ANALYSIS & PDE

Volume 9 No. 7 2016

msp

Analysis & PDE
msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2016 is US $235/year for the electronic version, and $430/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 9, No. 7, 2016
dx.doi.org/10.2140/apde.2016.9.1523

THE FINAL-STATE PROBLEM FOR THE CUBIC-QUINTIC NLS WITH NONVANISHING BOUNDARY CONDITIONS

ROWAN KILLIP, JASON MURPHY AND MONICA VISAN

We construct solutions with prescribed scattering state to the cubic-quintic NLS

$$(i\partial_t+\Delta)\psi=\alpha_1\psi-\alpha_3|\psi|^2\psi+\alpha_5|\psi|^4\psi$$

in three spatial dimensions in the class of solutions with $|\psi(x)|\to c>0$ as $|x|\to\infty$. This models disturbances in an infinite expanse of (quantum) fluid in its quiescent state — the limiting modulus c corresponds to a local minimum in the energy density.

Our arguments build on work of Gustafson, Nakanishi, and Tsai on the (defocusing) Gross–Pitaevskii equation. The presence of an energy-critical nonlinearity and changes in the geometry of the energy functional add several new complexities. One new ingredient in our argument is a demonstration that solutions of such (perturbed) energy-critical equations exhibit continuous dependence on the initial data with respect to the *weak* topology on H^1_x.

1. Introduction

We study the cubic-quintic nonlinear Schrödinger equation (NLS) with nonvanishing boundary conditions in three space dimensions:

$$\begin{cases}(i\partial_t+\Delta)\psi=\alpha_1\psi-\alpha_3|\psi|^2\psi+\alpha_5|\psi|^4\psi, & (t,x)\in\mathbb{R}\times\mathbb{R}^3,\\ \psi(0)=\psi_0.\end{cases}\tag{1-1}$$

We consider parameters $\alpha_1,\alpha_3,\alpha_5>0$ so that $\alpha_3^2-4\alpha_1\alpha_5>0$, which guarantees that the polynomial $\alpha_1-\alpha_3x+\alpha_5x^2$ has two distinct positive roots $r_0^2>r_1^2>0$. The boundary condition is given by

$$\lim_{|x|\to\infty}|\psi(t,x)|=r_0.\tag{1-2}$$

The choice of the larger root guarantees the energetic stability of the constant solution; it constitutes a local minimum of the energy functional (1-7).

Equation (1-1) appears in a great variety of physical problems. It is a model in superfluidity [Ginzburg and Pitaevskiĭ 1958; Ginzburg and Sobyanin 1976], descriptions of bosons [Barashenkov et al. 1989] and of defectons [Pushkarov and Kojnov 1978], the theory of ferromagnetic and molecular chains [Pushkarov and Primatarova 1984; 1986], and in nuclear hydrodynamics [Kartavenko 1984]. The popularity of this model can be explained by its simplicity combined with the fact that it captures an important phenomenology: the constituents of most fluids experience an attractive interaction at low densities and a repulsion at high

MSC2010: 35Q55.
Keywords: final-state problem, wave operators, cubic-quintic NLS, nonvanishing boundary conditions.

densities. The recent paper [Killip et al. 2014] focuses on the analogous problem with data decaying at infinity, which constitutes a model for the dynamics of a finite body of fluid; the model (1-1) describes the behavior of a localized disturbance in an infinite expanse of fluid that is otherwise quiescent.

By rescaling both space-time and the values of ψ, it suffices to consider the case $r_0^2 = 1$ and $\alpha_5 = 1$. This leaves one free parameter

$$\gamma := 1 - r_1^2 \in (0, 1), \tag{1-3}$$

in terms of which equation (1-1) becomes

$$\begin{cases} (i\partial_t + \Delta)\psi = (|\psi|^2 - 1)(|\psi|^2 - 1 + \gamma)\psi, \\ \psi(0) = \psi_0, \end{cases} \tag{1-4}$$

with the boundary condition

$$\lim_{|x|\to\infty} \psi(t, x) = 1. \tag{1-5}$$

As discussed in [Gérard 2006] (albeit in the context of the Gross–Pitaevskii equation), finite energy functions obeying (1-2) have a unique limiting phase as $|x| \to \infty$, which we can normalize to be zero, yielding (1-5). Furthermore, the dynamics of (1-1) preserve the value of this phase, so that the boundary condition is independent of time, as well. This breaks the gauge invariance of (1-1) and prohibits using a phase factor to remove the linear term in this equation. The presence of the linear term leads to weaker dispersion at low frequencies, which presents a key challenge in understanding the long-time behavior of solutions.

We are interested in perturbations of the constant solution $\psi \equiv 1$, and thus it is natural to introduce the function $u = u_1 + iu_2$ defined via $\psi = 1 + u$. Using (1-4), we arrive at the following equation for u:

$$\begin{cases} (i\partial_t + \Delta)u - 2\gamma u_1 = N(u), \\ u(0) = u_0, \end{cases} \tag{1-6}$$

where $N(u) = \sum_{j=2}^5 N_j(u)$, with

$$\begin{aligned} N_2(u) &= (3\gamma + 4)u_1^2 + \gamma u_2^2 + 2i\gamma u_1 u_2, \\ N_3(u) &= (\gamma + 8)u_1^3 + (\gamma + 4)u_1 u_2^2 + i[(\gamma + 4)u_1^2 u_2 + \gamma u_2^3], \\ N_4(u) &= 5u_1^4 + 6u_1^2 u_2^2 + u_2^4 + i[4u_1^3 u_2 + 4u_1 u_2^3], \\ N_5(u) &= |u|^4 u = u_1^5 + 2u_1^3 u_2^2 + u_2^4 u_1 + i[u_1^4 u_2 + 2u_1^2 u_2^3 + u_2^5]. \end{aligned}$$

The Hamiltonian for (1-4) is given by

$$E(\psi) = \frac{1}{2}\int_{\mathbb{R}^3} |\nabla\psi|^2 \, dx + \frac{\gamma}{4}\int_{\mathbb{R}^3} (|\psi|^2 - 1)^2 \, dx + \frac{1}{6}\int_{\mathbb{R}^3} (|\psi|^2 - 1)^3 \, dx. \tag{1-7}$$

Introducing the notation

$$q(u) := |\psi|^2 - 1 = 2u_1 + |u|^2,$$

we may write

$$2\gamma u_1 + N(u) = [\gamma q(u) + q(u)^2](1 + u)$$

and

$$E(1 + u) = \frac{1}{2}\int_{\mathbb{R}^3} |\nabla u|^2 \, dx + \frac{\gamma}{4}\int_{\mathbb{R}^3} q(u)^2 \, dx + \frac{1}{6}\int_{\mathbb{R}^3} q(u)^3 \, dx. \tag{1-8}$$

In the sequel we will write $E(u)$ for $E(1+u)$; when there is no risk of confusion we will simply write $q(u)=q$. Note that q represents density fluctuations relative to the constant background. The quantity $\int q(t,x)\,dx$, which represents the total surplus/deficit of matter relative to the constant background, is conserved in time; in this work we do not rely on this conservation law.

Well-posedness in the energy space. We define the energy space for (1-6) to be

$$\mathcal{E} := \{u \in \dot{H}^1_x(\mathbb{R}^3) : q(u) \in L^2_x(\mathbb{R}^3)\}, \tag{1-9}$$

with associated metric

$$[d_{\mathcal{E}}(u,v)]^2 := \|u-v\|^2_{\dot{H}^1_x} + \|q(u)-q(v)\|^2_{L^2_x},$$

and we let $\|u\|_{\mathcal{E}} := d_{\mathcal{E}}(u,0)$ denote the energy-norm.

To justify our choice of energy space, we first note that functions with finite energy-norm have finite energy. Indeed, using Sobolev embedding and the fact that $(L^3_x + L^6_x) \cap L^2_x \subset L^3_x$, it is not hard to see that if $u \in \mathcal{E}$ then $q(u) \in L^3_x$, and so $|E(u)| < \infty$. In fact,

$$|E(u)| \lesssim \|u\|^2_{\mathcal{E}} + \|u\|^3_{\mathcal{E}}.$$

On the other hand, in Lemma 3.1 we will show that for $\gamma \in \left[\frac{2}{3}, 1\right)$, functions with finite energy have finite energy-norm. When $\gamma \in \left(0, \frac{2}{3}\right)$, the energy is not coercive unless we impose an additional smallness assumption (see Lemma 3.2).

When the energy is not coercive, there is no unique candidate for the name "energy space". The authors of [Killip et al. 2012] worked with the following notion of energy space:

$$\mathcal{E}_{\text{KOPV}} := \{u \in \dot{H}^1_x(\mathbb{R}^3) \cap L^4_x(\mathbb{R}^3) : \operatorname{Re} u \in L^2_x(\mathbb{R}^3)\}.$$

Note that $\mathcal{E}_{\text{KOPV}} \subset \mathcal{E}$. In the same work, they also proved that (1-6) is globally well-posed for data $u_0 \in \mathcal{E}_{\text{KOPV}}$; in particular, solutions are unconditionally unique in $C(\mathbb{R}; \mathcal{E}_{\text{KOPV}})$.

In Section 3, we prove global well-posedness and unconditional uniqueness for (1-6) in the energy space $\mathcal{E}$ (see Theorem 3.3). As in [Killip et al. 2012; Tao et al. 2007; Zhang 2006], our approach is to regard the equation as a perturbation of the defocusing energy-critical NLS

$$(i\partial_t + \Delta)u = |u|^4 u, \tag{1-10}$$

which was proven to be globally well-posed, first in the radial case and then for general data in the celebrated papers [Bourgain 1999; Colliander et al. 2008]. Proving well-posedness for a Schrödinger equation in three dimensions that contains a quintic nonlinearity requires control over the $\dot{H}^1_x$-norm of the solution. As the energy (1-8) is not necessarily coercive for $\gamma \in \left(0, \frac{2}{3}\right)$, conservation of the Hamiltonian does not supply the requisite a priori bound. To resolve this issue we will require that both the energy and the kinetic energy of the data are small when $\gamma \in \left(0, \frac{2}{3}\right)$.

Statement of the main result. The stability of the equilibrium solution $\psi \equiv 1$ to (1-4) is equivalent to the small-data problem for (1-6). In this direction, there are two natural problems to consider, namely, the initial-value and the final-state problems for (1-6). For the former, the question is whether small and localized initial data lead to solutions that are global and decay as $|t| \to \infty$. For the latter, the question is whether one can construct a solution that scatters to a prescribed asymptotic state. In this paper we prove two results related to the final-state problem. We will address the initial-value problem in a forthcoming work.

To fit (1-6) into the standard framework of dispersive equations it is convenient to diagonalize the equation. Setting

$$U = |\nabla|\langle\nabla\rangle^{-1} \quad \text{and} \quad H = |\nabla|\langle\nabla\rangle, \quad \text{with } \langle\nabla\rangle := \sqrt{2\gamma - \Delta} \text{ and } |\nabla| = (-\Delta)^{\frac{1}{2}},$$

we arrive at the following equation for $v := Vu := u_1 + iUu_2$:

$$\begin{cases} (i\partial_t - H)v = N_v(u) := U\,\mathrm{Re}[N(u)] + i\,\mathrm{Im}[N(u)], \\ v(0) = Vu_0. \end{cases} \tag{1-11}$$

Note that $u_{\mathrm{lin}}(t) := V^{-1}e^{-itH}Vu_+$ solves the equation

$$(i\partial_t + \Delta)u_{\mathrm{lin}} - 2\gamma\,\mathrm{Re}\,u_{\mathrm{lin}} = 0 \quad \text{with } u_{\mathrm{lin}}(0) = u_+; \tag{1-12}$$

this is the linearization of (1-6) about $u = 0$.

