
\end{aligned}$$

as well as

$$|G^\mu_{2(j)}(\xi)| \lesssim \mu^2\mathrm{e}^{-c^\star/\mu}.$$

We additionally use the estimate (2.60) for the projections onto the stable and unstable subspaces of L^μ_{sh} to conclude that

$$\begin{aligned}
|Z_{(j+1)}(\xi)| &\le \int_{-\mathrm{e}^{c^\star/2\mu}}^{\xi} \|\mathrm{e}^{(\xi-t)L^\mu_{\mathrm{sh}}}P_{\mathrm{ss}}\|_{\mathcal{L}(\mathcal{X}_{\mathrm{sh}})}|G^\mu_{1(j)}(t)+G^\mu_{2(j)}(t)|\,\mathrm{d}t\\
&\qquad + \int_{\xi}^{\mathrm{e}^{c^\star/2\mu}} \|\mathrm{e}^{(\xi-t)L^\mu_{\mathrm{sh}}}P_{\mathrm{su}}\|_{\mathcal{L}(\mathcal{X}_{\mathrm{sh}})}|G^\mu_{1(j)}(t)+G^\mu_{2(j)}(t)|\,\mathrm{d}t,\\
&\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}\left(\int_{-\mathrm{e}^{c^\star/2\mu}}^{\xi}\mathrm{e}^{\omega_1 t}\,\mathrm{d}t\,\mathrm{e}^{-\omega_1\xi} + \int_{\xi}^{\mathrm{e}^{c^\star/2\mu}}\mathrm{e}^{-\omega_2 t}\,\mathrm{d}t\,\mathrm{e}^{\omega_2\xi}\right)\\
&\le \mu^2\mathrm{e}^{-c^\star/2\mu}\left(\frac{1}{\omega_1}+\frac{1}{\omega_2}\right)\\
&\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu},
\end{aligned}$$

which is inequality (4.31).

The next step is the estimate (4.32) for $r_{(j+1)}$ on $\xi \in [0,\mathrm{e}^{c^\star/2\mu}]$. Proposition 2.18 implies that

$$|f^\mu_{3(j)}(\xi)| = |f^\mu_3(r_{(j)}(\xi))| \lesssim |r_{(j)}(\xi)|^2 \le \mathrm{e}^{-c^\star/\mu}$$

and that

$$\begin{aligned}
|f^\mu_{4(j)}(\xi)| &= |f^\mu_4(Z_{(j)}(\xi),r_{(j)}(\xi),(U^\mu)^{-1}Q^\mu W_{(j)}(\xi))|\\
&\lesssim \mu^3\|(Z_{(j)},r_{(j)},(U^\mu)^{-1}Q^\mu W_{(j)})\|\\
&\le \mu^3\mathrm{e}^{-c^\star/2\mu}.
\end{aligned}$$

The estimates for $\{S_1^\mu, S_2^\mu, U_1^\mu, U_2^\mu\}$ (see Lemma 2.19) therefore yield

$$
\begin{aligned}
|r_{(j+1)}(\xi)| &\leq \sum_{i=1}^{2} \int_0^\xi \Big(|f_{3(j)}^\mu(t)| + |f_{4(j)}^\mu(t)|\Big)|(S_i^\mu)^*(t)| \,\mathrm{d}t\, |S_i^\mu(\xi)| \\
&\qquad\qquad + \int_\xi^{\mathrm{e}^{c^\star/2\mu}} \Big(|f_{3(j)}^\mu(t)| + |f_{4(j)}^\mu(t)|\Big)|(U_i^\mu)^*(t)| \,\mathrm{d}t\, |U_i^\mu(\xi)| \\
&\lesssim \mu^3 \mathrm{e}^{-c^\star/2\mu}\Big(\int_0^\xi \mathrm{e}^{\lambda^\mu t}\,\mathrm{d}t\,\mathrm{e}^{-\lambda^\mu \xi} + \int_\xi^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\lambda^\mu t}\,\mathrm{d}t\,\mathrm{e}^{\lambda^\mu\xi}\Big) \\
&\leq \mu \mathrm{e}^{-c^\star/2\mu}(2 - \mathrm{e}^{-\lambda^\mu\xi}) \\
&\lesssim \mu \mathrm{e}^{-c^\star/2\mu},
\end{aligned}
$$

and we conclude inequality (4.32).

Next we establish the estimate (4.33) for $W_{(j+1)}$ on $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$. Here it is sufficient to assume that $\|W_0\|_{s+1} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}$, which follows from the inequality

$$
\|W_0\|_{s+1} \leq \|(Q^\mu)^{-1}U^\mu\|^{-1}\|w_0\|_{\mathcal{X}_c^{s+1}} \leq \|(Q^\mu)^{-1}U^\mu\|^{-1}\mu^2\mathrm{e}^{-c^\star/2\mu}
$$

(see the assumption on the initial value w_0 below Theorem 4.26). The stronger requirement $\|w_0\|_{\mathcal{X}_c^{s+2}} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}$ is however needed in showing that $\{(Z_{(j)}, r_{(j)}, W_{(j)})\}_{j\in\mathbb{N}_0}$ is a Cauchy sequence below. We first take the complex inner product of equation (4.30) with $\partial_\tau^k W_{(j+1)}(\xi)$ and integrate over $(0,2\pi)$ with respect to τ to obtain

$$
\begin{aligned}
&\int_0^{2\pi} \Big\langle \Big(I + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\Big)\partial_\xi\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau \\
&= \int_0^{2\pi} \langle (Q^\mu)^{-1}L^\mu Q^\mu \partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \\
&\quad + \int_0^{2\pi} \langle (T^\mu)^* \underline{g}_{1(j)}^\mu(\xi) T^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau \\
&\quad + \int_0^{2\pi} \Big\langle (T^\mu)^* \underline{\hat{g}}_{1(j)}^\mu(\xi) R_1^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau \\
&\quad + \int_0^{2\pi} \Big\langle R_2^\mu \underline{\hat{g}}_{1(j)}^\mu(\xi) Q^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau \\
&\quad + \int_0^{2\pi} \Big\langle (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{(j)}^\mu(\xi) \\
&\qquad\qquad \times \Big(\partial_\tau^k\Big(\hat{A}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau W_{(j+1)}(\xi)\Big) \\
&\qquad\qquad\qquad - \hat{A}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau^{k+1} W_{(j+1)}(\xi)\Big), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau \\
&\quad + \int_0^{2\pi} \langle (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\xi)\partial B^\mu[0](\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \\
&\quad - \int_0^{2\pi} \Big\langle (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}) \\
&\qquad\qquad \times \partial_\tau^k U^\mu P_{\mathrm{h}}\Big(\hat{g}_{1(j)}^\mu(\xi)(U^\mu)^{-1}Q^\mu\partial_\tau W_{(j+1)}(\xi)\Big), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau \\
&\quad + \mu\int_0^{2\pi} \langle (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}) \\
&\qquad\qquad \times \partial_\tau^k U^\mu b(z_{(j)}, (U^\mu)^{-1}Q^\mu W_{(j+1)})(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \\
&\quad + \int_0^{2\pi} \langle (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu \hat{g}_{2(j)}^\mu(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \\
&\quad + \int_0^{2\pi} \langle (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu \hat{h}_{(j)}^\mu(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau. \qquad (4.37)
\end{aligned}
$$

We proceed by estimating the right-hand side of equation (4.37) term by term.

- First we note that $W \in U^\mu \mathcal{X}_{\mathrm{c}}^t$, $t = 0, \ldots, s+1$, can be written as

  $$W(\tau) = \sum_{m\in\mathbb{Z}} W_m \mathrm{e}^{\mathrm{i}m\tau}$$

  with $W_m = (W_{m,1}, \ldots, W_{m,r_m-\sigma_m}, 0, \ldots, 0)^\top$ (see the comments below equation (4.30)). It follows that

  $$\int_0^{2\pi} \langle (Q^\mu)^{-1}L^\mu Q^\mu W, W\rangle \,\mathrm{d}\tau = 2\pi \sum_{m\in\mathbb{Z}} \langle \mathrm{diag}(\lambda_{m,1}^\mu, \ldots, \lambda_{m,r_m-\sigma_m}^\mu, 0, \ldots, 0)W_m, W_m\rangle.$$

The reality conditions (2.19) imply that

$$\begin{aligned}
&\sum_{m\in\mathbb{Z}\setminus\{0\}} \langle \operatorname{diag}(\lambda^\mu_{m,1},\ldots,\lambda^\mu_{m,r_m-\sigma_m},0,\ldots,0)W_m, W_m\rangle \\
&\quad = \sum_{m\in\mathbb{N}} 2\operatorname{Re}\langle \operatorname{diag}(\lambda^\mu_{m,1},\ldots,\lambda^\mu_{m,r_m-\sigma_m},0,\ldots,0)W_m, W_m\rangle \\
&\quad = 0,
\end{aligned}$$

since the entries of the matrix $\operatorname{diag}(\lambda^\mu_{m,1},\ldots,\lambda^\mu_{m,r_m-\sigma_m},0,\ldots,0)$ are purely imaginary so that it is skew-Hermitian. A similar computation, which takes the reality conditions (2.21) into account, yields

$$\langle \operatorname{diag}(\lambda^\mu_{0,1},\ldots,\lambda^\mu_{0,n-\sigma_0},0,\ldots,0)W_0, W_0\rangle = 0$$

and we conclude that

$$\int_0^{2\pi} \langle (Q^\mu)^{-1}L^\mu Q^\mu \partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau = 0. \tag{4.38}$$

- Next we use integration by parts and the fact that the operator $(T^\mu)^*\underline{\hat{g}}^\mu_{1(j)}(\xi)T^\mu$ is self-adjoint to compute

$$\begin{aligned}
&\int_0^{2\pi} \langle (T^\mu)^*\underline{\hat{g}}^\mu_{1(j)}(\xi)T^\mu\partial_\tau^{k+1}W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \\
&\quad = -\frac{1}{2}\int_0^{2\pi} \langle \partial_\tau^k W_{(j+1)}(\xi), (T^\mu)^*\partial_\tau\underline{\hat{g}}^\mu_{1(j)}(\xi)T^\mu\partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau.
\end{aligned}$$

Hölder's inequality, the Sobolev embedding theorem and the estimate for $\underline{\hat{g}}_1$ given by equation (4.27) thus imply that

$$\begin{aligned}
&\left|\int_0^{2\pi} \langle \partial_\tau^k W_{(j+1)}(\xi), (T^\mu)^*\partial_\tau\underline{\hat{g}}^\mu_{1(j)}(\xi)T^\mu\partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau\right| \\
&\quad \lesssim \|\partial_\tau\underline{\hat{g}}^\mu_{1(j)}(\xi)\|_{L^\infty}\|\partial_\tau^k W_{(j+1)}(\xi)\|_0^2 \\
&\quad \lesssim \|\underline{\hat{g}}^\mu_{1(j)}(\xi)\|_{s+1}\|W_{(j+1)}(\xi)\|_{s+1}^2 \\
&\quad \lesssim \mu\|(Z_{(j)}, p^\mu + r_{(j)}, (U^\mu)^{-1}Q^\mu W_{(j)})\|\|W_{(j+1)}(\xi)\|_{s+1}^2 \\
&\quad \lesssim \mu(\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-\eta\lambda^\mu\xi})\|W_{(j+1)}(\xi)\|_{s+1}^2.
\end{aligned} \tag{4.39}$$

- The next step is to compute an estimate for the third and the fourth term on the right-hand side of equation (4.37) which involve the smoothing operators R_1^μ and R_2^μ. We first write

$$\begin{aligned}
&\int_0^{2\pi} \langle (T^\mu)^*\underline{\hat{g}}^\mu_{1(j)}(\xi)R_1^\mu\partial_\tau^{k+1}W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \\
&\quad = -\int_0^{2\pi} \Big\langle \partial_\tau^k\Big((T^\mu)^*\underline{\hat{g}}^\mu_{1(j)}(\xi)R_1^\mu\partial_\tau W_{(j+1)}(\xi)\Big) \\
&\qquad\qquad - (T^\mu)^*\underline{\hat{g}}^\mu_{1(j)}(\xi)R_1^\mu\partial_\tau^{k+1}W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau \\
&\qquad + \int_0^{2\pi} \Big\langle \partial_\tau^k\Big((T^\mu)^*\underline{\hat{g}}^\mu_{1(j)}(\xi)R_1^\mu\partial_\tau W_{(j+1)}(\xi)\Big), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle \,\mathrm{d}\tau.
\end{aligned}$$

Using Lemma 4.5, we find that

$$
\begin{aligned}
&\Big| \int_0^{2\pi} \Big\langle \partial_\tau^k \Big((T^\mu)^* \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \partial_\tau W_{(j+1)}(\xi) \Big) \\
&\qquad\qquad - (T^\mu)^* \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi) \Big\rangle \, \mathrm{d}\tau \Big| \\
&\quad \leq \Big\| \partial_\tau^k \Big((T^\mu)^* \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \partial_\tau W_{(j+1)}(\xi) \Big) - (T^\mu)^* \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi) \Big\|_0 \\
&\qquad \times \| \partial_\tau^k W_{(j+1)}(\xi) \|_0 \\
&\quad \lesssim \Big(\| \partial_\tau W_{(j+1)}(\xi) \|_{L^\infty} \| (T^\mu)^* \partial_\tau^k \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \|_0 \\
&\qquad + \| \partial_\tau^k W_{(j+1)}(\xi) \|_0 \| (T^\mu)^* \partial_\tau \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \|_{L^\infty} \Big) \| \partial_\tau^k W_{(j+1)}(\xi) \|_0 \\
&\quad \lesssim \| \underline{\hat{g}}^\mu_{1(j)}(\xi) \|_{s+1} \| W_{(j+1)}(\xi) \|_{s+1}^2 \\
&\quad \lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \| W_{(j+1)}(\xi) \|_{s+1}^2
\end{aligned}
$$

and the smoothing property of the operator R_1^μ implies that

$$
\begin{aligned}
&\Big| \int_0^{2\pi} \Big\langle \partial_\tau^k \Big((T^\mu)^* \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \partial_\tau W_{(j+1)}(\xi) \Big), \partial_\tau^k W_{(j+1)}(\xi) \Big\rangle \, \mathrm{d}\tau \Big| \\
&\quad \lesssim \| (T^\mu)^* \underline{\hat{g}}^\mu_{1(j)}(\xi) R_1^\mu \partial_\tau W_{(j+1)}(\xi) \|_{s+1} \| W_{(j+1)}(\xi) \|_{s+1} \\
&\quad \lesssim \| \underline{\hat{g}}^\mu_{1(j)}(\xi) \|_{s+1} \| R_1^\mu \partial_\tau W_{(j+1)}(\xi) \|_{s+1} \| W_{(j+1)}(\xi) \|_{s+1} \\
&\quad \lesssim \| \underline{\hat{g}}^\mu_{1(j)}(\xi) \|_{s+1} \| \partial_\tau W_{(j+1)}(\xi) \|_{s} \| W_{(j+1)}(\xi) \|_{s+1} \\
&\quad \lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \| W_{(j+1)}(\xi) \|_{s+1}^2.
\end{aligned} \tag{4.40}
$$

A similar computation shows that

$$
\begin{aligned}
&\Big| \int_0^{2\pi} \Big\langle R_2^\mu \underline{\hat{g}}^\mu_{1(j)}(\xi) Q^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi) \Big\rangle \, \mathrm{d}\tau \Big| \\
&\quad \lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \| W_{(j+1)}(\xi) \|_{s+1}^2.
\end{aligned} \tag{4.41}
$$

- Next we use Lemma 4.5 together with the estimate for $\hat{A}_1$ given by equation (4.27) to conclude that

$$\begin{aligned}
&\Big| \int_0^{2\pi} \Big\langle (Q^\mu)^{-1} (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{(j)}^\mu(\xi) \\
&\qquad \times \Big(\partial_\tau^k \Big(\hat{A}_{1(j)}^\mu(\xi) (\hat{S}_0^\mu)^{-\frac{1}{2}} Q^\mu \partial_\tau W_{(j+1)}(\xi) \Big) \\
&\qquad\qquad - \hat{A}_{1(j)}^\mu(\xi) (\hat{S}_0^\mu)^{-\frac{1}{2}} Q^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi) \Big), \partial_\tau^k W_{(j+1)}(\xi) \Big\rangle \, \mathrm{d}\tau \Big| \\
&\lesssim \Big\| \partial_\tau^k \Big(\hat{A}_{1(j)}^\mu(\xi) (\hat{S}_0^\mu)^{-\frac{1}{2}} Q^\mu \partial_\tau W_{(j+1)}(\xi) \Big) - \hat{A}_{1(j)}^\mu(\xi) (\hat{S}_0^\mu)^{-\frac{1}{2}} Q^\mu \partial_\tau^{k+1} W_{(j+1)}(\xi) \Big\|_0 \\
&\qquad \times \|\partial_\tau^k W_{(j+1)}(\xi)\|_0 \\
&\lesssim \Big(\|\partial_\tau W_{(j+1)}(\xi)\|_{L^\infty} \|\hat{A}_{1(j)}^\mu(\xi)\|_0 + \|\partial_\tau^k W_{(j+1)}(\xi)\|_0 \|\hat{A}_{1(j)}^\mu(\xi)\|_{L^\infty} \Big) \\
&\qquad \times \|\partial_\tau^k W_{(j+1)}(\xi)\|_0 \\
&\lesssim \|\hat{A}_{1(j)}^\mu(\xi)\|_{s+1} \|W_{(j+1)}(\xi)\|_{s+1}^2 \\
&\lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \|W_{(j+1)}(\xi)\|_{s+1}^2. \qquad (4.42)
\end{aligned}$$

- Similarly, the estimate for $\hat{S}_1$ given by equation (4.27) yields

$$\begin{aligned}
&\Big| \int_0^{2\pi} \langle (Q^\mu)^{-1} (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu(\xi) \partial B^\mu[0] (\hat{S}_0^\mu)^{-\frac{1}{2}} Q^\mu \partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi) \rangle \, \mathrm{d}\tau \Big| \\
&\lesssim \|\hat{S}_{1(j)}^\mu(\xi)\|_{s+1} \|W_{(j+1)}(\xi)\|_{s+1}^2 \\
&\lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \|W_{(j+1)}(\xi)\|_{s+1}^2. \qquad (4.43)
\end{aligned}$$

- The projection P_h has finite-dimensional range and is therefore infinitely smoothing, that is

$$\|P_\mathrm{h} u\|_{t_1} \lesssim \|u\|_{t_2}, \quad t_1 \geq t_2.$$

We additionally use the estimate for the nonlinearity $\hat{g}_1$ in Proposition 2.17 to find that

$$\begin{aligned}
&\Big| \int_0^{2\pi} \Big\langle (Q^\mu)^{-1} (I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu(\xi) (\hat{S}_0^\mu)^{-\frac{1}{2}}) \\
&\qquad \times \partial_\tau^k U^\mu P_\mathrm{h} \Big(\hat{g}_{1(j)}^\mu(\xi) (U^\mu)^{-1} Q^\mu \partial_\tau W_{(j+1)}(\xi) \Big), \partial_\tau^k W_{(j+1)}(\xi) \Big\rangle \, \mathrm{d}\tau \Big| \\
&\lesssim \Big\| P_\mathrm{h} \Big(\hat{g}_{1(j)}^\mu(\xi) (U^\mu)^{-1} Q^\mu \partial_\tau W_{(j+1)}(\xi) \Big) \Big\|_{s+1} \|W_{(j+1)}(\xi)\|_{s+1} \\
&\lesssim \|\hat{g}_{1(j)}^\mu(\xi) (U^\mu)^{-1} Q^\mu \partial_\tau W_{(j+1)}(\xi)\|_s \|W_{(j+1)}(\xi)\|_{s+1} \\
&\lesssim \|\hat{g}_{1(j)}^\mu(\xi)\|_{s+1} \|\partial_\tau W_{(j+1)}(\xi)\|_s \|W_{(j+1)}(\xi)\|_{s+1} \\
&\lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \|W_{(j+1)}(\xi)\|_{s+1}^2. \qquad (4.44)
\end{aligned}$$

- Next we note that the bilinear mapping b satisfies

$$\|b(z,w)\|_{s+1} \lesssim |z|\|w\|_{s+1},$$

so that

$$\begin{aligned}
&\left|\mu \int_0^{2\pi} \langle (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}) \right. \\
&\qquad\qquad \left. \times\, \partial_\tau^k U^\mu b(z_{(j)}, (U^\mu)^{-1} Q^\mu W_{(j+1)})(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \right| \\
&\quad \lesssim \mu \|b(z_{(j)}, (U^\mu)^{-1} Q^\mu W_{(j+1)})(\xi)\|_{s+1} \|W_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \mu |z_{(j)}(\xi)| \|(U^\mu)^{-1} Q^\mu W_{(j+1)}(\xi)\|_{s+1} \|W_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \mu |p^\mu(\xi) + r_{(j)}(\xi)| \|W_{(j+1)}(\xi)\|_{s+1}^2 \\
&\quad \lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \|W_{(j+1)}(\xi)\|_{s+1}^2.
\end{aligned} \tag{4.45}$$

- It follows from the estimate for $\hat{g}_2$ given by Proposition 2.17 that

$$\begin{aligned}
&\left| \int_0^{2\pi} \langle (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}) \partial_\tau^k U^\mu \hat{g}_{2(j)}^\mu(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \right| \\
&\quad \lesssim \|\hat{g}_{2(j)}^\mu(\xi)\|_{s+1} \|W_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \Big(|p^\mu(\xi) + r_{(j)}(\xi)| |(p^\mu(\xi) + r_{(j)}(\xi), \mu)|^2 \|(Z_{(j)}(\xi), (U^\mu)^{-1} Q^\mu W_{(j)}(\xi))\| \\
&\qquad\qquad + \mu \|(Z_{(j)}(\xi), (U^\mu)^{-1} Q^\mu W_{(j)}(\xi))\|^2 \Big) \|W_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \Big(\mu^2 (\mu \mathrm{e}^{-\eta\lambda^\mu\xi} + \mathrm{e}^{-c^\star/2\mu}) \mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-c^\star/\mu} \Big) \|W_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \mu (\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-\eta\lambda^\mu\xi} \mathrm{e}^{-c^\star/2\mu}) \|W_{(j+1)}(\xi)\|_{s+1}.
\end{aligned} \tag{4.46}$$

- Finally

$$\begin{aligned}
&\left| \int_0^{2\pi} \langle (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}}) \partial_\tau^k U^\mu \hat{h}_{(j)}^\mu(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \right| \\
&\quad \lesssim \|\hat{h}_{(j)}^\mu(\xi)\|_{s+1} \|W_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \mu^2 \mathrm{e}^{-c^\star/\mu} \|W_{(j+1)}(\xi)\|_{s+1} \\
&\quad \leq \mu (\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-\eta\lambda^\mu\xi} \mathrm{e}^{-c^\star/2\mu}) \|W_{(j+1)}(\xi)\|_{s+1}
\end{aligned} \tag{4.47}$$

due to the fact that $\hat{h}$ is exponentially small (see Proposition 2.18).

Inserting the estimates (4.38)–(4.47) into (4.37) we find that

$$\begin{aligned}
&\int_0^{2\pi} \langle (I + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}} \hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac{1}{2}} Q^\mu) \partial_\xi \partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle \,\mathrm{d}\tau \\
&\quad \lesssim \mu (\mathrm{e}^{-c^\star/\mu} + \mu^2 \mathrm{e}^{-\eta\lambda^\mu\xi} \mathrm{e}^{-c^\star/2\mu}) \|W_{(j+1)}(\xi)\|_{s+1} + \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu\xi}) \|W_{(j+1)}(\xi)\|_{s+1}^2.
\end{aligned} \tag{4.48}$$

In the next step we derive the estimate

$$\|W_{(j+1)}(\xi)\|_{s+1}^2 \lesssim \|W_{(j+1)}(0)\|_{s+1}^2 + \int_0^\xi \Big(\mu(\mathrm{e}^{-c^\star/\mu} + \mu^2\mathrm{e}^{-\eta\lambda^\mu t}\mathrm{e}^{-c^\star/2\mu})\|W_{(j+1)}(t)\|_{s+1} + \mu(\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-\eta\lambda^\mu t})\|W_{(j+1)}(t)\|_{s+1}^2\Big)\,\mathrm{d}t \tag{4.49}$$

from inequality (4.48). Note that

$$\begin{aligned}\int_0^{2\pi} &\langle (I + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac12}Q^\mu)\partial_\xi\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\ &= \frac12\frac{\mathrm{d}}{\mathrm{d}\xi}\|W_{j+1}(\xi)\|_{s+1}^2\\ &\qquad + \int_0^{2\pi}\langle (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac12}Q^\mu\partial_\xi\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau.\end{aligned}$$

Our strategy is therefore to appropriately reformulate the remainder term

$$\int_0^{2\pi}\langle (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac12}Q^\mu\partial_\xi\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau$$

in order to obtain an estimate for $\frac{\mathrm{d}}{\mathrm{d}\xi}\|W_{j+1}(\xi)\|_{s+1}^2$ and thus, via integration over $(0,\xi)$, for $\|W_{j+1}(\xi)\|_{s+1}^2$. Here we make use of the inequality

$$\|W\|_t^2 \lesssim \int_0^{2\pi} |\langle \hat{S}_{(j)}^\mu W, W\rangle|\,\mathrm{d}\tau \lesssim \|W\|_t^2$$

for $W \in \mathcal{X}^t$, $t = 0,\dots,s+1$ (see Assumptions 1(iii)).

We first note that the operator $(Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac12}Q^\mu$ decomposes into a sum of a self-adjoint and a smoothing operator, namely

$$(Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac12}Q^\mu = (T^\mu)^*(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac12}T^\mu + R_{(j)}^\mu,$$

where

$$R_{(j)}^\mu = R_2^\mu(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac12}Q^\mu + (T^\mu)^*(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac12}R_1^\mu$$

satisfies

$$\|R_{(j)}^\mu u\|_{t+1} \lesssim \mu(\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-\eta\lambda^\mu\xi})\|u\|_t \tag{4.50}$$

due to the smoothing property of the operators R_1^μ, R_2^μ and the estimate for $\hat{S}_1$ (equation (4.27)). We compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\xi}&\int_0^{2\pi}\langle (T^\mu)^*(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac12}T^\mu\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\ &= \int_0^{2\pi}\Big\langle \partial_\xi\Big((T^\mu)^*(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac12}T^\mu\Big)\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\Big\rangle\,\mathrm{d}\tau\\ &\qquad + 2\int_0^{2\pi}\langle (T^\mu)^*(\hat{S}_0^\mu)^{-\frac12}\hat{S}_{1(j)}^\mu(\xi)(\hat{S}_0^\mu)^{-\frac12}T^\mu\partial_\xi\partial_\tau^k W_{(j+1)}(\xi), \partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\end{aligned} \tag{4.51}$$

and conclude that

$$\begin{aligned}
&\Big|\int_0^{2\pi}\Big\langle \partial_\xi\Big((T^\mu)^*(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}T^\mu\Big)\partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\Big\rangle\,\mathrm{d}\tau\Big|\\
&\quad\lesssim \Big\|\partial_\xi\Big((T^\mu)^*(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}T^\mu\Big)\Big\|_{L^\infty}\|W_{(j+1)}\|_{s+1}^2\\
&\quad\lesssim \Big\|\partial_\xi\Big((T^\mu)^*(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}T^\mu\Big)\Big\|_{s}\|W_{(j+1)}\|_{s+1}^2\\
&\quad= \Big\|\mathrm{d}\Big((T^\mu)^*(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}T^\mu\Big)(\partial_\xi Z_{(j)},\partial_\xi p^\mu+\partial_\xi r_{(j)},(U^\mu)^{-1}Q^\mu\partial_\xi W_j)\Big\|_s\\
&\qquad\quad\times\|W_{(j+1)}\|_{s+1}^2\\
&\quad\lesssim \mu(\mathrm{e}^{-c^\star/2\mu}+\mu\mathrm{e}^{-\eta\lambda^\mu\xi})\|W_{(j+1)}(\xi)\|_{s+1}^2.
\end{aligned}\tag{4.52}$$

Here we have made use of the estimate for $\mathrm{d}\hat S_1$ given in equation (4.28) which implies that

$$\Big\|\mathrm{d}\Big((T^\mu)^*(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}T^\mu\Big)\Big\|_{\mathcal{L}(\mathcal{X}^{s+1};H^{s+1}_{\mathrm{per}})}\lesssim\mu,$$

and the inequalities

$$|\partial_\xi Z_{(j)}|\lesssim \mathrm{e}^{-c^\star/2\mu},\quad |\partial_\xi r_{(j)}|\lesssim \mathrm{e}^{-c^\star/2\mu},\quad \|\partial_\xi W_{(j)}\|_s\lesssim \mathrm{e}^{-c^\star/2\mu},$$

which follow from the induction hypotheses (4.34)–(4.36) together with equation (4.30) with j replaced by $j-1$. Moreover we have that

$$\begin{aligned}
&\int_0^{2\pi}\langle (T^\mu)^*(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}T^\mu\partial_\xi\partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\
&= \int_0^{2\pi}\langle (I+(Q^\mu)^{-1}(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}Q^\mu)\partial_\xi\partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\
&\quad-\int_0^{2\pi}\langle R^\mu_{(j)}\partial_\xi\partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\
&\quad-\frac12\frac{\mathrm{d}}{\mathrm{d}\xi}\int_0^{2\pi}\langle \partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau.
\end{aligned}\tag{4.53}$$

Applying the bounded linear operator $(I+(Q^\mu)^{-1}(\hat S_0^\mu)^{-\frac12}\hat S^\mu_{1(j)}(\xi)(\hat S_0^\mu)^{-\frac12}Q^\mu)^{-1}$ to equation (4.30) and using the same arguments as in the estimates (4.38)–(4.47) together with the smoothing property (4.50) of the operator $R^\mu_{(j)}$ to handle the terms which contain $(k+1)$-th derivates with respect to τ, we find that

$$\begin{aligned}
&\Big|\int_0^{2\pi}\langle R^\mu_{(j)}\partial_\xi\partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\Big|\\
&\quad\lesssim \mu(\mathrm{e}^{-c^\star/\mu}+\mu^2\mathrm{e}^{-\eta\lambda^\mu\xi}\mathrm{e}^{-c^\star/2\mu})\|W_{(j+1)}(\xi)\|_{s+1}+\mu(\mathrm{e}^{-c^\star/2\mu}+\mu\mathrm{e}^{-\eta\lambda^\mu\xi})\|W_{(j+1)}(\xi)\|_{s+1}^2.
\end{aligned}\tag{4.54}$$

Using formulae (4.51), (4.53) and inserting estimates (4.52), (4.54) in inequality (4.48) we find that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\xi}\int_0^{2\pi}\langle \hat{S}^\mu_{(j)}(\hat{S}^\mu_0)^{-\frac{1}{2}}T^\mu\partial_\tau^k W_{(j+1)}(\xi),(\hat{S}^\mu_0)^{-\frac{1}{2}}T^\mu\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\
&=\frac{\mathrm{d}}{\mathrm{d}\xi}\int_0^{2\pi}\langle \partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\
&\qquad+\frac{\mathrm{d}}{\mathrm{d}\xi}\int_0^{2\pi}\langle (T^\mu)^*(\hat{S}^\mu_0)^{-\frac{1}{2}}\hat{S}^\mu_{1(j)}(\xi)(\hat{S}^\mu_0)^{-\frac{1}{2}}T^\mu\partial_\tau^k W_{(j+1)}(\xi),\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\
&\lesssim \mu(\mathrm{e}^{-c^\star/\mu}+\mu^2\mathrm{e}^{-\eta\lambda^\mu\xi}\mathrm{e}^{-c^\star/2\mu})\|W_{(j+1)}(\xi)\|_{s+1}+\mu(\mathrm{e}^{-c^\star/2\mu}+\mu\mathrm{e}^{-\eta\lambda^\mu\xi})\|W_{(j+1)}(\xi)\|_{s+1}^2.
\end{aligned}$$

Integration over $(0,\xi)$, the inequalities

$$\begin{aligned}
\|\partial_\tau^k W_{(j+1)}(\xi)\|_0^2 &\lesssim \|(\hat{S}^\mu_0)^{-\frac{1}{2}}T^\mu\partial_\tau^k W_{(j+1)}(\xi)\|_0^2\\
&\lesssim \int_0^{2\pi}\langle \hat{S}^\mu_{(j)}(\hat{S}^\mu_0)^{-\frac{1}{2}}T^\mu\partial_\tau^k W_{(j+1)}(\xi),(\hat{S}^\mu_0)^{-\frac{1}{2}}T^\mu\partial_\tau^k W_{(j+1)}(\xi)\rangle\,\mathrm{d}\tau\\
&\lesssim \|(\hat{S}^\mu_0)^{-\frac{1}{2}}T^\mu\partial_\tau^k W_{(j+1)}(\xi)\|_0^2\\
&\lesssim \|\partial_\tau^k W_{(j+1)}(\xi)\|_0^2
\end{aligned}$$

and summation over $k\in\{0,\ldots,s+1\}$ yield (4.49).

Our final step in the proof of the estimate (4.33) is the following two-step estimation technique for inequality (4.49) devised by Groves & Schneider [10]. The direct calculation

$$\mu^2\int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t\lesssim \xi\sup_{t\in[0,\mathrm{e}^{c^\star/2\mu}]}\|W_{(j+1)}(t)\|_{s+1}^2$$

leads to a term which is neither dominated by the right-hand side of inequality (4.36) nor has a sufficient power of μ to be absorbed to the left-hand side of inequality (4.49). We therefore split the range of integration into $[0,\xi^\star]$ and $[\xi^\star,\mathrm{e}^{c^\star/2\mu}]$, where

$$\xi^\star=\frac{\alpha|\log\mu|}{\eta\lambda^\mu}$$

that is $\mathrm{e}^{-\eta\lambda^\mu\xi^\star}=\mu^\alpha$ for an appropriately chosen positive constant α. In the first step we prove the desired estimate on the 'short' interval $[0,\xi^\star]$ using Gronwall's inequality. Satisfactory estimates for the integrals over $[\xi^\star,\mathrm{e}^{c^\star/2\mu}]$ are obtained by an optimal choice of α (and hence $\xi^\star$) in the second step.