Our main result in this paper is the following theorem:

Theorem 1.1. *Suppose $\gamma \in \big[\frac{2}{3}, 1\big)$. For any $u_+ \in H^1_{\mathrm{real}} + i\dot{H}^1_{\mathrm{real}}$, there exists a global solution $u \in C(\mathbb{R}; \mathcal{E})$ to (1-6) such that*

$$\lim_{t\to\infty} \|u(t) - u_{\mathrm{lin}}(t)\|_{\dot{H}^1_x} = 0, \tag{1-13}$$

where $u_{\mathrm{lin}}(t) := V^{-1}e^{-itH}Vu_+$. Moreover, we have modified asymptotics in the energy space, in the sense that this same solution u obeys

$$\lim_{t\to\infty} d_{\mathcal{E}}\big(u(t), u_{\mathrm{lin}}(t) - \gamma\langle\nabla\rangle^{-2}|u_{\mathrm{lin}}(t)|^2\big) = 0. \tag{1-14}$$

In the case $\gamma \in \big(0, \frac{2}{3}\big)$, both conclusions still hold if additionally $\|u_+\|_{H^1_{\mathrm{real}} + i\dot{H}^1_{\mathrm{real}}}$ is sufficiently small.

Remark 1.2. The hypotheses on u_+ are not sufficient to guarantee that $u_{\mathrm{lin}}(t) \in \mathcal{E}$ at any time t; correspondingly, one cannot hope to say that u is close to u_{lin} in the energy space. Nonetheless, (1-13) does show that the modification in (1-14) only plays a role at very low frequencies. Indeed, simple computations show that the modification can be omitted, for example, when u_+ is a Schwartz function.

We do not guarantee uniqueness of the solution u in Theorem 1.1. Later, we will show uniqueness within a restricted class of solutions u for suitable scattering states u_+; see Theorem 1.4 and Corollary 1.7 below.

Discussion of relevant past results. To give proper context to our work, we need to discuss prior work of Gustafson, Nakanishi, and Tsai [Gustafson et al. 2006; 2007; 2009] on the Gross–Pitaevskii equation

$$\begin{cases} (i\partial_t + \Delta)\psi = (|\psi|^2 - 1)\psi, \\ \psi(0) = \psi_0, \\ \lim\limits_{|x|\to\infty} \psi(t,x) = 1. \end{cases} \tag{1-15}$$

Note that unlike in (1-4), the cubic nonlinearity here is defocusing. Writing $\psi = 1 + u$, this equation preserves the energy

$$E_{\mathrm{GP}}(u) = \frac{1}{2}\int_{\mathbb{R}^3} |\nabla u|^2\,dx + \frac{1}{4}\int_{\mathbb{R}^3} q(u)^2\,dx. \tag{1-16}$$

In contrast to (1-8), this energy density is lacking the sign-indefinite $q(u)^3$-term. Correspondingly, the energy is coercive and the nonlinearity is energy-subcritical.

The final-state problem for the Gross–Pitaevskii equation was addressed by Gustafson et al. [2007; 2009] in two and three dimensions and in [Gustafson et al. 2006] in higher dimensions. They also considered the initial-value problem in dimensions $d \geq 3$ in [Gustafson et al. 2006; 2009].

The jumping-off point for Theorem 1.1 is an analogous result appearing in [Gustafson et al. 2009] for the Gross–Pitaevskii equation, which in turn builds on earlier work of Nakanishi [2001] on the (gauge-invariant) NLS. As our strategy is modeled closely on his, it is worth discussing in detail the following result:

Theorem 1.3 [Nakanishi 2001]. *Given $u_+ \in H^1_x(\mathbb{R}^3)$ and $\frac{2}{3} < p < \frac{4}{3}$, there is a solution to*

$$(i\partial_t + \Delta)u = |u|^p u \tag{1-17}$$

that obeys $e^{-it\Delta}u(t) \to u_+$ in $H^1_x(\mathbb{R}^3)$.

Sketch of proof. Nakanishi first defines solutions u^T to (1-17) with $u^T(T) = e^{iT\Delta}u_+$. As the problem is L^2_x-subcritical, these solutions are easily seen to be global with uniformly bounded H^1_x-norm (even in the focusing case).

By writing (1-17) in Duhamel form and exploiting the dispersive estimate (2-2), it is not difficult to show that for each $\phi \in C_c^\infty(\mathbb{R}^3)$, the collection of functions

$$\bigl\{t \mapsto \langle \phi, e^{-it\Delta}u^T(t)\rangle : T \in \mathbb{R}\bigr\} \tag{1-18}$$

forms an equicontinuous family on a compactification $\mathbb{R} \cup \{\pm\infty\}$ of the real line. In particular, each such function has limiting values as $t \to \pm\infty$. Applying Arzelà–Ascoli and the Cantor diagonal argument (H^1_x is separable), one can find a sequence $T_n \to \infty$ and a function $u^\infty \in L^\infty_t H^1_x$ so that

$$e^{-it\Delta}u^{T_n}(t) \rightharpoonup e^{-it\Delta}u^\infty(t) \quad \text{weakly in } H^1_x \text{ for each } t \in \mathbb{R}.$$

This construction guarantees that u^∞ has two further properties. First, the function $t \mapsto e^{-it\Delta}u^\infty(t)$ is weakly H^1_x-continuous on $\mathbb{R} \cup \{\pm\infty\}$, that is, when H^1_x is endowed with the weak topology. Secondly, for any $\phi \in C_c^\infty(\mathbb{R}^3)$,

$$\langle \phi, e^{-it\Delta}u^{T_n}(t)\rangle \to \langle \phi, e^{-it\Delta}u^\infty(t)\rangle \quad \text{as } n \to \infty, \text{ uniformly in } t \in \mathbb{R}.$$

Using these properties it is elementary to verify that $e^{-it\Delta}u^\infty(t) \rightharpoonup u_+$ as $t \to \infty$. This leaves two obligations: firstly, one must show that u^∞ is actually a solution to (1-17) and secondly, one must upgrade weak convergence to norm convergence.

Due to the H^1_x-subcriticality of the nonlinearity, the Rellich–Kondrashov theorem allows one to show that u^∞ is a weak solution to (1-17). For this problem, weak solutions with values in H^1_x are necessarily strong solutions and so we may conclude that u^∞ is a solution to (1-17).

Lastly, to upgrade weak convergence to strong convergence, one exploits conservation of mass and energy and the Radon–Riesz theorem. For example, one may argue as follows: The quantity

$$F(u) := \int_{\mathbb{R}^3} |\nabla u|^2 + \frac{2}{p+2}|u|^{p+2} + |u|^2\,dx \tag{1-19}$$

is conserved under the flow (1-17). Exploiting this, dispersion of the linear flow, and weak lower-semicontinuity of norms, we deduce that

$$\overline{\lim_{t\to\infty}}\,\|e^{-it\Delta}u^\infty(t)\|^2_{H^1_x} \le F(u^\infty) \le \varliminf_{n\to\infty} F(u^{T_n}(0)) = \varliminf_{n\to\infty} F(u^{T_n}(T_n)) = \|u_+\|^2_{H^1_x}.$$

Given that $e^{-it\Delta}u^\infty(t) \rightharpoonup u_+$, we deduce that $e^{-it\Delta}u^\infty(t) \to u_+$ in H^1_x. □

In order to adapt this beautiful argument to the Gross–Pitaevskii setting, the authors of [Gustafson et al. 2009] had to overcome two significant obstacles: (i) One needs to make the (conserved) energy (1-16) associated to (1-15) play the role of F in the argument above. It is far from obvious that this has the requisite convexity. (ii) The simple arguments used to prove equicontinuity of the family (1-18) no longer work. This failure stems from lower-power terms in the nonlinearity combined with the fact that energy conservation gives poor a priori spatial decay of solutions; while it guarantees $q(u)\in L^2_x$, it only yields $u_1\in L^3_x$ and no better than $u_2\in L^6_x$. This is not sufficient decay to allow direct access to any of the integrable-in-time dispersive estimates obeyed by the propagator.

The key to obtaining equicontinuity of the analogue of the family (1-18) in the Gross–Pitaevskii setting is to exploit certain nonresonant structures in the nonlinearity that allow one to integrate by parts in time. In implementing this approach, one sees that it is necessary to exhibit such nonresonance in both the quadratic and cubic terms of the nonlinearity. Such a brute force attack is rather messy. The burden can be significantly reduced by using test functions whose Fourier support excludes the origin. We will demonstrate this (primarily expository) improvement over the arguments from [Gustafson et al. 2009] in the proof of Proposition 6.2 below. One particular virtue of this approach is that it makes clear from the start that the argument is inherently completely immune to the poor dispersion manifested by the propagator (2-4) at low frequencies.

In [Gustafson et al. 2009], the authors exploit the quadratic nonresonant structure in a more elegant way through the use of a normal form transformation

$$z = [u_1 + (2-\Delta)^{-1}|u|^2] + i\sqrt{-\Delta/(2-\Delta)}\,u_2. \tag{1-20}$$

In this work they also observe (and then utilize) the further nonresonant structure at the cubic level (akin to (6-30)). There is some flexibility in the choice of normal form that witnesses the requisite nonresonance; however, the particular one employed in [Gustafson et al. 2009] has the dramatic additional benefit of overcoming obstacle (i) described above. The necessary convexity of the energy functional becomes clearer when written in their new variables: with u and z related by (1-20),

$$E_{\mathrm{GP}}(u) = \tfrac{1}{2}\big\|\sqrt{2-\Delta}\,z\big\|_{L_x^2}^2 + \tfrac{1}{4}\big\|\sqrt{-\Delta/(2-\Delta)}\,|u|^2\big\|_{L_x^2}^2. \tag{1-21}$$

The virtue of this identity is best understood in the context of (6-8). Because the right-most term in (1-21) is nonnegative, combining (1-21) with (6-8) yields

$$\overline{\lim_{t\to\infty}}\ \tfrac{1}{2}\big\|\sqrt{2-\Delta}\,z(t)\big\|_{L_x^2}^2 = \tfrac{1}{2}\big\|\sqrt{2-\Delta}\,z_+\big\|_{L_x^2}^2,$$

where $z(t)$ and z_+ represent a particular solution and its putative scattering state, both in terms of the normal form variable. This is just what is needed as input for the Radon–Riesz theorem.

Discussion of the main result. In order to prove Theorem 1.1 we will need to capitalize on all of the ideas introduced in [Gustafson et al. 2009] to prove the analogous result for the Gross–Pitaevskii equation. In particular, we will exploit a normal form transformation modeled closely on (1-20), namely,

$$z = M(u) := \big[u_1 + \gamma(2\gamma-\Delta)^{-1}|u|^2\big] + i\sqrt{-\Delta/(2\gamma-\Delta)}\,u_2. \tag{1-22}$$

However, several new difficulties arise above and beyond those overcome in [Gustafson et al. 2009].
(i) The first group of new difficulties is associated to the presence of energy-critical terms in the nonlinearity.
(ii) The second group of difficulties stems from the shape of the energy functional.

(i) We begin by discussing the difficulties that arise from the energy-critical terms. As discussed earlier in the introduction, we already need to give consideration to the energy-critical terms in the proof of Theorem 3.3, which states that (1-6) admits global solutions for initial data in the energy space $\mathcal{E}$. A more significant challenge involves establishing a form of well-posedness with respect to the weak $\dot{H}_x^1$ topology (see Theorem 4.1), as we will now explain.

In the argument of Nakanishi described above, it was used that weak limits (in the H_x^1 topology pointwise in time) of strong solutions to (1-17) are themselves strong solutions. In the subcritical case, one sees relatively easily that such limits are weak solutions (via Rellich–Kondrashov) and can then exploit earlier work (see [Cazenave 2003, Chapters 3–4]) showing that weak solutions with values in H_x^1 are strong solutions. In particular, solutions converging weakly to zero (in H_x^1) by concentrating will actually converge to zero in the space-time norms used to construct such solutions. In a similar way, we see that increasingly concentrated parts of a solution (which will drop out under taking a weak limit) do not affect parts of the solution living at unit scale.

These arguments break down in the presence of the quintic nonlinearity, which is energy-critical. In particular, initial data that converge weakly to zero in H_x^1 by concentrating at a point lead to solutions that do not go to zero in the space-time norms needed for well-posedness. Correspondingly, highly concentrated parts of a solution may have large norm and so, naively at least, have a nontrivial effect on the remainder of the solution. Thus, it is not clear that weak limits of solutions should even be continuous in time! The key to escaping this nightmare is to show that two parts of a solution have little effect on one another if they live at widely separated scales. We will achieve this by employing concentration compactness techniques.

Before tackling the full equation (1-6), one should first ensure that one can prove that weak limits of solutions are themselves solutions in the case of the energy-critical NLS equation (1-10). Questions of this type appear to have been studied before only in the case of the energy-critical wave equation [Bahouri and Gérard 1999]. As there, we proceed by harnessing the full power of the associated concentration compactness ideas. Specifically, one starts with a nonlinear profile decomposition and then further exploits some of the decoupling ideas used in its proof. In this paper, we will implement this strategy in the setting of (1-6); this is ample guidance for anyone seeking to reconstruct the argument for (1-10).

As a precursor to the nonlinear profile decomposition needed to prove that weak limits of solutions to (1-6) are themselves solutions, we must first develop a linear profile decomposition adapted to (1-6); see Proposition 4.3. Despite the fact that the linear equation underlying (1-6) differs from that underlying (1-10), we are able to adapt the profile decomposition for the linear Schrödinger equation to our setting, rather than proceeding ab initio. To develop the *nonlinear* profile decomposition, we need to construct solutions to (1-6) associated to each linear profile. For profiles living at unit scale, existence of these solutions (and all requisite bounds) follows from Theorem 3.3. Profiles whose characteristic length scale diverges can be approximated by linear solutions on bounded time intervals and so require no special attention. However, highly concentrated profiles require independent treatment; this is the content of Proposition 4.5. There are two subtle points here: (a) Such profiles are merely $\dot H^1_x$ and so do not have finite energy. (b) The characteristic time scale associated to such profiles is very short; thus, understanding such solutions even on a bounded interval essentially requires an understanding of their infinite time behavior.

The nonlinear profile decomposition posits that the nonlinear evolution of the initial data can be approximated by the sum of the nonlinear evolutions of its constituent profiles. This is verified by demonstrating decoupling of the profiles inside the nonlinearity (see Lemma 4.7) and exploiting a suitable stability theory for the equation (see Proposition 3.5). The latter requires certain a priori bounds, which are shown to hold in Lemma 4.6. Once it is known that the nonlinear profile decomposition faithfully represents the true solution, it is relatively elementary to complete the proof of well-posedness in the weak topology, that is, the proof of Theorem 4.1.

This completes our discussion of the new difficulties (relative to [Gustafson et al. 2009]) associated to the presence of energy-critical nonlinear terms.

(ii) We turn to the second main group of difficulties mentioned above, which stem from the shape of the energy functional. First, the lack of coercivity when $\gamma \in \big(0, \frac{2}{3}\big)$ was discussed already as an obstacle to proving global well-posedness. In this case, we restore coercivity by imposing a smallness condition on the initial data.

As also discussed above, convexity of the energy functional plays a key role in upgrading weak convergence to strong convergence in the argument of Nakanishi, via an argument of Radon–Riesz type. The analogue of (1-21) for our equation is as follows: For $z = M(u)$ as in (1-22),

$$E(u) = \tfrac{1}{2}\|\langle\nabla\rangle z\|_{L^2_x}^2 + \tfrac{1}{4}\gamma\|U|u|^2\|_{L^2_x}^2 + \int \tfrac{1}{6}q(u)^3\,dx. \tag{1-23}$$

Unlike its analogue (1-21), this does not yield an inequality between the energy and the H^1_x-norm of z. Indeed, the leading-order correction is the sign-indefinite term $\frac{4}{3}\int (u_1)^3\,dx$. Correspondingly, we will need to be concerned with the structure of our solution $u^\infty(t)$ as $t \to \infty$ to ensure that it does not contain surplus energy beyond that needed for its (putative) scattering state. Recall that $u^\infty(t)$ is merely constructed as a weak limit of solutions $u^{T_n}(t)$ defined by their values at $t = T_n$, which gives very little a priori information on its structure.

The resolution of this dilemma is to prove a form of energy decoupling between the part of the solution matching the scattering state and any residual part; see Lemma 6.3. Ultimately, this energy decoupling shows that any residual part of the solution must converge to zero in norm, which in fact obviates any explicit implementation of the Radon–Riesz-style argument described above.

Existence of wave operators. Recall that in Theorem 1.1, we cannot guarantee uniqueness of the nonlinear solution with prescribed scattering state. However, we are able to guarantee uniqueness under stronger hypotheses. Specifically, for scattering states with good linear decay, we can guarantee that there is only one nonlinear solution scattering to it with comparable decay. The decay of such solutions will be measured in the norm

$$\|u\|_{X_T} := \sup_{t\geq T} t^{\frac{1}{2}} \|u(t)\|_{H_x^{1,3}(\mathbb{R}^3)}.$$

Theorem 1.4. *Fix $\gamma \in (0,1)$. There exists $\eta > 0$ so that if $u_+ \in H^1_{\text{real}} + i\dot{H}^1_{\text{real}}$ satisfies*

$$\|V^{-1}e^{-itH}Vu_+\|_{X_1} \leq \eta, \tag{1-24}$$

then there exists a global solution $u \in C(\mathbb{R};\mathcal{E})$ to (1-6) such that

$$\lim_{t\to\infty} \|u(t) - V^{-1}e^{-itH}Vu_+\|_{H^1_{\text{real}}+i\dot{H}^1_{\text{real}}} = 0. \tag{1-25}$$

Moreover u is unique in the class of solutions with $\|u\|_{X_T} \leq 4\eta$ for some $T \geq 1$.

Remark 1.5. The proof of this theorem gives a quantitative rate in (1-25), namely,

$$\|u(t) - V^{-1}e^{-itH}Vu_+\|_{H^1_{\text{real}}+i\dot{H}^1_{\text{real}}} \lesssim t^{-\frac{1}{4}}. \tag{1-26}$$

Remark 1.6. Writing $u_{\text{lin}}(t) = V^{-1}e^{-itH}Vu_+$, we note that $u_+ \in H^1_{\text{real}} + i\dot{H}^1_{\text{real}}$ and $\|u_{\text{lin}}\|_{X_1} < \infty$ guarantee that u_{lin} is uniformly bounded in the energy space $\mathcal{E}$ for $t \geq 1$.

Finally, we observe that we can guarantee the smallness condition (1-24) by assuming control over weighted norms.

Corollary 1.7. *Let $\gamma \in (0,1)$ and $u_+ \in H^1_{\text{real}} + i\dot{H}^1_{\text{real}}$. If*

$$\|\langle x\rangle^{\frac{1}{2}+}\langle\nabla\rangle u_+\|_{L^2_x} + \|\langle x\rangle^{\frac{4}{3}+}\langle\nabla\rangle^{\frac{5}{6}}\operatorname{Re} u_+\|_{L^2_x}$$

is sufficiently small, then there exists a global solution $u \in C(\mathbb{R};\mathcal{E})$ to (1-6) such that (1-25) holds.

We prove Theorem 1.4 and Corollary 1.7 in Section 7. The proof, which relies primarily on dispersive and Strichartz estimates, consists of a contraction mapping argument that simultaneously solves the requisite PDE for $z = M(u)$ and inverts the normal form transformation. The argument differs little from that used to prove Theorem 1.1 in [Gustafson et al. 2007].

Outline of the paper. In Section 2 we set some notation and collect several useful lemmas.

Section 3 concerns the well-posedness of (1-6) in the energy space. We prove Theorem 3.3, giving global well-posedness and unconditional uniqueness in the energy space for (1-6). We also prove a stability result, Proposition 3.5.

The proof of the main result, Theorem 1.1, is ultimately carried out in Section 6. The strategy is modeled on the proof of Theorem 1.3 sketched above. Recalling that proof, we can broadly describe the three main steps as follows: (a) weak convergence uniformly in time, (b) well-posedness in the weak topology, and (c) strong convergence. As discussed above, new difficulties in our setting prevent a naive implementation of Nakanishi's strategy. Thus, we need to establish some preliminary results before launching into the proof of Theorem 1.1.

In Section 4, we consider step (b) and prove Theorem 4.1; briefly, this theorem states that if $u_n(0) \rightharpoonup u_0$ in $\dot{H}^1_x$, then $u_n(t) \rightharpoonup u(t)$ in $\dot{H}^1_x$ for all t, where u_n and u are solutions to (1-6) with initial data $u_n(0)$ and u_0, respectively. As described above, ingredients include (i) a linear profile decomposition adapted to (1-6) and (ii) a way to construct nonlinear solutions associated to the linear profiles. We prove the linear profile decomposition Proposition 4.3 by adapting the energy-critical linear profile decomposition for the Schrödinger propagator. For linear profiles living at unit length scales, we use Theorem 3.3 to construct the corresponding nonlinear profiles. The construction of nonlinear profiles in the case of highly concentrated linear profiles is more delicate and relies on the main result of [Colliander et al. 2008]. Specifically, we approximate such solutions to (1-6) by solutions to the energy-critical NLS and invoke the stability result, Proposition 3.5. The details are carried out in Proposition 4.5.

In Section 5, we discuss the normal form transformation, which is needed for steps (a) and (c). As discussed in the subsection on page 1526, low powers in the nonlinearity and poor spatial decay are problematic for establishing the equicontinuity needed to prove weak convergence. To remedy this, we exploit nonresonant structure in the equation via the normal form transformation M defined in (1-22). We prove some continuity and invertibility properties of this transformation in Proposition 5.1. We also prove Lemma 5.3 relating the energy and the inverse of the normal form transformation, which plays a role in step (c).

With the results of Section 4 and Section 5 in place, we are in a position to prove Theorem 1.1 in Section 6. Following the strategy of Nakanishi and using the normal form transformation and Theorem 4.1, we first construct the putative scattering solution u^∞. Working with the variables $z^\infty = M(u^\infty)$, we then prove a weak convergence result, Proposition 6.2. Having removed the worst quadratic terms via normal form transformation, establishing the requisite equicontinuity is a more feasible prospect; as in the work of [Gustafson et al. 2009], however, we still need to exhibit additional nonresonance at the cubic level.

We next upgrade to strong convergence, still at the level of z^∞. This relies largely on an energy decoupling lemma, Lemma 6.3. Finally, to complete the proof of Theorem 1.1, we show that strong convergence for z^∞ implies the desired convergence properties for u^∞. For this, we make use of results proved in Section 5 concerning the inverse of the normal form transformation (e.g., Lemma 5.3).

Finally, in Section 7 we prove Theorem 1.4 and Corollary 1.7. These results are much simpler than Theorem 1.1; they follow from a contraction mapping argument and rely primarily on Strichartz/dispersive estimates.

2. Notation and useful lemmas

Some notation. We write $A \lesssim B$ or $A = O(B)$ to indicate that $A \le CB$ for some constant $C > 0$. Dependence of implicit constants on various parameters will be indicated with subscripts. For example, $A \lesssim_\varphi B$ means that $A \le CB$ for some $C = C(\varphi)$. The dependence of implicit constants on the parameter γ defined in (1-3) will not be explicitly indicated. We write $A \sim B$ if $A \lesssim B$ and $B \lesssim A$. We write $A \ll B$ if $A \le cB$ for some small $c > 0$.