Applying Young's inequality to the first term under the integral in inequality (4.49) and using the facts that

$$\mu^2\lesssim\lambda^\mu,\quad \mathrm{e}^{-c^\star/2\mu}\lesssim\mu,\quad \mathrm{e}^{-\eta\lambda^\mu\xi}\le 1,\quad \mathrm{e}^{-c^\star/\mu}\lesssim\mu^2\mathrm{e}^{-c^\star/2\mu}$$

for sufficiently small values of μ, we find that

$$\|W_{(j+1)}(\xi)\|_{s+1}^2\le\|W_{(j+1)}(0)\|_{s+1}^2+\int_0^\xi c_1\lambda^\mu\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t+c_2\mu^4\mathrm{e}^{-c^\star/\mu}\xi$$

for some multiplicative constants $c_1, c_2 > 0$ and Gronwall's inequality implies that

$$\|W_{(j+1)}(\xi)\|_{s+1}^2 \le \Big(\|W_{(j+1)}(0)\|_{s+1}^2 + c_2\mu^4 \mathrm{e}^{-c^\star/\mu}\xi\Big)\mathrm{e}^{c_1\lambda^\mu\xi}.$$

Now we restrict to $\xi \in [0, \xi^\star]$ and choose $\alpha = \eta/c_1$, so that

$$\mathrm{e}^{c_1\lambda^\mu\xi^\star} = \mathrm{e}^{|\log\mu|} = \frac{1}{\mu},$$

and therefore

$$\begin{aligned}
\|W_{(j+1)}(\xi)\|_{s+1}^2 &\le \frac{1}{\mu}(\|W_{(j+1)}(0)\|_{s+1}^2 + c_2\mu^4\mathrm{e}^{-c^\star/\mu}\xi^\star) \\
&= \frac{1}{\mu}\Big(\|W_0\|_{s+1}^2 + \frac{c_2\mu^4}{c_1\lambda^\mu}\mathrm{e}^{-c^\star/\mu}|\log\mu|\Big) \\
&\lesssim \frac{1}{\mu}(\|W_0\|_{s+1}^2 + \mu^2\mathrm{e}^{-c^\star/\mu}|\log\mu|) \\
&\lesssim \frac{1}{\mu}(\mu^2\mathrm{e}^{-c^\star/\mu} + \mu^2\mathrm{e}^{-c^\star/\mu}|\log\mu|) \\
&\lesssim \mu|\log\mu|\mathrm{e}^{-c^\star/\mu}.
\end{aligned} \tag{4.55}$$

Inequality (4.55) gives the desired estimate for $\|W_{(j+1)}(\xi)\|_{s+1}$ on the short interval $[0, \xi^\star]$ and we use this to obtain a corresponding estimate on the whole interval $[0, \mathrm{e}^{c^\star/2\mu}]$.

From inequality (4.49) we conclude that

$$
\begin{aligned}
\|W_{(j+1)}(\xi)\|_{s+1}^2 \lesssim \|W_{(j+1)}(0)\|_{s+1}^2 &+ \mu \mathrm{e}^{-c^\star/\mu} \int_0^\xi \|W_{(j+1)}(t)\|_{s+1}\,\mathrm{d}t \\
&+ \mu^3 \mathrm{e}^{-c^\star/2\mu} \int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}\,\mathrm{d}t \\
&+ \mu \mathrm{e}^{-c^\star/2\mu} \int_0^\xi \|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
&+ \mu^2 \int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
\lesssim \mu^2 \mathrm{e}^{-c^\star/\mu} &+ \mu \mathrm{e}^{-c^\star/\mu}\xi \sup_{t\in[0,\xi]} \|W_{(j+1)}(t)\|_{s+1} \\
&+ \mu^3 \mathrm{e}^{-c^\star/2\mu} \int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t \sup_{t\in[0,\xi]} \|W_{(j+1)}(t)\|_{s+1} \\
&+ \mu \mathrm{e}^{-c^\star/2\mu}\xi \sup_{t\in[0,\xi]} \|W_{(j+1)}(t)\|_{s+1}^2 \\
&+ \mu^2 \int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
\lesssim \mu^2 \mathrm{e}^{-c^\star/\mu} &+ \mu \mathrm{e}^{-c^\star/2\mu} \sup_{t\in[0,e^{c^\star/2\mu}]} \|W_{(j+1)}(t)\|_{s+1} \\
&+ \mu^3 \mathrm{e}^{-c^\star/2\mu} \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t \sup_{t\in[0,e^{c^\star/2\mu}]} \|W_{(j+1)}(t)\|_{s+1} \\
&+ \mu \sup_{t\in[0,e^{c^\star/2\mu}]} \|W_{(j+1)}(t)\|_{s+1}^2 \\
&+ \mu^2 \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t.
\end{aligned}
\tag{4.56}
$$

Observe that

$$
\begin{aligned}
&\mu^2 \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
&= \mu^2 \int_0^{\xi^\star} \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t + \mu^2 \int_{\xi^\star}^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\|W_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
&\le \mu^2\xi^\star \sup_{t\in[0,\xi^\star]} \|W_{(j+1)}(t)\|_{s+1}^2 + \mu^2 \int_{\xi^\star}^\infty \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t \sup_{t\in[0,e^{c^\star/2\mu}]} \|W_{(j+1)}(t)\|_{s+1}^2 \\
&\lesssim \mu^2\xi^\star \sup_{t\in[0,\xi^\star]} \|W_{(j+1)}(t)\|_{s+1}^2 + \frac{\mathrm{e}^{-\eta\lambda^\mu\xi^\star}}{\eta} \sup_{t\in[0,e^{c^\star/2\mu}]} \|W_{(j+1)}(t)\|_{s+1}^2 \\
&\lesssim \mu|\log\mu|^2\mathrm{e}^{-c^\star/\mu} + \mu^\alpha \sup_{t\in[0,e^{c^\star/2\mu}]} \|W_{(j+1)}(t)\|_{s+1}^2,
\end{aligned}
$$

in which we have used inequality (4.55) as well as the definitions $\xi^\star = |\log\mu|/c_1\lambda^\mu$ and $\mu^\alpha = \mathrm{e}^{-\eta\lambda^\mu\xi^\star}$ in the last step. Inserting this estimate and the inequality

$$
\int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t \le \int_0^\infty \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t = \frac{1}{\eta\lambda^\mu} \lesssim \frac{1}{\mu^2},
$$

into (4.56), we conclude that

$$\sup_{t\in[0,e^{c^\star/2\mu}]} \|W_{(j+1)}(t)\|_{s+1} \lesssim \sqrt{\mu}|\log\mu|\mathrm{e}^{-c^\star/2\mu} + \sqrt{\mu+\mu^\alpha}\sup_{t\in[0,e^{c^\star/2\mu}]}\|W_{(j+1)}(t)\|_{s+1} \quad (4.57)$$

so that

$$\|W_{(j+1)}(\xi)\|_{s+1} \lesssim \sqrt{\mu}|\log\mu|\mathrm{e}^{-c^\star/2\mu},$$

that is inequality (4.33).

Convergence of the iteration scheme

Finally we prove that the sequence $\{(Z_{(j)}, r_{(j)}, W_{(j)})\}_{j\in\mathbb{N}_0}$ is convergent by considering the sequence $\{(Z_{(j+1)} - Z_{(j)}, r_{(j+1)} - r_{(j)}, W_{(j+1)} - W_{(j)})\}_{j\in\mathbb{N}_0}$ and showing that it is contractive. To this end we use the estimates for the derivatives of the nonlinearities given in Propositions 2.17 and 2.18 together with Lemma 4.30. The result is formulated in the following lemma. We use the notation introduced above Lemma 4.30.

Lemma 4.31. *Let $j\in\mathbb{N}_0$ and define*

$$\tilde{Z}_{(j+1)} = Z_{(j+1)} - Z_{(j)}, \quad \tilde{r}_{(j+1)} = r_{(j+1)} - r_{(j)}, \quad \tilde{W}_{(j+1)} = W_{(j+1)} - W_{(j)}.$$

The estimate

$$\|(\tilde{Z}_{(j+1)}, \tilde{r}_{(j+1)}, \tilde{W}_{(j+1)})\| \le \frac{1}{2}\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|$$

holds for all sufficiently small values of μ.

In order to prove Lemma 4.31 we proceed as in the proof of Lemma 4.30 and show that

$$|\tilde{Z}_{(j+1)}(\xi)| \lesssim \mu^2\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi\in[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}], \quad (4.58)$$

$$|\tilde{r}_{(j+1)}(\xi)| \lesssim \mu\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi\in[0, \mathrm{e}^{c^\star/2\mu}] \quad (4.59)$$

$$\|\tilde{W}_{(j+1)}(\xi)\|_{s+1} \lesssim \sqrt{\mu}|\log\mu|\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi\in[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}] \quad (4.60)$$

for all $j\in\mathbb{N}_0$.

We begin with the hyperbolic part $(\tilde{Z}_{(j+1)}, \tilde{r}_{(j+1)})$. From the definitions (4.14), (4.15) we obtain the formulae

$$\begin{aligned}\tilde{Z}_{(j+1)}(\xi) = &\int_{-\mathrm{e}^{c^\star/2\mu}}^{\xi} \mathrm{e}^{(\xi-t)L^\mu_{\mathrm{sh}}} P_{\mathrm{ss}}(G^\mu_{1(j)}(t) - G^\mu_{1(j-1)}(t) + G^\mu_{2(j)}(t) - G^\mu_{2(j-1)}(t))\,\mathrm{d}t\\ &- \int_{\xi}^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{(\xi-t)L^\mu_{\mathrm{sh}}} P_{\mathrm{su}}(G^\mu_{1(j)}(t) - G^\mu_{1(j-1)}(t) + G^\mu_{2(j)}(t) - G^\mu_{2(j-1)}(t))\,\mathrm{d}t\end{aligned}$$

and

$$\begin{aligned}\tilde{r}_{(j+1)}(\xi) = \sum_{i=1}^{2}\bigg(&\int_0^{\xi}\langle f^\mu_{3(j)}(t) - f^\mu_{3(j-1)}(t) + f^\mu_{4(j)}(t) - f^\mu_{4(j-1)}(t), (S^\mu_i)^*\rangle\,\mathrm{d}t\bigg) S^\mu_i(\xi)\\ &- \bigg(\int_{\xi}^{\mathrm{e}^{c^\star/2\mu}}\langle (f^\mu_{3(j)} - f^\mu_{3(j-1)} + f^\mu_{4(j)} - f^\mu_{4(j-1)})(t), (U^\mu_i)^*(t)\rangle\,\mathrm{d}t\bigg) U^\mu_i(\xi).\end{aligned}$$

We find using the estimates for $\mathrm{d}G_1$, $\mathrm{d}G_2$, $\mathrm{d}f_3$, $\mathrm{d}f_4$ in Propositions 2.17 and 2.18 together with Lemma 4.30 that

$$\begin{aligned}
|G^{\mu}_{1(j)}(\xi) - G^{\mu}_{1(j-1)}(\xi)| &\lesssim \mu^2 \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}],\\
|G^{\mu}_{2(j)}(\xi) - G^{\mu}_{2(j-1)}(\xi)| &\lesssim \mu^2 \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}],\\
|f^{\mu}_{3(j)}(\xi) - f^{\mu}_{3(j-1)}(\xi)| &\lesssim \mu^3 \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi \in [0, \mathrm{e}^{c^\star/2\mu}],\\
|f^{\mu}_{4(j)}(\xi) - f^{\mu}_{4(j-1)}(\xi)| &\lesssim \mu^3 \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi \in [0, \mathrm{e}^{c^\star/2\mu}].
\end{aligned}$$

The calculations in the proof of inequalities (4.31), (4.32) yield

$$\begin{aligned}
|\tilde{Z}_{(j+1)}(\xi)| &\lesssim \mu^2 \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}],\\
|\tilde{r}_{(j+1)}(\xi)| &\lesssim \mu \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|, \quad \xi \in [0, \mathrm{e}^{c^\star/2\mu}],
\end{aligned}$$

thus proving inequalities (4.58), (4.59).

Next we establish the estimate for $\tilde{W}_{(j+1)}$ on $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$. We find from equation (4.30) that

$$
\begin{aligned}
&\Big(I + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\Big)\partial_\xi\partial_\tau^k\tilde{W}_{(j+1)} \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{1(j)}^\mu - \hat{S}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\xi\partial_\tau^k W_{(j)} \\
&= (Q^\mu)^{-1}L^\mu Q^\mu\partial_\tau^k\tilde{W}_{(j+1)} + (T^\mu)^*\underline{\hat{g}}_{1(j)}^\mu T^\mu\partial_\tau^{k+1}\tilde{W}_{(j+1)} + (T^\mu)^*(\underline{\hat{g}}_{1(j)}^\mu - \underline{\hat{g}}_{1(j-1)}^\mu)T^\mu\partial_\tau^{k+1}W_{(j)} \\
&\quad + (T^\mu)^*\underline{\hat{g}}_{1(j)}^\mu R_1^\mu\partial_\tau^{k+1}\tilde{W}_{(j+1)} + (T^\mu)^*(\underline{\hat{g}}_{1(j)}^\mu - \underline{\hat{g}}_{1(j-1)}^\mu)R_1^\mu\partial_\tau^{k+1}W_{(j)} \\
&\quad + R_2^\mu\underline{\hat{g}}_{1(j)}^\mu Q^\mu\partial_\tau^{k+1}\tilde{W}_{(j+1)} + R_2^\mu(\underline{\hat{g}}_{1(j)}^\mu - \underline{\hat{g}}_{1(j-1)}^\mu)Q^\mu\partial_\tau^{k+1}W_{(j)} \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{(j)}^\mu\Big(\partial_\tau^k\Big(\hat{A}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau\tilde{W}_{(j+1)}\Big) - \hat{A}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau^{k+1}\tilde{W}_{(j+1)}\Big) \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{(j)}^\mu - \hat{S}_{(j-1)}^\mu)\Big(\partial_\tau^k\Big(\hat{A}_{1(j-1)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau W_{(j)}\Big) \\
&\qquad\qquad - \hat{A}_{1(j-1)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau^{k+1}W_{(j)}\Big) \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{(j)}^\mu\Big(\partial_\tau^k\Big((\hat{A}_{1(j)}^\mu - \hat{A}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau W_{(j)}\Big) \\
&\qquad\qquad - (\hat{A}_{1(j)}^\mu - \hat{A}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau^{k+1}W_{(j)}\Big) \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu\partial B^\mu[0](\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau^k\tilde{W}_{(j+1)} \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{1(j)}^\mu - \hat{S}_{1(j-1)}^\mu)\partial B^\mu[0](\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\tau^k W_{(j)} \\
&\quad - (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu P_{\mathrm{h}}\Big(\hat{g}_{1(j)}^\mu(U^\mu)^{-1}Q^\mu\partial_\tau\tilde{W}_{(j+1)}\Big) \\
&\quad - (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu P_{\mathrm{h}}\Big((\hat{g}_{1(j)}^\mu - \hat{g}_{1(j-1)}^\mu)(U^\mu)^{-1}Q^\mu\partial_\tau W_{(j)}\Big) \\
&\quad - (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{1(j)}^\mu - \hat{S}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}\partial_\tau^k U^\mu P_{\mathrm{h}}\Big(\hat{g}_{1(j-1)}^\mu(U^\mu)^{-1}Q^\mu\partial_\tau W_{(j)}\Big) \\
&\quad + \mu(Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu b(z_{(j)}, (U^\mu)^{-1}Q^\mu\tilde{W}_{(j+1)}) \\
&\quad + \mu(Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{1(j)}^\mu - \hat{S}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}\partial_\tau^k U^\mu b(z_{(j)}, (U^\mu)^{-1}Q^\mu W_{(j)}) \\
&\quad + \mu(Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu b(\tilde{r}_{(j)}, (U^\mu)^{-1}Q^\mu\tilde{W}_{(j)}) \\
&\quad + (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu(\hat{g}_{2(j)}^\mu - \hat{g}_{2(j-1)}^\mu) \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{1(j)}^\mu - \hat{S}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}\partial_\tau^k U^\mu\hat{g}_{2(j-1)}^\mu \\
&\quad + (Q^\mu)^{-1}(I + (\hat{S}_0^\mu)^{-\frac{1}{2}}\hat{S}_{1(j)}^\mu(\hat{S}_0^\mu)^{-\frac{1}{2}})\partial_\tau^k U^\mu(\hat{h}_{(j)}^\mu - \hat{h}_{(j-1)}^\mu) \\
&\quad + (Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{1(j)}^\mu - \hat{S}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}\partial_\tau^k U^\mu\hat{h}_{(j-1)}^\mu.
\end{aligned}
\tag{4.61}
$$

We take the complex inner product of equation (4.61) with $\partial_\tau^k\tilde{W}_{(j+1)}$, integrate over $(0, 2\pi)$ with respect to τ and estimate each term individually. To estimate the terms

$$
\int_0^{2\pi}\Big\langle(Q^\mu)^{-1}(\hat{S}_0^\mu)^{-\frac{1}{2}}(\hat{S}_{1(j)}^\mu - \hat{S}_{1(j-1)}^\mu)(\hat{S}_0^\mu)^{-\frac{1}{2}}Q^\mu\partial_\xi\partial_\tau^k W_{(j)}, \partial_\tau^k\tilde{W}_{(j+1)}(\xi)\Big\rangle\,\mathrm{d}\tau
$$

and

$$\int_0^{2\pi} \left\langle (T^\mu)^*(\underline{\hat{g}}^\mu_{1(j)}(\xi) - \underline{\hat{g}}^\mu_{1(j-1)}(\xi)) T^\mu \partial_\tau^{k+1} W_{(j)}(\xi), \partial_\tau^k \tilde{W}_{(j+1)}(\xi) \right\rangle \mathrm{d}\tau$$

we use the stronger regularity assumption $\|w_0\|_{s+2} \leq \mu \mathrm{e}^{-c^\star/2\mu}$ und suppose that $W_{(0)}$ satisfies $\|W_{(0)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_c^{s+2})} \leq \mathrm{e}^{-c^\star/2\mu}$. Lemma 4.30 with s replaced by $s+1$ implies that

$$\sup_{j\in\mathbb{N}_0} \|W_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+2})} \leq \mathrm{e}^{-c^\star/2\mu}$$

as well as

$$\sup_{j\in\mathbb{N}_0} \|\partial_\xi W_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^{s+1})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

We conclude that

$$\begin{aligned}
&\left| \int_0^{2\pi} \left\langle (Q^\mu)^{-1} (\hat{S}^\mu_0)^{-\frac{1}{2}} (\hat{S}^\mu_{1(j)} - \hat{S}^\mu_{1(j-1)}) (\hat{S}^\mu_0)^{-\frac{1}{2}} Q^\mu \partial_\xi \partial_\tau^k W_{(j)}, \partial_\tau^k \tilde{W}_{(j+1)}(\xi) \right\rangle \mathrm{d}\tau \right| \\
&\quad \lesssim \|(\hat{S}^\mu_{1(j)}(\xi) - \hat{S}^\mu_{1(j-1)}(\xi))\|_{L^\infty} \|\partial_\xi W_{(j)}(\xi)\|_{s+1} \|\tilde{W}_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \mu \mathrm{e}^{-c^\star/2\mu} \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\| \|\tilde{W}_{(j+1)}(\xi)\|_{s+1}
\end{aligned}$$

and

$$\begin{aligned}
&\left| \int_0^{2\pi} \left\langle (T^\mu)^*(\underline{\hat{g}}^\mu_{1(j)}(\xi) - \underline{\hat{g}}^\mu_{1(j-1)}(\xi)) T^\mu \partial_\tau^{k+1} W_{(j)}(\xi), \partial_\tau^k \tilde{W}_{(j+1)}(\xi) \right\rangle \mathrm{d}\tau \right| \\
&\quad \lesssim \|(\underline{\hat{g}}^\mu_{1(j)}(\xi) - \underline{\hat{g}}^\mu_{1(j-1)}(\xi))\|_{L^\infty} \|W_{(j)}(\xi)\|_{s+2} \|\tilde{W}_{(j+1)}(\xi)\|_{s+1} \\
&\quad \lesssim \mu \mathrm{e}^{-c^\star/2\mu} \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\| \|\tilde{W}_{(j+1)}(\xi)\|_{s+1}
\end{aligned}$$

in which we made use of the estimates for $\mathrm{d}\hat{g}_1$ and $\mathrm{d}\hat{S}_1$ given in Proposition 2.17 and equation (4.28).

The remaining terms on the right-hand side are estimated using the same methods as in the proof of inequalities (4.38)–(4.47) together with Lemma 4.30 and the estimates for the derivatives of the other nonlinearities given in Propositions 2.17 and 2.18 and equation (4.28). The result is

$$\begin{aligned}
&\int_0^{2\pi} \langle (I + (Q^\mu)^{-1} (\hat{S}^\mu_0)^{-\frac{1}{2}} \hat{S}^\mu_{1(j)}(\xi) (\hat{S}^\mu_0)^{-\frac{1}{2}} Q^\mu) \partial_\xi \partial_\tau^k \tilde{W}_{(j+1)}(\xi), \partial_\tau^k \tilde{W}_{(j+1)}(\xi) \rangle \, \mathrm{d}\tau \\
&\quad \lesssim \mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu \xi}) \|\tilde{W}_{(j+1)}(\xi)\|_{s+1}^2 \\
&\qquad\qquad + \mu (\mathrm{e}^{-c^\star/2\mu} + \mu^2 \mathrm{e}^{-\eta\lambda^\mu \xi}) \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\| \|\tilde{W}_{(j+1)}(\xi)\|_{s+1}.
\end{aligned} \tag{4.62}$$

and from inequality (4.62) we deduce that

$$\begin{aligned}
\|\tilde{W}_{(j+1)}(\xi)\|_{s+1}^2 \lesssim \int_0^\xi \Big(&\mu (\mathrm{e}^{-c^\star/2\mu} + \mu \mathrm{e}^{-\eta\lambda^\mu t}) \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 \\
&+ \mu (\mathrm{e}^{-c^\star/2\mu} + \mu^2 \mathrm{e}^{-\eta\lambda^\mu t}) \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\| \|\tilde{W}_{(j+1)}(t)\|_{s+1} \Big) \, \mathrm{d}t
\end{aligned} \tag{4.63}$$

in the same way as we deduced inequality (4.49) from the estimate (4.48), where we also have used the fact that $\tilde{W}(0) = 0$.

We handle inequality (4.63) using a two-step estimation technique similar to that employed in the proof of inequality (4.33). Young's inequality implies that

$$\|\tilde{W}_{(j+1)}(\xi)\|_{s+1}^2 \lesssim \int_0^\xi c_3\lambda^\mu \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 \,\mathrm{d}t + c_4\mu^4\Big(\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|\Big)^2 \xi$$

for some multiplicative constants $c_3, c_4 > 0$ and applying Gronwall's inequality and proceeding as in the proof of inequality (4.55), we find that

$$\|\tilde{W}_{(j+1)}(\xi)\|_{s+1}^2 \lesssim \mu|\log\mu|\Big(\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|\Big)^2 \tag{4.64}$$

for each $\xi \in [0, \xi^\star]$ with

$$\xi^\star = \frac{\alpha|\log\mu|}{\eta\lambda^\mu}, \qquad \alpha = \frac{\eta}{c_3}.$$

For $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$, it follows from inequality (4.63) that

$$\begin{aligned}
\|\tilde{W}_{(j+1)}(\xi)\|_{s+1}^2 &\lesssim \mu\mathrm{e}^{-c^\star/2\mu}\int_0^\xi \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t + \mu^2\int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\|\tilde{W}_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
&\quad + \mu\mathrm{e}^{-c^\star/2\mu}\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|\int_0^\xi \|\tilde{W}_{(j+1)}(t)\|_{s+1}\,\mathrm{d}t \\
&\quad + \mu^3\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|\int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\|\tilde{W}_{(j+1)}(t)\|_{s+1}\,\mathrm{d}t. \\
&\lesssim \mu\mathrm{e}^{-c^\star/2\mu}\xi\sup_{t\in[0,\xi]}\|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 + \mu^2\int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\|\tilde{W}_{(j+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
&\quad + \mu\mathrm{e}^{-c^\star/2\mu}\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|\xi\sup_{t\in[0,\xi]}\|\tilde{W}_{(j+1)}(t)\|_{s+1} \\
&\quad + \mu^3\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|\int_0^\xi \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t\sup_{t\in[0,\xi]}\|\tilde{W}_{(j+1)}(t)\|_{s+1} \\
&\leq \mu\sup_{t\in[0,\mathrm{e}^{c^\star/2\mu}]}\|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 + \mu^2\int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\|\tilde{W}_{(m+1)}(t)\|_{s+1}^2\,\mathrm{d}t \\
&\quad + \mu\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|\sup_{t\in[0,\mathrm{e}^{c^\star/2\mu}]}\|\tilde{W}_{(j+1)}(t)\|_{s+1} \\
&\quad + \mu^3\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|\int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t\sup_{t\in[0,\mathrm{e}^{c^\star/2\mu}]}\|\tilde{W}_{(j+1)}(t)\|_{s+1}.
\end{aligned} \tag{4.65}$$

Next observe that

$$
\begin{aligned}
&\mu^2 \int_0^{e^{c^\star/2\mu}} e^{-\eta\lambda^\mu t} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 \, dt \\
&\quad = \mu^2 \int_0^{\xi^\star} e^{-\eta\lambda^\mu t} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 \, dt + \mu^2 \int_{\xi^\star}^{e^{c^\star/2\mu}} e^{-\eta\lambda^\mu t} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 \, dt \\
&\quad \leq \mu^2 \xi^\star \sup_{t\in[0,\xi^\star]} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 + \mu^2 \int_{\xi^\star}^{\infty} e^{-\eta\lambda^\mu t} \, dt \sup_{t\in[0,e^{c^\star/2\mu}]} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 \\
&\quad \lesssim \mu^3 |\log\mu| \xi^\star \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|^2 + \frac{e^{-\eta\lambda^\mu \xi^\star}}{\eta} \sup_{t\in[0,e^{c^\star/2\mu}]} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2 \\
&\quad \lesssim \mu |\log\mu|^2 \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|^2 + \mu^\alpha \sup_{t\in[0,e^{c^\star/2\mu}]} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2, \qquad (4.66)
\end{aligned}
$$

in which we have used inequality (4.64) as well as $\xi^\star = |\log\mu|/c_3\lambda^\mu$ and $\mu^\alpha = e^{-\eta\lambda^\mu\xi^\star}$. Inserting the estimate (4.66) and the inequality

$$
\int_0^{e^{c^\star/2\mu}} e^{-\eta\lambda^\mu t} \, dt \leq \frac{1}{\eta\lambda^\mu} \lesssim \frac{1}{\mu^2}
$$

into (4.65), we conclude that

$$
\begin{aligned}
\|\tilde{W}_{(j+1)}(\xi)\|_{s+1}^2 \lesssim \mu \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\| &\sup_{t\in[0,e^{c^\star/2\mu}]} \|\tilde{W}_{(j+1)}(t)\|_{s+1} \\
&+ \mu |\log\mu|^2 \Big(\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\| \Big)^2 \\
&+ (\mu + \mu^\alpha) \sup_{t\in[0,e^{c^\star/2\mu}]} \|\tilde{W}_{(j+1)}(t)\|_{s+1}^2,
\end{aligned}
$$

from which it follows that

$$
\|\tilde{W}_{(j+1)}(\xi)\|_{s+1} \lesssim \sqrt{\mu} |\log\mu| \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)})\|,
$$

for sufficiently small values of $\mu \geq 0$ and each $\xi \in [0, e^{c^\star/2\mu}]$, which is inequality (4.60).

Altogether, it follows from inequalities (4.58), (4.59) and (4.60) that

$$
\begin{aligned}
\|\tilde{Z}_{(j+1)}, \tilde{r}_{(j+1)}, \tilde{W}_{(j+1)}\| &\lesssim (\mu^2 + \mu + \sqrt{\mu}\log(\mu)) \|\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)}\| \\
&\lesssim \mu \|\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{W}_{(j)}\|
\end{aligned}
$$

and this implies the assertion of Lemma 4.31. The existence of a limit $(Z^\star, r^\star, W^\star)$ of the sequence $\{(Z_{(j)}, r_{(j)}, W_{(j)})\}_{j\in\mathbb{N}_0}$ such that

$$
\begin{aligned}
\|Z^\star\|_{C([-e^{c^\star/2\mu}, e^{c^\star/2\mu}];\mathcal{X}_{\mathrm{sh}})} &\lesssim \mu^2 e^{-c^\star/2\mu} \\
\|r^\star\|_{C([0,e^{c^\star/2\mu}];\mathcal{X}_{\mathrm{wh}})} &\lesssim \mu e^{-c^\star/2\mu} \\
\|W^\star\|_{C([0,e^{c^\star/2\mu}];\mathcal{X}_{\mathrm{c}}^{s+1})} &\lesssim \mu^\delta e^{-c^\star/2\mu}, \quad \delta \in [0, \tfrac{1}{2}),
\end{aligned}
$$

now follows from Lemmata 4.30 and 4.31. The statement of Theorem 4.26 is obtained by observing that $\xi \mapsto (R_{\mathrm{sh}} Z^\star(-\xi), r^\star(\xi), W^\star(\xi))$ is also a limit of $\{(Z_{(j)}, r_{(j)}, W_{(j)})\}_{j\in\mathbb{N}_0}$, that is $Z^\star(\xi) = R_{\mathrm{sh}} Z^\star(-\xi)$, extending $r^\star$, $W^\star$ by reversibility to $[-e^{c^\star/2\mu}, e^{c^\star/2\mu}]$, and using the facts that the operators $U^\mu \in \mathcal{L}(\mathcal{X}^{s+1})$ and $Q^\mu \in \mathcal{L}(\mathcal{X}^{s+1})$ (see Lemma 4.29) are invertible.

4.3 Assorted proofs

Here we present the proof of Lemma 4.29.

Lemma 4.29. *There exists an invertible bounded linear operator $Q^\mu \colon \mathcal{X}^t \to \mathcal{X}^t$ (for all $t > \frac{1}{2}$) such that*

$$((Q^\mu)^{-1} L^\mu Q^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} (D_m^\mu u_m) \mathrm{e}^{\mathrm{i} m \tau},$$

where D_m^μ are diagonal matrices depending analytically upon μ in a neighbourhood of the origin which is independent of $m \in \mathbb{Z}$. Moreover

$$Q^\mu = T^\mu + R_1^\mu, \quad (Q^\mu)^{-1} = (T^\mu)^{-1} + R_2^\mu,$$

where $T^\mu \colon \mathcal{X}^t \to \mathcal{X}^t$ is unitary and $R_1^\mu, R_2^\mu \colon \mathcal{X}^t \to \mathcal{X}^{t+1}$ are invertible bounded linear operators. The operators T^μ and R_1^μ, R_2^μ depend analytically upon μ.

Proof. Consider the ordered bases $\{e_{m,1}\mathrm{e}^{\mathrm{i}m\tau}, \ldots, e_{m,n}\mathrm{e}^{\mathrm{i}m\tau}\}$ for E_m $(m \in Z \setminus \{\pm 1\})$, $\{e_{1,1}\mathrm{e}^{\mathrm{i}\tau}, \ldots, e_{1,n-2}\mathrm{e}^{\mathrm{i}\tau}, u\mathrm{e}^{\mathrm{i}\tau}, v\mathrm{e}^{\mathrm{i}\tau}\}$ for E_1 and $\{\overline{e}_{1,1}\mathrm{e}^{-\mathrm{i}\tau}, \ldots, \overline{e}_{1,n-2}\mathrm{e}^{-\mathrm{i}\tau}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}, \overline{v}\mathrm{e}^{-\mathrm{i}\tau}\}$ for E_{-1} defined in Section 2.1. The m-th Fourier component $P_m u$ of a function $u \in \mathcal{X}^t$ is represented by a coordinate vector of the form $(u_{m,1}, \ldots, u_{m,n})^\top$ and $U^\mu P_m u = P_m U^\mu u$ has the same coordinate vector representation with respect to the ordered bases

$$\{U_m^\mu(e_{m,1}\mathrm{e}^{\mathrm{i}m\tau}), \ldots, U_m^\mu(e_{m,n}\mathrm{e}^{\mathrm{i}m\tau})\}$$

for E_m $(m \in Z \setminus \{\pm 1\})$,

$$\{U_1^\mu(e_{1,1}\mathrm{e}^{\mathrm{i}\tau}), \ldots, U_1^\mu(e_{1,n-2}\mathrm{e}^{\mathrm{i}\tau}), U_1^\mu(u\mathrm{e}^{\mathrm{i}\tau}), U_1^\mu(v\mathrm{e}^{\mathrm{i}\tau})\}$$

for E_1 and

$$\{U_{-1}^\mu(\overline{e}_{1,1}\mathrm{e}^{-\mathrm{i}\tau}), \ldots, U_{-1}^\mu(\overline{e}_{1,n-2}\mathrm{e}^{-\mathrm{i}\tau}), U_{-1}^\mu(\overline{u}\mathrm{e}^{-\mathrm{i}\tau}), U_{-1}^\mu(\overline{v}\mathrm{e}^{-\mathrm{i}\tau})\}$$

for E_{-1}.

We write

$$U_1^\mu(u\mathrm{e}^{\mathrm{i}\tau}) = u_1^\mu \mathrm{e}^{\mathrm{i}\tau}, \quad U_1^\mu(\overline{u}\mathrm{e}^{-\mathrm{i}\tau}) = u_{-1}^\mu \mathrm{e}^{-\mathrm{i}\tau}, \quad U_1^\mu(v\mathrm{e}^{\mathrm{i}\tau}) = v_1^\mu \mathrm{e}^{\mathrm{i}\tau}, \quad U_1^\mu(\overline{v}\mathrm{e}^{-\mathrm{i}\tau}) = v_{-1}^\mu \mathrm{e}^{-\mathrm{i}\tau}$$

$$U_m^\mu(e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}) = e_{m,r}^\mu \mathrm{e}^{\mathrm{i}m\tau}, \quad m \in \mathbb{Z}, \ r = 1, \ldots, r_m,$$

where the vectors $u_{\pm 1}^\mu$, $v_{\pm 1}^\mu$, $e_{m,r}^\mu$ depend analytically upon μ,

$$\langle u_1^\mu, v_1^\mu \rangle = \Big\langle (1, \pm\lambda^\mu, 0, 0)^\top \Big\rangle, \quad \langle u_{-1}^\mu, v_{-1}^\mu \rangle = \Big\langle (0, 0, 1, \pm\lambda^\mu)^\top \Big\rangle$$

(see Assumptions 3) and $e_{m,r}^\mu$ are geometrically simple eigenvectors of the matrix

$$\mathrm{i}m\tilde{A}^\mu(0) + \partial\tilde{B}^\mu[0] = m\Big(\mathrm{i}\tilde{A}^\mu(0) + \frac{1}{m}\partial\tilde{B}^\mu[0]\Big).$$

The corresponding eigenvalues $\lambda_{m,r}^\mu$ depend analytically upon μ, where $\lambda_{m,r}^0 = \lambda_{m,r}$ are defined in Section 2.1.