We write a complex-valued function u as $u = u_1 + iu_2$. When X is a monomial, we use the notation $\emptyset(X)$ to denote a finite linear combination of products of the factors of X, where Mikhlin multipliers (e.g., Littlewood–Paley projections) and/or complex conjugation may be additionally applied in each factor. We extend $\emptyset$ to polynomials via $\emptyset(X+Y) = \emptyset(X) + \emptyset(Y)$.

For a time interval I we write $L_t^q L_x^r(I \times \mathbb{R}^3)$ for the Banach space of functions $u : I \times \mathbb{R}^3 \to \mathbb{C}$ equipped with the norm

$$\|u\|_{L_t^q L_x^r(I\times\mathbb{R}^3)} = \left(\int_I \|u(t)\|_{L_x^r(\mathbb{R}^3)}^q \, dt\right)^{\frac{1}{q}},$$

with the usual adjustments when q or r is infinity. If $q = r$ we write $L_t^q L_x^q = L_{t,x}^q$. We often abbreviate $\|u\|_{L_t^q L_x^r(I\times\mathbb{R}^3)} = \|u\|_{L_t^q L_x^r}$ and $\|u\|_{L_x^r(\mathbb{R}^3)} = \|u\|_{L_x^r}$. We also write $C(I; X)$ to denote the space of continuous functions on I taking values in X.

We use the following convention for the Fourier transform on $\mathbb{R}^3$:

$$\hat{f}(\xi) = (2\pi)^{-\frac{3}{2}} \int_{\mathbb{R}^3} e^{-ix\cdot\xi} f(x)\, dx \quad \text{so that} \quad f(x) = (2\pi)^{-\frac{3}{2}} \int_{\mathbb{R}^3} e^{ix\cdot\xi} \hat{f}(\xi)\, d\xi.$$

The fractional differential operator $|\nabla|^s$ is defined by $\widehat{|\nabla|^s f}(\xi) = |\xi|^s \hat{f}(\xi)$. We will also make use of the following Fourier multiplier operators (and powers thereof):

$$\begin{aligned}
\langle \xi \rangle &= \sqrt{2\gamma + |\xi|^2}, & \langle \nabla \rangle &= \sqrt{2\gamma - \Delta}, \\
U(\xi) &= \sqrt{|\xi|^2(2\gamma + |\xi|^2)^{-1}}, & U &= \sqrt{(-\Delta)(2\gamma - \Delta)^{-1}}, \\
H(\xi) &= \sqrt{|\xi|^2(2\gamma + |\xi|^2)}, & H &= \sqrt{(-\Delta)(2\gamma - \Delta)}.
\end{aligned}$$

Fix $\gamma \in (0, 1)$ as in (1-3). We define homogeneous and inhomogeneous Sobolev norms $\dot{H}_x^{s,r}$ and $H_x^{s,r}$ as the completion of Schwartz functions under the norms

$$\|f\|_{\dot{H}_x^{s,r}} := \|(-\Delta)^{\frac{s}{2}} f\|_{L_x^r} \quad \text{and} \quad \|f\|_{H_x^{s,r}} := \|(2\gamma - \Delta)^{\frac{s}{2}} f\|_{L_x^r},$$

respectively. When $r = 2$ we abbreviate $\dot{H}_x^{s,2} = \dot{H}_x^s$ and $H_x^{s,2} = H_x^s$. Note that this definition of the H_x^s-norm is equivalent (up to constants depending on γ) to the standard one, which uses the operator $(1-\Delta)^{\frac{s}{2}}$.

Basic harmonic analysis. We employ the standard Littlewood–Paley theory. Let ϕ be a radial bump function supported in $\{|\xi| \le \frac{11}{10}\}$ and equal to 1 on the unit ball. For $N \in 2^{\mathbb{Z}}$ we define the Littlewood–Paley projections

$$\widehat{P_{\leq N}u}(\xi) = \phi\big(\tfrac{1}{N}\xi\big)\hat{u}(\xi), \quad \widehat{P_N u}(\xi) = \big[\phi\big(\tfrac{1}{N}\xi\big) - \phi\big(\tfrac{1}{2N}\xi\big)\big]\hat{u}(\xi), \quad \text{and} \quad P_{>N} = \mathrm{Id} - P_{\leq N}.$$

These operators commute with all other Fourier multiplier operators. They are self-adjoint and bounded on every L_x^p and H_x^s space for $1 \leq p \leq \infty$ and $s \geq 0$. We write $P_{\mathrm{lo}} = P_{\leq 1}$ and $P_{\mathrm{hi}} = P_{>1}$.

The Littlewood–Paley projections obey the following standard estimates.

Lemma 2.1 (Bernstein estimates). *For $1 \leq r \leq q \leq \infty$ and $s \geq 0$ we have*

$$\big\| |\nabla|^s P_{\leq N} u \big\|_{L_x^r(\mathbb{R}^3)} \lesssim N^s \| P_{\leq N} u\|_{L_x^r(\mathbb{R}^3)},$$
$$\| P_{>N} u\|_{L_x^r(\mathbb{R}^3)} \lesssim N^{-s} \big\| |\nabla|^s P_{>N} u \big\|_{L_x^r(\mathbb{R}^3)},$$
$$\| P_{\leq N} u\|_{L_x^q(\mathbb{R}^3)} \lesssim N^{\frac{3}{r} - \frac{3}{q}} \| P_{\leq N} u\|_{L_x^r(\mathbb{R}^3)}.$$

We will need the following:

Lemma 2.2 (fractional chain rule, [Christ and Weinstein 1991]). *Suppose $G \in C^1(\mathbb{C})$ and $s \in (0,1]$. Let $1 < r, r_2 < \infty$ and $1 < r_1 \leq \infty$ satisfy $1/r_1 + 1/r_2 = 1/r$. Then*

$$\big\| |\nabla|^s G(u) \big\|_{L_x^r} \lesssim \|G'(u)\|_{L_x^{r_1}} \big\| |\nabla|^s u \big\|_{L_x^{r_2}}.$$

We will also need the following result concerning bilinear Fourier multipliers. For a real-valued function $B(\xi_1, \xi_2)$ we define the operator $B[f,g]$ via

$$\widehat{B[f,g]}(\xi) := (2\pi)^{\frac{3}{2}} \int_{\mathbb{R}^3} B(\eta, \xi - \eta) \hat{f}(\eta) \hat{g}(\xi - \eta)\, d\eta. \tag{2-1}$$

Lemma 2.3 (Coifman–Meyer bilinear estimate, [Coifman and Meyer 1978; Meyer and Coifman 1991]). *If the symbol $B(\xi_1, \xi_2)$ satisfies*

$$|\partial_{\xi_1}^\alpha \partial_{\xi_2}^\beta B(\xi_1, \xi_2)| \lesssim_{\alpha,\beta} (|\xi_1| + |\xi_2|)^{-(|\alpha| + |\beta|)}$$

for all multi-indices α, β up to sufficiently high order, then

$$\|B[f,g]\|_{L_x^r} \lesssim \|f\|_{L_x^{r_1}} \|g\|_{L_x^{r_2}}$$

for all $1 < r < \infty$ and $1 < r_1, r_2 < \infty$ satisfying $1/r = 1/r_1 + 1/r_2$.

Linear estimates. We record here the dispersive and Strichartz estimates for the propagators $e^{it\Delta}$ and e^{-itH}.

As is well known, the linear Schrödinger propagator in three space dimensions can be written as

$$[e^{it\Delta} f](x) = (4\pi i t)^{-\frac{3}{2}} \int_{\mathbb{R}^3} e^{\frac{i|x-y|^2}{4t}} f(y)\, dy$$

for $t \neq 0$. This yields the dispersive estimates

$$\|e^{it\Delta} f\|_{L_x^r(\mathbb{R}^3)} \lesssim |t|^{-(\frac{3}{2} - \frac{3}{r})} \|f\|_{L_x^{r'}(\mathbb{R}^3)} \tag{2-2}$$

for $t \neq 0$, where $2 \leq r \leq \infty$ and $1/r + 1/r' = 1$. This estimate can be used to prove the standard Strichartz estimates for $e^{it\Delta}$. We state the result we need in three space dimensions.

Proposition 2.4 (Strichartz estimates for $e^{it\Delta}$, [Ginibre and Velo 1992; Keel and Tao 1998; Strichartz 1977]). *For a space-time slab $I\times\mathbb{R}^3$ and $2\le q,\tilde q\le\infty$ with $2/q+3/r=2/\tilde q+3/\tilde r=\frac{3}{2}$, we have*

$$\left\| e^{it\Delta}\varphi+\int_0^t e^{i(t-s)\Delta}F(s)\,ds\right\|_{L_t^qL_x^r(I\times\mathbb{R}^3)}\lesssim\|\varphi\|_{L_x^2}+\|F\|_{L_t^{\tilde q'}L_x^{\tilde r'}(I\times\mathbb{R}^3)}.$$

Using stationary phase, one can prove a similar dispersive estimate for e^{-itH} (see [Gustafson et al. 2006; 2009]). In fact, there is a small gain at low frequencies compared to the estimates for the linear Schrödinger propagator; while the dispersion relation for this propagator has less curvature in the radial direction than that for Schrödinger, this is more than compensated for by the increased curvature in the angular directions.

Proposition 2.5 (estimates for e^{-itH}, [Gustafson et al. 2006, 2009]). *For $2\le r\le\infty$ we have*

$$\|e^{-itH}f\|_{L_x^r(\mathbb{R}^3)}\lesssim|t|^{-(\frac{3}{2}-\frac{3}{r})}\|U^{\frac{1}{2}-\frac{1}{r}}f\|_{L_x^{r'}(\mathbb{R}^3)}\tag{2-3}$$

for $t\neq0$. In particular, for a space-time slab $I\times\mathbb{R}^3$ and $2\le q,\tilde q\le\infty$ with $2/q+3/r=2/\tilde q+3/\tilde r=\frac{3}{2}$, we have

$$\left\| e^{-itH}\varphi+\int_0^t e^{-i(t-s)H}F(s)\,ds\right\|_{L_t^qL_x^r(I\times\mathbb{R}^3)}\lesssim\|\varphi\|_{L_x^2}+\|F\|_{L_t^{\tilde q'}L_x^{\tilde r'}(I\times\mathbb{R}^3)}.$$

For an interval I and $s\ge0$ we define the Strichartz norm by

$$\|u\|_{\dot S^s(I)}=\sup\left\{\left\||\nabla|^s u\right\|_{L_t^qL_x^r(I\times\mathbb{R}^3)}:2\le q\le\infty,\ \frac{2}{q}+\frac{3}{r}=\frac{3}{2}\right\}.$$

The Strichartz space $\dot S^s(I)$ is then defined to be the closure of test functions under this norm. We let $\dot N^s(I)$ denote the corresponding dual Strichartz space.

In several places it will be more convenient to work with (1-6) rather than the diagonalized (1-11). The linear propagator associated with (1-6) takes the form

$$V^{-1}e^{-itH}V\begin{bmatrix}f_1\\f_2\end{bmatrix}=\begin{bmatrix}\cos(tH)&U\sin(tH)\\-U^{-1}\sin(tH)&\cos(tH)\end{bmatrix}\begin{bmatrix}f_1\\f_2\end{bmatrix}\tag{2-4}$$

for any function $f=f_1+if_2$. We will make use of the following Strichartz estimates for this propagator:

Lemma 2.6. *Fix $T>0$. Given $2<q,\tilde q\le\infty$ with $2/q+3/r=2/\tilde q+3/\tilde r=\frac{3}{2}$, we have*

$$\left\|V^{-1}e^{-itH}V\varphi+\int_0^t V^{-1}e^{-i(t-s)H}VF(s)\,ds\right\|_{L_t^qL_x^r}\lesssim_T\|\varphi\|_{L_x^2}+\|F\|_{L_t^{\tilde q'}L_x^{\tilde r'}},\tag{2-5}$$

where all space-time norms are over $[-T,T]\times\mathbb{R}^3$.

Proof. As we are excluding the endpoint, it suffices (via a TT^* argument) to prove the result when $F\equiv0$; moreover, it clearly suffices to consider each entry in the matrix (2-4) separately. In view of the boundedness of U, three out of four of these matrix elements obey the same Strichartz estimates as e^{-itH}; see Proposition 2.5. As $P_{\mathrm{hi}}U^{-1}$ is also bounded, we need only prove Strichartz estimates for

$P_{\mathrm{lo}}U^{-1}\sin(tH)$. However, this is easily done via Hölder and Bernstein's inequality:

$$\begin{aligned}\|P_{\mathrm{lo}}U^{-1}\sin(tH)\varphi\|_{L_t^qL_x^r([-T,T]\times\mathbb{R}^3)} &\lesssim T^{\frac{1}{q}}\|P_{\mathrm{lo}}U^{-1}\sin(tH)\varphi\|_{L_t^\infty L_x^2([-T,T]\times\mathbb{R}^3)}\\ &\lesssim T^{1+\frac{1}{q}}\|\varphi\|_{L_x^2(\mathbb{R}^3)}.\end{aligned}\tag{2-6}$$

This completes the proof of the lemma. □

At high frequencies, the operator e^{-itH} closely resembles the Schrödinger propagator (on bounded time intervals); specifically, we have

$$\sqrt{|\xi|^2(2\gamma+|\xi|^2)}=|\xi|^2+\gamma+m(\xi)\quad\text{with } |m(\xi)|\lesssim\langle\xi\rangle^{-2}.\tag{2-7}$$

Indeed, it is not difficult to verify that $m(\xi)$ defines a Mikhlin multiplier. This observation will play a key role in our treatment of highly concentrated profiles in Section 4. For the moment, however, we simply use it to obtain a crude local smoothing estimate.

Lemma 2.7 (local smoothing). *Given $T>0$ and $R>0$,*

$$\big\||\nabla|^{\frac{1}{2}}V^{-1}e^{-itH}V\varphi\big\|_{L^2_{t,x}(\{|t|\leq T\}\times\{|x|\leq R\})}\lesssim_{R,T}\|\varphi\|_{L^2_x}.\tag{2-8}$$

Proof. We treat high and low frequencies separately. In the low-frequency regime, we exploit (2-4) and argue as in (2-6) to deduce that

$$\big\||\nabla|^{\frac{1}{2}}P_{\mathrm{lo}}V^{-1}e^{-itH}V\varphi\big\|_{L^2_{t,x}(\{|t|\leq T\}\times\{|x|\leq R\})}\lesssim T^{\frac{1}{2}}(1+T)\|\varphi\|_{L^2_x}.$$

In the high-frequency regime, we can use the usual local smoothing estimate for the Schrödinger equation together with

$$\big\||\nabla|^{\frac{1}{2}}P_{\mathrm{hi}}V^{-1}[e^{-itH}-e^{-it(\gamma-\Delta)}]V\varphi\big\|_{L_t^\infty L_x^2(\{|t|\leq T\}\times\mathbb{R}^3)}\lesssim T\|\varphi\|_{L^2_x},$$

which follows from (2-7). □

In practice, we will use the following corollary.

Corollary 2.8. *Let K be a compact subset of $I\times\mathbb{R}^3$ for some interval $I\subset\mathbb{R}$. Then the following estimates hold*:

$$\|\nabla e^{it\Delta}f\|_{L^2_{t,x}(K)}\lesssim_K\|e^{it\Delta}f\|^{\frac{1}{3}}_{L^{10}_{t,x}(I\times\mathbb{R}^3)}\|f\|^{\frac{2}{3}}_{\dot H^1_x},$$

$$\|\nabla V^{-1}e^{-itH}Vf\|_{L^2_{t,x}(K)}\lesssim_K\|V^{-1}e^{-itH}Vf\|^{\frac{1}{3}}_{L^{10}_{t,x}(I\times\mathbb{R}^3)}\|f\|^{\frac{2}{3}}_{\dot H^1_x}.$$

Proof. Fix $N>0$. By the Bernstein and Hölder inequalities,

$$\|\nabla P_{\leq N}e^{it\Delta}f\|_{L^2_{t,x}(K)}\lesssim_K N\|e^{it\Delta}f\|_{L^{10}_{t,x}(I\times\mathbb{R}^3)}.$$

By the local smoothing estimate for $e^{it\Delta}$ and Bernstein, we also have

$$\|\nabla P_{>N}e^{it\Delta}f\|_{L^2_{t,x}(K)}\lesssim_K\big\||\nabla|^{\frac{1}{2}}P_{>N}f\big\|_{L^2_x}\lesssim_K N^{-\frac{1}{2}}\|\nabla f\|_{L^2_x}.$$

Optimizing in the choice of N yields the first estimate.

To obtain the second estimate one argues in exactly the same way, making use of Lemma 2.7. □

3. Global well-posedness in the energy space

In this section we discuss the well-posedness of (1-6) in the energy space. We begin by justifying the name "energy space" given to the set $\mathcal{E}$ defined in (1-9). Recall from the Introduction that if $u \in \mathcal{E}$, then $|E(u)| < \infty$. The following two lemmas prove that if the energy of u is finite, then $u \in \mathcal{E}$; when $\gamma \in \left(0, \frac{2}{3}\right)$, this requires an additional smallness condition.

Lemma 3.1. *If $\gamma \in \left(\frac{2}{3}, 1\right)$ and $E(u) < \infty$, then $u \in \mathcal{E}$ with $\|u\|_{\mathcal{E}}^2 \lesssim E(u)$. If $\gamma = \frac{2}{3}$ and $E(u) < \infty$, then $u \in \mathcal{E}$ with*

$$\|\nabla u\|_{L_x^2}^2 \lesssim E(u) \quad \text{and} \quad \|q\|_{L_x^2}^2 \lesssim E(u) + [E(u)]^3.$$

Proof. When $\gamma > \frac{2}{3}$ we use the fact that $q \geq -1$ in (1-8) to write

$$E(u) \geq \frac{1}{2} \int |\nabla u|^2 \, dx + \frac{\gamma}{4} \int \left(1 - \frac{2}{3\gamma}\right) q^2 \, dx,$$

which immediately implies the result.

We now turn to the case when $\gamma = \frac{2}{3}$. In this case, the energy takes the form

$$E(u) = \frac{1}{2} \int |\nabla u|^2 \, dx + \frac{1}{6} \int q^2(q+1) \, dx.$$

As $q \geq -1$, we have $q^2(q+1) \geq 0$. Thus $u \in \dot{H}_x^1(\mathbb{R}^3)$ and $\|\nabla u\|_{L_x^2}^2 \lesssim E(u)$.

To estimate the L_x^2-norm of q, we note that

$$\int_{\{q \geq -\frac{1}{2}\}} q^2 \, dx \leq 2 \int q^2(q+1) \, dx \lesssim E(u).$$

On the other hand, if $q < -\frac{1}{2}$ then $|u_1| > \frac{1}{4}$; thus, by Chebyshev's inequality and Sobolev embedding,

$$\int_{\{q < -\frac{1}{2}\}} q^2 \, dx \leq 4^6 \|u_1\|_{L_x^6}^6 \lesssim \|\nabla u\|_{L_x^2}^6 \lesssim [E(u)]^3. \qquad \square$$

We next consider the full range $\gamma \in (0, 1)$. In this case, we can guarantee coercivity of the energy under an appropriate smallness assumption.

Lemma 3.2. *For any $\gamma \in (0, 1)$ there exists $\delta_\gamma > 0$ so that the following hold*:

(i) *If $E(u) < \infty$ and $\|\nabla u_1\|_{L_x^2}^2 \leq \delta_\gamma$, then $u \in \mathcal{E}$ with $\|u\|_{\mathcal{E}}^2 \lesssim E(u)$.*

(ii) *For any ball B,*

$$\|\nabla u_1\|_{L_x^2(\mathbb{R}^3)}^2 \leq \delta_\gamma \implies \int_{B^c} \tfrac{1}{2}|\nabla u|^2 + \tfrac{1}{4}\gamma q^2 + \tfrac{1}{6} q^3 \, dx \geq 0. \tag{3-1}$$

(iii) *If $u : I \times \mathbb{R}^3 \to \mathbb{C}$ is a solution to (1-6) with $E(u) \leq \frac{1}{4}\delta_\gamma$ and $\|\nabla \operatorname{Re} u(t_0)\|_{L_x^2}^2 \leq \delta_\gamma$ for some $t_0 \in I$, then*

$$\|\nabla \operatorname{Re} u\|_{L_t^\infty L_x^2(I \times \mathbb{R}^3)}^2 \leq \delta_\gamma \quad \text{and} \quad \|u\|_{L^\infty(I;\mathcal{E})}^2 \lesssim E(u).$$

Proof. We begin by writing

$$E(u) = \int \tfrac{1}{2}|\nabla u|^2 + \tfrac{1}{8}\gamma q^2 + \tfrac{1}{6}q^2\big(q + \tfrac{3}{4}\gamma\big)\,dx$$
$$\geq \int \tfrac{1}{2}|\nabla u|^2 + \tfrac{1}{8}\gamma q^2\,dx + \int_{\{q<-\frac{3}{4}\gamma\}} \tfrac{1}{6}q^2\big(q + \tfrac{3}{4}\gamma\big)\,dx.$$

For $q < -\frac{3}{4}\gamma$ we have $|u_1| > \frac{3}{8}\gamma$. Thus, by Chebyshev's inequality and Sobolev embedding, we have

$$\big|\{q < -\tfrac{3}{4}\gamma\}\big| \leq \Big(\frac{8}{3\gamma}\Big)^6 \|u_1\|_{L_x^6}^6 \lesssim \gamma^{-6}\|\nabla u_1\|_{L_x^2}^6.$$

Recalling that $q \geq -1$, we find that for $\|\nabla u_1\|_{L_x^2}^2 \ll \gamma^{\frac{3}{2}}$ we have

$$\bigg|\int_{\{q<-\frac{3}{4}\gamma\}} \tfrac{1}{6}q^2\big(q + \tfrac{3}{4}\gamma\big)\,dx\bigg| \lesssim \gamma^{-6}\|\nabla u_1\|_{L_x^2}^6 \leq \tfrac{1}{4}\|\nabla u_1\|_{L_x^2}^2.$$

Thus

$$E(u) \geq \int \tfrac{1}{4}|\nabla u|^2 + \tfrac{1}{8}\gamma q^2\,dx,$$

which yields conclusion (i) of the lemma. Claim (iii) also follows from this and a continuity argument.

To obtain (ii), we repeat the argument above, using the fact that Sobolev embedding holds in the exterior of any ball B. □

We next turn to the question of global well-posedness for (1-6) with initial data $u_0 \in \mathcal{E}$. From the lemmas above we see that $u(t) \in \mathcal{E}$ and $\|\nabla u(t)\|_{L_x^2}^2 \lesssim E(u_0)$ for all times of existence, whenever (1) $\gamma \in \big[\frac{2}{3}, 1\big)$ or (2) $\gamma \in \big(0, \frac{2}{3}\big)$ and $E(u_0)$ and $\|\nabla \operatorname{Re} u_0\|_{L_x^2}$ are sufficiently small. This a priori bound on $\|\nabla u(t)\|_{L_x^2}$ allows us to treat (1-6) as a perturbation of the defocusing energy-critical NLS, which was proven to be globally well-posed with finite space-time bounds in [Colliander et al. 2008]. See also [Killip et al. 2012; Tao et al. 2007] for similar perturbative arguments.