Proposition 2.2 implies that the invertible matrices

$$Q_{\pm 1}^\mu = (e_{\pm 1,1}^\mu | \cdots | e_{\pm 1,n-2}^\mu | u_{\pm 1}^\mu | v_{\pm 1}^\mu), \qquad Q_m^\mu = (e_{m,1}^\mu | \cdots | e_{m,n}^\mu), \quad m \in \mathbb{Z} \setminus \{\pm 1\},$$

depend analytically upon μ in a neighbourhood of the origin which is independent of m and that

$$Q_m^\mu = Q_\infty^\mu + \frac{1}{m} R_{1,m}^\mu, \quad (Q_m^\mu)^{-1} = (Q_\infty^\mu)^{-1} + \frac{1}{m} R_{2,m}^\mu.$$

Here Q_∞^μ is the unitary matrix whose columns are given by the eigenvectors of the skew-Hermitian and hence normal matrix $\mathrm{i}\tilde{A}^\mu(0)$ and $R_{1,m}^\mu$, $R_{2,m}^\mu$ satisfy

$$\frac{1}{m} R_{i,m}^\mu \to 0$$

as $|m| \to \infty$ in $\mathcal{L}(\mathbb{C}^n)$ uniformly over μ.

The formula

$$(Q^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} (Q_m^\mu u_m) \mathrm{e}^{\mathrm{i}m\tau}$$

defines an invertible bounded linear operator $Q^\mu \colon \mathcal{X}^t \to \mathcal{X}^t$ with inverse

$$((Q^\mu)^{-1} u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} ((Q_m^\mu)^{-1} u_m) \mathrm{e}^{\mathrm{i}m\tau}$$

and since

$$(L^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} \Big((\mathrm{i}m\tilde{A}^\mu(0) + \partial \tilde{B}^\mu[0]) u_m \Big) \mathrm{e}^{\mathrm{i}m\tau}$$

we conclude that

$$((Q^\mu)^{-1} L^\mu Q^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} (D_m^\mu u_m) \mathrm{e}^{\mathrm{i}m\tau}$$

with $D_m^\mu = \operatorname{diag}(\lambda_{m,1}^\mu, \ldots, \lambda_{m,n}^\mu)$. Moreover

$$Q^\mu = T^\mu + R_1^\mu, \quad (Q^\mu)^{-1} = (T^\mu)^{-1} + R_2^\mu,$$

where the linear operators R_1^μ, R_2^μ and T^μ on $\mathcal{X}^t$ are defined by the formulae

$$(T^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} (Q_\infty^\mu u_m) \mathrm{e}^{\mathrm{i}m\tau}, \quad ((T^\mu)^{-1} u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m \in \mathbb{Z}} ((Q_\infty^\mu)^{-1} u_m) \mathrm{e}^{\mathrm{i}m\tau},$$

$$(R_1^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m=-M}^{M} \Big((Q_m^\mu - Q_\infty^\mu) u_m \Big) \mathrm{e}^{\mathrm{i}m\tau} + \frac{1}{\sqrt{2\pi}} \sum_{|m|>M} \frac{1}{m} (R_{1,m}^\mu u_m) \mathrm{e}^{\mathrm{i}m\tau}$$

and

$$(R_2^\mu u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m=-M}^{M} \Big(((Q_m^\mu)^{-1} - (Q_\infty^\mu)^{-1}) u_m \Big) \mathrm{e}^{\mathrm{i}m\tau} + \frac{1}{\sqrt{2\pi}} \sum_{|m|>M} \frac{1}{m} (R_{2,m}^\mu u_m) \mathrm{e}^{\mathrm{i}m\tau},$$

so that R_1^μ, R_2^μ and T^μ depend analytically upon μ due to Lemma 2.4. The operator $T^\mu \colon \mathcal{X}^t \to \mathcal{X}^t$ is unitary, since the matrix Q_∞^μ is unitary and the computation

$$\begin{aligned}
\|R_1^\mu u\|_{t+1} &= \sum_{m=-M}^{M} (1+m^2)^{t+1} |(Q_m^\mu - Q_\infty^\mu) u_m|^2 + \sum_{|m|>M} \frac{(1+m^2)^{t+1}}{m^2} |R_{1,m}^\mu u_m|^2 \\
&\lesssim \sum_{m=-M}^{M} (1+m^2)^t |u_m|^2 + \sum_{|m|>M} (1+m^2)^t |u_m|^2 \\
&= \|u\|_t
\end{aligned}$$

(with a similar computation for R_2^μ) shows that R_1^μ, R_2^μ are bounded linear operators $\mathcal{X}^t \to \mathcal{X}^{t+1}$. ■

5 Existence theory for the semilinear problem

In this chapter we present an alternative approach to the theory in Chapter 4 for the semilinear case (that is $\tilde{g}_1^\mu = 0$ in equation (2.54)). Recall that Theorem 4.26 yields the existence of a solution $(Z^\star, r^\star, w^\star)$ to equations (2.55)–(2.57) such that $|Z^\star(\xi)|, |r^\star(\xi)|, \|w^\star(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \leq \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$. We improve the result of Theorem 4.26 under the additional hypothesis of the existence of a suitable Lyapunov functional $\mathcal{I}^\varepsilon$ in Assumptions 5 which, in scaled coordinates, reads as follows.

Assumptions 5. There exist $s \geq 0$ and a conserved quantity $\mathcal{I}^\mu(Z, z, w)$ for the system (2.55)–(2.57) which is analytic at the origin in $\mathbb{R} \times \mathcal{X}_\mathrm{sh} \times \mathcal{X}_\mathrm{wh} \times \mathcal{X}_\mathrm{c}^s$ and satisfies the estimate

$$\mathcal{I}^\mu(Z, z, w) = \mathcal{O}(\mu^4 \|(Z, z, w)\|^2_{\mathcal{X}_\mathrm{sh} \times \mathcal{X}_\mathrm{wh} \times \mathcal{X}_\mathrm{c}^s}) \tag{5.1}$$

(the factor μ^4 arises from the scaling $q' = (\lambda^\mu)^{-1} q$ in Section 2.3). Furthermore,

$$\mathcal{I}^\mu(0, 0, w) = \mu^4 Q(w) + \mathcal{O}(\mu^6 \|w\|^3_{\mathcal{X}_\mathrm{c}^s}), \tag{5.2}$$

in which $Q\colon \mathcal{X}_\mathrm{c}^s \to \mathbb{R}$ is a quadratic form satisfying

$$\|w\|^2_{\mathcal{X}_\mathrm{c}^s} \lesssim \|Q(w)\|_{\mathcal{X}_\mathrm{c}^s} \lesssim \|w\|^2_{\mathcal{X}_\mathrm{c}^s}.$$

In particular we construct solutions to equations (2.55)–(2.57) which are exponentially small on the entire real line, in other words reversible solutions to equations (2.53), (2.54) whose pointwise distance to the approximate pulse p^μ does not exceed $\mathrm{e}^{-c^\star/2\mu}$ on the entire real line. The result is given by the following theorem.

Theorem 5.1. *Suppose that $\tilde{g}_1^\mu = 0$ and Assumptions 5 hold. Equations (2.55)–(2.57) admit reversible solutions $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ in*

$$C_\mathrm{b}(\mathbb{R}; \mathcal{X}_\mathrm{sh}) \times C_\mathrm{b}(\mathbb{R}; \mathcal{X}_\mathrm{wh}) \times C_\mathrm{b}(\mathbb{R}; \mathcal{X}_\mathrm{c}^s),$$

parameterised by $w_0 \in \mathcal{X}_\mathrm{c}^s$ with $\|w_0\|_{\mathcal{X}_\mathrm{c}^s} \leq \mu \mathrm{e}^{-c^\star/2\mu}$ and $R_\mathrm{c} w_0 = w_0$, such that

$$\|Z^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})}, \|r^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})}, \|w^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{c}^s)} \leq \mathrm{e}^{-c^\star/2\mu}.$$

These solutions satisfy

$$\begin{aligned}\|Z^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} &\lesssim \mu^2 \mathrm{e}^{-c^\star/2\mu},\\ \|r^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})} &\lesssim \mu \mathrm{e}^{-c^\star/2\mu},\\ \|w^\star_{w_0}\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{c}^s)} &\lesssim \mu^\delta \mathrm{e}^{-c^\star/2\mu}\end{aligned}$$

for $\delta \in [0, \frac{1}{2})$.

We use the strategy devised by Groves & Schneider [8] (see also Hewer [13, Chapter 5]) and employ the following steps to prove Theorem 5.1. In Section 5.1 we consider the modified system

$$Z_\xi = L^\mu_\mathrm{sh} Z + G^\mu_1(Z, p^\mu + r, \phi(w)) + G^\mu_2(r), \tag{5.3}$$
$$r_\xi = L^\mu_\mathrm{wh} r + \mathcal{L}^\mu r + f^\mu_3(r) + f^\mu_4(Z, r, \phi(w)), \tag{5.4}$$
$$w_\xi = L^\mu_\mathrm{c} w + \mu b(p^\mu + r, \phi(w)) + \hat{g}^\mu(Z, p^\mu + r, \phi(w)) + \hat{h}^\mu(r). \tag{5.5}$$

which is obtained from equations (2.55)–(2.57) by applying a cut-off function ϕ supported on $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$ to the w-component of the nonlinearities (note that we have written $\hat{g}^\mu$ instead of $\hat{g}^\mu_2$ for notational simplicity, since $\hat{g}_1 = 0$ in equation (2.57)). This modification enables us to control the linear growth of the centre part w on the entire real line in the estimates below (see inequality (5.33)). We construct solutions $(Z^\star, r^\star, w^\star)$ to equations (5.3)–(5.5) which exist globally and have the properties that

(i) their hyperbolic part $(Z^\star, r^\star)$ satisfies $\|(Z^\star(\xi), r^\star(\xi))\|_{\mathcal{X}_\mathrm{sh}\times\mathcal{X}_\mathrm{wh}} \leq \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$ (see Figure 5.1),

(ii) their centre part $w^\star$ satisfies $\|w^\star\|_{\mathcal{X}_\mathrm{c}^s} \leq \mathrm{e}^{-c^\star/2\mu}$ for $\xi \in [-\mathrm{e}^{-c^\star/2\mu}, \mathrm{e}^{-c^\star/2\mu}]$.

In particular we obtain solutions $(p^\mu + r^\star, Z^\star + w^\star)$ to equations (2.53), (2.54) which are exponentially close to p^μ for $\xi \in [-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$ (see Figure 5.2). These solutions are parametrised by their initial values, the set of which forms a local centre-stable manifold $W^\mathrm{cs}_\mathrm{loc}$ (see Definition 5.3 below).

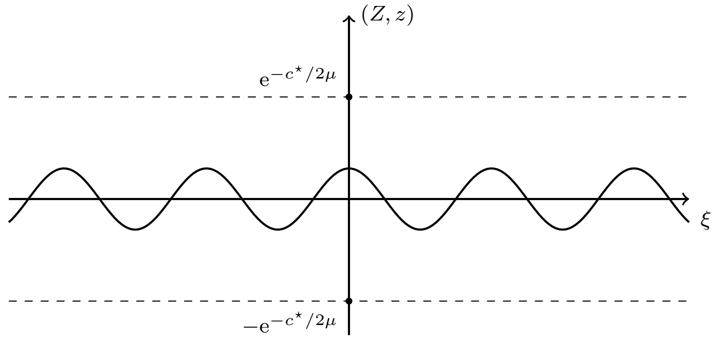

Figure 5.1: The hyperbolic part $(Z^\star, r^\star)$ of solutions to equations (5.3)–(5.5) with initial values on $W^\mathrm{cs}_\mathrm{loc}$ satisfies $\|(Z^\star(\xi), r^\star(\xi))\|_{\mathcal{X}_\mathrm{sh}\times\mathcal{X}_\mathrm{wh}} \leq \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$.

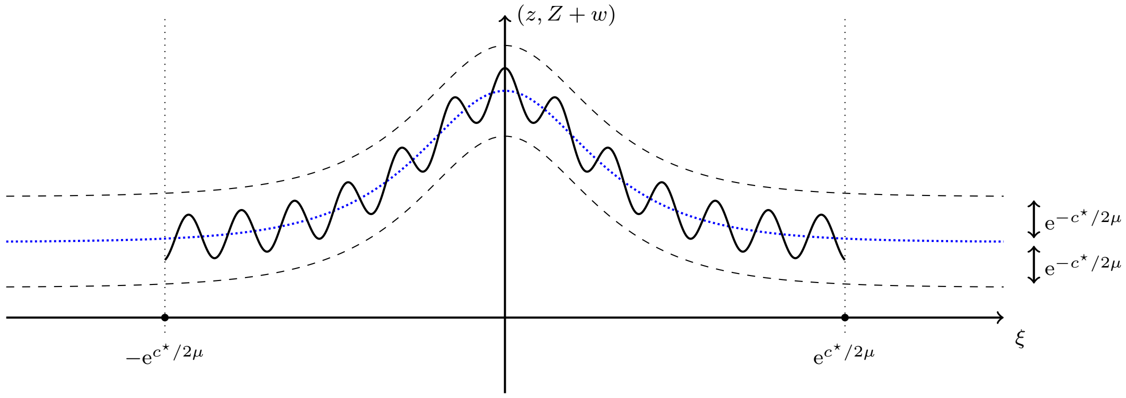

Figure 5.2: A generalised pulse solution to equations (2.53), (2.54) of the form $(p^\mu + r^\star, Z^\star + w^\star)$ lies in an exponentially thin tubular neighbourhood of $(p^\mu, 0)$ on the exponentially large interval $[-\mathrm{e}^{c^\star/2\mu}, \mathrm{e}^{c^\star/2\mu}]$. We make no statement about its existence or behaviour outside this interval.

In Sections 5.2 and 5.3 we prove that in fact $\|w^\star\|_{\mathcal{X}_c^s} \le \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$, from which it follows that $W^{\mathrm{cs}}_{\mathrm{loc}}$ is actually a *global* centre-stable manifold (meaning that its elements are initial values of generalised pulse solutions which exist on the entire real line), as follows. In Section 5.2 we first consider the modified equations

$$Z_\xi = L^\mu_{\mathrm{sh}} Z + G^\mu_1(Z, z, \phi(w)) + G^\mu_2(z), \tag{5.6}$$

$$z_\xi = L^\mu_{\mathrm{wh}} z + \tilde{f}^\mu_1(r) + \tilde{f}^\mu_2(Z, z, \phi(w)), \tag{5.7}$$

$$w_\xi = L^\mu_{\mathrm{c}} w + \mu b(z, \phi(w)) + \hat{g}^\mu(Z, z, \phi(w)) + \hat{h}^\mu(z), \tag{5.8}$$

which are obtained from the original equations (2.53), (2.54) by decoupling the equation for $q = Z + w$ and applying the cut-off function ϕ (see above) to the w-component of the nonlinearities. Note that solutions (Z, z, w) to equations (5.6)–(5.8) with $\|w(\xi)\|_{\mathcal{X}_c^s} \le \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$ correspond to solutions $(z, Z + w)$ of equations (2.53), (2.54) where the cut-off function is not used. Here we construct solutions $(Z^{\star\star}, z^{\star\star}, w^{\star\star})$ to equations (5.6)–(5.8) whose hyperbolic part satisfies $\|(Z^{\star\star}(\xi), z^{\star\star}(\xi))\|_{\mathcal{X}_{\mathrm{sh}} \times \mathcal{X}_{\mathrm{wh}}} \le \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$. The initial values of these solutions form a *centre manifold* W^{c}, which is an invariant manifold in the following sense: any solution (Z, z, w) to equations (5.6)–(5.8) with $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W^{\mathrm{c}}$ satisfies $(Z(\xi), z(\xi), w(\xi)) \in W^{\mathrm{c}}$ for all $\xi \ge \xi_0$. The *local centre manifold* $W^{\mathrm{c}}_{\mathrm{loc}} = W^{\mathrm{c}} \cap \{w\colon \|w\|_{\mathcal{X}_c^s} \le \mathrm{e}^{-c^\star/2\mu}\}$ is on the other hand *locally invariant*: any solution (Z, z, w) with $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W^{\mathrm{c}}_{\mathrm{loc}}$ remains on $W^{\mathrm{c}}_{\mathrm{loc}}$ (and hence on W^{c}) as long as $\|w(\xi)\|_{\mathcal{X}_c^s} \le \mathrm{e}^{-c^\star/2\mu}$. In fact the existence of the Lyapunov functional $\mathcal{I}^\mu$ (Assumptions 5) further guarantees that a solution (Z, z, w) with $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W^{\mathrm{c}}_{\mathrm{loc}}$ indeed satisfies $\|w(\xi)\|_{\mathcal{X}_c^s} \le \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \ge \xi_0$.

Our strategy in Section 5.3 is therefore to show that a solution $(Z^\star, p^\mu + r^\star, w^\star)$ to equations (2.55)–(2.57), where $(Z^\star, r^\star, w^\star)$ solves equations (5.3)–(5.5), converges exponentially fast to a solution $(Z^{\star\star}, z^{\star\star}, w^{\star\star})$ to equations (5.6)–(5.8) with $(Z^{\star\star}(\xi), z^{\star\star}(\xi), w^{\star\star}(\xi)) \in W^{\mathrm{c}}$ (that is $\|(Z^{\star\star}(\xi), z^{\star\star}(\xi))\|_{\mathcal{X}_{\mathrm{sh}}\times\mathcal{X}_{\mathrm{wh}}} \le \mathrm{e}^{-c^\star/2\mu}$) for all $\xi \in \mathbb{R}$. In particular we prove that there exists $\xi^\star < \mathrm{e}^{c^\star/2\mu}$ such that $|p^\mu(\xi^\star)| \lesssim \mu\mathrm{e}^{-c^\star/2\mu}$, and thus

$$\|(Z^\star(\xi^\star), p^\mu(\xi^\star) + r^\star(\xi^\star), w^\star(\xi^\star)) - (Z^{\star\star}(\xi^\star), z^{\star\star}(\xi^\star), w^{\star\star}(\xi^\star))\|_s \lesssim \mu\mathrm{e}^{-c^\star/2\mu}$$

(see Figure 5.3). Since by construction $\|w^\star(\xi^\star)\|_{\mathcal{X}_{\mathrm{c}}^s} \lesssim \mu^\delta\mathrm{e}^{-c^\star/2\mu}$, it follows that $\|w^{\star\star}(\xi^\star)\|_{\mathcal{X}_{\mathrm{c}}^s} \le \mathrm{e}^{-c^\star/2\mu}$ and since $(Z^{\star\star}(\xi^\star), z^{\star\star}(\xi^\star), w^{\star\star}(\xi^\star)) \in W^{\mathrm{c}}$ we conclude that $(Z^{\star\star}(\xi), z^{\star\star}(\xi), w^{\star\star}(\xi)) \in W^{\mathrm{c}}$ with $\|w^{\star\star}(\xi)\|_{\mathcal{X}_{\mathrm{c}}^s} \le \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \ge \xi^\star$ (see above). This observation finally implies that $\phi(w^\star(\xi)) = w^\star(\xi)$ for all $\xi \ge \xi^\star$, so that $(Z^\star(\xi^\star), p^\mu(\xi^\star) + r^\star(\xi^\star), w^\star(\xi^\star))$ is exponentially small for all $\xi \ge \xi^\star$ and thus a global generalised pulse solution to the original equations (2.53), (2.54) (without cut-off function).

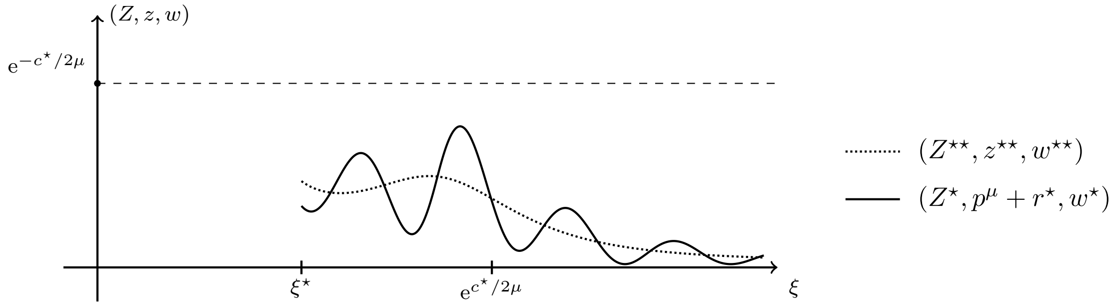

Figure 5.3: A solution $(Z^\star, p^\mu + r^\star, w^\star)$ with initial values on $W^{\mathrm{cs}}_{\mathrm{loc}}$ converges exponentially fast to a solution $(Z^{\star\star}, z^{\star\star}, w^{\star\star})$ on W^{c}. Observe that $(Z^\star, p^\mu{+}r^\star, w^\star)$ is (i) exponentially small, (ii) exponentially close to $(Z^{\star\star}, z^{\star\star}, w^{\star\star})$ on $[\xi^\star, \mathrm{e}^{c^\star/2\mu})$. The (exponential) convergence of $(Z^\star, p^\mu{+}r^\star, w^\star)$ to $(Z^{\star\star}, z^{\star\star}, w^{\star\star})$ and the fact that $(Z^{\star\star}, z^{\star\star}, w^{\star\star})$ remains on W^{c} with $\|w^{\star\star}(\xi)\|_{\mathcal{X}_{\mathrm{c}}^s} \le \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \ge \xi^\star$ imply that $(Z^\star, p^\mu + r^\star, w^\star)$ remains exponentially small for all $\xi \ge \xi^\star$.

5.1 Construction of a local centre-stable manifold

In this section we solve the modified equations (5.3)–(5.5), namely

$$Z_\xi = L^\mu_{\mathrm{sh}} Z + \underline{G}^\mu_1(Z, p^\mu + r, w) + G^\mu_2(r), \tag{5.9}$$
$$r_\xi = L^\mu_{\mathrm{wh}} r + \mathcal{L}^\mu r + f^\mu_3(r) + \underline{f}^\mu_4(Z, r, w), \tag{5.10}$$
$$w_\xi = L^\mu_{\mathrm{c}} w + \mu\underline{b}(p^\mu + r, w) + \underline{\hat{g}}^\mu(Z, p^\mu + r, w) + \hat{h}^\mu(r), \tag{5.11}$$

which are obtained from equations (2.55)–(2.57) by defining

$$\underline{G}^\mu_1(Z, z, w) = G^\mu_1(Z, z, \phi(w)),$$

and

$$\underline{f}^\mu_4(Z, z, w) = f^\mu_4(Z, z, \phi(w)), \quad \underline{b}(z, w) = b(z, \phi(w)), \quad \underline{\hat{g}}^\mu(Z, z, w) = \hat{g}^\mu(Z, z, \phi(w)).$$

Here $\phi\colon \mathcal{X}_\mathrm{c}^s \to \mathcal{X}_\mathrm{c}^s$ is defined as $\phi(w) = \psi(\mathrm{e}^{c^\star/2\mu}\|w\|_{\mathcal{X}_\mathrm{c}^s})w$, in which $\psi \in C^\infty(\mathbb{R})$ is a cut-off function satisfying

$$\psi(x) = \begin{cases} 1, & |x| \leq 1, \\ 0, & |x| \geq 2, \end{cases}$$

and $|\psi^{(l)}(x)| \leq 2^l$ for all $x \in \mathbb{R}$, $l \in \mathbb{N}_0$. Reversible solutions to equations (5.9)–(5.11) with $\|w(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \leq \mathrm{e}^{-c^\star/2\mu}$ correspond to reversible solutions of the original equations (2.55)–(2.57) (which are a reformulation of equations (2.53), (2.54)), where no cut-off function is used (see Theorem 4.26). To establish the following global existence result for the modified equations (5.9)–(5.11) we work in exponentially weighted spaces

$$C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{sh}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{wh}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{c}^s),$$

where $\eta \in [0,1)$ (see Lemma 2.13).

Theorem 5.2. *Equations (5.9)–(5.11) admit reversible solutions $(Z_{w_0}^\star, r_{w_0}^\star, w_{w_0}^\star)$ in*

$$C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh}) \times C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{c}^s),$$

parameterised by $w_0 \in \mathcal{X}_\mathrm{c}^s$ with $\|w_0\|_{\mathcal{X}_\mathrm{c}^s} \leq \mu\mathrm{e}^{-c^\star/2\mu}$, $R_\mathrm{c}w_0 = w_0$, such that

$$\|Z_{w_0}^\star\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})}, \|r_{w_0}^\star\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})}, \|w_{w_0}^\star\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{wh})} \leq \mathrm{e}^{-c^\star/2\mu}$$

and

$$\|w_{w_0}^\star\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{c}^s)} \leq \mathrm{e}^{-c^\star/2\mu}.$$

These solutions satisfy

$$\begin{aligned} \|Z_{w_0}^\star\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} &\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}, \\ \|r_{w_0}^\star\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})} &\lesssim \mu\mathrm{e}^{-c^\star/2\mu}, \\ \|w_{w_0}^\star\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{c}^s)} &\lesssim \mu\mathrm{e}^{-c^\star/2\mu}, \\ \|w_{w_0}^\star\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{c}^s)} &\lesssim \mu^\delta\mathrm{e}^{-c^\star/2\mu} \end{aligned}$$

for arbitrary but fixed $\delta \in [0,\frac{1}{2})$.

Definition 5.3. The set

$$W_\mathrm{loc}^\mathrm{cs} = \left\{\left(Z_{w_0}^\star(0), r_{w_0}^\star(0), w_{w_0}^\star(0)\right) : w_0 \in \mathcal{X}_\mathrm{c}^s,\ R_\mathrm{c}w_0 = w_0,\ \|w_0\|_{\mathcal{X}_\mathrm{c}^s} \leq \mu\mathrm{e}^{-c^\star/2\mu}\right\}$$

is called a *local centre-stable manifold for solutions to equations (2.55)–(2.57) at $\xi = 0$.*

Remark 5.4. Definition 5.3 imposes a weaker regularity assumption on w_0 than Definition 4.27, so that our local centre-stable manifold in the semilinear case is ‘larger’ than that in Definition 4.27 for the general quasilinear problem.

The iteration scheme

In order to prove Theorem 5.2 we first reformulate equations (5.9), (5.10) as integral equations by replacing G_1, G_2 and f_4 in equations (2.58), (2.59) with respectively $\underline{G}_1$, $\underline{G}_2$ and $\underline{f}_4$. We obtain

$$\begin{aligned} Z(\xi) = {} & \int_{-\infty}^{\xi} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{ss}} \Big(\underline{G}_1^{\mu}(Z, p^{\mu} + \mathcal{E}_{\mathrm{wh}} r, \mathcal{E}_{\mathrm{c}} w)(t) + \underline{G}_2^{\mu}(\mathcal{E}_{\mathrm{wh}} r)(t) \Big) \, \mathrm{d}t \\ & - \int_{\xi}^{\infty} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{su}} \Big(\underline{G}_1^{\mu}(Z, p^{\mu} + \mathcal{E}_{\mathrm{wh}} r, \mathcal{E}_{\mathrm{c}} w)(t) + \underline{G}_2^{\mu}(\mathcal{E}_{\mathrm{wh}} r)(t) \Big) \, \mathrm{d}t, \end{aligned} \tag{5.12}$$

$$\begin{aligned} r(\xi) = {} & \sum_{i=1}^{2} \Big(\int_0^{\xi} \Big\langle \Big(f_3^{\mu}(r)(t) + \underline{f}_4^{\mu}(Z, r, w)(t) \Big), (S_i^{\mu})^*(t) \Big\rangle \, \mathrm{d}t \Big) S_i^{\mu}(\xi) \\ & - \Big(\int_{\xi}^{\infty} \Big\langle \Big(f_3^{\mu}(r)(t) + \underline{f}_4^{\mu}(Z, r, w)(t) \Big), (U_i^{\mu})^*(t) \Big\rangle \, \mathrm{d}t \Big) U_i^{\mu}(\xi) \end{aligned} \tag{5.13}$$

(recall that the projections P_{ss}, P_{su} are defined below (2.58), (2.59); the fundamental system $\{S_1^{\mu}, S_2^{\mu}, U_1^{\mu}, U_2^{\mu}\}$ is constructed in Lemma 2.19). Next we write equation (5.11) as the fixed-point point problem

$$w(\xi) = S_{\mathrm{c}}^{\mu}(\xi) w(0) + \int_0^{\xi} S_{\mathrm{c}}^{\mu}(\xi - t) \Big(\mu \underline{b}(p^{\mu} + r, w)(t) + \underline{\hat{g}}^{\mu}(Z, p^{\mu} + r, w)(t) + \hat{h}^{\mu}(r)(t) \Big) \, \mathrm{d}t, \tag{5.14}$$

where $\{S_{\mathrm{c}}^{\mu}(\xi)\}_{\xi \in \mathbb{R}} \subseteq \mathcal{L}(\mathcal{X}_{\mathrm{c}}^{s})$ is the C_0-group generated by the operator L_{c}^{μ} (see Lemma 5.5 below).

We consider the following iteration scheme based on equations (5.12)–(5.14). Choose $w_0 \in \mathcal{X}_{\mathrm{c}}^{s}$ with $\|w_0\|_{\mathcal{X}_{\mathrm{c}}^{s}} \leq \mu \mathrm{e}^{-c^{\star}/2\mu}$, $R_{\mathrm{c}} w_0 = w_0$, let $j \in \mathbb{N}_0$ and suppose that

$$Z_{(j)} \in C_{\eta\lambda^{\mu}}(\mathbb{R}; \mathcal{X}_{\mathrm{sh}}), \quad r_{(j)} \in C_{\eta\lambda^{\mu}}([0,\infty); \mathcal{X}_{\mathrm{wh}}), \quad w_{(j)} \in C_{\eta\lambda^{\mu}}([0,\infty); \mathcal{X}_{\mathrm{c}}^{s})$$

satisfy

$$\|Z_{(j)}\|_{C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}_{\mathrm{sh}})} \leq \mathrm{e}^{-c^{\star}/2\mu}, \quad \|r_{(j)}\|_{C_{\mathrm{b}}([0,\infty); \mathcal{X}_{\mathrm{wh}})} \leq \mathrm{e}^{-c^{\star}/2\mu}, \quad \|w_{(j)}\|_{C([0, \mathrm{e}^{c^{\star}/2\mu}]; \mathcal{X}_{\mathrm{c}}^{s})} \leq \mathrm{e}^{-c^{\star}/2\mu}$$

with $R_{\mathrm{wh}} r_{(j)}(0) = r_{(j)}(0)$ and $R_{\mathrm{c}} w_{(j)}(0) = w_{(j)}(0)$. We define

$$Z_{(j+1)} \in C_{\eta\lambda^{\mu}}(\mathbb{R}; \mathcal{X}_{\mathrm{sh}})$$

by

$$\begin{aligned} Z_{(j+1)}(\xi) = {} & \int_{-\infty}^{\xi} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{ss}} (\underline{G}_{1(j)}^{\mu}(t) + \underline{G}_{2(j)}^{\mu}(t)) \, \mathrm{d}t \\ & - \int_{\xi}^{\infty} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{su}} (\underline{G}_{1(j)}^{\mu}(t) + \underline{G}_{2(j)}^{\mu}(t)) \, \mathrm{d}t. \end{aligned} \tag{5.15}$$

Similarly

$$r_{(j+1)} \in C_{\eta\lambda^{\mu}}([0,\infty); \mathcal{X}_{\mathrm{wh}})$$

is defined by

$$r_{(j+1)}(\xi) = \sum_{i=1}^{2}\Big(\int_0^\xi \Big\langle \Big(f_{3(j)}^\mu(t) + \underline{f}_{4(j)}^\mu(t)\Big), (S_i^\mu)^*(t)\Big\rangle \,\mathrm{d}t\Big) S_i^\mu(\xi) \\ - \Big(\int_\xi^\infty \Big\langle \Big(f_{3(j)}^\mu(t) + \underline{f}_{4(j)}^\mu(t)\Big), (U_i^\mu)^*(t)\Big\rangle \,\mathrm{d}t\Big) U_i^\mu(\xi), \tag{5.16}$$

and

$$w_{(j+1)} \in C_{\eta\lambda^\mu}([0,\infty); \mathcal{X}_\mathrm{c}^s)$$

by

$$w_{(j+1)}(\xi) = S_\mathrm{c}^\mu(\xi) w_0 + \int_0^\xi S_\mathrm{c}^\mu(\xi - t)\Big(\mu \underline{b}(z_{(j)}, w_{(j+1)})(t) + \underline{\hat{g}}_{(j)}^\mu(t) + \hat{h}_{(j)}^\mu(t)\Big)\,\mathrm{d}t. \tag{5.17}$$

Here we have abbreviated $z_{(j)} = p^\mu + r_{(j)}$ and $\underline{G}_{1(j)}^\mu = \underline{G}_1^\mu(Z_{(j)}, p^\mu + \mathcal{E}_\mathrm{wh} r_{(j)}, w_{(j)})$ with similar abbreviations for the remaining nonlinearities. Note that $R_\mathrm{wh} r_{(j+1)}(0) = r_{(j+1)}(0)$, since the functions U_1^μ, U_2^μ are reversible, and $R_\mathrm{c} w_{(j+1)}(0) = w_{(j+1)}(0)$ by construction.

In Lemmata 5.6 and 5.7 below we prove the convergence of the iteration scheme (5.15)–(5.17) and show that its limit $(Z^\star, r^\star, w^\star)$ lies in

$$\Big\{(Z, r, w) \in C_{\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_\mathrm{sh}) \times C_{\eta\lambda^\mu}\Big([0,\infty); \mathcal{X}_\mathrm{wh}\Big) \times C_{\eta\lambda^\mu}\Big([0,\infty); \mathcal{X}_\mathrm{c}^s\Big) : \\ \|Z\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})},\ \|r\|_{C_\mathrm{b}([0,\infty;\mathcal{X}_\mathrm{wh})},\ \|w\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_\mathrm{c}^s)} \le \mathrm{e}^{-c^\star/2\mu}, \\ \|w\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{c}^s)} \le \mathrm{e}^{-c^\star/2\mu},\ R_\mathrm{wh} r(0) = r(0),\ R_\mathrm{c} w(0) = w(0)\Big\}.$$

We verify *a posteriori* that $Z^\star(\xi) = R_\mathrm{sh} Z^\star(-\xi)$; extending $r^\star$ and $w^\star$ by reversibility thus establishes Theorem 5.2. The C_0-group generated by L_c^μ is constructed in the following lemma.

Lemma 5.5. *Consider the Riesz basis $\{e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\}_{m\in\mathbb{Z}}^{r=1,\ldots,r_m-\sigma_m}$ for $\mathcal{X}_\mathrm{c}^s$ (see the comments above Remark 2.9) with the associated purely imaginary eigenvalues $\lambda_{m,r}^\mu = \mathrm{i}\nu_{m,r}$ and denote the corresponding dual Riesz basis by $\{f_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\}_{m\in\mathbb{Z}}^{r=1,\ldots,r_m-\sigma_m}$. The formula*

$$\Big(S_\mathrm{c}^\mu(\xi) w\Big)(\tau) = \sum_{\substack{m\in\mathbb{Z}\\ 1\le r\le r_m-\sigma_m}} \mathrm{e}^{\xi\lambda_{m,r}^\mu} \langle w, f_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau},$$

where $\langle\cdot,\cdot\rangle$ denotes the usual scalar-product on $\mathcal{X}^s$, defines a strongly continuous group $\{S_\mathrm{c}^\mu(\xi)\}_{\xi\in\mathbb{R}}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}\xi}(S_\mathrm{c}^\mu(\xi) w) = L_\mathrm{c}^\mu S_\mathrm{c}^\mu(\xi) w$$

for all $w \in \mathcal{X}_\mathrm{c}^{s+1}$. The operator norm of $S_\mathrm{c}^\mu(\xi)$ is bounded independently of μ and ξ.

Proof. By definition $S_\mathrm{c}^\mu(0) w = w$ and $S_\mathrm{c}^\mu(\xi_1 + \xi_2) w = S_\mathrm{c}^\mu(\xi_1) S_\mathrm{c}^\mu(\xi_2) w$ for all $w \in \mathcal{X}_\mathrm{c}^s$. Moreover, since

$$L_\mathrm{c}^\mu w = \sum_{\substack{m\in\mathbb{Z}\\ 1\le r\le r_m-\sigma_m}} \lambda_{m,r}^\mu \langle w, f_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau},$$

we have that $\frac{\mathrm{d}}{\mathrm{d}\xi}(S_\mathrm{c}^\mu(\xi)w) = L_\mathrm{c}^\mu S_\mathrm{c}^\mu(\xi)w$ for all $w \in \mathcal{X}_\mathrm{c}^{s+1}$.

Finally, since $\{e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\}_{m\in\mathbb{Z}}^{r=1,\dots,r_m-\sigma_m}$ is a Riesz basis for $\mathcal{X}_\mathrm{c}^s$, there exists a scalar product $\langle\langle\cdot,\cdot\rangle\rangle$ on $\mathcal{X}_\mathrm{c}^s$ which induces a norm $|||\cdot|||_{\mathcal{X}_\mathrm{c}^s}$ which is equivalent to the norm $\|\cdot\|_{\mathcal{X}_\mathrm{c}^s}$ induced by $\langle\cdot,\cdot\rangle$ such that $\{e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\}_{m\in\mathbb{Z}}^{r=1,\dots,r_m-\sigma_m}$ is an orthonormal basis for $\mathcal{X}_\mathrm{c}^s$ with respect to $\langle\langle\cdot,\cdot\rangle\rangle$. It follows that

$$\begin{aligned}
\|S_\mathrm{c}^\mu(\xi)w\|_{\mathcal{X}_\mathrm{c}^s}^2 &= \Bigg\| \sum_{\substack{m\in\mathbb{Z}\\ 1\le r\le r_m-\sigma_m}} \mathrm{e}^{\xi\lambda_{m,r}^\mu}\langle w, f_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\Bigg\|_{\mathcal{X}_\mathrm{c}^s}^2 \\
&\lesssim \Bigg|\Bigg|\Bigg| \sum_{\substack{m\in\mathbb{Z}\\ 1\le r\le r_m-\sigma_m}} \mathrm{e}^{\xi\lambda_{m,r}^\mu}\langle w, f_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\Bigg|\Bigg|\Bigg|_{\mathcal{X}_\mathrm{c}^s}^2 \\
&= \sum_{\substack{m\in\mathbb{Z}\\ 1\le r\le r_m-\sigma_m}} |\mathrm{e}^{\mathrm{i}\xi\nu_{m,r}^\mu}|\,|||\langle w, f_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}|||_{\mathcal{X}_\mathrm{c}^s}^2 \\
&= \Bigg|\Bigg|\Bigg| \sum_{\substack{m\in\mathbb{Z}\\ 1\le r\le r_m-\sigma_m}} \langle w, f_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\rangle e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}\Bigg|\Bigg|\Bigg|_{\mathcal{X}_\mathrm{c}^s}^2 \\
&\lesssim \|w\|_{\mathcal{X}_\mathrm{c}^s}^2,
\end{aligned}$$

that is $S_\mathrm{c}^\mu(\xi)$ is bounded uniformly over μ and ξ. ■

Estimates for the iteration scheme

To prove that the sequence $\{(Z_{(j)}, r_{(j)}, w_{(j)})\}_{j\in\mathbb{N}_0}$ converges in

$$C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{sh}) \times C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_\mathrm{wh}) \times C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_\mathrm{c}^s)$$

to a limit

$$\begin{aligned}
(Z^\star, r^\star, w^\star) &\in \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \\
&\subseteq C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh}) \times C_\mathrm{b}([0,\infty);\mathcal{X}_\mathrm{wh}) \times C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_\mathrm{c}^s)
\end{aligned}$$

we show that it lies in $\overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0) \times \overline{B}_{\mathrm{e}^{-c^\star/2\mu}}(0)$ for appropriately chosen initial values. The result is summarised in the following lemma.

Lemma 5.6. *Suppose that the functions $Z_{(0)} \in C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})$, $r_{(0)} \in C_\mathrm{b}([0,\infty);\mathcal{X}_\mathrm{wh})$ and $w_{(0)} \in C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_\mathrm{c}^s)$ satisfy*

$$\begin{aligned}
\|Z_{(0)}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} &\le \mathrm{e}^{-c^\star/2\mu}, \\
\|r_{(0)}\|_{C_\mathrm{b}([0,\infty);\mathcal{X}_\mathrm{wh})} &\le \mathrm{e}^{-c^\star/2\mu}, \\
\|w_{(0)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_\mathrm{c}^s)} &\le \mathrm{e}^{-c^\star/2\mu}
\end{aligned}$$

and $\|w_0\|_{\mathcal{X}_\mathrm{c}^s} \le \mu\mathrm{e}^{-c^\star/2\mu}$. The estimates

$$\begin{aligned}
\sup_{j\in\mathbb{N}_0} \|Z_{(j)}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} &\le \mathrm{e}^{-c^\star/2\mu}, \\
\sup_{j\in\mathbb{N}_0} \|r_{(j)}\|_{C_\mathrm{b}([0,\infty);\mathcal{X}_\mathrm{wh})} &\le \mathrm{e}^{-c^\star/2\mu}
\end{aligned}$$

and

$$\sup_{j\in\mathbb{N}_0} \|w_{(j)}\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_c^s)} \le \mathrm{e}^{-c^\star/2\mu}, \qquad \sup_{j\in\mathbb{N}_0} \|w_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_c^s)} \le \mathrm{e}^{-c^\star/2\mu}$$

hold for all sufficiently small values of μ.

In order to prove Lemma 5.6 we work inductively and demonstrate that

$$\|Z_{(j+1)}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} \lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}, \tag{5.18}$$

$$\|r_{(j+1)}\|_{C_\mathrm{b}([0,\infty);\mathcal{X}_\mathrm{wh})} \lesssim \mu\mathrm{e}^{-c^\star/2\mu}, \tag{5.19}$$

$$\|w_{(j+1)}\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_c^s)} \lesssim \mu\mathrm{e}^{-c^\star/2\mu}, \tag{5.20}$$

$$\|w_{(j+1)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_c^s)} \lesssim \sqrt{\mu}|\log(\mu)|^2\mathrm{e}^{-c^\star/2\mu} \tag{5.21}$$

whenever

$$\|Z_{(j)}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} \le \mathrm{e}^{-c^\star/2\mu}, \quad \|r_{(j)}\|_{C_\mathrm{b}([0,\infty);\mathcal{X}_\mathrm{wh})} \le \mathrm{e}^{-c^\star/2\mu}, \quad \|w_{(j)}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_c^s)} \le \mathrm{e}^{-c^\star/2\mu}$$

for any $j \in \mathbb{N}_0$.

We begin with the result (5.18) for $Z_{(j+1)}$. Using the estimates for G_1, G_2 in Propositions 2.17 and 2.18 together with the fact that by construction $\|\phi(w)\|_{\mathcal{X}_c^s} \lesssim \mathrm{e}^{-c^\star/2\mu}$ for each $w \in \mathcal{X}_c^s$, we find that

$$\begin{aligned}
|\underline{G}_{1(j)}^\mu(\xi)| &= |G_1^\mu(Z_{(j)}(\xi), (p^\mu + \mathcal{E}_\mathrm{wh} r_{(j)})(\xi), \phi(\mathcal{E}_\mathrm{c} w_{(j)}(\xi)))| \\
&\lesssim \mu|p^\mu + \mathcal{E}_\mathrm{wh} r_{(j)}|\|(Z_{(j)}(\xi), \phi(\mathcal{E}_\mathrm{c} w_{(j)}(\xi)))\|_s + \mu\|(Z_{(j)}(\xi), \phi(\mathcal{E}_\mathrm{c} w_{(j)}(\xi)))\|_s^2 \\
&\qquad + |p^\mu + \mathcal{E}_\mathrm{wh} r_{(j)}|^3\|(Z_{(j)}(\xi), \phi(\mathcal{E}_\mathrm{c} w_{(j)}(\xi)))\|_s \\
&\lesssim \mu\mathrm{e}^{-c^\star/\mu} + \mu^2\mathrm{e}^{-c^\star/2\mu} \\
&\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}
\end{aligned}$$

as well as

$$|\underline{G}_{2(j)}^\mu(\xi)| \lesssim \mu^2\mathrm{e}^{-c^\star/\mu}.$$

We additionally use the estimate (2.60) for the projections onto the stable and unstable subspaces of L_sh^μ to conclude that

$$\begin{aligned}
|Z_{(j+1)}(\xi)| &\le \int_{-\infty}^{\xi} \|\mathrm{e}^{(\xi-t)L_\mathrm{sh}^\mu} P_\mathrm{ss}\|_{\mathcal{L}(\mathcal{X}_\mathrm{sh})} |\underline{G}_{1(j)}^\mu(t) + \underline{G}_{2(j)}^\mu(t)| \,\mathrm{d}t \\
&\qquad + \int_{\xi}^{\infty} \|\mathrm{e}^{(\xi-t)L_\mathrm{sh}^\mu} P_\mathrm{su}\|_{\mathcal{L}(\mathcal{X}_\mathrm{sh})} |\underline{G}_{1(j)}^\mu(t) + \underline{G}_{2(j)}^\mu(t)| \,\mathrm{d}t, \\
&\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}\left(\int_{-\infty}^{\xi} \mathrm{e}^{\omega_1 t} \,\mathrm{d}t\, \mathrm{e}^{-\omega_1\xi} + \int_{\xi}^{\infty} \mathrm{e}^{-\omega_2 t} \,\mathrm{d}t\, \mathrm{e}^{\omega_2\xi}\right) \\
&\le \mu^2\mathrm{e}^{-c^\star/2\mu}\left(\frac{1}{\omega_1} + \frac{1}{\omega_2}\right) \\
&\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu},
\end{aligned}$$

which is inequality (5.18).

The next step is the estimate (5.19) for $r_{(j+1)}$ on $\xi \in [0,\infty)$. Proposition 2.18 implies that

$$|f^\mu_{3(j)}(\xi)| = |f^\mu_3(r_{(j)}(\xi))| \lesssim |r_{(j)}(\xi)|^2 \leq \mathrm{e}^{-c^\star/\mu}$$

and that

$$\begin{aligned}|\underline{f}^\mu_{4(j)}(\xi)| &= |\underline{f}^\mu_4(Z_{(j)}(\xi), r_{(j)}(\xi), \phi(w_{(j)}(\xi)))|\\ &\lesssim \mu^3 \|(Z_{(j)}(\xi), r_{(j)}(\xi), \phi(w_{(j)}(\xi)))\|_s\\ &\leq \mu^3 \mathrm{e}^{-c^\star/2\mu}.\end{aligned}$$

The estimates for $\{S^\mu_1, S^\mu_2, U^\mu_1, U^\mu_2\}$ (see Lemma 2.19) therefore yield

$$\begin{aligned}|r_{(j+1)}(\xi)| &\leq \sum_{i=1}^2 \int_0^\xi \Big(|f^\mu_{3(j)}(t)| + |\underline{f}^\mu_{4(j)}(t)|\Big)|(S^\mu_i)^*(t)|\,\mathrm{d}t\,|S^\mu_i(\xi)|\\ &\qquad + \int_\xi^\infty \Big(|f^\mu_{3(j)}(t)| + |\underline{f}^\mu_{4(j)}(t)|\Big)|(U^\mu_i)^*(t)|\,\mathrm{d}t\,|U^\mu_i(\xi)|\\ &\lesssim \mu^3 \mathrm{e}^{-c^\star/2\mu}\Big(\int_0^\xi \mathrm{e}^{\lambda^\mu t}\,\mathrm{d}t\,\mathrm{e}^{-\lambda^\mu\xi} + \int_\xi^\infty \mathrm{e}^{-\lambda^\mu t}\,\mathrm{d}t\,\mathrm{e}^{\lambda^\mu \xi}\Big)\\ &\leq \mu \mathrm{e}^{-c^\star/2\mu}(2 - \mathrm{e}^{-\lambda^\mu\xi})\\ &\lesssim \mu \mathrm{e}^{-c^\star/2\mu},\end{aligned}$$

thus establishing inequality (5.19).

Next we establish the estimates (5.21), (5.20) for $w_{(j+1)}$. We first note that

$$\begin{aligned}\|w_{(j+1)}(\xi)\|_{\mathcal{X}^s_\mathrm{c}} &\lesssim \|w_0\|_{\mathcal{X}^s_\mathrm{c}} + \int_0^\xi \mu\|\underline{b}(p^\mu + r_{(j)}, w_{(j+1)})(t)\|_{\mathcal{X}^s_\mathrm{c}} + \|\underline{\hat{g}}^\mu_{(j)}(t)\|_{\mathcal{X}^s_\mathrm{c}} + \|\hat{h}^\mu_{(j)}(t)\|_{\mathcal{X}^s_\mathrm{c}}\,\mathrm{d}t\\ &\lesssim \mu \mathrm{e}^{-c^\star/2\mu} + \int_0^\xi \Big(\mu^3 \mathrm{e}^{-\eta\lambda^\mu t}\mathrm{e}^{-c^\star/2\mu} + \mu^3\mathrm{e}^{-\eta\lambda^\mu t}\|\phi(w_{(j)}(t))\|_{\mathcal{X}^s_\mathrm{c}} + \mu\mathrm{e}^{-c^\star/\mu}\\ &\qquad + \mu\mathrm{e}^{-c^\star/2\mu}\|\phi(w_{(j)}(t))\|_{\mathcal{X}^s_\mathrm{c}} + \mu\|\phi(w_{(j)}(t))\|^2_{\mathcal{X}^s_\mathrm{c}}\\ &\qquad + \mu^2\mathrm{e}^{-\eta\lambda^\mu t}\|\phi(w_{(j+1)}(t))\|_{\mathcal{X}^s_\mathrm{c}}\\ &\qquad + \mu\mathrm{e}^{-c^\star/2\mu}\|\phi(w_{(j+1)}(t))\|_{\mathcal{X}^s_\mathrm{c}}\Big)\,\mathrm{d}t \end{aligned} \tag{5.22}$$

due to the estimates for the nonlinearities given in Propositions 2.17 and 2.18 and the fact that the group $\{S^\mu_\mathrm{c}(\xi)\}_{\xi\in\mathbb{R}}$ is bounded independently of μ and ξ (Lemma 5.5). For $\xi \in [0,\infty)$ we use the estimate $\|\phi(w)\|_{\mathcal{X}^s_\mathrm{c}} \lesssim \mathrm{e}^{-c^\star/2\mu}$ for each $w \in \mathcal{X}^s_\mathrm{c}$ and deduce from (5.22) that

$$\begin{aligned}\mathrm{e}^{-\eta\lambda^\mu\xi}\|w_{(j+1)}(\xi)\|_{\mathcal{X}^s_\mathrm{c}} &\lesssim \mu\mathrm{e}^{-\eta\lambda^\mu\xi}\mathrm{e}^{-c^\star/2\mu} + \mathrm{e}^{-\eta\lambda^\mu\xi}\int_0^\xi \Big(\mu^2\mathrm{e}^{-\eta\lambda^\mu t}\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-c^\star/\mu}\Big)\,\mathrm{d}t\\ &\leq \mu\mathrm{e}^{-c^\star/2\mu} + \mathrm{e}^{-\eta\lambda^\mu\xi}\xi\Big(\mu^2\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-c^\star/\mu}\Big)\\ &\lesssim \mu\mathrm{e}^{-c^\star/2\mu},\end{aligned}$$

since $\mathrm{e}^{-\eta\lambda^\mu\xi}\xi \lesssim 1$, thus establishing inequality (5.20). For $\xi \in [0, \mathrm{e}^{c^\star/2\mu}]$ we use the estimate $\|\phi(w)\|_{\mathcal{X}_c^s} \lesssim \|w\|_{\mathcal{X}_c^s}$ in (5.22) to obtain

$$\|w_{(j+1)}(\xi)\|_{\mathcal{X}_c^s} \lesssim \mu\mathrm{e}^{-c^\star/2\mu} + \int_0^\xi \Big(\mu(\mathrm{e}^{-c^\star/\mu} + \mu^2\mathrm{e}^{-\eta\lambda^\mu t}\mathrm{e}^{-c^\star/2\mu}) + \mu(\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-\eta\lambda^\mu t})\|w_{(j+1)}(t)\|_{\mathcal{X}_c^s}\Big)\,\mathrm{d}t. \tag{5.23}$$

The proof of inequality (5.20) requires a two-step estimation method for inequality (5.23) (see the comments below the derivation of inequality (4.49)). We begin by establishing an estimate for $\|w_{(j+1)}(\xi)\|_{\mathcal{X}_c^s}$ on the short interval $[0, \xi^\star]$, where

$$\xi^\star = \frac{\alpha|\log\mu|}{\eta\lambda^\mu}$$

for an appropriately chosen constant $\alpha > 0$, so that $\mathrm{e}^{-\eta\lambda^\mu\xi^\star} = \mu^\alpha$. Using the facts that

$$\mu^2 \lesssim \lambda^\mu, \quad \mathrm{e}^{-c^\star/2\mu} \lesssim \mu, \quad \mathrm{e}^{-\eta\lambda^\mu\xi} \le 1, \quad \mathrm{e}^{-c^\star/\mu} \lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}$$

for sufficiently small values of μ, we find that

$$\|w_{(j+1)}(\xi)\|_{\mathcal{X}_c^s} \le \mu\mathrm{e}^{-c^\star/2\mu} + \int_0^\xi c_1\lambda^\mu\|w_{(j+1)}(t)\|_{\mathcal{X}_c^s}\,\mathrm{d}t + c_2\mu^3\mathrm{e}^{-c^\star/\mu}\xi$$

for some multiplicative constants $c_1, c_2 > 0$ and Gronwall's inequality implies that

$$\|w_{(j+1)}(\xi)\|_{\mathcal{X}_c^s} \le \Big(\mu\mathrm{e}^{-c^\star/2\mu} + c_2\mu^3\mathrm{e}^{-c^\star/\mu}\xi\Big)\mathrm{e}^{c_1\lambda^\mu\xi}.$$

Now we restrict to $\xi \in [0, \xi^\star]$ and choose $\alpha = \eta/2c_1$, so that

$$\mathrm{e}^{c_1\lambda^\mu\xi^\star} = \mathrm{e}^{\frac{|\log\mu|}{2}} = \frac{1}{\sqrt{\mu}},$$

and therefore

$$\begin{aligned}
\|w_{(j+1)}(\xi)\|_{\mathcal{X}_c^s} &\le \frac{1}{\sqrt{\mu}}(\mu\mathrm{e}^{-c^\star/2\mu} + c_2\mu^3\mathrm{e}^{-c^\star/\mu}\xi^\star) \\
&= \frac{1}{\sqrt{\mu}}\Big(\mu\mathrm{e}^{-c^\star/2\mu} + \frac{c_2\mu^3}{2c_1\lambda^\mu}\mathrm{e}^{-c^\star/\mu}|\log\mu|\Big) \\
&\lesssim \frac{1}{\sqrt{\mu}}(\mu\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-c^\star/\mu}|\log\mu|) \\
&\lesssim \frac{1}{\sqrt{\mu}}(\mu\mathrm{e}^{-c^\star/2\mu} + \mu\mathrm{e}^{-c^\star/\mu}|\log\mu|) \\
&\lesssim \sqrt{\mu}|\log\mu|\mathrm{e}^{-c^\star/2\mu}.
\end{aligned} \tag{5.24}$$

Inequality (5.24) gives the desired estimate for $\|w_{(j+1)}(\xi)\|_{\mathcal{X}_c^s}$ on the short interval $[0, \xi^\star]$ and we use this estimate to obtain a corresponding estimate on the interval $[0, \mathrm{e}^{c^\star/2\mu}]$.

From inequality (5.23) we conclude that

$$\|w_{(j+1)}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \lesssim \mu \mathrm{e}^{-c^\star/2\mu} + \mu \sup_{\xi\in[0,\mathrm{e}^{c^\star/2\mu}]} \|w_{(j+1)}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} + \mu^2 \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s}\,\mathrm{d}t \tag{5.25}$$

and observe that

$$\begin{aligned}
&\mu^2 \int_0^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s}\,\mathrm{d}t \\
&\quad= \mu^2 \int_0^{\xi^\star} \mathrm{e}^{-\eta\lambda^\mu t}\|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s}\,\mathrm{d}t + \mu^2 \int_{\xi^\star}^{\mathrm{e}^{c^\star/2\mu}} \mathrm{e}^{-\eta\lambda^\mu t}\|w_{(j+1)}(t)\|_s\,\mathrm{d}t \\
&\quad\leq \mu^2\xi^\star \sup_{t\in[0,\xi^\star]} \|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s} + \mu^2 \int_{\xi^\star}^{\infty} \mathrm{e}^{-\eta\lambda^\mu t}\,\mathrm{d}t \sup_{t\in[0,e^{c^\star/2\mu}]} \|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s} \\
&\quad\lesssim \mu^2\xi^\star \sup_{t\in[0,\xi^\star]} \|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s} + \frac{\mathrm{e}^{-\eta\lambda^\mu\xi^\star}}{\eta} \sup_{t\in[0,e^{c^\star/2\mu}]} \|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s} \\
&\quad\lesssim \sqrt{\mu}|\log\mu|^2\mathrm{e}^{-c^\star/2\mu} + \mu^\alpha \sup_{t\in[0,e^{c^\star/2\mu}]} \|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s},
\end{aligned}$$

in which we have used inequality (5.24) as well as the definitions $\xi^\star = |\log\mu|/2c_1\lambda^\mu$ and $\mu^\alpha = \mathrm{e}^{-\eta\lambda^\mu\xi^\star}$ in the last step. Inserting this estimate into (5.25), we conclude that

$$\sup_{t\in[0,e^{c^\star/2\mu}]} \|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s} \lesssim \sqrt{\mu}|\log\mu|^2\mathrm{e}^{-c^\star/2\mu} + (\mu+\mu^\alpha) \sup_{t\in[0,e^{c^\star/2\mu}]} \|w_{(j+1)}(t)\|_{\mathcal{X}_\mathrm{c}^s} \tag{5.26}$$

so that

$$\|w_{(j+1)}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \lesssim \sqrt{\mu}|\log(\mu)|^2\mathrm{e}^{-c^\star/2\mu},$$

that is inequality (5.21).

Convergence of the iteration scheme

Finally we prove that the sequence $\{(Z_{(j)}, r_{(j)}, w_{(j)})\}_{j\in\mathbb{N}_0}$ is convergent by considering the sequence $\{(Z_{(j+1)} - Z_{(j)}, r_{(j+1)} - r_{(j)}, w_{(j+1)} - w_{(j)})\}_{j\in\mathbb{N}_0}$ and showing that it is contractive. To this end we use the estimates for the derivatives of the nonlinearities given in Propositions 2.17 and 2.18 together with Lemma 5.6. The result is formulated in the following lemma. We use the abbreviation

$$\|(Z,r,w)\| = \|Z\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_\mathrm{sh})} + \|r\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_\mathrm{wh})} + \|w\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_\mathrm{c}^s)},$$

so that in particular

$$\|(Z(\xi), r(\xi), w(\xi))\|_s \leq \mathrm{e}^{\eta\lambda^\mu\xi}\|(Z,r,w)\|$$

and

$$\sup_{j\in\mathbb{N}_0} \|(Z_{(j)}, r_{(j)}, w_{(j)})\| \lesssim \mathrm{e}^{-c^\star/2\mu}$$

due to Lemma 5.6.

Lemma 5.7. *Let $j \in \mathbb{N}_0$ and define*

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

The estimate

$$\|(\tilde{Z}_{(j+1)}, \tilde{r}_{(j+1)}, \tilde{w}_{(j+1)})\| \leq \frac{1}{2}\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|$$

holds for all sufficiently small values of μ.

In order to prove Lemma 5.7 we proceed as in the proof of Lemma 5.6 and show that

$$\|\tilde{Z}_{(j+1)}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu^2\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|, \tag{5.27}$$

$$\|\tilde{r}_{(j+1)}\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_{\mathrm{wh}})} \lesssim \mu\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|. \tag{5.28}$$

$$\|\tilde{w}_{(j+1)}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s)} \lesssim \mu\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\| \tag{5.29}$$

for all $j \in \mathbb{N}_0$.

We begin with the hyperbolic part $(\tilde{Z}_{(j+1)}, \tilde{r}_{(j+1)})$. From the definitions (5.15), (5.16) we obtain the formulae

$$\begin{aligned}\tilde{Z}_{(j+1)}(\xi) = &\int_{-\infty}^{\xi} \mathrm{e}^{(\xi-t)L^\mu_{\mathrm{sh}}} P_{\mathrm{ss}}(\underline{G}^\mu_{1(j)}(t) - \underline{G}^\mu_{1(j-1)}(t) + \underline{G}^\mu_{2(j)}(t) - \underline{G}^\mu_{2(j-1)}(t))\,\mathrm{d}t \\ &- \int_{\xi}^{\infty} \mathrm{e}^{(\xi-t)L^\mu_{\mathrm{sh}}} P_{\mathrm{su}}(\underline{G}^\mu_{1(j)}(t) - \underline{G}^\mu_{1(j-1)}(t) + \underline{G}^\mu_{2(j)}(t) - \underline{G}^\mu_{2(j-1)}(t))\,\mathrm{d}t\end{aligned}$$

and

$$\begin{aligned}\tilde{r}_{(j+1)}(\xi) = \sum_{i=1}^{2}\Bigg(&\left(\int_0^\xi \langle f^\mu_{3(j)}(t) - f^\mu_{3(j-1)}(t) + \underline{f}^\mu_{4(j)}(t) - \underline{f}^\mu_{4(j-1)}(t), (S^\mu_i)^*(t)\rangle\,\mathrm{d}t\right) S^\mu_i(\xi) \\ &- \left(\int_\xi^\infty \langle (f^\mu_{3(j)} - f^\mu_{3(j-1)} + \underline{f}^\mu_{4(j)} - \underline{f}^\mu_{4(j-1)})(t), (U^\mu_i)^*(t)\rangle\,\mathrm{d}t\right) U^\mu_i(\xi)\Bigg).\end{aligned}$$

Using the fact that $\|\phi(w_1) - \phi(w_2)\|_{\mathcal{X}_{\mathrm{c}}^s} \lesssim \|w_1 - w_2\|_{\mathcal{X}_{\mathrm{c}}^s}$ uniformly over $w_1, w_2 \in \mathcal{X}_{\mathrm{c}}^s$ we conclude that

$$\begin{aligned}&\Big\|\Big(\tilde{Z}_{(j)}(t), (\mathcal{E}_{\mathrm{wh}}\tilde{r}_{(j)})(t), \phi\Big((\mathcal{E}_{\mathrm{c}}w_{(j)})(t)\Big) - \phi\Big((\mathcal{E}_{\mathrm{c}}w_{(j-1)})(t)\Big)\Big)\Big\|_s \\ &\quad\lesssim \Big\|\Big(\tilde{Z}_{(j)}(t), (\mathcal{E}_{\mathrm{wh}}\tilde{r}_{(j)})(t), (\mathcal{E}_{\mathrm{c}}\tilde{w}_{(j)})(t)\Big)\Big\|_s \\ &\quad\lesssim \mathrm{e}^{\eta\lambda^\mu t}\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|\end{aligned}$$

and hence the estimates for $\mathrm{d}G_1$ and $\mathrm{d}G_2$ given in Propositions 2.17 and 2.18 and Lemma 4.30 together with Lemma 5.6 imply that

$$\begin{aligned}|\underline{G}^\mu_{1(j)}(\xi) - \underline{G}^\mu_{1(j-1)}(\xi)| &\lesssim \mu^2\mathrm{e}^{\eta\lambda^\mu t}\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|, \\ |\underline{G}^\mu_{2(j)}(\xi) - \underline{G}^\mu_{2(j-1)}(\xi)| &\lesssim \mu^2\mathrm{e}^{\eta\lambda^\mu t}\|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|.\end{aligned}$$

The calculation in the proof of (5.18) yield

$$\|\tilde{Z}_{(j+1)}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu^2 \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|,$$

thus proving inequality (5.27).

In the same fashion we obtain

$$\begin{aligned}|f^\mu_{3(j)}(\xi) - f^\mu_{3(j-1)}(\xi)| &\lesssim \mu^3 \mathrm{e}^{\eta\lambda^\mu t} \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|,\\ |\underline{f}^\mu_{4(j)}(\xi) - \underline{f}^\mu_{4(j-1)}(\xi)| &\lesssim \mu^3 \mathrm{e}^{\eta\lambda^\mu t} \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|\end{aligned}$$

and hence

$$\begin{aligned}\mathrm{e}^{-\eta\lambda^\mu \xi} |\tilde{r}_{(j+1)}(\xi)| &\lesssim \mu^3 \Big(\mathrm{e}^{-\eta\lambda^\mu\xi} \int_0^\xi \mathrm{e}^{(\eta+1)\lambda^\mu t} \,\mathrm{d}t \, \mathrm{e}^{-\lambda^\mu \xi} \\ &\qquad + \mathrm{e}^{-\eta\lambda^\mu\xi} \int_\xi^\infty \mathrm{e}^{(\eta-1)\lambda^\mu t} \,\mathrm{d}t \, \mathrm{e}^{\lambda^\mu\xi} \Big) \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\| \\ &= \frac{\mu^3}{\lambda^\mu} \Big(\frac{1}{\eta+1}(1 - \mathrm{e}^{-(\eta+1)\lambda^\mu \xi}) + \frac{1}{1-\eta} \Big) \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\| \\ &\lesssim \mu \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|,\end{aligned}$$

where we have used the fact that $\eta < 1$, so that the integral

$$\int_\xi^\infty \mathrm{e}^{(\eta-1)\lambda^\mu t} \,\mathrm{d}t$$

converges. It follows that

$$\|\tilde{r}_{(j+1)}\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_{\mathrm{wh}})} \lesssim \mu \|(\tilde{Z}_{(j)}, \tilde{r}_{(j)}, \tilde{w}_{(j)})\|,$$

which is inequality (5.28).