Theorem 3.3 (global well-posedness and unconditional uniqueness). *For $\gamma \in \big[\frac{2}{3}, 1\big)$ and $u_0 \in \mathcal{E}$, there exists a unique global solution $u \in C(\mathbb{R}; \mathcal{E})$ to (1-6).*

For $\gamma \in \big(0, \frac{2}{3}\big)$, if $u_0 \in \mathcal{E}$ satisfies $\|\nabla \operatorname{Re} u_0\|_{L_x^2}^2 \leq \delta_\gamma$ and $E(u_0) \leq \frac{1}{4}\delta_\gamma$, then there exists a unique global solution $u \in C(\mathbb{R}; \mathcal{E})$ to (1-6). Here δ_γ is as in Lemma 3.2.

In both cases the solution remains uniformly bounded in $\mathcal{E}$ and for any $T > 0$,

$$\|u\|_{\dot{S}^1([-T,T])} \lesssim_T 1.$$

Remark 3.4. When $\gamma \in \big(0, \frac{2}{3}\big)$, smallness of the initial data is only exploited to prove global existence; the proof we present below guarantees uniqueness of any solution in $C(I; \mathcal{E})$ on any time interval $I \subseteq \mathbb{R}$.

Proof. As mentioned above, Lemmas 3.1 and 3.2 imply that under the hypotheses of Theorem 3.3 we have $\|\nabla u(t)\|_{L_x^2}^2 \lesssim E(u_0)$ for all times t of existence. This allows us to treat (1-6) as a perturbation of the defocusing energy-critical NLS. Indeed, we may rewrite (1-6) as

$$(i\partial_t + \Delta)u = |u|^4 u + \mathcal{R}(u),$$

where $\mathcal{R}(u) = 2\gamma \operatorname{Re} u + \sum_{j=2}^{4} N_j(u)$. Noting that the "error" $\mathcal{R}(u)$ is energy-subcritical, one may argue as in [Killip et al. 2012, Section 4.2] to construct a global solution $u \in C(\mathbb{R}; \mathcal{E}) \cap L_t^{10} \dot{H}_x^{1,\frac{30}{13}}(\mathbb{R} \times \mathbb{R}^3)$ to (1-6). A key ingredient in this argument is the main result in [Colliander et al. 2008], which guarantees that the defocusing energy-critical NLS is globally well-posed with finite $L_t^{10} \dot{H}_x^{1,\frac{30}{13}}(\mathbb{R} \times \mathbb{R}^3)$ norm. We omit the details of this argument. Instead, we present the proof of uniqueness of solutions in the energy space, because the choice of energy space in this paper does not allow for a direct implementation of the methods in [Killip et al. 2012, Section 4.3].

Fix a compact time interval $I = [0, \tau]$ with $\tau > 0$ small. Let $u \in C(\mathbb{R}; \mathcal{E}) \cap L_t^{10} \dot{H}_x^{1,\frac{30}{13}}(\mathbb{R} \times \mathbb{R}^3)$ be the solution to (1-6) constructed via the perturbative argument described above. Suppose $\tilde{u} \in C(I; \mathcal{E})$ is another solution such that $\tilde{u}(0) = u(0)$. We wish to show that $u = \tilde{u}$ almost everywhere on $I \times \mathbb{R}^3$.

To this end, we define $w = \tilde{u} - u$ and let $0 < \eta < 1$ be a small parameter to be determined below. As $w(0) = 0$ and $w \in C(I; \dot{H}_x^1)$, we can choose τ small enough so that

$$\|w\|_{L_t^\infty \dot{H}_x^1(I \times \mathbb{R}^3)} \leq \eta. \tag{3-2}$$

As $\nabla u \in L_t^{10} L_x^{\frac{30}{13}}(I \times \mathbb{R}^3)$, we may also use Sobolev embedding and choose τ possibly even smaller to guarantee that

$$\|u\|_{L_{t,x}^{10}(I \times \mathbb{R}^3)} \leq \eta. \tag{3-3}$$

We also note that as u and $\tilde{u}$ are bounded in $\mathcal{E}$, we have that $q(u)$, $q(\tilde{u})$ are bounded in L_x^2; u, $\tilde{u}$ are bounded in L_x^6; and u_1, $\tilde{u}_1$ are bounded in $L_x^3 \cap L_x^6$.

We will first show that w is bounded in Strichartz spaces on $I \times \mathbb{R}^3$. To see this, we write

$$(i\partial_t + \Delta) w = 2\gamma \tilde{u}_1 + N(\tilde{u}) - [2\gamma u_1 + N(u)],$$

where $N(u)$ is as in (1-6). We make use of $q(u)$ and $q(\tilde{u})$ to rewrite

$$\begin{aligned}(i\partial_t + \Delta) w = O(|\tilde{u}|^5 + |u|^5) + O(|\tilde{u}|^4 + |u|^4) + O(|\tilde{u}|^3 + |u|^3) \\ + \gamma q(\tilde{u}) + (2\gamma + 4)\tilde{u}_1^2 + 2i\gamma \tilde{u}_1 \tilde{u}_2 - [\gamma q(u) + (2\gamma + 4) u_1^2 + 2i\gamma u_1 u_2].\end{aligned}$$

As $w(0) = 0$, we can use Strichartz to estimate

$$\begin{aligned}\|w\|_{L_t^2 L_x^6} + \|w\|_{L_t^4 L_x^3} + \|w\|_{L_t^\infty L_x^2} \lesssim \|\tilde{u}^5\|_{L_t^2 L_x^{6/5}} + \|u^5\|_{L_t^2 L_x^{6/5}} + \|\tilde{u}^4\|_{L_t^{4/3} L_x^{3/2}} \\ + \|u^4\|_{L_t^{4/3} L_x^{3/2}} + \|\tilde{u}^3\|_{L_t^1 L_x^2} + \|u^3\|_{L_t^1 L_x^2} \\ + \|q(\tilde{u})\|_{L_t^1 L_x^2} + \|q(u)\|_{L_t^1 L_x^2} + \|\tilde{u}\tilde{u}_1\|_{L_t^1 L_x^2} + \|u u_1\|_{L_t^1 L_x^2},\end{aligned}$$

where all space-time norms are over $I \times \mathbb{R}^3$. Using Hölder, we find

$$\|u^5\|_{L_t^2 L_x^{6/5}} \lesssim \tau^{1/2} \|u\|_{L_t^\infty L_x^6}^5, \quad \|u^4\|_{L_t^{4/3} L_x^{3/2}} \lesssim \tau^{3/4} \|u\|_{L_t^\infty L_x^6}^4, \quad \|u^3\|_{L_t^1 L_x^2} \lesssim \tau \|u\|_{L_t^\infty L_x^6},$$
$$\|q(u)\|_{L_t^1 L_x^2} \lesssim \tau \|q(u)\|_{L_t^\infty L_x^2}, \quad \|u u_1\|_{L_t^1 L_x^2} \lesssim \tau \|u\|_{L_t^\infty L_x^6} \|u_1\|_{L_t^\infty L_x^3},$$

and we can estimate similarly for $\tilde{u}$. Thus we conclude

$$\|w\|_{L_t^2 L_x^6} + \|w\|_{L_t^4 L_x^3} + \|w\|_{L_t^\infty L_x^2} < \infty. \tag{3-4}$$

We will show that, in fact,

$$\|w\|_{L_t^2 L_x^6} + \|w\|_{L_t^4 L_x^3} + \|w\|_{L_t^\infty L_x^2} = 0, \tag{3-5}$$

which implies $w = 0$ almost everywhere, as desired. To this end, we again rewrite the equation for w, using z to indicate that either w or u may appear. We have

$$(i\partial_t + \Delta)w = O\big(|w||u|^4 + |w|^5 + |w||z|^3 + |w||z|^2 + |w||z| + |w|\big).$$

We now use Strichartz, (3-2) and (3-3) to estimate

$$\begin{aligned}
&\|w\|_{L_t^2 L_x^6} + \|w\|_{L_t^4 L_x^3} + \|w\|_{L_t^\infty L_x^2} \\
&\quad\lesssim \|wu^4\|_{L_t^{10/9} L_x^{30/17}} + \|w^5\|_{L_t^2 L_x^{6/5}} + \|wz^3\|_{L_t^2 L_x^{6/5}} + \|wz^2\|_{L_t^{4/3} L_x^{3/2}} + \|wz\|_{L_t^1 L_x^2} + \|w\|_{L_t^1 L_x^2} \\
&\quad\lesssim \|u\|_{L_{t,x}^{10}}^4 \|w\|_{L_t^2 L_x^6} + \|w\|_{L_t^\infty L_x^6}^4 \|w\|_{L_t^2 L_x^6} + \tau^{\frac14} \|z\|_{L_t^\infty L_x^6}^3 \|w\|_{L_t^4 L_x^3} \\
&\qquad\qquad + \tau^{\frac12} \|z\|_{L_t^\infty L_x^6}^2 \|w\|_{L_t^4 L_x^3} + \tau^{\frac34} \|z\|_{L_t^\infty L_x^6} \|w\|_{L_t^4 L_x^3} + \tau \|w\|_{L_t^\infty L_x^2} \\
&\quad\lesssim \eta^4 \|w\|_{L_t^2 L_x^6} + \tau^{\frac14} \|w\|_{L_t^4 L_x^3} + \tau \|w\|_{L_t^\infty L_x^2}.
\end{aligned}$$

Choosing η, τ sufficiently small and using (3-4), we conclude that (3-5) holds and so $u = \tilde u$ almost everywhere on $I \times \mathbb{R}^3$. As uniqueness is a local property, this yields uniqueness in the energy space for solutions to (1-6). □

Next we develop a stability theory for (1-6), which we will need in Section 4.

Proposition 3.5 (stability theory). *Fix $T > 0$ and let $\tilde u : [-T, T] \times \mathbb{R}^3 \to \mathbb{C}$ be a solution to the perturbed equation*

$$(i\partial_t + \Delta - 2\gamma \operatorname{Re})\tilde u = N(\tilde u) + e$$

for some function e. Suppose that

$$\|\tilde u\|_{L_t^\infty \dot H_x^1([-T,T]\times\mathbb{R}^3)} + \|\nabla \tilde u\|_{L_t^{10} L_x^{30/13}([-T,T]\times\mathbb{R}^3)} \le L \tag{3-6}$$

for some constant $L > 0$. Let $u_0 \in \dot H_x^1(\mathbb{R}^3)$ and assume that

$$\|\tilde u(0) - u_0\|_{\dot H_x^1} + \left\| \int_0^t e^{i(t-s)\Delta} \nabla e(s)\, ds \right\|_{L_t^\infty L_x^2 \cap L_{t,x}^{10/3}([-T,T]\times\mathbb{R}^3)} \le \varepsilon \tag{3-7}$$

for some $\varepsilon \le \varepsilon_0(L, T)$. Then for $\varepsilon_0(L, T)$ sufficiently small there exists a solution $u : [-T, T] \times \mathbb{R}^3 \to \mathbb{C}$ to (1-6) with data $u(0) = u_0$ and

$$\|\nabla(\tilde u - u)\|_{L_t^\infty L_x^2 \cap L_{t,x}^{10/3}([-T,T]\times\mathbb{R}^3)} \le C(L, T)\varepsilon, \tag{3-8}$$

$$\|u\|_{\dot S^1([-T,T])} \le C(L, T). \tag{3-9}$$

Proof. The existence of the solution u on a small neighborhood of $t = 0$ follows from the arguments described in Theorem 3.3. In that setting, the solution could be extended globally due to energy control. That argument does not apply here as $u_0 \in \dot H_x^1$ by itself does not guarantee finiteness of the energy;

furthermore, we permit here large data even when $\gamma < \frac{2}{3}$, in which case the energy need not be coercive. However, these earlier arguments do show that if a solution should blow up in finite time, then the $\dot{S}^1$-norm must diverge. Consequently, we can prove that the solution exists and obeys (3-8) and (3-9) on the whole interval $[-T, T]$ by showing that it obeys (3-8) and (3-9) on any subinterval $0 \ni I \subseteq [-T, T]$ on which it does exist. This is what we do.

For brevity, we define the following norm: given a time interval $[a, b] \subset \mathbb{R}$,

$$\|u\|_{Y([a,b])} := \|\nabla u\|_{L_t^\infty L_x^2 \cap L_{t,x}^{10/3}([a,b]\times\mathbb{R}^3)}.$$

Given $0 < \eta < 1$ to be chosen later, we divide I into intervals J where

$$|J| \le \eta \quad \text{and} \quad \|\nabla\tilde{u}\|_{L_t^{10} L_x^{30/13}(J\times\mathbb{R}^3)} \le \eta. \tag{3-10}$$

The number K of such intervals depends only on L, T, and η. Below we will show that for η sufficiently small,

$$\inf_{t_0\in J} \|\tilde{u}(t_0) - u(t_0)\|_{\dot{H}_x^1} \le \eta \implies \|\tilde{u} - u\|_{Y(J)} \le A \inf_{t_0\in J} \|\tilde{u}(t_0) - u(t_0)\|_{\dot{H}_x^1} \tag{3-11}$$

for some absolute constant A on such intervals J. Iterating this completes the proof of (3-8) and yields constants

$$\varepsilon_0 = A^{-K(L,T,\eta)}\eta \quad \text{and} \quad C(L, T) = K(L, T, \eta)A^{K(L,T,\eta)}.$$

We now verify (3-11). Writing $u = \tilde{u} + v$, we use Strichartz and (3-7) to estimate

$$\|v\|_{Y(J)} \lesssim \inf_{t_0\in J} \|v(t_0)\|_{\dot{H}_x^1} + \big\|\nabla[N(\tilde{u} + v) - N(\tilde{u})]\big\|_{\dot{N}^0(J)} + |J|\,\|v\|_{L_t^\infty \dot{H}_x^1} + \varepsilon,$$

where $N(\cdot)$ denotes the nonlinearity, as in (1-6). Moreover,

$$\begin{aligned}\big\|\nabla[N(\tilde{u} + v) - N(\tilde{u})]\big\|_{\dot{N}^0(J)} \lesssim \|\nabla\tilde{u}\|_{L_t^{10} L_x^{30/13}} \|v\|_{L_{t,x}^{10}} \sum_{k=2}^{5} |J|^{\frac{5-k}{4}} \Big(\|\tilde{u}\|_{L_{t,x}^{10}}^{k-2} + \|v\|_{L_{t,x}^{10}}^{k-2}\Big)&\\ + \|\nabla v\|_{L_t^{10} L_x^{30/13}} \sum_{k=2}^{5} |J|^{\frac{5-k}{4}} \Big(\|\tilde{u}\|_{L_{t,x}^{10}}^{k-1} + \|v\|_{L_{t,x}^{10}}^{k-1}\Big)&,\end{aligned}$$

where all space-time norms are over $J \times \mathbb{R}^3$. Using Sobolev embedding and (3-10), we therefore obtain

$$\|v\|_{Y(J)} \lesssim \inf_{t_0\in J} \|v(t_0)\|_{\dot{H}_x^1} + \sum_{k=1}^{5} \eta^{\frac{5-k}{4}} \|v\|_{Y(J)}^k + \varepsilon.$$

Choosing η sufficiently small, a simple bootstrap argument yields (3-11).

Using the fact that u is a solution to (1-6), a further application of the Strichartz inequality gives (3-9). □

We also record the following corollary.

Corollary 3.6 (small-data space-time bounds). *Given $T > 0$ there exists $\eta(T) > 0$ such that*

$$\|u_0\|_{\dot{H}_x^1} \le \eta(T) \implies \|u\|_{\dot{S}^1([-T,T])} \lesssim_T \|u_0\|_{\dot{H}_x^1},$$

where u denotes the solution to (1-6) with data $u(0) = u_0$.

Proof. We apply Proposition 3.5 with $\tilde{u} = e^{it\Delta}u_0$. By the Strichartz inequality,

$$\|\tilde{u}\|_{L_t^\infty \dot{H}_x^1([-T,T]\times\mathbb{R}^3)} + \|\nabla\tilde{u}\|_{L_t^{10} L_x^{30/13}([-T,T]\times\mathbb{R}^3)} \lesssim \|u_0\|_{\dot{H}_x^1},$$

while a little computation yields

$$\left\| \int_0^t e^{i(t-s)\Delta} e(s)\, ds \right\|_{\dot{S}^1([-T,T])} \lesssim \sum_{k=1}^{5} T^{\frac{5-k}{4}} \|u_0\|_{\dot{H}_x^1}^k.$$

Proposition 3.5 now gives the claim, provided $\eta(T)$ is taken sufficiently small. □

4. Well-posedness in the weak topology

In this section we prove the following well-posedness result in the weak $\dot{H}_x^1$ topology. As described in the Introduction, this theorem will play a key role in the proof of Theorem 1.1 in Section 6.

Theorem 4.1 (weak topology well-posedness). *Let $\gamma \in (0,1)$ and let $\{u_n(0)\}_{n\geq 1}$ be a bounded sequence in $\mathcal{E}$. Assume that $u_n(0) \rightharpoonup u_0$ weakly in $\dot{H}_x^1(\mathbb{R}^3)$. If $\gamma \in \big(0, \frac{2}{3}\big)$ we assume additionally that*

$$\|\nabla \operatorname{Re} u_n(0)\|_{L_x^2} \leq \delta_\gamma \quad \textit{and} \quad E(u_n(0)) \leq \tfrac{1}{4}\delta_\gamma,$$

where δ_γ is as in Theorem 3.3. Then there exists a unique solution $u \in C(\mathbb{R};\mathcal{E})$ to (1-6) with $u(0) = u_0$, and for all $t \in \mathbb{R}$ we have

$$u_n(t) \rightharpoonup u(t) \quad \textit{weakly in } \dot{H}_x^1(\mathbb{R}^3), \tag{4-1}$$

where $u_n \in C(\mathbb{R};\mathcal{E})$ denotes the solution to (1-6) with initial data $u_n(0)$, whose existence is guaranteed by Theorem 3.3.

We begin with the following lemma, which guarantees that the limit u_0 belongs to the energy space and obeys the necessary smallness conditions when $\gamma \in \big(0, \frac{2}{3}\big)$, so that the existence and uniqueness of the solution $u \in C(\mathbb{R};\mathcal{E})$ follow from Theorem 3.3.

Lemma 4.2. *Fix $\gamma \in (0,1)$ and suppose $\{u_n\}_{n\geq 1}$ is a bounded sequence in $\mathcal{E}$ that satisfies $u_n(x - x_n) \rightharpoonup u_0(x)$ weakly in $\dot{H}_x^1(\mathbb{R}^3)$ for some sequence $\{x_n\}_{n\geq 1} \subseteq \mathbb{R}^3$. Then $u_0 \in \mathcal{E}$. Moreover, if $\gamma \geq \frac{2}{3}$, then*

$$E(u_0) \leq \liminf_{n\to\infty} E(u_n). \tag{4-2}$$

If $\gamma \in \big(0, \frac{2}{3}\big)$ and $\|\nabla \operatorname{Re} u_n\|_{L_x^2}^2 \leq \delta_\gamma$, then $\|\nabla \operatorname{Re} u_0\|_{L_x^2}^2 \leq \delta_\gamma$ and (4-2) holds. Here δ_γ is as in Theorem 3.3.

Proof. Without loss of generality, we may assume that $x_n \equiv 0$.

To prove that $u_0 \in \mathcal{E}$, it suffices to show that $q(u_0) \in L_x^2$. As $u_n \rightharpoonup u_0$ weakly in $\dot{H}_x^1(\mathbb{R}^3)$, invoking Rellich–Kondrashov and passing to a subsequence, we deduce that $u_n \to u_0$ in $L_x^p(K)$ for any $2 \leq p < 6$ and any compact set $K \subset \mathbb{R}^3$. Therefore, for any ball $B \subset \mathbb{R}^3$,

$$\int_B |q(u_0(x))|^2\, dx = \lim_{n\to\infty} \int_B |q(u_n(x))|^2\, dx \leq \liminf_{n\to\infty} \int_{\mathbb{R}^3} |q(u_n(x))|^2\, dx < \infty.$$

As the bound does not depend on B, this proves $q(u_0) \in L_x^2$.

Proceeding similarly and using (weak) lower semicontinuity of the $\dot{H}^1_x$- and L^6_x-norms, we obtain

$$\int_B \tfrac{1}{2}|\nabla u_0|^2 + \tfrac{1}{4}\gamma q(u_0)^2 + \tfrac{1}{6}q(u_0)^3\,dx \le \liminf_{n\to\infty} \int_B \tfrac{1}{2}|\nabla u_n|^2 + \tfrac{1}{4}\gamma q(u_n)^2 + \tfrac{1}{6}q(u_n)^3\,dx$$

for any ball B. It is crucial here that the sextic term in the energy appears with a positive coefficient.

When $\gamma \in \left[\frac{2}{3}, 1\right)$, the energy density is positive and so the right-hand side above is majorized by $\liminf E(u_n)$. When $\gamma \in \left(0, \frac{2}{3}\right)$, we use instead (3-1) to reach the same conclusion. As $u_0 \in \mathcal{E}$, the dominated convergence theorem yields (4-2). □

We next prove a linear profile decomposition adapted to (1-12) for $\dot{H}^1_x$-bounded sequences. Beginning with the profile decomposition for the linear Schrödinger equation, we group the profiles according to the behavior of their associated parameters. We also show that the error term vanishes in the limit under propagation by $V^{-1}e^{-itH}V$ (in addition to propagation by $e^{it\Delta}$).

Proposition 4.3 (linear profile decomposition). *Suppose $\{f_n\}_{n\ge 1}$ is a bounded sequence in $\dot{H}^1_x(\mathbb{R}^3)$ and let $T > 0$. Passing to a subsequence, there exists $J^* \in \{0, 1, 2, \dots\} \cup \{\infty\}$ and for each finite $1 \le j \le J^*$ there exist a nonzero profile $\phi^j \in \dot{H}^1_x(\mathbb{R}^3)$, scales $\{\lambda^j_n\}_{n\ge 1} \subset (0,\infty)$, and positions $\{(t^j_n, x^j_n)\}_{n\ge 1} \subset \mathbb{R}\times\mathbb{R}^3$ conforming to one of the following two scenarios*:

- $\lambda^j_n \equiv 1$ *and* $t^j_n \equiv 0$,
- $\lambda^j_n \to 0$ *as* $n \to \infty$ *and either* $t^j_n \equiv 0$ *or* $t^j_n(\lambda^j_n)^{-2} \to \pm\infty$ *as* $n \to \infty$,

so that for any finite $0 \le J \le J^*$ we have the decomposition

$$f_n(x) = \sum_{j=1}^{J} e^{-it^j_n\Delta}\left[(\lambda^j_n)^{-\frac{1}{2}}\phi^j\left(\frac{x - x^j_n}{\lambda^j_n}\right)\right] + w^J_n(x)$$

satisfying the following properties:

$$(\lambda^j_n)^{\frac{1}{2}}\big(e^{it^j_n\Delta} f_n\big)(\lambda^j_n x + x^j_n) \rightharpoonup \phi^j \quad \text{weakly in } \dot{H}^1_x, \tag{4-3}$$

$$\lim_{J\to J^*}\limsup_{n\to\infty}\big[\|V^{-1}e^{-itH}V w^J_n\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)} + \|e^{it\Delta}w^J_n\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)}\big] = 0, \tag{4-4}$$

$$\sup_J \limsup_{n\to\infty}\left[\|f_n\|^2_{\dot{H}^1_x} - \sum_{j=1}^{J}\|\phi^j\|^2_{\dot{H}^1_x} - \|w^J_n\|^2_{\dot{H}^1_x}\right] = 0, \tag{4-5}$$

$$(\lambda^j_n)^{\frac{1}{2}}(e^{it^j_n\Delta} w^J_n)(\lambda^j_n x + x^j_n) \rightharpoonup 0 \quad \text{weakly in } \dot{H}^1_x \text{ for all } 1 \le j \le J, \tag{4-6}$$

$$\lim_{n\to\infty} \frac{\lambda^j_n}{\lambda^l_n} + \frac{\lambda^l_n}{\lambda^j_n} + \frac{|x^j_n - x^l_n|^2}{\lambda^j_n\lambda^l_n} + \frac{|t^j_n - t^l_n|}{\lambda^j_n\lambda^l_n} = \infty \quad \text{for all } j \ne l. \tag{4-7}$$