Next we establish the estimate (5.29) for $\tilde{w}_{(j+1)}$. From definition (5.17) we obtain

$$\begin{aligned}\tilde{w}_{(j+1)}(\xi) = \int_0^\xi S^\mu_{\mathrm{c}}(\xi - t) \Big(&\mu \underline{b}(p^\mu + r_{(j)}, w_{(j+1)})(t) - \mu \underline{b}(p^\mu + r_{(j-1)}, w_{(j)})(t) \\ &+ \underline{\hat{g}}^\mu_{(j)}(t) - \underline{\hat{g}}^\mu_{(j-1)}(t) + \hat{h}^\mu_{(j)}(t) - \hat{h}^\mu_{(j-1)}(t) \Big) \,\mathrm{d}t.\end{aligned}$$

The estimates for the derivatives $\mathrm{d}\hat{g}$ and $\mathrm{d}\hat{h}$ given in Propositions 2.17 and 2.18 together with Lemma 5.6 imply

$$\|\underline{\hat{g}}^\mu_{(j)}(t) - \underline{\hat{g}}^\mu_{(j-1)}(t)\|_{\mathcal{X}^s_{\mathrm{c}}} \lesssim (\mu \mathrm{e}^{-c^\star/2\mu} + \mu^3 \mathrm{e}^{-\eta\lambda^\mu t}) \|(\tilde{Z}_{(j)}(t), \tilde{r}_{(j)}(t), \tilde{w}_{(j)}(t))\|_s,$$

$$\|\hat{h}^\mu_{(j)}(t) - \hat{h}^\mu_{(j-1)}(t)\|_{\mathcal{X}^s_{\mathrm{c}}} \lesssim \mu \mathrm{e}^{-c^\star/\mu} \|(\tilde{Z}_{(j)}(t), \tilde{r}_{(j)}(t), \tilde{w}_{(j)}(t))\|_s$$

and

$$\begin{aligned}&\|\underline{b}(p^\mu + r_{(j)}, w_{(j+1)})(t) - \underline{b}(p^\mu + r_{(j-1)}, w_{(j)})(t)\|_{\mathcal{X}^s_{\mathrm{c}}} \\ &\quad \le \|\underline{b}(p^\mu + r_{(j)}, \tilde{w}_{(j+1)})(t)\|_{\mathcal{X}^s_{\mathrm{c}}} + \|\underline{b}(\tilde{r}_{(j)}, w_{(j)})(t)\|_{\mathcal{X}^s_{\mathrm{c}}} \\ &\quad \lesssim (\mu \mathrm{e}^{-\eta\lambda^\mu t} + \mathrm{e}^{-c^\star/2\mu}) \|\tilde{w}_{(j+1)}\|_{\mathcal{X}^s_{\mathrm{c}}} + \mathrm{e}^{-c^\star/2\mu} \|(\tilde{Z}_{(j)}(t), \tilde{r}_{(j)}(t), \tilde{w}_{(j)}(t))\|_s.\end{aligned}$$

It follows that

$$\begin{aligned}
\mathrm{e}^{-\eta\lambda^\mu\xi}\|\tilde{w}_{(j+1)}(\xi)\|_{\mathcal{X}_c^s} &\lesssim \mathrm{e}^{-\eta\lambda^\mu\xi}\int_0^\xi \Big(\mu(\mu\mathrm{e}^{-\eta\lambda^\mu t}+\mathrm{e}^{-c^\star/2\mu})\|\tilde{w}_{(j+1)}(t)\|_{\mathcal{X}_c^s} \\
&\qquad\qquad + \mu\mathrm{e}^{-c^\star/2\mu}\|(\tilde{Z}_{(j)}(t),\tilde{r}_{(j)}(t),\tilde{w}_{(j)}(t))\|_s\Big)\,\mathrm{d}t \\
&\leq \mu\mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{-\eta\lambda^\mu\xi}\int_0^\xi \mathrm{e}^{\eta\lambda^\mu t}\,\mathrm{d}t\,\|(\tilde{Z}_{(j)},\tilde{r}_{(j)},\tilde{w}_{(j)})\| \\
&\qquad + \left(\mathrm{e}^{-\eta\lambda^\mu\xi}\int_0^\xi\Big(\mu^2+\mu\mathrm{e}^{-c^\star/2\mu}\mathrm{e}^{\eta\lambda^\mu t}\Big)\,\mathrm{d}t\right)\|\tilde{w}_{(j+1)}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_c^s)} \\
&\leq \frac{\mu\mathrm{e}^{-c^\star/2\mu}}{\eta\lambda^\mu}(1-\mathrm{e}^{-\eta\lambda^\mu\xi})\|(\tilde{Z}_{(j)},\tilde{r}_{(j)},\tilde{w}_{(j)})\| \\
&\qquad + \left(\mu^2\mathrm{e}^{-\eta\lambda^\mu\xi}\xi+\frac{\mu\mathrm{e}^{-c^\star/2\mu}}{\eta\lambda^\mu}(1-\mathrm{e}^{-\eta\lambda^\mu\xi})\right)\|\tilde{w}_{(j+1)}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_c^s)} \\
&\lesssim \mu\|(\tilde{Z}_{(j)},\tilde{r}_{(j)},\tilde{w}_{(j)})\|+\mu\|\tilde{w}_{(j+1)}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_c^s)}
\end{aligned}$$

and thus

$$\|\tilde{w}_{(j+1)}\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_c^s)} \lesssim \mu\|(\tilde{Z}_{(j)},\tilde{r}_{(j)},\tilde{w}_{(j)})\|$$

were we again have used the fact that the group $\{S_c^\mu(\xi)\}_{\xi\in\mathbb{R}}$ is uniformly bounded over μ and ξ. This calculation establishes inequality (5.29).

The existence of a limit $(Z^\star, r^\star, w^\star)$ of the sequence $\{(Z_{(j)}, r_{(j)}, w_{(j)})\}_{j\in\mathbb{N}_0}$ such that

$$\begin{aligned}
\|Z^\star\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} &\lesssim \mu^2\mathrm{e}^{-c^\star/2\mu}, \\
\|r^\star\|_{C_\mathrm{b}([0,\infty);\mathcal{X}_\mathrm{wh})} &\lesssim \mu\mathrm{e}^{-c^\star/2\mu}, \\
\|w^\star\|_{C_{\eta\lambda^\mu}([0,\infty);\mathcal{X}_c^s)} &\lesssim \mu\mathrm{e}^{-c^\star/2\mu} \\
\|w^\star\|_{C([-\mathrm{e}^{c^\star/2\mu},\mathrm{e}^{c^\star/2\mu}];\mathcal{X}_c^s)} &\lesssim \mu^\delta\mathrm{e}^{-c^\star/2\mu}, \qquad \delta\in[0,\tfrac{1}{2}),
\end{aligned}$$

now follows from Lemmata 5.6 and 5.7. The statement of Theorem 5.2 is obtained by observing that $\xi\mapsto(R_\mathrm{sh}Z^\star(-\xi), r^\star(\xi), w^\star(\xi))$ is also a limit of $\{(Z_{(j)}, r_{(j)}, w_{(j)})\}_{j\in\mathbb{N}_0}$, that is $Z^\star(\xi)=R_\mathrm{sh}Z^\star(-\xi)$, and extending $r^\star$, $w^\star$ by reversibility to the entire real line.

5.2 Construction of a local centre manifold

In this section we solve the modified equations (5.6)–(5.8), namely

$$Z_\xi = L_\mathrm{sh}^\mu Z + \underline{G}_1^\mu(Z,z,w) + G_2^\mu(z), \tag{5.30}$$

$$z_\xi = L_\mathrm{wh}^\mu z + f_1^\mu(z) + \underline{f}_2^\mu(Z,z,w), \tag{5.31}$$

$$w_\xi = L_c^\mu w + \mu\underline{b}(z,w) + \underline{\hat{g}}^\mu(Z,z,w) + \hat{h}^\mu(z), \tag{5.32}$$

which are obtained from equations (2.53), (2.54) by decomposing equation (2.54) for $q=Z+w$ into the corresponding component equations for $Z=P_\mathrm{sh}q$, $w=P_\mathrm{c}q$ and applying the cut-off function ϕ (see the comments below equations (5.9)–(5.11)) to the w-component of the nonlinearities. Here we have dropped the tildes in the formulae for the functions $\tilde{f}_1^\mu$ and $\tilde{f}_2^\mu$ in equation (2.53) and defined

$$\underline{f}_2^\mu(Z,z,w) = f_2^\mu(Z,z,\phi(w)).$$

Exponentially small solutions to equations (5.30)–(5.32) (satisfying in particular $\|w(\xi)\|_{\mathcal{X}_c^s} \leq \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \in \mathbb{R}$) correspond to exponentially small solutions to the original equations (2.53), (2.54). Here we construct solutions to (5.30)–(5.32) whose

(i) hyperbolic part is exponentially small on the entire real line,

(ii) centre part satisfies $\|w(\xi)\|_{\mathcal{X}_c^s} \leq \mathrm{e}^{-c^\star/2\mu}$ for all $\xi \geq \xi^\star$ if $\|w(\xi^\star)\|_{\mathcal{X}_c^s} \leq \frac{1}{2}\mathrm{e}^{-c^\star/2\mu}$ for some $\xi^\star > 0$.

The exact result is given in Lemma 5.11 below.

We begin with the construction of a (local) centre manifold for equations (5.30)–(5.32). This manifold is defined as the graph of a reduction function Ψ which maps the initial value of the centre part of a solution to (5.30)–(5.32) to the initial value of the corresponding exponentially small hyperbolic part.

We reformulate equations (5.30), (5.31) as integral equations and (5.32) as fixed-point problem by using the fundamental solution set $\{S_3^\mu, S_4^\mu, U_3^\mu, U_4^\mu\}$ (see Lemma 2.21) in equation (5.31). Employing the corresponding iteration scheme

$$\begin{aligned}
Z_{(j+1)}(\xi) &= \int_{-\infty}^{\xi} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{ss}}(\underline{G}_{1(j)}^\mu(t) + \underline{G}_{2(j)}^\mu(t))\,\mathrm{d}t \\
&\qquad - \int_{\xi}^{\infty} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{su}}(\underline{G}_{1(j)}^\mu(t) + \underline{G}_{2(j)}^\mu(t))\,\mathrm{d}t, \\
z_{(j+1)}(\xi) &= \sum_{i=3}^{4}\Big(\int_{-\infty}^{\xi} \Big\langle \Big(f_{1(j)}^\mu(t) + \underline{f}_{2(j)}^\mu(t)\Big), S_i^*(t)\Big\rangle\,\mathrm{d}t\Big) S_i(\xi) \\
&\qquad\qquad - \Big(\int_{\xi}^{\infty} \Big\langle \Big(f_{1(j)}^\mu(t) + \underline{f}_{2(j)}^\mu(t)\Big), U_i^*(t)\Big\rangle\,\mathrm{d}t\Big) U_i(\xi), \\
w_{(j+1)}(\xi) &= S_{\mathrm{c}}^\mu(\xi)w_0 + \int_0^{\xi} S_{\mathrm{c}}^\mu(\xi-t)\Big(\mu\underline{b}(z_{(j)}, w_{(j+1)})(t) + \underline{\hat{g}}_{(j)}^\mu(t) + \hat{h}_{(j)}^\mu(t)\Big)\,\mathrm{d}t
\end{aligned}$$

and repeating the arguments in Section 5.1 for arbitrary $w_0 \in \mathcal{X}_c^s$ yields the following result.

Theorem 5.8. *For each $w_0 \in \mathcal{X}_c^s$ equations (5.30)–(5.32) admit a unique solution*

$$(Z_{w_0}^{\star\star}, z_{w_0}^{\star\star}, w_{w_0}^{\star\star}) \in C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_c^s)$$

satisfying

$$\|Z_{w_0}^{\star\star}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|z_{w_0}^{\star\star}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

In the context of Theorem 5.8 we define the mapping

$$\Psi\colon \mathcal{X}_{\mathrm{c}}^{s} \to \mathcal{X}_{\mathrm{sh}} \times \mathcal{X}_{\mathrm{wh}}, \quad \Psi(w_0) = \Big(Z_{w_0}^{\star\star}(0), z_{w_0}^{\star\star}(0)\Big).$$

Taking any solution

$$(Z, z, w) \in C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}_{\mathrm{wh}}) \times C(\mathbb{R}; \mathcal{X}_{\mathrm{c}}^{s})$$

to equations (5.30)–(5.32) with

$$\|Z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \leq \mathrm{e}^{-c^\star/2\mu}$$

we find that

$$w(\xi) = S_{\mathrm{c}}^{\mu}(\xi)w(0) + \int_0^\xi S_{\mathrm{c}}^{\mu}(\xi - t)\Big(\mu \underline{b}(z, w) + \underline{\hat{g}}^{\mu}(Z, z, w) + \hat{h}^{\mu}(z)\Big)\,\mathrm{d}t$$

and thus

$$\begin{aligned} \|w(\xi)\|_{\mathcal{X}_{\mathrm{c}}^{s}} &\lesssim \|w(0)\|_{\mathcal{X}_{\mathrm{c}}^{s}} + |\xi| \|\mu \underline{b}(z, w) + \underline{\hat{g}}^{\mu}(Z, z, w) + \hat{h}^{\mu}(z)\|_{\mathcal{X}_{\mathrm{c}}^{s}} \\ &\lesssim \|w(0)\|_{\mathcal{X}_{\mathrm{c}}^{s}} + |\xi|\mathrm{e}^{-c^\star/\mu}. \end{aligned} \tag{5.33}$$

We conclude using the inequality $\mathrm{e}^{-\eta\lambda^\mu|\xi|}|\xi| \lesssim 1$ that $\|w\|_{C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^{s})} < \infty$, so that in fact

$$(Z, z, w) \in C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R}; \mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_{\mathrm{c}}^{s}).$$

Theorem 5.8 therefore implies that $(Z(0), z(0)) = \Psi(w(0))$ and this fact allows us to make the following definition.

Definition 5.9. The set

$$W^{\mathrm{c}} = \Big\{\Big(Z_{w_0}^{\star\star}(0), z_{w_0}^{\star\star}(0), w_{w_0}^{\star\star}(0)\Big) : w_0 \in \mathcal{X}_{\mathrm{c}}^{s}\Big\} = \Big\{\Big(\Psi(w_0), w_0\Big) : w_0 \in \mathcal{X}_{\mathrm{c}}^{s}\Big\}$$

is called a *centre manifold for solutions to equations* (5.30)–(5.32) *at* $\xi = 0$.

Since the system (5.30)–(5.32) is invariant under translations in ξ, the function

$$(\tilde{Z}(\xi), \tilde{z}(\xi), \tilde{w}(\xi)) = (Z(\xi + \xi_0), z(\xi + \xi_0), w(\xi + \xi_0)), \quad \xi_0 \in \mathbb{R},$$

solves (5.30)–(5.32) whenever $(Z(\xi), z(\xi), w(\xi))$ is a solution. It follows that

$$(\tilde{Z}(0), \tilde{z}(0)) = \Psi(\tilde{w}(0))$$

(see the comments above Definition 5.9), so that

$$(Z(\xi_0), z(\xi_0)) = \Psi(w(\xi_0)).$$

We conclude that W^{c} is an invariant manifold, that is $(Z(\xi), z(\xi), w(\xi)) \in W^{\mathrm{c}}$ for all $\xi \in \mathbb{R}$ and in particular

$$\|Z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \leq \mathrm{e}^{-c^\star/2\mu},$$

if $(Z(0), z(0), w(0)) \in W^{\mathrm{c}}$.

Next we restrict the domain of Ψ to $\{\|w_0\|_{\mathcal{X}_{\mathrm{c}}^s} \leq \mathrm{e}^{-c^\star/2\mu}\}$ and consider the *local centre manifold*

$$W_{\mathrm{loc}}^{\mathrm{c}} = \{(Z, z, w) \in \mathcal{X}^s \colon (Z, z) = \Psi(w),\ w \in \mathcal{X}_{\mathrm{c}}^s,\ \|w\|_{\mathcal{X}_{\mathrm{c}}^s} \leq \mathrm{e}^{-c^\star/2\mu}\} \subseteq W^{\mathrm{c}}$$

consisting of those points on W^{c} which have exponentially small centre part. Note that in general $W_{\mathrm{loc}}^{\mathrm{c}}$ is not an invariant manifold. However, each solution (Z, z, w) of equations (5.30)–(5.32) with

$$\|Z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})}, \|w\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s)} \leq \mathrm{e}^{-c^\star/2\mu}$$

satisfies $(Z(\xi), z(\xi), w(\xi)) \in W_{\mathrm{loc}}^{\mathrm{c}}$ for all $\xi \in \mathbb{R}$. On the other hand, each solution (Z, z, w) of equations (5.30)–(5.32) with $(Z(\xi_0), z(\xi_0), w(\xi_0)) \in W_{\mathrm{loc}}^{\mathrm{c}}$ for some $\xi_0 \in \mathbb{R}$ remains on $W_{\mathrm{loc}}^{\mathrm{c}}$ as long as

$$|Z(\xi)|, |z(\xi)|, \|w(\xi)\|_{\mathcal{X}_{\mathrm{c}}^s} \leq \mathrm{e}^{-c^\star/2\mu}$$

(in this case we have that $(Z(\xi), z(\xi), w(\xi)) \in W^{\mathrm{c}}$ for all $\xi \in \mathbb{R}$, since W^{c} is an invariant manifold, and $\|w(\xi)\|_{\mathcal{X}_{\mathrm{c}}^s} \leq \mathrm{e}^{-c^\star/2\mu}$ implies that $(Z(\xi), z(\xi), w(\xi)) \in W_{\mathrm{loc}}^{\mathrm{c}}$).

The next step is to show that $\|w(\xi_0)\|_{\mathcal{X}_{\mathrm{c}}^s} \leq \frac{1}{2}\mathrm{e}^{-c^\star/2\mu}$ implies that

$$(Z(\xi), z(\xi), w(\xi)) \in W_{\mathrm{loc}}^{\mathrm{c}}$$

for all $\xi \geq \xi_0$. The proof requires the following estimate for the function Ψ.

Lemma 5.10. *The function* $\Psi \colon \mathcal{X}_{\mathrm{c}}^s \to \mathcal{X}_{\mathrm{sh}} \times \mathcal{X}_{\mathrm{wh}}$ *satisfies the estimate*

$$|\Psi(w_0)| \lesssim \mu \|w_0\|_{\mathcal{X}_{\mathrm{c}}^s}$$

for each $w_0 \in \mathcal{X}_{\mathrm{c}}^s$ *with* $\|w_0\|_{\mathcal{X}_{\mathrm{c}}^s} \leq \mathrm{e}^{-c^\star/2\mu}$.

Proof. Fix $w_0 \in \mathcal{X}_{\mathrm{c}}^s$ with $\|w_0\|_{\mathcal{X}_{\mathrm{c}}^s} \leq \mathrm{e}^{-c^\star/2\mu}$ and define the mappings

$$\begin{aligned}
\mathcal{F}_{\mathrm{sh}} &\colon C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s) \to C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})\\
\mathcal{F}_{\mathrm{wh}} &\colon C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s) \to C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})\\
\mathcal{F}_{\mathrm{c}} &\colon C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s) \to C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s)
\end{aligned}$$

by

$$\begin{aligned}
\big(\mathcal{F}_{\mathrm{sh}}(Z,z,w)\big)(\xi) &= \int_{-\infty}^{\xi} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{ss}}(\underline{G}_1^{\mu}(Z,z,w)(t) + \underline{G}_2^{\mu}(z)(t))\,\mathrm{d}t \\
&\qquad - \int_{\xi}^{\infty} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{su}}(\underline{G}_1^{\mu}(Z,z,w)(t) + \underline{G}_2^{\mu}(z)(t))\,\mathrm{d}t, \\
\big(\mathcal{F}_{\mathrm{wh}}(Z,z,w)\big)(\xi) &= \sum_{i=3}^{4}\Big(\int_{-\infty}^{\xi} \Big\langle \Big(f_1^{\mu}(z)(t) + \underline{f}_2^{\mu}(Z,z,w)(t)\Big), S_i^*(t)\Big\rangle\,\mathrm{d}t\Big) S_i(\xi) \\
&\qquad - \Big(\int_{\xi}^{\infty} \Big\langle \Big(f_1^{\mu}(z)(t) + \underline{f}_2^{\mu}(Z,z,w)(t)\Big), U_i^*(t)\Big\rangle\,\mathrm{d}t\Big) U_i(\xi), \\
\big(\mathcal{F}_{\mathrm{c}}(Z,z,w)\big)(\xi) &= S_{\mathrm{c}}^{\mu}(\xi)w_0 + \int_0^{\xi} S_{\mathrm{c}}^{\mu}(\xi-t)\Big(\underline{\hat{g}}^{\mu}(Z,z,w)(t) + \mu\underline{b}(z,w)(t) + \hat{h}^{\mu}(z)(t)\Big)\,\mathrm{d}t.
\end{aligned}$$

It follows from Lemma 5.6 that the mappings $\mathcal{F}_{\mathrm{sh}}$ and $\mathcal{F}_{\mathrm{wh}}$ take values in respectively $C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})$ and $C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})$. Furthermore the arguments in the proof of Lemma 5.7 show that

$$\|\mathcal{F}_{\mathrm{sh}}(Z_1,z_1,w_1) - \mathcal{F}_{\mathrm{sh}}(Z_2,z_2,w_2)\|_{C_{\eta\lambda^{\mu}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu^2 \|(Z_1-Z_2, z_1-z_2, w_1-w_2)\|$$

and

$$\|\mathcal{F}_{\mathrm{wh}}(Z_1,z_1,w_1) - \mathcal{F}_{\mathrm{wh}}(Z_2,z_2,w_2)\|_{C_{\eta\lambda^{\mu}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu \|(Z_1-Z_2, z_1-z_2, w_1-w_2)\|$$

as well as

$$\begin{aligned}
&\Big\|\Big(\mathcal{F}_{\mathrm{c}}(Z_1,z_1,w_1)\Big)(\xi) - \Big(\mathcal{F}_{\mathrm{c}}(Z_2,z_2,w_2)\Big)(\xi)\Big\|_s \\
&\quad\lesssim \frac{\mu \mathrm{e}^{-c^{\star}/2\mu}}{\eta\lambda^{\mu}}(1-\mathrm{e}^{-\eta\lambda^{\mu}|\xi|})\|(Z_1-Z_2, z_1-z_2, w_1-w_2)\| \\
&\quad\lesssim \mu\|(Z_1-Z_2, z_1-z_2, w_1-w_2)\|
\end{aligned}$$

for all $\xi\in\mathbb{R}$, where we have used the notation introduced above Lemma 5.7.

The mapping $\mathcal{F} = (\mathcal{F}_{\mathrm{sh}}, \mathcal{F}_{\mathrm{wh}}, \mathcal{F}_{\mathrm{c}})$ is thus a contraction with Lipschitz constant less than or equal to $\frac{1}{2}$ and each solution $(Z_{w_0}^{\star\star}, z_{w_0}^{\star\star}, w_{w_0}^{\star\star})$ of equations (5.30)–(5.32) is a fixed point of $\mathcal{F}$, that is $\mathcal{F}(Z_{w_0}^{\star\star}, z_{w_0}^{\star\star}, w_{w_0}^{\star\star}) = (Z_{w_0}^{\star\star}, z_{w_0}^{\star\star}, w_{w_0}^{\star\star})$. It follows that

$$\|(Z,z,w) - (Z_{w_0}^{\star\star}, z_{w_0}^{\star\star}, w_{w_0}^{\star\star})\| \le 2\|(Z,z,w) - \mathcal{F}(Z,z,w)\|$$

for all

$$(Z,z,w) \in C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^{\mu}}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^{s})$$

with $\|Z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \le \mathrm{e}^{-c^{\star}/2\mu}$. This estimate for the choice $Z=0$, $z=0$ and

$w = S_{\mathrm{c}}^{\mu}(\cdot)w_0$ yields

$$\begin{aligned}
|\Psi(w_0)| &= \|(Z_{w_0}^{\star\star}(0), z_{w_0}^{\star\star}(0), w_0) - (0,0,w_0)\|_s \\
&= \|(Z_{w_0}^{\star\star}(0), z_{w_0}^{\star\star}(0), w_{w_0}^{\star\star}(0)) - (0,0,S_{\mathrm{c}}^{\mu}(0)w_0)\|_s \\
&\le \|(Z_{w_0}^{\star\star}, z_{w_0}^{\star\star}, w_{w_0}^{\star\star}) - (0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0)\| \\
&\le 2\Big\|\mathcal{F}\Big(0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big) - (0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0)\Big\| \\
&\le 2\Big\|\mathcal{F}_{\mathrm{sh}}\Big(0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big)\Big\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} + 2\Big\|\mathcal{F}_{\mathrm{wh}}\Big(0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big)\Big\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \\
&\qquad + 2\Big\|\mathcal{F}_{\mathrm{c}}\Big(0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big) - S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big\|_{C_{\eta\lambda^{\mu}}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^{s})}.
\end{aligned}$$

Using Propositions 2.17 and 2.18, inequality (2.60), Lemma 2.21 and Lemma 5.5 we find that

$$\Big|\underline{G}_1^{\mu}\Big(0,0,S_{\mathrm{c}}^{\mu}(t)w_0\Big) + G_2^{\mu}(0)\Big| \lesssim \mu\Big\|\phi\Big(S_{\mathrm{c}}^{\mu}(t)w_0\Big)\Big\|_s^2 \lesssim \mu\|w_0\|_{\mathcal{X}_{\mathrm{c}}^{s}}^2,$$

so that

$$\Big\|\mathcal{F}_{\mathrm{sh}}\Big(0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big)\Big\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu\|w_0\|_{\mathcal{X}_{\mathrm{c}}^{s}}^2,$$

and

$$\begin{aligned}
\Big|\mathcal{F}_{\mathrm{wh}}\Big(0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big)(\xi)\Big| &\lesssim \int_{-\infty}^{\xi} \mathrm{e}^{\lambda^{\mu}(t-\xi)}\Big\|\phi\Big(S_{\mathrm{c}}^{\mu}(t)w_0\Big)\Big\|_{\mathcal{X}_{\mathrm{c}}^{s}}^2\,\mathrm{d}t \\
&\qquad + \int_{\xi}^{\infty} \mathrm{e}^{\lambda^{\mu}(\xi-t)}\Big\|\phi\Big(S_{\mathrm{c}}^{\mu}(t)w_0\Big)\Big\|_{\mathcal{X}_{\mathrm{c}}^{s}}^2\,\mathrm{d}t \\
&\lesssim \frac{\mathrm{e}^{-c^\star/2\mu}}{\lambda^{\mu}}\|w_0\|_{\mathcal{X}_{\mathrm{c}}^{s}} \\
&\lesssim \mu\|w_0\|_{\mathcal{X}_{\mathrm{c}}^{s}}.
\end{aligned}$$

Moreover

$$\mathrm{e}^{-\eta\lambda^{\mu}|\xi|}\Big\|\mathcal{F}_{\mathrm{c}}\Big(0,0,S_{\mathrm{c}}^{\mu}(\cdot)w_0\Big)(\xi) - S_{\mathrm{c}}^{\mu}(\xi)w_0\Big\|_{\mathcal{X}_{\mathrm{c}}^{s}} \lesssim \mu\mathrm{e}^{-\eta\lambda^{\mu}|\xi|}|\xi|\|w_0\|_{\mathcal{X}_{\mathrm{c}}^{s}}^2 \lesssim \mu\|w_0\|_{\mathcal{X}_{\mathrm{c}}^{s}}^2.$$

■

By applying the Lyapunov functional $\mathcal{I}^{\mu}$ (see Assumptions 5) to a solution (Z,z,w) to equations (5.30)–(5.32) with $(Z(\xi_0),z(\xi_0),w(\xi_0)) \in W_{\mathrm{loc}}^{\mathrm{c}}$ for some ξ_0 we conclude that $(Z(\xi),z(\xi),w(\xi)) \in W_{\mathrm{loc}}^{\mathrm{c}}$ for all $\xi \ge \xi_0$. The result is summarised in the following lemma.

Lemma 5.11. *Suppose that*

$$(Z,z,w) \in C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}}) \times C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}}) \times C_{\eta\lambda^{\mu}}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^{s})$$

is a solution of equations (5.30)–(5.32) and $(Z(\xi_0),z(\xi_0),w(\xi_0)) \in W^{\mathrm{c}}$ *with*

$$\|w(\xi_0)\|_{\mathcal{X}_{\mathrm{c}}^{s}} \le \tfrac{1}{2}\mathrm{e}^{-c^\star/2\mu}$$

(so that $(Z(\xi_0),z(\xi_0),w(\xi_0)) \in W_{\mathrm{loc}}^{\mathrm{c}}$*). It follows that*

$$\|w(\xi)\|_{\mathcal{X}_{\mathrm{c}}^{s}} \le \tfrac{7}{8}\mathrm{e}^{-c^\star/2\mu}$$

(and therefore $(Z(\xi),z(\xi),w(\xi)) \in W_{\mathrm{loc}}^{\mathrm{c}}$*) for all* $\xi \ge \xi_0$.

Proof. Using the estimates (5.1), (5.2) together with Lemma 5.10 we find that

$$\frac{1}{2}\mu^4\|w\|^2_{\mathcal{X}^s_c} \leq \mathcal{I}^\mu(Z,z,w) = \mathcal{I}^\mu(\Psi(w),w) \leq \frac{3}{2}\mu^4\|w\|^2_{\mathcal{X}^s_c}$$

for each $(Z,z,w) \in W^{\mathrm{c}}_{\mathrm{loc}}$ and sufficiently small values of μ.

Suppose that $\|w(\xi)\|_{\mathcal{X}^s_c} > \frac{7}{8}\mathrm{e}^{-c^\star/2\mu}$ for some $\xi > \xi_0$. Since $w \in C_{\eta\lambda^\mu}(\mathbb{R};\mathcal{X}^s_{\mathrm{c}})$ is continuous, we conclude that there exists $\xi^\star \in (\xi_0,\xi)$ such that $\|w(\xi^\star)\|_{\mathcal{X}^s_c} = \frac{7}{8}\mathrm{e}^{-c^\star/2\mu}$. It follows that

$$\begin{aligned}
\frac{49}{128}\mathrm{e}^{-c^\star/\mu} &= \frac{1}{2}\|w(\xi^\star)\|^2_{\mathcal{X}^s_c} \\
&\leq \frac{1}{\mu^4}\mathcal{I}^\mu\Big(\Psi(w(\xi^\star)),w(\xi^\star)\Big) \\
&= \frac{1}{\mu^4}\mathcal{I}^\mu\Big(\Psi(w(\xi_0)),w(\xi_0)\Big) \\
&\leq \frac{3}{2}\|w(\xi_0)\|^2_{\mathcal{X}^s_c} \\
&\leq \frac{3}{8}\mathrm{e}^{-c^\star/\mu},
\end{aligned}$$

a contradiction. ■

5.3 The global centre-stable manifold

In this section we prove that the local centre-stable manifold $W^{\mathrm{cs}}_{\mathrm{loc}}$ constructed in Section 5.1 (see Definition 5.3) is a global centre-stable manifold, that is solutions $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ to equations (5.9)–(5.11) whose initial values lie on $W^{\mathrm{cs}}_{\mathrm{loc}}$, satisfy

$$\|w^\star_{w_0}(\xi)\|_{\mathcal{X}^s_c} \leq \mathrm{e}^{-c^\star/2\mu}$$

for all $\xi \geq 0$ (and hence by reversibility for all $\xi \in \mathbb{R}$). To this end we prove that solutions found in Theorem 5.2 whose initial values lie on $W^{\mathrm{cs}}_{\mathrm{loc}}$ converge exponentially to solutions of the modified equations (5.30)–(5.32) found in Theorem 5.8 on W^{c}. This result is summarised in the following theorem.

Theorem 5.12. *Suppose that $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ is a solution of the modified equations (5.9)–(5.11). It follows that $(Z^\star_{w_0}, p^\mu + r^\star_{w_0}, w^\star_{w_0})$ converges exponentially to a solution (Z,z,w) of equations (5.30)–(5.32) on W^{c}, that is*

$$\lim_{\xi\to\infty} \mathrm{e}^{\eta\lambda^\mu\xi}\|(Z^\star_{w_0}(\xi), p^\mu(\xi) + r^\star_{w_0}(\xi), w^\star_{w_0}(\xi)) - (Z(\xi),z(\xi),w(\xi))\|_s = 0.$$

In particular there exists $\xi^\star = \xi^\star(\mu) \in (0, \mathrm{e}^{c^\star/2\mu}]$ such that

$$\sup_{\xi\in[\xi^\star,\infty)} \|w^\star_{w_0}(\xi) - w(\xi)\|_{\mathcal{X}^s_c} \lesssim \mu\mathrm{e}^{-c^\star/2\mu}.$$

In order to prove Theorem 5.12 we fix

$$\xi^\star = \frac{c^\star}{\mu\eta\lambda^\mu}, \tag{5.34}$$

so that $\xi^\star \in (0, \mathrm{e}^{c^\star/2\mu}]$ with $\mathrm{e}^{-\eta\lambda^\mu\xi^\star/2} = \mathrm{e}^{-c^\star/2\mu}$ and

$$|p^\mu(\xi)| \lesssim \mu \mathrm{e}^{-c^\star/2\mu}$$

for all $\xi \geq \frac{\xi^\star}{2}$. Next we consider a cut-off function $\chi \in C^\infty(\mathbb{R})$ with

$$\chi(\xi) = \begin{cases} 0, & |\xi| \leq \frac{\xi^\star}{2}, \\ 1, & |\xi| \geq \xi^\star, \end{cases}$$

$|\chi'(\xi)| \leq \frac{1}{\xi^\star}$ for all $\xi \in \mathbb{R}$ and define $(\hat{Z}, \hat{z}, \hat{w}) = \chi(Z^\star_{w0}, p^\mu + r^\star_{w0}, w^\star_{w0})$, such that $(\hat{Z}, \hat{z}, \hat{w})$ satisfies the modified equations

$$\hat{Z}_\xi = L^\mu_{\mathrm{sh}}\hat{Z} + \underline{G}^\mu_1(\hat{Z}, \hat{z}, \hat{w}) + G^\mu_2(\hat{z}) + Q^\mu_{\mathrm{sh}}, \tag{5.35}$$

$$\hat{z}_\xi = L^\mu_{\mathrm{wh}}\hat{z} + f^\mu_1(\hat{z}) + \underline{f}^\mu_2(\hat{Z}, \hat{z}, \hat{w}) + Q^\mu_{\mathrm{wh}}, \tag{5.36}$$

$$\hat{w}_\xi = L^\mu_{\mathrm{c}}\hat{w} + \underline{\hat{g}}^\mu(\hat{Z}, \hat{z}, \hat{w}) + \mu\underline{b}(\hat{z}, \hat{w}) + \hat{h}^\mu(\hat{z}) + Q^\mu_{\mathrm{c}}, \tag{5.37}$$

with

$$\begin{aligned}
Q^\mu_{\mathrm{sh}} &= \chi' Z^\star_{w0} + \chi\underline{G}^\mu_1(Z^\star_{w0}, p^\mu + r^\star_{w0}, w^\star_{w0}) - \underline{G}^\mu_1\Big(\chi(Z^\star_{w0}, p^\mu + r^\star_{w0}, w^\star_{w0})\Big) \\
&\quad + \chi G^\mu_2(p^\mu + r^\star_{w0}) - G^\mu_2\Big(\chi(p^\mu + r^\star_{w0})\Big), \\
Q^\mu_{\mathrm{wh}} &= \chi'(p^\mu + r^\star_{w0}) + \chi f^\mu_1(p^\mu + r^\star_{w0}) - f^\mu_1\Big(\chi(p^\mu + r^\star_{w0})\Big) \\
&\quad + \chi\underline{f}^\mu_2(Z^\star_{w0}, p^\mu + r^\star_{w0}, w^\star_{w0}) - \underline{f}^\mu_2\Big(\chi(Z^\star_{w0}, p^\mu + r^\star_{w0}, w^\star_{w0})\Big), \\
Q^\mu_{\mathrm{c}} &= \chi' w^\star_{w0} + \chi\underline{\hat{g}}^\mu(Z^\star_{w0}, p^\mu + r^\star_{w0}, w^\star_{w0}) - \underline{\hat{g}}^\mu\Big(\chi(Z^\star_{w0}, p^\mu + r^\star_{w0}, w^\star_{w0})\Big) \\
&\quad + \mu\chi\underline{b}(p^\mu + r^\star_{w0}, w^\star_{w0}) - \mu\underline{b}\Big(\chi(p^\mu + r^\star_{w0}, w^\star_{w0})\Big) \\
&\quad + \chi\hat{h}^\mu(p^\mu + r^\star_{w0}) - \hat{h}^\mu\Big(\chi(p^\mu + r^\star_{w0})\Big).
\end{aligned}$$

To show that $(\hat{Z}, \hat{z}, \hat{w})$ converges exponentially to a solution (Z, z, w) of equations (5.30)–(5.32) on W^c, we make the *Ansatz*

$$(Z, z, w) = (\hat{Z}, \hat{z}, \hat{w}) - (\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}}),$$

so that $(\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})$ satisfies the equations

$$\begin{aligned}
\partial_\xi\Delta_{\mathrm{sh}} &= L^\mu_{\mathrm{sh}}\Delta_{\mathrm{sh}} + \underline{G}^\mu_1(\hat{Z}, \hat{z}, \hat{w}) + G^\mu_2(\hat{z}) \\
&\quad - \underline{G}^\mu_1(\hat{Z} - \Delta_{\mathrm{sh}}, \hat{z} - \Delta_{\mathrm{wh}}, \hat{w} - \Delta_{\mathrm{c}}) - G^\mu_2(\hat{z} - \Delta_{\mathrm{wh}}) + Q^\mu_{\mathrm{sh}}, \\
\partial_\xi\Delta_{\mathrm{wh}} &= L^\mu_{\mathrm{wh}}\Delta_{\mathrm{wh}} + f^\mu_1(\hat{z}) + \underline{f}^\mu_2(\hat{Z}, \hat{z}, \hat{w}) - f^\mu_1(\hat{z} - \Delta_{\mathrm{wh}}) \\
&\quad - \underline{f}^\mu_2(\hat{Z} - \Delta_{\mathrm{sh}}, \hat{z} - \Delta_{\mathrm{wh}}, \hat{w} - \Delta_{\mathrm{c}}) + Q^\mu_{\mathrm{wh}}, \\
\partial_\xi\Delta_{\mathrm{c}} &= L^\mu_{\mathrm{c}}\Delta_{\mathrm{c}} + \underline{\hat{g}}^\mu(\hat{Z}, \hat{z}, \hat{w}) + \mu\underline{b}(\hat{z}, \hat{w}) + \hat{h}^\mu(\hat{z}) - \underline{\hat{g}}^\mu(\hat{Z} - \Delta_{\mathrm{sh}}, \hat{z} - \Delta_{\mathrm{wh}}, \hat{w} - \Delta_{\mathrm{c}}) \\
&\quad - \mu\underline{b}(\hat{z} - \Delta_{\mathrm{wh}}, \hat{w} - \Delta_{\mathrm{c}}) - \hat{h}^\mu(\hat{z} - \Delta_{\mathrm{wh}}) + Q^\mu_{\mathrm{c}},
\end{aligned}$$

and establish that $(\Delta_{\mathrm{wh}}, \Delta_{\mathrm{sh}}, \Delta_{\mathrm{c}})$ decays exponentially to zero as $\xi \to \infty$. To this end we consider the iteration scheme

$$\begin{aligned}\Delta_{\mathrm{sh}}^{(j+1)}(\xi) = \int_{-\infty}^{\xi} & \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{ss}}\Big(\underline{G}_1^{\mu}(\hat{Z},\hat{z},\hat{w}) + G_2^{\mu}(\hat{z}) - G_2^{\mu}(\hat{z}-\Delta_{\mathrm{wh}}^{(j)}) \\ & - \underline{G}_1^{\mu}(\hat{Z}-\Delta_{\mathrm{sh}}^{(j)}, \hat{z}-\Delta_{\mathrm{wh}}^{(j)}, \hat{w}-\Delta_{\mathrm{c}}^{(j)}) + Q_{\mathrm{sh}}^{\mu}(t)\Big)\,\mathrm{d}t \\ - \int_{\xi}^{\infty} & \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^{\mu}} P_{\mathrm{su}}\Big(\underline{G}_1^{\mu}(\hat{Z},\hat{z},\hat{w}) + G_2^{\mu}(\hat{z}) - G_2^{\mu}(\hat{z}-\Delta_{\mathrm{wh}}^{(j)}) \\ & - \underline{G}_1^{\mu}(\hat{Z}-\Delta_{\mathrm{sh}}^{(j)}, \hat{z}-\Delta_{\mathrm{wh}}^{(j)}, \hat{w}-\Delta_{\mathrm{c}}^{(j)}) + Q_{\mathrm{sh}}^{\mu}(t)\Big)\,\mathrm{d}t,\end{aligned} \tag{5.38}$$

$$\begin{aligned}\Delta_{\mathrm{wh}}^{(j+1)}(\xi) = \sum_{i=3}^{4}\Bigg(& \int_{-\infty}^{\xi} \Big\langle\Big(f_1^{\mu}(\hat{z})(t) + \underline{f}_2^{\mu}(\hat{Z},\hat{z},\hat{w})(t) - f_1^{\mu}(\hat{z}-\Delta_{\mathrm{wh}}^{(j)})(t) \\ & - \underline{f}_2^{\mu}(\hat{Z}-\Delta_{\mathrm{sh}}^{(j)}, \hat{z}-\Delta_{\mathrm{wh}}^{(j)}, \hat{w}-\Delta_{\mathrm{c}}^{(j)})(t) \\ & + Q_{\mathrm{wh}}^{\mu}(t)\Big), S_i^*(t)\Big\rangle\,\mathrm{d}t\Bigg) S_i(\xi) \\ & - \int_{\xi}^{\infty} \Big\langle\Big(f_1^{\mu}(\hat{z})(t) + \underline{f}_2^{\mu}(\hat{Z},\hat{z},\hat{w})(t) - f_1^{\mu}(\hat{z}-\Delta_{\mathrm{wh}}^{(j)})(t) \\ & - \underline{f}_2^{\mu}(\hat{Z}-\Delta_{\mathrm{sh}}^{(j)}, \hat{z}-\Delta_{\mathrm{wh}}^{(j)}, \hat{w}-\Delta_{\mathrm{c}}^{(j)})(t) \\ & + Q_{\mathrm{wh}}^{\mu}(t)\Big), U_i^*(t)\Big\rangle\,\mathrm{d}t\Bigg) U_i(\xi),\end{aligned} \tag{5.39}$$

$$\begin{aligned}\Delta_{\mathrm{c}}^{(j+1)}(\xi) = \int_{\xi}^{\infty} S_{\mathrm{c}}^{\mu}(\xi-t)\Big(& \underline{\hat{g}}^{\mu}(\hat{Z},\hat{z},\hat{w})(t) + \mu\underline{b}(\hat{z},\hat{w})(t) + \hat{h}^{\mu}(\hat{z})(t) \\ & - \underline{\hat{g}}^{\mu}(\hat{Z}-\Delta_{\mathrm{sh}}^{(j)}, \hat{z}-\Delta_{\mathrm{wh}}^{(j)}, \hat{w}-\Delta_{\mathrm{c}}^{(j)})(t) \\ & - \mu\underline{b}(\hat{z}-\Delta_{\mathrm{wh}}^{(j)}, \hat{w}-\Delta_{\mathrm{c}}^{(j)})(t) - \hat{h}^{\mu}(\hat{z}-\Delta_{\mathrm{wh}}^{(j)})(t) + Q_{\mathrm{c}}^{\mu}(t)\Big)\,\mathrm{d}t\end{aligned} \tag{5.40}$$

and show that the sequence $\{(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \Delta_{\mathrm{c}}^{(j)})\}_{j\in\mathbb{N}}$ is convergent in

$$C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}_{\mathrm{sh}}) \times C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}_{\mathrm{wh}}) \times C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}_{\mathrm{c}}^{s})$$

for an appropriately chosen initial value $(\Delta_{\mathrm{sh}}^{(0)}, \Delta_{\mathrm{wh}}^{(0)}, \Delta_{\mathrm{c}}^{(0)})$. First we need estimates for the remainder terms appearing in the equations (5.35)–(5.37).

Lemma 5.13. *The remainder terms Q_{wh}^{μ}, Q_{sh}^{μ} and Q_{c}^{μ} satisfy the estimates*

$$\|Q_{\mathrm{sh}}^{\mu}\|_{C_{\mathrm{b}}(\mathbb{R},\mathcal{X}_{\mathrm{sh}})}, \|Q_{\mathrm{wh}}^{\mu}\|_{C_{\mathrm{b}}(\mathbb{R},\mathcal{X}_{\mathrm{wh}})}, \|Q_{\mathrm{c}}^{\mu}\|_{C_{\mathrm{b}}(\mathbb{R},\mathcal{X}_{\mathrm{c}}^{s})} \lesssim \mu^4 \mathrm{e}^{-c^\star/2\mu}.$$

and

$$\|Q_{\mathrm{sh}}^{\mu}\|_{C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}_{\mathrm{sh}})}, \|Q_{\mathrm{wh}}^{\mu}\|_{C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}_{\mathrm{wh}})}, \|Q_{\mathrm{c}}^{\mu}\|_{C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}_{\mathrm{c}}^{s})} < \infty.$$

Proof. First note that by definition $Q_{\mathrm{sh}}^{\mu}(\xi), Q_{\mathrm{wh}}^{\mu}(\xi), Q_{\mathrm{c}}^{\mu}(\xi) = 0$ if $|\xi| \notin [\frac{\xi^\star}{2}, \xi^\star]$. For $|\xi| \in [\frac{\xi^\star}{2}, \xi^\star]$ we have that $|\chi'(\xi)| \leq \frac{1}{\xi^\star} \lesssim \mu^3$ so that the terms in Q_{sh}^{μ}, Q_{wh}^{μ} and Q_{c}^{μ} which are linear in respectively $Z_{w_0}^\star$, $p^\mu + r_{w_0}^\star$ and $w_{w_0}^\star$ can be estimated by

$$\|\chi'(\xi)(Z_{w_0}^\star(\xi), p^\mu(\xi) + r_{w_0}^\star(\xi), w_{w_0}^\star(\xi))\|_s \lesssim \mu^4 \mathrm{e}^{-c^\star/2\mu},$$

since

$$|p^\mu(\xi)+r^\star_{w_0}(\xi)|, \|w^\star_{w_0}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}, \quad |\xi| \in [\tfrac{\xi^\star}{2}, \xi^\star].$$

The remaining nonlinear terms are at least quadratic in $(Z^\star_{w_0}, p^\mu + r^\star_{w_0}, w^\star_{w_0})$ so that

$$\begin{aligned}|Q^\mu_\mathrm{sh}(\xi)|, |Q^\mu_\mathrm{wh}(\xi)|, \|Q^\mu_\mathrm{c}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} &\lesssim \|(Z^\star_{w_0}(\xi), p^\mu(\xi)+r^\star_{w_0}(\xi), w^\star_{w_0}(\xi))\|_s^2 \\ &\qquad + \|\chi'(\xi)(Z^\star_{w_0}(\xi), p^\mu(\xi)+r^\star_{w_0}(\xi), w^\star_{w_0}(\xi))\|_s \\ &\lesssim \mu^2 \mathrm{e}^{-c^\star/\mu} + \mu^4 \mathrm{e}^{-c^\star/2\mu} \\ &\lesssim \mu^4 \mathrm{e}^{-c^\star/2\mu}\end{aligned}$$

which verifies the first assertion. Moreover, this estimate yields

$$\begin{aligned}|Q^\mu_\mathrm{sh}(\xi)|, |Q^\mu_\mathrm{wh}(\xi)|, \|Q^\mu_\mathrm{c}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} &\lesssim \mu^4 \mathrm{e}^{-c^\star/2\mu} \\ &= \mu^4 \mathrm{e}^{-\eta\lambda^\mu \xi^\star/2} \\ &= \mu^4 \mathrm{e}^{\eta\lambda^\mu\xi^\star/2} \mathrm{e}^{-\eta\lambda^\mu\xi^\star} \\ &= \mu^4 \mathrm{e}^{c^\star/2\mu} \mathrm{e}^{-\eta\lambda^\mu\xi^\star} \\ &\leq \mu^4 \mathrm{e}^{c^\star/2\mu} \mathrm{e}^{-\eta\lambda^\mu\xi}\end{aligned}$$

so that

$$\mathrm{e}^{\eta\lambda^\mu\xi}|Q^\mu_\mathrm{sh}(\xi)|, \mathrm{e}^{\eta\lambda^\mu\xi}|Q^\mu_\mathrm{wh}(\xi)|, \mathrm{e}^{\eta\lambda^\mu\xi}\|Q^\mu_\mathrm{c}(\xi)\|_{\mathcal{X}_\mathrm{c}^s} < \infty, \quad \xi \in \mathbb{R}.$$

■

The next step is to show that the weakly and strongly hyperbolic parts of the sequence $\{(\Delta^{(j)}_\mathrm{sh}, \Delta^{(j)}_\mathrm{wh}, \Delta^{(j)}_\mathrm{c})\}_{j\in\mathbb{N}}$ are exponentially small with respect to μ.

Lemma 5.14. *Suppose that $j \in \mathbb{N}_0$ and*

$$\Delta^{(j)}_\mathrm{sh} \in C_{-\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_\mathrm{sh}), \quad \Delta^{(j)}_\mathrm{wh} \in C_{-\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_\mathrm{wh}), \quad \Delta^{(j)}_\mathrm{c} \in C_{-\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_\mathrm{c}^s)$$

satisfy

$$\|\Delta^{(j)}_\mathrm{sh}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})}, \|\Delta^{(j)}_\mathrm{wh}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

It follows that

$$\Delta^{(j+1)}_\mathrm{sh} \in C_{-\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_\mathrm{sh}), \quad \Delta^{(j+1)}_\mathrm{wh} \in C_{-\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_\mathrm{wh}), \quad \Delta^{(j+1)}_\mathrm{c} \in C_{-\eta\lambda^\mu}(\mathbb{R}; \mathcal{X}_\mathrm{c}^s)$$

with

$$\|\Delta^{(j+1)}_\mathrm{sh}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{sh})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}, \quad \|\Delta^{(j+1)}_\mathrm{wh}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}_\mathrm{wh})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}$$

for all sufficiently small values of μ.

Proof. By construction we have that

$$|\hat{z}(\xi)| \lesssim |\chi(\xi)(p^\mu(\xi)+r^\star_{w_0}(\xi))| \lesssim \mu \mathrm{e}^{-c^\star/2\mu}, \quad |\hat{Z}(\xi)| \lesssim |Z^\star_{w_0}(\xi)| \lesssim \mu^2 \mathrm{e}^{-c^\star/2\mu}$$

and

$$\|\phi(\hat{w}(\xi))\|_{\mathcal{X}_c^s}, \|\phi(\hat{w}(\xi) - \Delta_{\mathrm{c}}^{(j)}(\xi))\|_{\mathcal{X}_c^s} \lesssim \mathrm{e}^{-c^\star/2\mu}$$

for all $\xi \in \mathbb{R}$. Using these inequalities together with the estimates for the derivatives of the nonlinearities given in Propositions 2.17 and 2.18 and Lemma 4.30 we find that

$$\begin{aligned}
&|\underline{G}_1^\mu(\hat{Z}, \hat{z}, \hat{w})(\xi) - \underline{G}_1^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})(\xi)| \\
&\qquad \lesssim \mu \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}(\xi), \Delta_{\mathrm{wh}}^{(j)}(\xi), \phi(\hat{w}) - \phi(\hat{w}(\xi) - \Delta_{\mathrm{c}}^{(j)}(\xi)) \Big) \Big\|_s
\end{aligned}$$

and

$$|\underline{G}_2^\mu(\hat{z})(\xi) - \underline{G}_2^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j)})(\xi)| \lesssim \mu \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}(\xi), \Delta_{\mathrm{wh}}^{(j)}(\xi), \phi(\hat{w}) - \phi(\hat{w}(\xi) - \Delta_{\mathrm{c}}^{(j)}(\xi)) \Big) \Big\|_s$$

so that

$$\begin{aligned}
&\mathrm{e}^{\eta\lambda^\mu\xi} |\Delta_{\mathrm{sh}}^{(j+1)}(\xi)| \\
&\quad \lesssim \mu \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \phi(\hat{w}) - \phi(\hat{w} - \Delta_{\mathrm{c}}^{(j)}) \Big) \Big\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}^s)} + \|Q_{\mathrm{sh}}^\mu\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}_{\mathrm{sh}})} \\
&\quad < \infty.
\end{aligned}$$

Similarly

$$\begin{aligned}
&\mathrm{e}^{\eta\lambda^\mu\xi} |\Delta_{\mathrm{wh}}^{(j+1)}(\xi)| \\
&\lesssim \Big(\mu^3 \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \phi(\hat{w}) - \phi(\hat{w} - \Delta_{\mathrm{c}}^{(j)}) \Big) \Big\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}^s)} + \|Q_{\mathrm{wh}}^\mu\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}_{\mathrm{wh}})} \Big) \\
&\qquad \times \Big(\int_{-\infty}^\xi \mathrm{e}^{(1-\eta)\lambda^\mu t} \,\mathrm{d}t \, \mathrm{e}^{-(1-\eta)\lambda^\mu \xi} + \int_\xi^\infty \mathrm{e}^{-(1+\eta)\lambda^\mu t} \,\mathrm{d}t \, \mathrm{e}^{(1+\eta)\lambda^\mu \xi} \Big) \\
&\lesssim \frac{\mu^3}{(1-\eta)\lambda^\mu} \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \phi(\hat{w}) - \phi(\hat{w} - \Delta_{\mathrm{c}}^{(j)}) \Big) \Big\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}^s)} \\
&\qquad + \frac{1}{(1-\eta)\lambda^\mu} \|Q_{\mathrm{wh}}^\mu\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}_{\mathrm{wh}})} \\
&\lesssim \mu \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \phi(\hat{w}) - \phi(\hat{w} - \Delta_{\mathrm{c}}^{(j)}) \Big) \Big\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}^s)} + \frac{1}{(1-\eta)\lambda^\mu} \|Q_{\mathrm{wh}}^\mu\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}_{\mathrm{wh}})} \\
&< \infty
\end{aligned}$$

(pointwise in μ). We further have that

$$\begin{aligned}
&\mathrm{e}^{\eta\lambda^\mu\xi} \|\Delta_{\mathrm{c}}^{(j+1)}(\xi)\|_{\mathcal{X}_c^s} \\
&\quad \lesssim \Big(\mu^3 \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \phi(\hat{w}) - \phi(\hat{w} - \Delta_{\mathrm{c}}^{(j)}) \Big) \Big\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}^s)} + \|Q_{\mathrm{c}}^\mu\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}_c^s)} \Big) \\
&\qquad \times \int_\xi^\infty \mathrm{e}^{-\eta\lambda^\mu t} \,\mathrm{d}t \, \mathrm{e}^{\eta\lambda^\mu \xi} \\
&\quad \lesssim \mu \Big\| \Big(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \phi(\hat{w}) - \phi(\hat{w} - \Delta_{\mathrm{c}}^{(j)}) \Big) \Big\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}^s)} + \frac{1}{\eta\lambda^\mu} \|Q_{\mathrm{c}}^\mu\|_{C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}_c^s)} \\
&\quad < \infty.
\end{aligned}$$

Finally, we use the above estimates on $|\Delta_{\mathrm{sh}}^{(j+1)}(\xi)|$ and $|\Delta_{\mathrm{wh}}^{(j+1)}(\xi)|$ (without any exponential weight) together with the hypothesis

$$\|\Delta_{\mathrm{sh}}^{(j)}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|\Delta_{\mathrm{wh}}^{(j)}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \leq \mathrm{e}^{-c^\star/2\mu}$$

and Lemma 5.13 to obtain

$$\|\Delta_{\mathrm{sh}}^{(j+1)}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}, \quad \|\Delta_{\mathrm{wh}}^{(j+1)}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}.$$

■

Next we conclude that the iteration scheme (5.38)–(5.40) is contractive. To this end we define

$$\begin{aligned}
\tilde{\Delta}_{\mathrm{sh}}^{(j+1)}(\xi) &= \Delta_{\mathrm{sh}}^{(j+1)}(\xi) - \Delta_{\mathrm{sh}}^{(j)}(\xi) \\
&= -\int_{-\infty}^{\xi} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{ss}} \Big(\underline{G}_1^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})(t) \\
&\qquad\qquad - \underline{G}_1^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j-1)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j-1)}, \hat{w} - \Delta_{\mathrm{c}}^{(j-1)})(t) \\
&\qquad\qquad + G_2^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j)})(t) - G_2^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j-1)})(t) \Big)\, \mathrm{d}t \\
&\quad + \int_{\xi}^{\infty} \mathrm{e}^{(\xi-t)L_{\mathrm{sh}}^\mu} P_{\mathrm{su}} \Big(\underline{G}_1^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})(t) \\
&\qquad\qquad - \underline{G}_1^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j-1)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j-1)}, \hat{w} - \Delta_{\mathrm{c}}^{(j-1)})(t) \\
&\qquad\qquad + G_2^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j)})(t) - G_2^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j-1)})(t) \Big)\, \mathrm{d}t,
\end{aligned}$$

$$\begin{aligned}
&\tilde{\Delta}_{\mathrm{wh}}^{(j+1)}(\xi) \\
&\quad = \Delta_{\mathrm{wh}}^{(j+1)}(\xi) - \Delta_{\mathrm{wh}}^{(j)}(\xi) \\
&\quad = \sum_{i=3}^{4} \Big(-\int_{-\infty}^{\xi} \Big\langle \Big(f_1^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j)})(t) - f_1^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j-1)})(t) \\
&\qquad\qquad + \underline{f}_2^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})(t) \\
&\qquad\qquad - \underline{f}_2^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j-1)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j-1)}, \hat{w} - \Delta_{\mathrm{c}}^{(j-1)})(t) \Big), S_i^*(t) \Big\rangle\, \mathrm{d}t \Big) S_i(\xi) \\
&\qquad + \int_{\xi}^{\infty} \Big\langle \Big(f_1^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j)})(t) - f_1^\mu(\hat{z} - \Delta_{\mathrm{wh}}^{(j-1)})(t) \\
&\qquad\qquad + \underline{f}_2^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})(t) \\
&\qquad\qquad - \underline{f}_2^\mu(\hat{Z} - \Delta_{\mathrm{sh}}^{(j-1)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j-1)}, \hat{w} - \Delta_{\mathrm{c}}^{(j-1)})(t) \Big), U_i^*(t) \Big\rangle\, \mathrm{d}t \Big) U_i(\xi),
\end{aligned}$$

$$\begin{aligned}\tilde{\Delta}_{\mathrm{c}}^{(j+1)}(\xi) &= \Delta_{\mathrm{c}}^{(j+1)}(\xi) - \Delta_{\mathrm{c}}^{(j)}(\xi)\\ &= -\int_{\xi}^{\infty} S_{\mathrm{c}}^{\mu}(\xi - t)\Big(\underline{\hat{g}}^{\mu}(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})(t)\\ &\qquad\qquad - \underline{\hat{g}}^{\mu}(\hat{Z} - \Delta_{\mathrm{sh}}^{(j-1)}, \hat{z} - \Delta_{\mathrm{wh}}^{(j-1)}, \hat{w} - \Delta_{\mathrm{c}}^{(j-1)})(t)\\ &\qquad\qquad + \mu\underline{b}(\hat{z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})(t)\\ &\qquad\qquad - \mu\underline{b}(\hat{z} - \Delta_{\mathrm{wh}}^{(j-1)}, \hat{w} - \Delta_{\mathrm{c}}^{(j-1)})(t)\\ &\qquad\qquad + \hat{h}^{\mu}(\hat{z} - \Delta_{\mathrm{wh}}^{(j)})(t) - \hat{h}^{\mu}(\hat{z} - \Delta_{\mathrm{wh}}^{(j-1)})(t)\Big)\,\mathrm{d}t.\end{aligned}$$

The following result is established by using the same estimates as in the proof of Lemma 5.14 with respectively $(\hat{Z}, \hat{z}, \hat{w})$ and $(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{Z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})$ replaced by $(\hat{Z} - \Delta_{\mathrm{sh}}^{(j)}, \hat{Z} - \Delta_{\mathrm{wh}}^{(j)}, \hat{w} - \Delta_{\mathrm{c}}^{(j)})$ and $(\hat{Z} - \Delta_{\mathrm{sh}}^{(j-1)}, \hat{Z} - \Delta_{\mathrm{wh}}^{(j-1)}, \hat{w} - \Delta_{\mathrm{c}}^{(j-1)})$, and noting that the convergence is not affected by the remainder terms Q_{sh}, Q_{wh} and Q_{c}.

Lemma 5.15. *The estimate*

$$\|(\tilde{\Delta}_{\mathrm{sh}}^{(j+1)}, \tilde{\Delta}_{\mathrm{wh}}^{(j+1)}, \tilde{\Delta}_{\mathrm{c}}^{(j+1)})\|_{C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}^{s})} \leq \frac{1}{2}\|(\tilde{\Delta}_{\mathrm{sh}}^{(j)}, \tilde{\Delta}_{\mathrm{wh}}^{(j)}, \tilde{\Delta}_{\mathrm{c}}^{(j)})\|_{C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}^{s})}$$

holds for all sufficiently small values of μ.

Corollary 5.16. *Suppose that*

$$(\Delta_{\mathrm{sh}}^{(0)}, \Delta_{\mathrm{wh}}^{(0)}, \Delta_{\mathrm{c}}^{(0)}) \in C_{-\eta\lambda^{\mu}}(\mathbb{R}, \mathcal{X}^{s})$$

satisfies

$$\|(\Delta_{\mathrm{sh}}^{(0)}, \Delta_{\mathrm{wh}}^{(0)}, \Delta_{\mathrm{c}}^{(0)})\|_{C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}^{s})} \leq \mathrm{e}^{\eta\lambda^{\mu}\xi^{\star}/2}$$

and

$$\|\Delta_{\mathrm{sh}}^{(0)}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|\Delta_{\mathrm{wh}}^{(0)}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \leq \mathrm{e}^{-c^{\star}/2\mu}.$$

(i) The sequence $\{(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \Delta_{\mathrm{c}}^{(j)})\}_{j\in\mathbb{N}_0}$ is a Cauchy sequence and hence convergent in $C_{-\eta\lambda^{\mu}}(\mathbb{R}, \mathcal{X}^{s})$. Its limit $(\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})$ satisfies the estimates

$$\|\Delta_{\mathrm{sh}}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|\Delta_{\mathrm{wh}}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \lesssim \mu\mathrm{e}^{-c^{\star}/2\mu}$$

and

$$\|(\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})\|_{C_{-\eta\lambda^{\mu}}(\mathbb{R},\mathcal{X}^{s})} \lesssim \mu\mathrm{e}^{\eta\lambda^{\mu}\xi^{\star}/2}.$$

(ii) We have that

$$(Z, z, w) = (\hat{Z}, \hat{z}, \hat{w}) - (\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}}) \in W^{\mathrm{c}}.$$

Proof.

(i) It is a direct consequence of Lemmata 5.14 and 5.15 that $\{(\Delta_{\mathrm{sh}}^{(j)}, \Delta_{\mathrm{wh}}^{(j)}, \Delta_{\mathrm{c}}^{(j)})\}_{j\in\mathbb{N}_0}$ is a Cauchy sequence in $C_{-\eta\lambda^\mu}(\mathbb{R}, \mathcal{X}^s)$ with a limit $(\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})$ satisfying

$$\|\Delta_{\mathrm{sh}}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})}, \|\Delta_{\mathrm{wh}}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}.$$

Moreover, inspecting the proofs of Lemmata 5.13 and 5.14 we find that

$$\begin{aligned}\|\Delta_{\mathrm{c}}\|_{C_{-\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s)} &\lesssim \mu\|(\Delta_{\mathrm{sh}}^{(0)}, \Delta_{\mathrm{wh}}^{(0)}, \Delta_{\mathrm{c}}^{(0)})\|_{C_{-\eta\lambda^\mu}(\mathbb{R},\mathcal{X}^s)} + \frac{1}{\eta\lambda^\mu}\|Q_{\mathrm{c}}\|_{C_{-\eta\lambda^\mu}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s)} \\ &\lesssim \mu \mathrm{e}^{\eta\lambda^\mu\xi^\star/2} + \mu^2 \mathrm{e}^{\eta\lambda^\mu\xi^\star/2} \\ &\lesssim \mu \mathrm{e}^{\eta\lambda^\mu\xi^\star/2}\end{aligned}$$

and thus

$$\|(\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})\|_{C_{-\eta\lambda^\mu}(\mathbb{R},\mathcal{X}^s)} \lesssim \mu \mathrm{e}^{\eta\lambda^\mu\xi^\star/2}.$$

(ii) By construction

$$(Z, z, w) = (\hat{Z}, \hat{z}, \hat{w}) - (\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})$$

is a solution of equations (5.30)–(5.32) and since

$$\|Z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \leq \|\hat{Z}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} + \|\Delta_{\mathrm{sh}}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{sh}})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu}$$

as well as

$$\|z\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \leq \|\hat{z}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} + \|\Delta_{\mathrm{wh}}\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{wh}})} \lesssim \mu \mathrm{e}^{-c^\star/2\mu},$$

we conclude that $(Z, z, w) \in W^{\mathrm{c}}$.

■

Suppose that $\xi \geq \xi^\star$ (see (5.34)), so that $w_{w_0}^\star(\xi) = \hat{w}(\xi)$ and thus

$$\begin{aligned}\|w_{w_0}^\star(\xi) - w(\xi)\|_{\mathcal{X}_{\mathrm{c}}^s} &= \|\hat{w}(\xi) - w(\xi)\|_{\mathcal{X}_{\mathrm{c}}^s} \\ &\leq \|(\Delta_{\mathrm{sh}}(\xi), \Delta_{\mathrm{wh}}(\xi), \Delta_{\mathrm{c}}(\xi))\|_s \\ &\leq \mathrm{e}^{-\eta\lambda^\mu\xi}\|(\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})\|_{C_{-\eta\lambda^\mu}(\mathbb{R},\mathcal{X}^s)} \\ &\lesssim \mu \mathrm{e}^{-\eta\lambda^\mu\xi^\star/2} \\ &= \mu \mathrm{e}^{-c^\star/2\mu}\end{aligned}$$

due to Corollary 5.16. This observation completes the proof of Theorem 5.12. The last step in the proof of the existence of global generalised pulse solutions to equations (2.53) and (2.54) (that is global exponentially small solutions to equations (2.55)–(2.57), see Theorem 5.1) is to deduce that

$$\|w_{w_0}^\star\|_{C_{\mathrm{b}}(\mathbb{R};\mathcal{X}_{\mathrm{c}}^s)} \leq \mathrm{e}^{-c^\star/2\mu}.$$

In the following result (Z, z, w) and $(\Delta_{\mathrm{sh}}, \Delta_{\mathrm{wh}}, \Delta_{\mathrm{c}})$ denote the functions in Corollary 5.16.

Theorem 5.17. *The centre part $w^\star_{w_0}$ of a solution $(Z^\star_{w_0}, r^\star_{w_0}, w^\star_{w_0})$ to equations (5.9)–(5.11) satisfies the estimate*

$$\|w^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}^s_\mathrm{c})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

Proof. Since $\xi^\star \leq \mathrm{e}^{c^\star/2\mu}$ and

$$\|w^\star_{w_0}\|_{C([0,\mathrm{e}^{c^\star/2\mu}];\mathcal{X}^s_\mathrm{c})} \lesssim \mu^\delta \mathrm{e}^{-c^\star/2\mu},$$

we conclude that

$$\|w(\xi^\star)\|_{\mathcal{X}^s_\mathrm{c}} \lesssim (\mu + \mu^\delta)\mathrm{e}^{-c^\star/2\mu} \leq \frac{1}{2}\mathrm{e}^{-c^\star/2\mu}$$

so that Lemma 5.11 yields

$$\sup_{\xi\in[\xi^\star,\infty)} \|w(\xi)\|_{\mathcal{X}^s_\mathrm{c}} \leq \frac{7}{8}\mathrm{e}^{-c^\star/2\mu}$$

and hence

$$\sup_{\xi\in[\xi^\star,\infty)} \|w^\star_{w_0}(\xi)\|_{\mathcal{X}^s_\mathrm{c}} \leq \Big(\mu + \frac{7}{8}\Big)\mathrm{e}^{-c^\star/2\mu} \leq \mathrm{e}^{-c^\star/2\mu}.$$

Due to reversibility it follows that

$$\|w^\star_{w_0}\|_{C_\mathrm{b}(\mathbb{R};\mathcal{X}^s_\mathrm{c})} \leq \mathrm{e}^{-c^\star/2\mu}.$$

∎

6 Breather solutions to Klein-Gordon equations

In this chapter we study the quasilinear Klein-Gordon equation

$$\phi_{tt} = \phi_{xx} - \phi + f_1(\phi, \phi_t, \phi_x)\phi_{xx} + f_2(\phi, \phi_t, \phi_x), \tag{6.1}$$

where $f_1, f_2 \colon \mathbb{R}^3 \to \mathbb{R}$ are analytic functions with $f_1(0) = 0$, $f_2(0) = 0$, $\mathrm{d}f_2[0] = 0$ and

$$f_i(a, -b, -c) = f_i(a, b, c), \quad i = 1, 2.$$

We seek *generalised breather solutions* to equation (6.1) of the form

$$\phi(x, t) = v_1(\xi, \tau), \qquad \xi = x, \quad \tau = \omega t,$$

in which v_1 is 2π-periodic in its second argument, and introduce a bifurcation parameter ε by writing

$$\omega = \sqrt{1 - \varepsilon^2}.$$

Our result is summarised in the following theorem, which is proved using the theory in Chapter 4, or Chapter 5 in the special case $f_1 = 0$, $f_2(a, b, c) = f_2(a)$ (henceforth referred to as the 'special case').

Theorem 6.1. *There exist positive constants ε_0 and $c^\star$ with the property that for each $\varepsilon \in (0, \varepsilon_0]$ equation (6.1) admits an infinite-dimensional, continuous family of generalised breather solutions of the form*

$$\phi(x, t) = v_1(x, \omega t),$$

parameterised by the values $v_1(0, 0)$. These solutions satisfy

$$v_1(\xi, \tau) = v_1(-\xi, -\tau), \quad v_1(\xi, \tau + 2\pi) = v_1(\xi, \tau), \quad |v_1(\xi, \tau) - p^\varepsilon(\xi, \tau)| \le \mathrm{e}^{-c^\star/2\sqrt{\varepsilon}}$$

for all $\tau \in \mathbb{R}$ and respectively $\xi \in [-\mathrm{e}^{c^\star/2\sqrt{\varepsilon}}, \mathrm{e}^{c^\star/2\sqrt{\varepsilon}}]$ and $\xi \in \mathbb{R}$ if additionally $f_1 = 0$, $f_2(a, b, c) = f_2(a)$. Here

$$p^\varepsilon(\xi, \tau) = \pm 2\varepsilon \left(\frac{2}{\check{C}}\right)^{1/2} \operatorname{sech}(\varepsilon\xi)\cos(\tau) + O(\varepsilon^{3/2}\mathrm{e}^{-\varepsilon\eta|\xi|}), \quad 0 < \eta < 1$$

(so that $\lim_{\xi \to \pm\infty} p^\varepsilon(\xi, \tau) = 0$ uniformly in $\tau \in \mathbb{R}$), where $\check{C}$ is a normal-form coefficient which is computed in Section 6.2 (see equations (6.26), (6.28)) and required to be positive.

Groves & Schneider [10] previously constructed 'moving breather' solutions to equation (6.1) of the form

$$\phi(x,t) = w_1(\xi,\tau), \quad \xi = x - \tfrac{1}{c_p}t, \quad \tau = \omega(x - c_p t),$$

where w_1 is 2π-periodic in its second argument and $c_p = c_p' + \mathcal{O}(\varepsilon^2)$ is a perturbation of the linear phase velocity $c_p' = \omega(1+\omega^2)^{-1/2}$. The arguments presented in [10] rely on the structure of the Klein-Gordon equation (6.1), whereas we demonstrate how the general theory in this thesis also applies to these equations. Groves & Schneider [8] also considered the special case, obtaining moving breathers *algebraically* close to a pulse of the form $p(\xi,\tau)$; here we improve their result, obtaining an *exponentially* close approximation.

6.1 Existence theory

We begin by writing equation (6.1) as a first order system of the form (2.1). Inserting our *Ansatz*

$$\phi(x,t) = v_1(\xi,\tau), \qquad \xi = x, \quad \tau = \omega t,$$

with $\omega = \sqrt{1-\varepsilon^2}$ into equation (6.1) we obtain

$$v_{1\xi\xi} = \frac{(1-\varepsilon^2)v_{1\tau\tau} + v_1 - f_2(v_1, \sqrt{1-\varepsilon^2}v_{1\tau}, v_{1\xi})}{1 + f_1(v_1, \sqrt{1-\varepsilon^2}v_{1\tau}, v_{1\xi})}. \tag{6.2}$$

In order to write equation (6.2) as a first order system we define $v_2 = v_{1\xi}$ and $v_3 = v_{1\tau}$, so that

$$v_{1\xi} = v_2, \quad v_{2\xi} = \frac{(1-\varepsilon^2)v_{3\tau} + v_1 - f_2(v_1, \sqrt{1-\varepsilon^2}v_3, v_2)}{1 + f_1(v_1, \sqrt{1-\varepsilon^2}v_3, v_2)}, \quad v_{3\xi} = v_{2\tau}.$$

It follows that equation (6.2) can be written as

$$v_\xi = A^\varepsilon(v)v_\tau + B^\varepsilon(v), \tag{6.3}$$

where

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad A^\varepsilon(v) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & f_3^\varepsilon(v) \\ 0 & 1 & 0 \end{pmatrix}, \quad B^\varepsilon(v) = \begin{pmatrix} v_2 \\ f_4^\varepsilon(v) \\ 0 \end{pmatrix}$$

with

$$f_3^\varepsilon(v) = \frac{1-\varepsilon^2}{1 + f_1(v_1, \sqrt{1-\varepsilon^2}v_3, v_2)}, \quad f_4^\varepsilon(v) = \frac{v_1 - f_2(v_1, \sqrt{1-\varepsilon^2}v_3, v_2)}{1 + f_1(v_1, \sqrt{1-\varepsilon^2}v_3, v_2)}$$

(so that in particular $f_i^\varepsilon(a,-b,-c) = f_i^\varepsilon(a,b,c)$).

Remark 6.2. Solutions to equation (6.3) are not necessarily solutions to equation (6.2) since they do not automatically satisfy $v_3 = v_{1\tau}$; the solutions constructed below do however have this property. More generally, a solution to equation (6.3) which satisfies $v_3(0,\tau) = v_{1\tau}(0,\tau)$ for all $\tau \in [0, 2\pi)$ satisfies $v_3(\xi,\tau) = v_{1\tau}(\xi,\tau)$ for all ξ, since

$$(v_3 - v_{1\tau})_\xi = v_{3\xi} - v_{1\xi\tau} = v_{2\tau} - v_{2\tau} = 0.$$

Basic properties and symmetrisation

We first establish that the first order system (6.3) is a reversible, symmetrisable hyperbolic system with analytic coefficients (see Assumptions 1) and perform the symmetrisation of its linear part (see the comments below Assumptions 1). Clearly the functions $A\colon \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ and $B\colon \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$ are analytic at the origin with $B^\varepsilon(0) = 0$ (Assumptions 1(i)) and the matrix

$$A^\varepsilon(0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1-\varepsilon^2 \\ 0 & 1 & 0 \end{pmatrix}$$

is strictly hyperbolic (Assumptions 1(ii)).

To establish the symmetrisability of A (Assumptions 1(iii)) we note that the function f_1 in equation (6.1) is by assumption analytic at the origin and satisfies $f_1(0) = 0$. It follows that $f_3^0(0) = 1$ and thus

$$f_3^\varepsilon(v) = 1 + \mathcal{O}(|(\varepsilon, v)|).$$

The symmetric matrix

$$S^\varepsilon(v) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & f_3^\varepsilon(v) \end{pmatrix}$$

is therefore positive definite (in a neighbourhood of the origin) and satisfies

$$\Big(S^\varepsilon(v)A^\varepsilon(v)\Big)^\top = S^\varepsilon(v)A^\varepsilon(v).$$

Note that

$$S^\varepsilon(0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1-\varepsilon^2 \end{pmatrix},$$

so that

$$S^\varepsilon(0)^{\frac{1}{2}} A^\varepsilon(0) S^\varepsilon(0)^{-\frac{1}{2}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \sqrt{1-\varepsilon^2} \\ 0 & \sqrt{1-\varepsilon^2} & 0 \end{pmatrix}$$

and

$$S^\varepsilon(0)^{\frac{1}{2}} \partial B^\varepsilon[0] S^\varepsilon(0)^{-\frac{1}{2}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \partial B^\varepsilon[0].$$

The reversibility of equation (6.3) (Assumptions 1(iv)) follows from the fact that the matrix

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

satisfies $T^2 = I$, $A^\varepsilon(Tv)T = TA^\varepsilon(v)$ and $B^\varepsilon(Tv) = -TB^\varepsilon(v)$. We conclude that equation (6.3) is invariant with respect to the transformation $\xi \mapsto -\xi$, $v \mapsto Rv$ with

$$(Rv)(\tau) = Tv(-\tau).$$

We also note that T commutes with $S^\varepsilon(0)$ and hence with $S^\varepsilon(0)^{\frac{1}{2}}$ and $S^\varepsilon(0)^{-\frac{1}{2}}$.

The change of variable $u = S^\varepsilon(0)^{\frac{1}{2}}v$ yields the reversible first order system

$$u_\xi = \tilde{A}^\varepsilon(u)u_\tau + \tilde{B}^\varepsilon(u) = L^\varepsilon u + N^\varepsilon(u), \tag{6.4}$$

where

$$\tilde{A}^\varepsilon(u) = S^\varepsilon(0)^{\frac{1}{2}}A^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u)S^\varepsilon(0)^{-\frac{1}{2}}, \quad \tilde{B}^\varepsilon(u) = S^\varepsilon(0)^{\frac{1}{2}}B^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u).$$

Here the linear operator L^ε and the nonlinearity N^ε are given by

$$L^\varepsilon u = \tilde{A}^\varepsilon(0)u_\tau + \partial\tilde{B}^\varepsilon[0]u = \begin{pmatrix} u_2 \\ \sqrt{1-\varepsilon^2}u_{3\tau} + u_1 \\ \sqrt{1-\varepsilon^2}u_{2\tau} \end{pmatrix}$$

and

$$\begin{aligned} N^\varepsilon(u) &= \left(\tilde{A}^\varepsilon(u) - \tilde{A}^\varepsilon(0)\right)u_\tau + \tilde{B}^\varepsilon(u) - \partial\tilde{B}^\varepsilon[0]u \\ &= \begin{pmatrix} 0 \\ \dfrac{f_3^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u) - f_3^\varepsilon(0)}{\sqrt{1-\varepsilon^2}}u_{3\tau} + f_4^\varepsilon(S^\varepsilon(0)^{-\frac{1}{2}}u) - u_1 \\ 0 \end{pmatrix}, \end{aligned}$$

where

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \sqrt{1-\varepsilon^2}v_3 \end{pmatrix}.$$

In the special case where the functions f_1 and f_2 in equation (6.1) satisfy $f_1 = 0$, $f_2(a,b,c) = f_2(a)$ it follows that

$$f_3^\varepsilon(v) = 1-\varepsilon^2, \quad f_4^\varepsilon(v) = v_1 - f_2(v_1),$$

and thus

$$N^\varepsilon(u) = \begin{pmatrix} 0 \\ f_2(u_1) \\ 0 \end{pmatrix}.$$

In this case the function

$$\mathcal{H}^{\varepsilon}(u) = \int_0^{2\pi} \left\{ \frac{1}{2}u_2^2 + \frac{1}{2}u_3^2 - \frac{1}{2}u_1^2 - F_2(u_1) \right\} \mathrm{d}\tau, \tag{6.5}$$

in which F_2 is an antiderivative of f_2, is a conserved quantity for the system (6.4).

Remark 6.3. As indicated in Assumptions 1(iii) no symmetrisation is needed if the function f_3 in equation (6.3) is a constant function of v. However we include the similarity transformation $S^{\varepsilon}(0)^{\frac{1}{2}}$ for simplicity below. The formula for the normal-form coefficient $\check{C}$ in the hypothesis of Theorem 6.1, which is derived in Section 6.2, is not affected since $S^0(0) = I$.

Spectral analysis

We examine the spectrum of the linear operator $L^{\varepsilon} \in \mathcal{L}(\mathcal{X}^{t+1}, \mathcal{X}^t)$ with

$$\mathcal{X}^t = H^t_{\mathrm{per}}(0, 2\pi; \mathbb{C}^3), \quad t \geq 0.$$

For this purpose we consider each mode seperately and compute the spectrum of the operator

$$L^{\varepsilon}_m \colon E_m \to E_m, \quad L^{\varepsilon}_m(u_m \mathrm{e}^{\mathrm{i}m\tau}) = \left(\mathrm{i}m\tilde{A}^{\varepsilon}(0) + \partial\tilde{B}^{\varepsilon}[0]\right) u_m \mathrm{e}^{\mathrm{i}m\tau}$$

for each $m \in \mathbb{Z}$, where

$$\mathrm{i}m\tilde{A}^{\varepsilon}(0) + \partial\tilde{B}^{\varepsilon}[0] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \mathrm{i}m\sqrt{1-\varepsilon^2} \\ 0 & \mathrm{i}m\sqrt{1-\varepsilon^2} & 0 \end{pmatrix}.$$

$\boldsymbol{m = 0}$: The matrix

$$\partial\tilde{B}^{\varepsilon}[0] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

has, independently of ε, the geometrically simple eigenvalues -1, 0 and 1. The corresponding (ε-independent) eigenspaces are given by

$$E^{(0)}_{-1} = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\rangle, \quad E^{(0)}_{0} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle, \quad E^{(0)}_{1} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

$\boldsymbol{m = \pm 1}$: For $\varepsilon = 0$ the matrix

$$\pm\mathrm{i}\tilde{A}^0(0) + \partial\tilde{B}^0[0] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \pm\mathrm{i} \\ 0 & \pm\mathrm{i} & 0 \end{pmatrix}$$

has the algebraically triple, geometrically simple eigenvalue 0. The (generalised) eigenvectors are

$$\begin{pmatrix}1\\0\\\pm\mathrm{i}\end{pmatrix},\quad \begin{pmatrix}0\\1\\0\end{pmatrix},\quad \begin{pmatrix}1\\0\\0\end{pmatrix}.$$

For $\varepsilon > 0$ the matrix

$$\pm\mathrm{i}\tilde{A}^{\varepsilon}(0) + \partial\tilde{B}^{\varepsilon}[0] = \begin{pmatrix}0 & 1 & 0\\ 1 & 0 & \pm\mathrm{i}\sqrt{1-\varepsilon^2}\\ 0 & \pm\mathrm{i}\sqrt{1-\varepsilon^2} & 0\end{pmatrix}$$

has the geometrically simple eigenvalues $-\varepsilon$, 0 and ε. The corresponding eigenspaces are given by

$$E_{-\varepsilon}^{(\pm1)} = \left\langle \begin{pmatrix}1\\-\varepsilon\\\pm\mathrm{i}\sqrt{1-\varepsilon^2}\end{pmatrix}\right\rangle,\quad E_{0}^{(\pm1)} = \left\langle \begin{pmatrix}\sqrt{1-\varepsilon^2}\\0\\\pm\mathrm{i}\end{pmatrix}\right\rangle,\quad E_{\varepsilon}^{(\pm1)} = \left\langle \begin{pmatrix}1\\\varepsilon\\\pm\mathrm{i}\sqrt{1-\varepsilon^2}\end{pmatrix}\right\rangle.$$

$\boldsymbol{|m| > 1}$: The matrix

$$\mathrm{i}m\tilde{A}^{\varepsilon}(0) + \partial\tilde{B}^{\varepsilon}[0] = \begin{pmatrix}0 & 1 & 0\\ 1 & 0 & \mathrm{i}m\sqrt{1-\varepsilon^2}\\ 0 & \mathrm{i}m\sqrt{1-\varepsilon^2} & 0\end{pmatrix}$$

has the geometrically simple eigenvalues $-\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}$, 0 and $\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}$. The corresponding eigenspaces are given by

$$E^{(m)}_{-\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}} = \left\langle \begin{pmatrix}1\\-\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}\\\mathrm{i}m\sqrt{1-\varepsilon^2}\end{pmatrix}\right\rangle,\quad E_0^{(m)} = \left\langle \begin{pmatrix}-\mathrm{i}m\sqrt{1-\varepsilon^2}\\0\\1\end{pmatrix}\right\rangle,$$

$$E^{(m)}_{\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}} = \left\langle \begin{pmatrix}1\\\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}\\\mathrm{i}m\sqrt{1-\varepsilon^2}\end{pmatrix}\right\rangle.$$

Remark 6.4. The eigenvalues of the matrix $\mathrm{i}m\tilde{A}^{\varepsilon}(0) + \partial\tilde{B}^{\varepsilon}[0]$ are geometrically simple for each $m \in \mathbb{Z}$ (Assumptions 2(i)) and purely imaginary for $|m| > 1$ if ε is sufficiently small.

Observe the special role of the eigenvalues in modes ±1: while the only mode ±1 eigenvalue for $\varepsilon = 0$ is the algebraically triple, geometrically simple eigenvalue 0, one simple zero eigenvalue persits for $\varepsilon > 0$. The additional simple zero eigenvalues in modes $m \in \mathbb{Z} \setminus \{\pm1\}$ however remain simple zero eigenvalues for all values of ε. In the considerations below they are therefore seperated from the eigenvalues undergoing a non-semisimple collision in modes ±1.

Finally, we construct the spectral projections $P^\varepsilon_{\mathrm{wh}}$, $P^\varepsilon_{\mathrm{sh}}$, $P^\varepsilon_{\mathrm{c}}$ and P^ε_0 for the linear operator L^ε associated with respectively its eigenvalues $-\varepsilon$, 0, ε in modes ± 1, its (finitely many) remaining eigenvalues with non-zero real part, its purely imaginary eigenvalues, and its remaining zero eigenvalues; precise definitions are given by formula (2.7) and

$$(P^\varepsilon_0 u)(\tau) = \frac{1}{\sqrt{2\pi}} \sum_{m\in\mathbb{Z}\setminus\{\pm 1\}} P^{m,\varepsilon}_0 (u_m \mathrm{e}^{\mathrm{i}m\tau}),$$

where $P^{m,\varepsilon}_0 \colon E_m \to E_m$ is the spectral projection associated with the zero eigenvalues of $L^\varepsilon|_{E_m}$, $m \in \mathbb{Z} \setminus \{\pm 1\}$. We find that

$$\begin{aligned}
&P^0_{\mathrm{wh}}\mathcal{X}^t \\
&\quad= \left\langle \begin{pmatrix}1\\0\\\mathrm{i}\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}0\\1\\0\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}1\\0\\0\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}1\\0\\-\mathrm{i}\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix}0\\1\\0\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix}1\\0\\0\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau} \right\rangle \\
&\quad= E_{-1} \oplus E_1, \\
&P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t \\
&\quad= \left\langle \begin{pmatrix}1\\-\varepsilon\\\mathrm{i}\sqrt{1-\varepsilon^2}\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}\sqrt{1-\varepsilon^2}\\0\\\mathrm{i}\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}1\\\varepsilon\\\mathrm{i}\sqrt{1-\varepsilon^2}\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \right. \\
&\qquad\qquad \left. \begin{pmatrix}1\\-\varepsilon\\-\mathrm{i}\sqrt{1-\varepsilon^2}\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix}\sqrt{1-\varepsilon^2}\\0\\-\mathrm{i}\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix}1\\\varepsilon\\-\mathrm{i}\sqrt{1-\varepsilon^2}\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau} \right\rangle \\
&\quad= E_{-1} \oplus E_1, \quad \varepsilon > 0, \\
&P^\varepsilon_{\mathrm{sh}}\mathcal{X}^t \\
&\quad= \left\langle \begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix} \right\rangle \subseteq E_0, \quad \varepsilon \geq 0, \\
&P^\varepsilon_{\mathrm{c}}\mathcal{X}^t \\
&\quad= \left\{ \sum_{m\in\mathbb{Z}\setminus\{-1,0,1\}} w^\varepsilon_{m,1} \begin{pmatrix}1\\-\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}\\\mathrm{i}m\sqrt{1-\varepsilon^2}\end{pmatrix}\mathrm{e}^{\mathrm{i}m\tau} + w^\varepsilon_{m,2} \begin{pmatrix}1\\\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1}\\\mathrm{i}m\sqrt{1-\varepsilon^2}\end{pmatrix}\mathrm{e}^{\mathrm{i}m\tau} : \right. \\
&\qquad\qquad \left. w^\varepsilon_{m,1}, w^\varepsilon_{m,2} \in \mathbb{C} \right\} \\
&\quad\subseteq \bigoplus_{m\in\mathbb{Z}\setminus\{-1,0,1\}} E_m, \quad \varepsilon \geq 0,
\end{aligned}$$

and

$$
\begin{aligned}
&P_0^\varepsilon \mathcal{X}^t \\
&\quad = \left\{ \sum_{m\in\mathbb{Z}\setminus\{-1,1\}} X_m^\varepsilon \begin{pmatrix} -\mathrm{i}m\sqrt{1-\varepsilon^2} \\ 0 \\ 1 \end{pmatrix} \mathrm{e}^{\mathrm{i}m\tau} \colon X_m^\varepsilon \in \mathbb{C} \right\} \\
&\quad \subseteq \bigoplus_{m\in\mathbb{Z}\setminus\{-1,1\}} E_m, \quad \varepsilon \geq 0,
\end{aligned}
$$

where the series converge in $\mathcal{X}^t$.

Formulation as a coupled system

Note that particular components of functions in the spaces $P_\mathrm{c}^\varepsilon \mathcal{X}^t$ and $P_0^\varepsilon \mathcal{X}^t$ have increased regularity. Taking

$$
w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in P_\mathrm{c}^\varepsilon \mathcal{X}^t
$$

we observe that $w_2, w_3 \in H_\mathrm{per}^t(0,2\pi;\mathbb{C})$ in particular requires $w_1 \in H_\mathrm{per}^{t+1}(0,2\pi;\mathbb{C})$ because

$$
\begin{aligned}
P_\mathrm{c}^\varepsilon \mathcal{X}^t = \Bigg\{ &\begin{pmatrix} 1 \\ \mathrm{i}\sqrt{(1-\varepsilon^2)D^2-1} \\ \sqrt{1-\varepsilon^2}\partial_\tau \end{pmatrix} u_{1,1} + \begin{pmatrix} 1 \\ -\mathrm{i}\sqrt{(1-\varepsilon^2)D^2-1} \\ \sqrt{1-\varepsilon^2}\partial_\tau \end{pmatrix} u_{1,2} \colon \\
&u_{1,1}, u_{1,2} \in H_\mathrm{per}^{t+1}(0,2\pi;\mathbb{C}) \Bigg\}, \quad \varepsilon \geq 0.
\end{aligned}
$$

Similarly we conclude that

$$
P_0^\varepsilon \mathcal{X}^t = \left\{ \begin{pmatrix} -\partial_\tau \\ 0 \\ 1 \end{pmatrix} u_3 \colon u_3 \in H_\mathrm{per}^{t+1}(0,2\pi;\mathbb{C}) \right\}, \quad \varepsilon \geq 0.
$$

It is a direct consequence of this observation, and the fact that a function $u \in \mathcal{X}^t$ of the form

$$
u = \begin{pmatrix} 0 \\ u_2 \\ 0 \end{pmatrix}
$$

has by definition no components in the space $P_0^\varepsilon \mathcal{X}^t$, that the nonlinearity N^ε maps $P_\mathrm{wh}^\varepsilon \mathcal{X}^{t+1} \oplus P_\mathrm{sh,c}^\varepsilon \mathcal{X}^{t+1}$ analytically into $P_\mathrm{wh}^\varepsilon \mathcal{X}^t \oplus P_\mathrm{sh,c}^\varepsilon \mathcal{X}^t$ for each $t > \frac{1}{2}$ and each $\varepsilon \geq 0$, where $P_\mathrm{sh,c}^\varepsilon = P_\mathrm{sh}^\varepsilon + P_\mathrm{c}^\varepsilon$. In the special case, where

$$
N^\varepsilon(u) = \begin{pmatrix} 0 \\ f_2(u_1) \\ 0 \end{pmatrix},
$$

we further observe that N^ε maps $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^t$ analytically into itself (in fact into $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^{t+1} \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^{t+1}$) for each $t \geq 0$ and each $\varepsilon \geq 0$. These facts allow us to study equation (6.4) in respectively the phase space

$$P^\varepsilon_{\mathrm{wh}}\mathcal{X}^s \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^s$$

for fixed $s > \frac{1}{2}$ and the phase space

$$P^\varepsilon_{\mathrm{wh}}\mathcal{X}^0 \oplus P^\varepsilon_{\mathrm{sh,c}}\mathcal{X}^0$$

in the special case.

We continue by writing equation (6.4) as

$$\begin{aligned} z^\varepsilon &= L^\varepsilon z^\varepsilon + P^\varepsilon_{\mathrm{wh}} N^\varepsilon(z^\varepsilon + q^\varepsilon) &(6.6)\\ q^\varepsilon &= L^\varepsilon q^\varepsilon + P^\varepsilon_{\mathrm{sh,c}} N^\varepsilon(z^\varepsilon + q^\varepsilon), &(6.7)\end{aligned}$$

where $z^\varepsilon = P^\varepsilon_{\mathrm{wh}}u$, $q^\varepsilon = P^\varepsilon_{\mathrm{sh,c}}u$. The spaces $P^\varepsilon_{\mathrm{wh}}\mathcal{X}^t$ and $P^\varepsilon_{\mathrm{sh}}\mathcal{X}^t$ evidently do not depend upon ε, while the ε-dependence of the space $P^\varepsilon_{\mathrm{c}}\mathcal{X}^t$ can be removed using a Kato similarity transformation. According to Lemma 2.5 there exists a mapping $U^\varepsilon \in \mathcal{L}(\mathcal{X}^t)$ for all $t \geq 0$ such that

$$U^\varepsilon P_{\mathrm{wh}}(U^\varepsilon)^{-1} = P^\varepsilon_{\mathrm{wh}}, \quad U^\varepsilon P_{\mathrm{sh}}(U^\varepsilon)^{-1} = P^\varepsilon_{\mathrm{sh}}, \quad U^\varepsilon P_{\mathrm{c}}(U^\varepsilon)^{-1} = P^\varepsilon_{\mathrm{c}}, \quad U^\varepsilon P_0(U^\varepsilon)^{-1} = P^\varepsilon_0, \tag{6.8}$$

where $P_{\mathrm{wh}} = P^0_{\mathrm{wh}}$, $P_{\mathrm{sh}} = P^0_{\mathrm{sh}}$, $P_{\mathrm{c}} = P^0_{\mathrm{c}}$ and $P_0 = P^0_0$. Writing $q^\varepsilon = U^\varepsilon q$ with $q \in \mathcal{X}^{s+1}_{\mathrm{sh,c}} = P^0_{\mathrm{sh,c}}\mathcal{X}^{s+1}$ and $z^\varepsilon = z$ with $z \in \mathcal{X}_{\mathrm{wh}} = \mathcal{X}^0_{\mathrm{wh}}$, we find that

$$\begin{aligned} z_\xi &= L^\varepsilon_{\mathrm{wh}} z + P^\varepsilon_{\mathrm{wh}} N^\varepsilon(z + U^\varepsilon q), &(6.9)\\ q_\xi &= L^\varepsilon_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}(U^\varepsilon)^{-1} N^\varepsilon(z + U^\varepsilon q), &(6.10)\end{aligned}$$

where

$$L^\varepsilon_{\mathrm{sh,c}} = (U^\varepsilon)^{-1} L^\varepsilon U^\varepsilon|_{\mathcal{X}^{t+1}_{\mathrm{sh,c}}} \quad (t \geq 0)$$

and the matrix of $L^\varepsilon_{\mathrm{wh}} = L^\varepsilon|_{\mathcal{X}_{\mathrm{wh}}}$ with respect to the basis

$$\left\{ \begin{pmatrix}1\\0\\\mathrm{i}\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}0\\1\\0\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}1\\0\\0\end{pmatrix}\mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix}1\\0\\-\mathrm{i}\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix}0\\1\\0\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix}1\\0\\0\end{pmatrix}\mathrm{e}^{-\mathrm{i}\tau} \right\}$$

for $\mathcal{X}_{\mathrm{wh}}$ is

$$\begin{pmatrix} 0 & \sqrt{1-\varepsilon^2} & 0 & 0 & 0 & 0\\ 1-\sqrt{1-\varepsilon^2} & 0 & 1 & 0 & 0 & 0\\ 0 & 1-\sqrt{1-\varepsilon^2} & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & \sqrt{1-\varepsilon^2} & 0\\ 0 & 0 & 0 & 1-\sqrt{1-\varepsilon^2} & 0 & 1\\ 0 & 0 & 0 & 0 & 1-\sqrt{1-\varepsilon^2} & 0 \end{pmatrix}.$$

Equation (6.9) is not of the form discussed in Chapter 2 (see equation (2.10)) due to the presence of an additional zero eigenvalue in modes ± 1. We therefore obtain a six-dimensional rather than a four-dimensional weakly hyperbolic subspace $\mathcal{X}_{\mathrm{wh}}$. Because of the special structure of the nonlinearity (see equation (6.9)) the equations for the components z_1 and z_3 are in fact

$$\begin{aligned} z_{1\xi} &= \sqrt{1-\varepsilon^2} z_2, \\ z_{3\xi} &= (1-\sqrt{1-\varepsilon^2}) z_2. \end{aligned}$$

The quantity

$$I^\varepsilon_{\mathrm{wh}}(z) = z_3 - \frac{1-\sqrt{1-\varepsilon^2}}{\sqrt{1-\varepsilon^2}} z_1$$

is therefore conserved, so that we can study equation (6.9) in the invariant subspace $\{I^\varepsilon_{\mathrm{wh}}(z)=0\}$.

Remark 6.5. Suppose that (z,q) is a solution to equations (6.9), (6.10) with $I^\varepsilon_{\mathrm{wh}}(z)=0$ and note that

$$z(\tau) = z_1 \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau} + z_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau} + \frac{1-\sqrt{1-\varepsilon^2}}{\sqrt{1-\varepsilon^2}} z_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau} + \text{c.c.}$$

and

$$\begin{aligned} (U^\varepsilon q)(\tau) = Z_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + Z_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \sum_{m \in \mathbb{Z}\setminus\{-1,0,1\}} & w^\varepsilon_{m,1} \begin{pmatrix} 1 \\ -\mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1} \\ \mathrm{i}m\sqrt{1-\varepsilon^2} \end{pmatrix} \mathrm{e}^{\mathrm{i}m\tau} \\ & + w^\varepsilon_{m,2} \begin{pmatrix} 1 \\ \mathrm{i}\sqrt{m^2(1-\varepsilon^2)-1} \\ \mathrm{i}m\sqrt{1-\varepsilon^2} \end{pmatrix} \mathrm{e}^{\mathrm{i}m\tau} \end{aligned}$$

(by definition of the spaces $P^\varepsilon_{\mathrm{sh}}\mathcal{X}^t$ and $P^\varepsilon_{\mathrm{c}}\mathcal{X}^t$). It follows that $u = z + U^\varepsilon q$ solves the system (6.4) and satisfies $\sqrt{1-\varepsilon^2} u_{1\tau} = u_3$ and hence $v_{1\tau} = v_3$. According to Remark 6.2 a solution to the coupled system (6.9), (6.10) therefore automatically leads to a solution of equation (6.2) and thus of the original equation (6.1).

Next we write equation (6.9) as

$$z_\xi = \check{L}^\varepsilon_{\mathrm{wh}} z + \check{N}^\varepsilon_{\mathrm{wh}}(z,q), \tag{6.11}$$

where

$$\check{L}^\varepsilon_{\mathrm{wh}} z = L^\varepsilon_{\mathrm{wh}}|_{\{I^\varepsilon_{\mathrm{wh}}(z)=0\}}, \quad \check{N}^\varepsilon_{\mathrm{wh}}(\cdot,\cdot) = P^\varepsilon_{\mathrm{wh}} N^\varepsilon(\cdot + U^\varepsilon \cdot)|_{\{I^\varepsilon_{\mathrm{wh}}(z)=0\}},$$

and observe that $\{I^\varepsilon_{\mathrm{wh}}(z)=0\}$ can be identified with the ε-independent, four-dimensional subspace

$$\check{\mathcal{X}}_{\mathrm{wh}} = \left\langle \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix} 1 \\ 0 \\ -\mathrm{i} \end{pmatrix} \mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{-\mathrm{i}\tau} \right\rangle$$

of $\mathcal{X}_{\mathrm{wh}}$. With a slight abuse of notation we therefore regard $\check{L}^{\varepsilon}_{\mathrm{wh}}$ and $\check{N}^{\varepsilon}_{\mathrm{wh}}$ as operators $\check{\mathcal{X}}_{\mathrm{wh}} \to \check{\mathcal{X}}_{\mathrm{wh}}$ and $\check{\mathcal{X}}_{\mathrm{wh}} \times \mathcal{X}^t_{\mathrm{sh,c}} \to \check{\mathcal{X}}_{\mathrm{wh}}$ respectively. The matrix of $\check{L}^{\varepsilon}_{\mathrm{wh}}$ with respect to the basis

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix} 1 \\ 0 \\ -\mathrm{i} \end{pmatrix} \mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{-\mathrm{i}\tau} \right\}$$

for $\check{\mathcal{X}}_{\mathrm{wh}}$ is

$$\begin{pmatrix} 0 & \sqrt{1-\varepsilon^2} & 0 & 0 \\ \frac{\varepsilon^2}{\sqrt{1-\varepsilon^2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{1-\varepsilon^2} \\ 0 & 0 & \frac{\varepsilon^2}{\sqrt{1-\varepsilon^2}} & 0 \end{pmatrix}.$$

Finally, we identify $\check{\mathcal{X}}_{\mathrm{wh}}$ with $\mathbb{C}^4$ and introduce the scaling $z = D^{\varepsilon}\check{z}$, where

$$D^{\varepsilon} \begin{pmatrix} \check{z}_1 \\ \check{z}_2 \end{pmatrix} = \begin{pmatrix} \sqrt{1-\varepsilon^2}\,\check{z}_1 \\ \check{z}_2 \end{pmatrix}. \tag{6.12}$$

We thus arrive at the system

$$\check{z}_{\xi} = \check{L}^{\varepsilon}_{\mathrm{wh}}\check{z} + f^{\varepsilon}(\check{z}, q), \tag{6.13}$$

$$q_{\xi} = L^{\varepsilon}_{\mathrm{sh,c}} q + P_{\mathrm{sh,c}}\Big(g^{\varepsilon}_1(\check{z}, q) q_{\tau}\Big) + g^{\varepsilon}_2(\check{z}, q) + h^{\varepsilon}(\check{z}), \tag{6.14}$$

where

$$\begin{aligned} f^{\varepsilon}(\check{z}, q) &= (D^{\varepsilon})^{-1}\check{N}^{\varepsilon}_{\mathrm{wh}}(D^{\varepsilon}\check{z}, q), \\ g^{\varepsilon}_1(\check{z}, q) &= (U^{\varepsilon})^{-1}\Big(\tilde{A}^{\varepsilon}(z + U^{\varepsilon}q) - \tilde{A}^{\varepsilon}(0)\Big)U^{\varepsilon}, \\ g^{\varepsilon}_2(\check{z}, q) &= P_{\mathrm{sh,c}}(U^{\varepsilon})^{-1}\Big((\tilde{A}^{\varepsilon}(z + U^{\varepsilon}q) - \tilde{A}^{\varepsilon}(z))z_{\tau} + \tilde{B}^{\varepsilon}(z + U^{\varepsilon}q) \\ &\qquad\qquad - \tilde{B}^{\varepsilon}(z) - \partial\tilde{B}^{\varepsilon}[0]U^{\varepsilon}q\Big), \\ h^{\varepsilon}(\check{z}) &= P_{\mathrm{sh,c}}(U^{\varepsilon})^{-1}\Big(\tilde{A}^{\varepsilon}(z)z_{\tau} + \tilde{B}^{\varepsilon}(z)\Big) \end{aligned}$$

and all functions are evaluated at

$$z_1 = \sqrt{1-\varepsilon^2}\,\check{z}_1, \quad z_2 = \check{z}_2, \quad z_3 = (1 - \sqrt{1-\varepsilon^2})\check{z}_1$$

(the relevant estimates for the nonlinearities are given in (2.12)–(2.15) and hold in particular for all $t \geq 0$ in the special case). The scaling (6.12) converts the matrix of $\check{L}^{\varepsilon}_{\mathrm{wh}}$ with respect to the basis

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}\tau}, \begin{pmatrix} 1 \\ 0 \\ -\mathrm{i} \end{pmatrix} \mathrm{e}^{-\mathrm{i}\tau}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{-\mathrm{i}\tau} \right\}$$

for $\check{\mathcal{X}}_{\mathrm{wh}}$ into

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ \varepsilon^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \varepsilon^2 & 0 \end{pmatrix}$$

such that the operator $\check{L}^{\varepsilon}_{\mathrm{wh}}$ in equation (6.13) satisfies Assumptions 3.

In the special case we have that $g_1 = 0$ in equation (6.14). Recall that

$$\mathcal{H}^{\varepsilon}(u) = \int_0^{2\pi} \left\{ \frac{1}{2}u_2^2 + \frac{1}{2}u_3^2 - \frac{1}{2}u_1^2 - F_2(u_1) \right\} \mathrm{d}\tau$$

(see equation (6.5)) is a conserved quantity for the system (6.4). The formula

$$\mathcal{I}^{\varepsilon}(z,q) = \mathcal{H}^{\varepsilon}(z + U^{\varepsilon} q)$$

thus defines a conserved quantity for the system (6.13), (6.14) in $\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^0_{\mathrm{sh,c}}$, where $u_1 \in H^1_{\mathrm{per}}(0, 2\pi; \mathbb{C})$ for $u = z + U^{\varepsilon} q$ (see Remark 6.5). Clearly

$$|\mathcal{I}^{\varepsilon}(z,q)| = \mathcal{O}(\|(z,q)\|^2_{\mathcal{X}_{\mathrm{wh}} \times \mathcal{X}^0_{\mathrm{sh,c}}})$$

and taking $w \in \mathcal{X}^0_{\mathrm{c}}$, so that $w_1 \in H^1_{\mathrm{per}}(0, 2\pi; \mathbb{C})$, with

$$w(\tau) = \sum_{m \in \mathbb{Z} \setminus \{-1,0,1\}} w_{m,1} \begin{pmatrix} 1 \\ -\mathrm{i}\sqrt{m^2-1} \\ \mathrm{i}m \end{pmatrix} \mathrm{e}^{\mathrm{i}m\tau} + w_{m,2} \begin{pmatrix} 1 \\ \mathrm{i}\sqrt{m^2-1} \\ \mathrm{i}m \end{pmatrix} \mathrm{e}^{\mathrm{i}m\tau},$$

we compute

$$\begin{aligned} \|w\|^2_{\mathcal{X}^0_{\mathrm{c}}} &= 8\pi \sum_{m=2}^{\infty} \left\{ m^2 |w_{m,1}|^2 + m^2 |w_{m,2}|^2 + w_{m,1}\overline{w}_{m,2} + \overline{w}_{m,1} w_{m,2} \right\} \\ &\cong \sum_{m=2}^{\infty} (1+m^2) \Big(|w_{m,1}|^2 + |w_{m,2}|^2 \Big) \end{aligned}$$

and

$$\begin{aligned} \|w_1\|_1^2 &= 4\pi \sum_{m=2}^{\infty} (1+m^2) \Big(|w_{m,1}|^2 + |w_{m,2}|^2 + w_{m,1}\overline{w}_{m,2} + \overline{w}_{m,1} w_{m,2} \Big) \\ &\cong \sum_{m=2}^{\infty} (1+m^2) \Big(|w_{m,1}|^2 + |w_{m,2}|^2 \Big) \end{aligned}$$

by estimating

$$|2\,\mathrm{Re}(w_{m,1}\overline{w}_{m,2})| \leq |w_{m,1}|^2 + |w_{m,2}|^2.$$

For the quadratic part $Q(w)$ of $\mathcal{I}^0(0,w)$ we further calculate

$$\begin{aligned} Q(w) &= \int_0^{2\pi} \left\{ \frac{1}{2} w_2^2 + \frac{1}{2} w_3^2 - \frac{1}{2} w_1^2 \right\} \mathrm{d}\tau \\ &= 4\pi \sum_{m=2}^{\infty} (m^2-1)\Big(|w_{m,1}|^2 + |w_{m,2}|^2 \Big) \\ &\cong \|w_1\|_1^2, \end{aligned}$$

from which it follows that

$$\|w\|_{\mathcal{X}_c^0}^2 \lesssim Q(w) \lesssim \|w\|_{\mathcal{X}_c^0}^2$$

so that the coupled system (6.13), (6.14) satisfies Assumptions 5 with $s = 0$.

Theorem 6.1 now follows by applying the existence theory in Chapter 4 (see Remark 2.9) or Chapter 5 to equations (6.13), (6.14). When applying the theory in Chapter 4, note that the relationship

$$P_{\mathrm{sh,c}} = I - P_{\mathrm{wh}} - P_0$$

reduces to the required formula

$$P_{\mathrm{sh,c}} = I - P_{\mathrm{wh}}$$

when applied to the nonlinearities (see the comments above equations (6.6), (6.7)).

6.2 Calculation of the bifurcation coefficient

It remains to calculate the coefficient $\check{C}$ in the hypothesis of Theorem 6.1. To this end it is necessary to apply Lemma 2.10, that is a change of variable of the form $\tilde{q} = q + \Gamma(z)$ (with $\Gamma\colon \mathcal{X}_{\mathrm{wh}} \to E_{-2} \oplus E_0 \oplus E_2$) which removes the quadratic, ε-independent part in the Maclaurin expansion of $h^\varepsilon(z)$, to equation (6.14) (see the comments above Lemma 2.10). We obtain the transformed system

$$z_\xi = L_{\mathrm{wh}}^\varepsilon z + \tilde{f}^\varepsilon(z,q), \tag{6.15}$$

$$q_\xi = L_{\mathrm{sh,c}}^\varepsilon q + P_{\mathrm{sh,c}}\Big(\tilde{g}_1^\varepsilon(z,q) q_\tau\Big) + \tilde{g}_2^\varepsilon(z,q) + \tilde{h}^\varepsilon(z), \tag{6.16}$$

where we have written z, $L_{\mathrm{wh}}^\varepsilon$ instead of $\check{z}$, $\check{L}_{\mathrm{wh}}^\varepsilon$ in equations (6.13), (6.14) for notational simplicity. Here $\tilde{f}^\varepsilon(z,q) = f^\varepsilon(z, q - \Gamma(z))$ and the transformed nonlinearity $\tilde{h}$ satisfies the estimate

$$\|\tilde{h}^\varepsilon(z)\|_{s+1} = \mathcal{O}(|z|^2 |(z,\varepsilon)|).$$

The coefficient $\check{C}$ is the coefficient of the cubic, ε-independent part in the Maclaurin expansion of $\tilde{f}^\varepsilon(z,0)$ in the system

$$z_\xi = L_{\mathrm{wh}}^\varepsilon z + \tilde{f}^\varepsilon(z,0)$$

which we write as

$$z_{1\xi} = z_2 \tag{6.17}$$

$$z_{2\xi} = \varepsilon^2 z_1 - \check{C} z_1 |z_1|^2 + \mathcal{R}_2^\varepsilon(z_1, z_2, \overline{z}_1, \overline{z}_2). \tag{6.18}$$

This system is obtained from equations (6.15), (6.16) by neglecting $\tilde{h}^\varepsilon(z)$ and setting $q = 0$ (note that the nonlinearity $N_{\mathrm{wh}}^\varepsilon$ in equation (6.11) and thus $\tilde{f}^\varepsilon$ have no components in z_1-direction so that the corresponding remainder term $\mathcal{R}_1^\varepsilon$ in equation (6.17) vanishes).

We recall the formula

$$\begin{aligned}\check{C} = -P_{(2)}\Big(3f^{0;3,0}\big[\{u\mathrm{e}^{\mathrm{i}\tau}\}^{(2)}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}\big] + f^{0;1,1}\big[\overline{u}\mathrm{e}^{-\mathrm{i}\tau}, -\Gamma_{2,2000,1}e_{2,1}\mathrm{e}^{2\mathrm{i}\tau} - \Gamma_{2,2000,2}e_{2,2}\mathrm{e}^{2\mathrm{i}\tau}\big] \\ + f^{0;1,1}\big[u\mathrm{e}^{\mathrm{i}\tau}, -\Gamma_{0,1010,1}e_{0,1} - \Gamma_{0,1010,2}e_{0,2}\big]\Big),\end{aligned} \tag{6.19}$$

in which we have used the notation introduced above Assumptions 4 and have written

$$z(\tau) = z_1 u\mathrm{e}^{\mathrm{i}\tau} + z_2 v\mathrm{e}^{\mathrm{i}\tau} + \overline{z}_1\overline{u}\mathrm{e}^{-\mathrm{i}\tau} + \overline{z}_2\overline{v}\mathrm{e}^{-\mathrm{i}\tau},$$

(note that $z_3 = 0$ in the formula for f^0 below equations (6.13), (6.14)) and

$$\big(\Gamma(z)\big)(\tau) = \sum_{i+j+k+l=2}\ \sum_{\substack{m\in\{-2,0,2\}\\ 1\le r\le 2}} \Gamma_{m,ijkl,r} z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r}\mathrm{e}^{\mathrm{i}m\tau}$$

with

$$u = \begin{pmatrix}1\\0\\\mathrm{i}\end{pmatrix}, \quad v = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad e_{0,1} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}, \quad e_{0,2} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix},$$

$$e_{2,1} = \frac{1}{2\sqrt{2}}\begin{pmatrix}1\\-\mathrm{i}\sqrt{3}\\2\mathrm{i}\end{pmatrix}, \quad e_{2,2} = \frac{1}{2\sqrt{2}}\begin{pmatrix}1\\\mathrm{i}\sqrt{3}\\2\mathrm{i}\end{pmatrix}.$$

The first step is to compute the relevant coefficients of the polynomial $\Gamma(z)$, which is the solution of the homological equation $\mathcal{L}\Gamma = h_2^0$ (see equation (2.24)), where

$$(\mathcal{L}\Gamma)(z) = L_{\mathrm{sh,c}}^0\Gamma(z) - \mathrm{d}\Gamma[z](L_{\mathrm{wh}}^0 z)$$

and $h_2^0(z)$ denotes the term in the Maclaurin expansion of

$$h^\varepsilon(z) = P_{\mathrm{sh,c}}(U^\varepsilon)^{-1}\Big(\tilde{A}^\varepsilon(z)z_\tau + \tilde{B}^\varepsilon(z)\Big)$$

which is homogeneous of degree two in z and does not depend upon ε. Using the analyticity of the functions f_3 and f_4 (see equation (6.3)) we can write

$$f_3^0(v) = \sum_{i+j+k=0}^{\infty} c_{ijk}v_1^i v_2^j v_3^k, \quad f_4^0(v) = \sum_{i+j+k=0}^{\infty} d_{ijk}v_1^i v_2^j v_3^k$$

with

$$c_{ijk} = \frac{1}{i!j!k!}\partial_1^i \partial_2^j \partial_3^k f_3^0(0), \quad d_{ijk} = \frac{1}{i!j!k!}\partial_1^i \partial_2^j \partial_3^k f_4^0(0)$$

and thus

$$\begin{aligned} A^0(v) &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & f_3^0(v) \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & c_{100}v_1 \\ 0 & 0 & 0 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & c_{200}v_1^2 + c_{020}v_2^2 + c_{011}v_2v_3 + c_{002}v_3^2 \\ 0 & 0 & 0 \end{pmatrix} + \mathcal{O}(|v|^3) \end{aligned} \tag{6.20}$$

as well as

$$\begin{aligned} B^0(v) &= \begin{pmatrix} v_2 \\ f_4^0(v) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} v_2 \\ d_{100}v_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ d_{200}v_1^2 + d_{020}v_2^2 + d_{011}v_2v_3 + d_{002}v_3^2 \\ 0 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 \\ d_{300}v_1^3 + d_{120}v_1v_2^2 + d_{102}v_1v_3^2 + d_{111}v_1v_2v_3 \\ 0 \end{pmatrix} + \mathcal{O}(|v|^4), \end{aligned} \tag{6.21}$$

where we have made use of the fact that

$$c_{ijk} = d_{ijk} = 0$$

if $j + k$ is odd because of $f_i^0(a, -b, -c) = f_i^0(a, b, c)$.

Observe that

$$z(\tau) = z_1 u \mathrm{e}^{\mathrm{i}\tau} + z_2 v \mathrm{e}^{\mathrm{i}\tau} + \overline{z}_1 \overline{u} \mathrm{e}^{-\mathrm{i}\tau} + \overline{z}_2 \overline{v} \mathrm{e}^{-\mathrm{i}\tau} = \begin{pmatrix} z_1 \mathrm{e}^{\mathrm{i}\tau} + \overline{z}_1 \mathrm{e}^{-\mathrm{i}\tau} \\ z_2 \mathrm{e}^{\mathrm{i}\tau} + \overline{z}_2 \mathrm{e}^{-\mathrm{i}\tau} \\ \mathrm{i}z_1 \mathrm{e}^{\mathrm{i}\tau} - \mathrm{i}\overline{z}_1 \mathrm{e}^{-\mathrm{i}\tau} \end{pmatrix}$$

and thus

$$z_\tau(\tau) = \begin{pmatrix} \mathrm{i}z_1 \mathrm{e}^{\mathrm{i}\tau} - \mathrm{i}\overline{z}_1 \mathrm{e}^{-\mathrm{i}\tau} \\ \mathrm{i}z_2 \mathrm{e}^{\mathrm{i}\tau} - \mathrm{i}\overline{z}_2 \mathrm{e}^{-\mathrm{i}\tau} \\ -z_1 \mathrm{e}^{\mathrm{i}\tau} - \overline{z}_1 \mathrm{e}^{-\mathrm{i}\tau} \end{pmatrix}.$$

We conclude that the quadratic terms $h_2^0(z)$ in the Maclaurin expansion of

$$h^0(z) = P_{\mathrm{sh,c}}\Big(A^0(z)z_\tau + B^0(z)\Big)$$

are

$$
\begin{aligned}
h_2^0(z) &= P_{\mathrm{sh,c}}\left(\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & c_{100}(z_1\mathrm{e}^{\mathrm{i}\tau}+\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau}) \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \mathrm{i}z_1\mathrm{e}^{\mathrm{i}\tau}-\mathrm{i}\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau} \\ \mathrm{i}z_2\mathrm{e}^{\mathrm{i}\tau}-\mathrm{i}\overline{z}_2\mathrm{e}^{-\mathrm{i}\tau} \\ -z_1\mathrm{e}^{\mathrm{i}\tau}-\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau} \end{pmatrix}\right. \\
&\qquad + \begin{pmatrix} 0 \\ d_{200}(z_1\mathrm{e}^{\mathrm{i}\tau}+\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau})^2 + d_{020}(z_2\mathrm{e}^{\mathrm{i}\tau}+\overline{z}_2\mathrm{e}^{-\mathrm{i}\tau})^2 \\ 0 \end{pmatrix} \\
&\qquad \left. + \begin{pmatrix} 0 \\ d_{011}(z_2\mathrm{e}^{\mathrm{i}\tau}+\overline{z}_2\mathrm{e}^{-\mathrm{i}\tau})(\mathrm{i}z_1\mathrm{e}^{\mathrm{i}\tau}-\mathrm{i}\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau}) + d_{002}(\mathrm{i}z_1\mathrm{e}^{\mathrm{i}\tau}-\mathrm{i}\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau})^2 \\ 0 \end{pmatrix}\right) \\
&= P_{\mathrm{sh,c}}\left(\begin{pmatrix} 0 \\ (d_{200}-c_{100})(z_1\mathrm{e}^{\mathrm{i}\tau}+\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau})^2 + d_{020}(z_2\mathrm{e}^{\mathrm{i}\tau}+\overline{z}_2\mathrm{e}^{-\mathrm{i}\tau})^2 \\ 0 \end{pmatrix}\right. \\
&\qquad \left. + \begin{pmatrix} 0 \\ d_{011}(z_2\mathrm{e}^{\mathrm{i}\tau}+\overline{z}_2\mathrm{e}^{-\mathrm{i}\tau})(\mathrm{i}z_1\mathrm{e}^{\mathrm{i}\tau}-\mathrm{i}\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau}) + d_{002}(\mathrm{i}z_1\mathrm{e}^{\mathrm{i}\tau}-\mathrm{i}\overline{z}_1\mathrm{e}^{-\mathrm{i}\tau})^2 \\ 0 \end{pmatrix}\right) \\
&= P_{\mathrm{sh,c}}\left(\begin{pmatrix} 0 \\ (d_{200}-c_{100})z_1^2 + d_{020}z_2^2 + \mathrm{i}d_{011}z_1z_2 - d_{002}z_1^2 \\ 0 \end{pmatrix}\mathrm{e}^{2\mathrm{i}\tau}\right. \\
&\qquad + \begin{pmatrix} 0 \\ 2(d_{200}-c_{100})|z_1|^2 + 2d_{020}|z_2|^2 - \mathrm{i}d_{011}z_2\overline{z}_1 + \mathrm{i}d_{011}z_1\overline{z}_2 + 2d_{002}|z_1|^2 \\ 0 \end{pmatrix} \\
&\qquad \left. + \begin{pmatrix} 0 \\ (d_{200}-c_{100})\overline{z}_1^2 + d_{020}\overline{z}_2^2 - \mathrm{i}d_{011}\overline{z}_1\overline{z}_2 - d_{002}\overline{z}_1^2 \\ 0 \end{pmatrix}\mathrm{e}^{-2\mathrm{i}\tau}\right) \\
&= \frac{\mathrm{i}\sqrt{2}}{\sqrt{3}}\Big((d_{200}-c_{100})z_1^2 + d_{020}z_2^2 + \mathrm{i}d_{011}z_1z_2 - d_{002}z_1^2\Big)e_{2,1}\mathrm{e}^{2\mathrm{i}\tau} \\
&\quad - \frac{\mathrm{i}\sqrt{2}}{\sqrt{3}}\Big((d_{200}-c_{100})z_1^2 + d_{020}z_2^2 + \mathrm{i}d_{011}z_1z_2 - d_{002}z_1^2\Big)e_{2,2}\mathrm{e}^{2\mathrm{i}\tau} \\
&\quad - \frac{1}{\sqrt{2}}\Big(2(d_{200}-c_{100})|z_1|^2 + 2d_{020}|z_2|^2 - \mathrm{i}d_{011}z_2\overline{z}_1 + \mathrm{i}d_{011}z_1\overline{z}_2 + 2d_{002}|z_1|^2\Big)e_{0,1} \\
&\quad + \frac{1}{\sqrt{2}}\Big(2(d_{200}-c_{100})|z_1|^2 + 2d_{020}|z_2|^2 - \mathrm{i}d_{011}z_2\overline{z}_1 + \mathrm{i}d_{011}z_1\overline{z}_2 + 2d_{002}|z_1|^2\Big)e_{0,2} \\
&\quad - \frac{\mathrm{i}\sqrt{2}}{\sqrt{3}}\Big((d_{200}-c_{100})\overline{z}_1^2 + d_{020}\overline{z}_2^2 - \mathrm{i}d_{011}\overline{z}_1\overline{z}_2 - d_{002}\overline{z}_1^2\Big)\overline{e}_{2,1}\mathrm{e}^{-2\mathrm{i}\tau} \\
&\quad + \frac{\mathrm{i}\sqrt{2}}{\sqrt{3}}\Big((d_{200}-c_{100})\overline{z}_1^2 + d_{020}\overline{z}_2^2 - \mathrm{i}d_{011}\overline{z}_1\overline{z}_2 - d_{002}\overline{z}_1^2\Big)\overline{e}_{2,2}\mathrm{e}^{-2\mathrm{i}\tau},
\end{aligned} \tag{6.22}
$$

so that in particular $\Gamma_{m,ijkl,r} = 0$ if $m \neq i + j - k - l$.

It follows from the formula

$$\mathcal{L}(z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r} \mathrm{e}^{\mathrm{i}m\tau}) = \lambda_{m,r} z_1^i z_2^j \overline{z}_1^k \overline{z}_2^l e_{m,r} \mathrm{e}^{\mathrm{i}m\tau} - i z_1^{i-1} z_2^{j+1} \overline{z}_1^k \overline{z}_2^l e_{m,r} \mathrm{e}^{\mathrm{i}m\tau} \\ - k z_1^i z_2^j \overline{z}_1^{k-1} \overline{z}_2^{l+1} e_{m,r} \mathrm{e}^{\mathrm{i}m\tau} \tag{6.23}$$

that

$$\mathcal{L}(\Gamma_{2,2000,r} z_1^2 e_{2,r} \mathrm{e}^{2\mathrm{i}\tau}) = \Gamma_{2,2000,r} \lambda_{2,r} z_1^2 e_{2,r} \mathrm{e}^{2\mathrm{i}\tau} - 2\Gamma_{2,2000,r} z_1 z_2 e_{2,r} \mathrm{e}^{2\mathrm{i}\tau},$$
$$\mathcal{L}(\Gamma_{0,1010,r} z_1 \overline{z}_1 e_{0,r}) = \Gamma_{0,1010,r} \lambda_{0,r} z_1 \overline{z}_1 e_{0,r} - \Gamma_{0,1010,r} z_2 \overline{z}_1 e_{0,r} - \Gamma_{0,1010,r} z_1 \overline{z}_2 e_{0,r},$$

and equation (6.23) shows that these are the only terms in the polynomial $(\mathcal{L}\Gamma)(z)$ which contain respectively the monomials z_1^2 and $z_1\overline{z}_1$. By comparing coefficients with the formula for $h_2^0(z)$ given in equation (6.22) we find that

$$-\mathrm{i}\sqrt{3}\Gamma_{2,2000,1} = \frac{\mathrm{i}\sqrt{2}}{\sqrt{3}}(d_{200} - c_{100} - d_{002}), \quad \mathrm{i}\sqrt{3}\Gamma_{2,2000,2} = -\frac{\mathrm{i}\sqrt{2}}{\sqrt{3}}(d_{200} - c_{100} - d_{002}),$$

and

$$-\Gamma_{0,1010,1} = -\sqrt{2}(d_{200} - c_{100} + d_{002}), \quad \Gamma_{0,1010,2} = \sqrt{2}(d_{200} - c_{100} + d_{002}),$$

so that

$$\Gamma_{2,2000,1} = \Gamma_{2,2000,2} = \frac{\sqrt{2}}{3}(c_{100} - d_{200} + d_{002}), \quad \Gamma_{0,1010,1} = \Gamma_{0,1010,2} = \sqrt{2}(d_{200} - c_{100} + d_{002}).$$

The next step is to compute the terms in the formula (6.19) for $\check{C}$. Using the expansions (6.20), (6.21) for A^0 and B^0 we obtain

$$\begin{aligned} 3f^{0;3,0}\left[\{u\mathrm{e}^{\mathrm{i}\tau}\}^{(2)}, \overline{u}\mathrm{e}^{-\mathrm{i}\tau}\right] &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & c_{200}\mathrm{e}^{2\mathrm{i}\tau} - c_{002}\mathrm{e}^{2\mathrm{i}\tau} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\mathrm{i}\mathrm{e}^{-\mathrm{i}\tau} \\ 0 \\ -\mathrm{e}^{-\mathrm{i}\tau} \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2c_{200} + 2c_{002} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathrm{i}\mathrm{e}^{\mathrm{i}\tau} \\ 0 \\ -\mathrm{e}^{\mathrm{i}\tau} \end{pmatrix} + \begin{pmatrix} 0 \\ 3d_{300}\mathrm{e}^{\mathrm{i}\tau} + d_{102}\mathrm{e}^{\mathrm{i}\tau} \\ 0 \end{pmatrix} \\ &= (-3c_{200} - c_{002} + 3d_{300} + d_{102})v\mathrm{e}^{\mathrm{i}\tau}, \end{aligned}$$

$$
\begin{aligned}
&f^{0;1,1}\Big[\overline{u}\mathrm{e}^{-\mathrm{i}\tau}, -\Gamma_{2,2000,1}e_{2,1}\mathrm{e}^{2\mathrm{i}\tau} - \Gamma_{2,2000,2}e_{2,2}\mathrm{e}^{2\mathrm{i}\tau}\Big] \\
&= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}}c_{100}\Gamma_{2,2000,1}e^{2\mathrm{i}\tau} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\mathrm{i}\mathrm{e}^{-\mathrm{i}\tau} \\ 0 \\ -\mathrm{e}^{-\mathrm{i}\tau} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & c_{100}e^{-\mathrm{i}\tau} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\frac{\mathrm{i}}{\sqrt{2}}\Gamma_{2,2000,1}e^{2\mathrm{i}\tau} \\ 0 \\ 2\sqrt{2}\Gamma_{2,2000,1}e^{2\mathrm{i}\tau} \end{pmatrix} \\
&\qquad + 2\begin{pmatrix} 0 \\ -\frac{1}{\sqrt{2}}d_{200}\Gamma_{2,2000,1}\mathrm{e}^{\mathrm{i}\tau} - \sqrt{2}d_{002}\Gamma_{2,2000,1}\mathrm{e}^{\mathrm{i}\tau} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ \frac{5}{\sqrt{2}}c_{100}\Gamma_{2,2000,1}e^{\mathrm{i}\tau} - \sqrt{2}d_{200}\Gamma_{2,2000,1}\mathrm{e}^{\mathrm{i}\tau} - 2\sqrt{2}d_{002}\Gamma_{2,2000,1}\mathrm{e}^{\mathrm{i}\tau} \\ 0 \end{pmatrix} \\
&= \Big(\frac{5}{3}c_{100}^2 - \frac{7}{3}c_{100}d_{200} + \frac{1}{3}c_{100}d_{002} + \frac{2}{3}d_{200}^2 + \frac{2}{3}d_{200}d_{002} - \frac{4}{3}d_{002}^2\Big)v\mathrm{e}^{\mathrm{i}\tau}
\end{aligned}
$$

and

$$
\begin{aligned}
&f^{0;1,1}\Big[u\mathrm{e}^{\mathrm{i}\tau}, -\Gamma_{0,1010,1}e_{0,1} - \Gamma_{0,1010,2}e_{0,2}\Big] \\
&= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\sqrt{2}c_{100}\Gamma_{0,1010,1} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathrm{i}\mathrm{e}^{\mathrm{i}\tau} \\ 0 \\ -\mathrm{e}^{\mathrm{i}\tau} \end{pmatrix} + 2\begin{pmatrix} 0 \\ -\sqrt{2}d_{200}\Gamma_{0,1010,1}\mathrm{e}^{\mathrm{i}\tau} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ \sqrt{2}c_{100}\Gamma_{0,1010,1}\mathrm{e}^{\mathrm{i}\tau} - 2\sqrt{2}d_{200}\Gamma_{0,1010,1}\mathrm{e}^{\mathrm{i}\tau} \\ 0 \end{pmatrix} \\
&= (6c_{100}d_{200} - 2c_{100}^2 + 2c_{100}d_{002} - 4d_{200}^2 - 4d_{200}d_{002})v\mathrm{e}^{\mathrm{i}\tau}.
\end{aligned}
$$

The above calculations finally yield

$$
\begin{aligned}
\check{C} &= 3c_{200} + c_{002} - 3d_{300} - d_{102} - \frac{5}{3}c_{100}^2 + \frac{7}{3}c_{100}d_{200} - \frac{1}{3}c_{100}d_{002} - \frac{2}{3}d_{200}^2 - \frac{2}{3}d_{200}d_{002} \\
&\quad + \frac{4}{3}d_{002}^2 - 6c_{100}d_{200} + 2c_{100}^2 - 2c_{100}d_{002} + 4d_{200}^2 + 4d_{200}d_{002} \\
&= 3c_{200} + c_{002} - 3d_{300} - d_{102} + \frac{1}{3}c_{100}^2 - \frac{11}{3}c_{100}d_{200} - \frac{7}{3}c_{100}d_{002} + \frac{10}{3}d_{200}^2 \\
&\quad + \frac{10}{3}d_{200}d_{002} + \frac{4}{3}d_{002}^2,
\end{aligned} \tag{6.24}
$$

which simplifies to

$$
\check{C} = -3d_{300} + \frac{10}{3}d_{200}^2 \tag{6.25}
$$

in the special case.

In terms of the nonlinearities f_1 and f_2 in equation (6.1) we deduce from (6.24) the formula

$$\check{C} = \frac{(\partial_2^2 f_2(0))^2}{3} + \frac{(\frac{5}{2}\partial_1^2 f_2(0) - \partial_1 f_1(0))\partial_2^2 f_2(0)}{3} + \frac{5(\partial_1^2 f_2(0))^2}{6} \tag{6.26}$$

$$+ \frac{\partial_1^3 f_2(0)}{2} + \frac{\partial_1\partial_2^2 f_2(0)}{6} + \frac{\partial_2^2 f_1(0)}{3}, \tag{6.27}$$

so that the hypothesis of Theorem 6.1 is satisfied whenever

$$\frac{\partial_1 f_1(0)\partial_2^2 f_2(0)}{3} < \frac{(\partial_2^2 f_2(0))^2}{3} + \frac{5\partial_1^2 f_2(0)\partial_2^2 f_2(0)}{6} + \frac{5(\partial_1^2 f_2(0))^2}{6} + \frac{\partial_1^3 f_2(0)}{2} + \frac{\partial_1\partial_2^2 f_2(0)}{6} + \frac{\partial_2^2 f_1(0)}{3}.$$

In the special case, where equation (6.1) reduces to the equation

$$\phi_{tt} = \phi_{xx} - \phi + f_2(\phi)$$

with an analytic function $f_2 \colon \mathbb{R} \to \mathbb{R}$ satisfying $f_2(0) = 0$, $\mathrm{d}f_2[0] = 0$, we use the simplified formula (6.25) to obtain

$$\check{C} = \frac{f_2'''(0)}{2} + \frac{5}{6} f_2''(0)^2. \tag{6.28}$$

The hypothesis of Theorem 6.1 is thus satisfied if

$$\frac{5}{3} f_2''(0)^2 > -f_2'''(0);$$

in particular if $f_2'''(0) > 0$.

Example 6.6 (Groves & Schneider [8])**.** Consider the sine-Gordon equation

$$\phi_{tt} = \phi_{xx} - \sin(\phi)$$

and the ϕ^4 equation

$$\phi_{tt} = \phi_{xx} - \phi + \phi^3.$$

We compute respectively $\check{C} = \frac{1}{2}$ and $\check{C} = 3$, so that both equations satisfy the hypothesis of Theorem 6.1.

Bibliography

[1] S. Benzoni-Gavage and D. Serre: *Multidimensional Hyperbolic Partial Differential Equations.* Clarendon Press, 2007.

[2] B. Birnir, H. McKean, and A. Weinstein: *The rigidity of sine-Gordon breathers.* Commun. Pure Appl. Math, 47:1043–1051, 1994.

[3] H. Brezis: *Functional analysis, Sobolev spaces and partial differential equations.* Springer, New York, 2010.

[4] B. Buffoni and J. Toland: *Analytic Theory of Global Bifurcation.* Princeton University Press, 2003.

[5] O. Christensen: *An Introduction to Frames and Riesz Bases.* Birkhäuser, 2016.

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

[7] A. Friedman: *Partial Differential Equations.* Dover Publications Inc., Mineola (New York), 2008.

[8] M. D. Groves and G. Schneider: *Modulating pulse solutions for a class of nonlinear wave equations.* Commun. Math. Phys., 219:489–522, 2001.

[9] M. D. Groves and G. Schneider: *Modulating pulse solutions for quasilinear wave equations.* J. Diff. Eqs., 219:221–258, 2005.

[10] M. D. Groves and G. Schneider: *Modulating pulse solutions to quadratic quasilinear wave equations over exponentially long length scales.* Commun. Math. Phys., 278:567–625, 2008.

[11] N. Gutheil: *Das Cauchy-Problem für symmetrische hyperbolische Gleichungssysteme.* Bachelor thesis, Saarland University, 2016.

[12] N. Gutheil: *Pulse solutions to a class of quasilinear symmetric hyperbolic evolutionary equations over exponentially long time scales.* Master thesis, Saarland University, 2019.

[13] B. Hewer: *Generalised pulse solutions for a class of quasilinear evolutionary equations with applications to water waves.* PhD thesis, Saarland University, 2023.

[14] G. Iooss and E. Lombardi: *Polynomial normal forms with exponentially small remainder for analytic vector fields.* J. Diff. Eqs., 212:1–61, 2005.

[15] T. Kato: *Perturbation Theory for Linear Operators.* Springer, 1980.

[16] A. Kelley: *The stable, center-stable, center, center-unstable, unstable manifolds.* J. Diff. Eqs., 3:546–570, 1967.

[17] K. Kirchgässner: *Nonlinear resonant surface waves and homoclinic bifurcation.* Adv. Appl. Mech., 26:135–181, 1988.

[18] R. Kress: *Linear Integral Equations.* Springer, New York, 1999.

[19] C. D. Sogge: *Lectures on Non-linear Wave Equations*, volume 2 of *Monographs in Analysis.* International Press, Boston, 2008.