Proof. Using the linear profile decomposition for the Schrödinger propagator for bounded sequences in $\dot{H}^1_x$ (see, for example, [Keraani 2001] or [Visan 2014, Theorem 4.1]), we obtain a decomposition

$$f_n(x) = \sum_{j=1}^{J} e^{-it^j_n\Delta}\left[(\lambda^j_n)^{-\frac{1}{2}}\phi^j\left(\frac{x - x^j_n}{\lambda^j_n}\right)\right] + r^J_n(x) \tag{4-8}$$

satisfying (4-3), (4-5), (4-6), and (4-7) (with w_n^J replaced by r_n^J), as well as

$$\lim_{J\to J^*}\limsup_{n\to\infty}\|e^{it\Delta}r_n^J\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)}=0. \tag{4-9}$$

We will first show that we may assume the parameters conform to the two scenarios described above; in particular, we will show that we may absorb any other bubbles of concentration into the error r_n^J, while maintaining condition (4-9). To complete the proof of the proposition, we will show that condition (4-9) (for the new error term) suffices to prove (4-4). Note that it is essential in what follows that we work on a compact time interval.

We will use the notation

$$\phi_n^j(x):=e^{-it_n^j\Delta}\Big[(\lambda_n^j)^{-\frac12}\phi^j\Big(\frac{x-x_n^j}{\lambda_n^j}\Big)\Big].$$

We begin with the following lemma.

Lemma 4.4. *If* $|t_n^j|+\lambda_n^j\to\infty$ *as* $n\to\infty$, *then*

$$\lim_{n\to\infty}\|e^{it\Delta}\phi_n^j\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)}=0.$$

Proof. A direct computation gives

$$\|e^{it\Delta}\phi_n^j\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)}=\|e^{it\Delta}\phi^j\|_{L^{10}_{t,x}(I\times\mathbb{R}^3)},$$

where

$$I=\Big[\frac{-t_n^j-T}{(\lambda_n^j)^2},\frac{-t_n^j+T}{(\lambda_n^j)^2}\Big].$$

If $\lambda_n^j\to\infty$, then the lengths of the time intervals appearing on the right-hand side of the equality above shrink to zero; consequently, by the dominated convergence theorem combined with the Strichartz inequality, we deduce the claim.

Passing to a subsequence, we may henceforth assume that $\lambda_n^j\to\lambda^j\in[0,\infty)$. In this case, we have $|t_n^j|\to\infty$, and so the time intervals escape to infinity. Thus the claim follows once again from the dominated convergence theorem combined with the Strichartz inequality. □

Discarding the bubbles of concentration whose parameters satisfy the hypotheses of Lemma 4.4, we can now see that we may reduce attention to the two scenarios described in Proposition 4.3. Indeed, passing to a subsequence, we may assume that $\lambda_n^j\to\lambda^j\in[0,\infty)$ and $t_n^j\to t^j\in(-\infty,\infty)$. If $\lambda^j\neq0$, then we may assume that $\lambda_n^j\equiv1$ and $t_n^j\equiv0$ by redefining the corresponding profile to be $(\lambda^j)^{-\frac12}e^{-it^j\Delta}[\phi^j(\cdot/\lambda^j)]$. The error incurred by this modification can be absorbed into r_n^J; indeed, we have

$$\begin{aligned}\Big\|\phi_n^j-(\lambda^j)^{-\frac12}e^{-it^j\Delta}\Big[\phi^j\Big(\frac{x-x_n^j}{\lambda^j}\Big)\Big]\Big\|_{\dot H^1_x}&\\ \le\Big\|(\lambda_n^j)^{-\frac12}\phi^j\Big(\frac{x}{\lambda_n^j}\Big)-(\lambda^j)^{-\frac12}\phi^j\Big(\frac{x}{\lambda^j}\Big)\Big\|_{\dot H^1_x}&+\Big\|(e^{-it_n^j\Delta}-e^{-it^j\Delta})\Big[(\lambda^j)^{-\frac12}\phi^j\Big(\frac{x}{\lambda^j}\Big)\Big]\Big\|_{\dot H^1_x},\end{aligned}$$

which tends to zero as $n \to \infty$ by the strong convergence of the linear Schrödinger propagator. If instead $\lambda^j = 0$, then passing to a further subsequence we may assume that either $t_n^j \equiv 0$ or $t_n^j (\lambda_n^j)^{-2} \to \pm\infty$ as $n \to \infty$. Indeed, if there is a subsequence along which $t_n^j (\lambda_n^j)^{-2} \to \tau \in (-\infty, \infty)$, then we redefine the profile to be $e^{-i\tau\Delta}\phi^j$ and $t_n^j \equiv 0$. It is easy to see that the resulting error can be absorbed into r_n^J.

It remains to prove that the new error w_n^J (which consists of r_n^J plus the bubbles of concentration whose parameters satisfy the hypotheses of Lemma 4.4) obeys (4-4). This is a consequence of the following: if

$$\lim_{J\to J^*} \limsup_{n\to\infty} \|e^{it\Delta} w_n^J\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)} = 0,$$

then

$$\lim_{J\to J^*} \limsup_{n\to\infty} \| V^{-1} e^{-itH} V w_n^J \|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)} = 0.$$

To prove this final implication, we argue as follows: In view of the representation (2-4) and the boundedness of U and $P_{\mathrm{hi}} U^{-1}$, it suffices to verify that $e^{\mp itH} e^{\mp it\Delta}$ and $P_{\mathrm{lo}} U^{-1} \sin(tH) e^{-it\Delta}$ are Mikhlin multipliers with bounds that are uniform for $t \in [-T, T]$. In the former case, this follows from (2-7); with regard to the latter, see (2-6).

This completes the proof of Proposition 4.3. □

In the proof of Theorem 4.1, we will construct solutions to (1-6) associated to each ϕ_n^j. For profiles conforming to the first scenario in Proposition 4.3, we can achieve this by an application of Lemma 4.2 and Theorem 3.3. For profiles conforming to the second scenario, this is a more difficult problem, which we address in the following proposition.

Proposition 4.5 (highly concentrated nonlinear profiles). *Let $\phi \in \dot{H}_x^1(\mathbb{R}^3)$ and $T > 0$. Assume $\{\lambda_n\}_{n\ge1} \subset (0,\infty)$ and $\{(t_n, x_n)\}_{n\ge1} \subset \mathbb{R}\times\mathbb{R}^3$ satisfy $\lambda_n \to 0$ and either $t_n \equiv 0$ or $t_n\lambda_n^{-2} \to \pm\infty$. Then for n sufficiently large, there exists a solution u_n to (1-6) with initial data*

$$u_n(0, x) = \phi_n(x) := e^{-it_n\Delta}\left[\lambda_n^{-\frac{1}{2}} \phi\left(\frac{x - x_n}{\lambda_n}\right)\right]$$

satisfying

$$\|u_n\|_{\dot{S}^1([-T,T])} \le C(\|\phi\|_{\dot{H}_x^1}). \tag{4-10}$$

Moreover, for all $\varepsilon > 0$ there exist $\phi_\varepsilon, \psi_\varepsilon \in C_c^\infty([-T, T]\times\mathbb{R}^3)$ such that

$$\limsup_{n\to\infty} \left\| u_n(t, x) - e^{-i\gamma t} \lambda_n^{-\frac{1}{2}} \phi_\varepsilon\left(\frac{t - t_n}{\lambda_n^2}, \frac{x - x_n}{\lambda_n}\right)\right\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)} \le \varepsilon, \tag{4-11}$$

$$\limsup_{n\to\infty} \left\| \nabla u_n(t, x) - e^{-i\gamma t} \lambda_n^{-\frac{3}{2}} \psi_\varepsilon\left(\frac{t - t_n}{\lambda_n^2}, \frac{x - x_n}{\lambda_n}\right)\right\|_{L^{\frac{10}{3}}_{t,x}([-T,T]\times\mathbb{R}^3)} \le \varepsilon. \tag{4-12}$$

Proof. As (1-6) is space-translation invariant, without loss of generality we may assume that $x_n \equiv 0$.

We proceed via a perturbative argument. Specifically, using a solution to the defocusing energy-critical NLS, we will construct an approximate solution $\tilde{u}_n$ to (1-6) with initial data asymptotically matching ϕ_n. This approximate solution will have good space-time bounds inherited from the solution to the defocusing energy-critical NLS. Using the stability result Proposition 3.5, we will then deduce that for n sufficiently large, there exist true solutions u_n to (1-6) with $u_n(0)=\phi_n$ that inherits the space-time bounds of $\tilde u_n$, thus proving (4-10). We turn to the details.

If $t_n \equiv 0$, let v be the solution to the defocusing energy-critical NLS

$$(i\partial_t + \Delta)v = |v|^4 v \tag{4-13}$$

with initial data $v(0)=\phi$. If $t_n\lambda_n^{-2} \to \pm\infty$, let v be the solution to (4-13) which scatters in $\dot H^1_x$ to $e^{it\Delta}\phi$ as $t\to\pm\infty$. By the main result in [Colliander et al. 2008], we have

$$\|v\|_{\dot S^1(\mathbb{R})} \le C(\|\phi\|_{\dot H^1_x}).$$

We are now in a position to introduce the approximate solutions $\tilde u_n$ to (1-6). For $n\ge 1$, we define

$$\tilde u_n(t,x) := e^{-i\gamma t}\lambda_n^{-\frac12} v\Big(\frac{t-t_n}{\lambda_n^2}, \frac{x}{\lambda_n}\Big).$$

The phase factor $e^{-i\gamma t}$ is necessary. It replaces the linear factor in (1-6) by a nonresonant term; see (4-15).

Note that

$$\|\tilde u_n\|_{\dot S^1(\mathbb{R})} = \|v\|_{\dot S^1(\mathbb{R})} \le C(\|\phi\|_{\dot H^1_x}). \tag{4-14}$$

Moreover, $\tilde u_n(0)$ asymptotically matches the initial data $u_n(0)=\phi_n$; indeed, by construction, we have

$$\|\tilde u_n(0)-\phi_n\|_{\dot H^1_x} = \Big\| v\Big(-\frac{t_n}{\lambda_n^2}\Big) - e^{-i(t_n/\lambda_n^2)\Delta}\phi\Big\|_{\dot H^1_x} \to 0 \quad \text{as } n\to\infty.$$

To invoke the stability result Proposition 3.5 and deduce claim (4-10), it remains to show that $\tilde u_n$ is an approximate solution to (1-6) on the interval $[-T,T]$ as $n\to\infty$. A computation yields

$$e_n := (i\partial_t + \Delta - 2\gamma\,\mathrm{Re})\tilde u_n - N(\tilde u_n) = -\gamma\overline{\tilde u_n} - \sum_{j=2}^{4} N_j(\tilde u_n). \tag{4-15}$$

To establish (4-10), we have to verify that the error e_n satisfies the smallness condition in (3-7) for n sufficiently large.

Let $\delta>0$ to be chosen later. There exist $T_1, T_2>0$ sufficiently large so that

$$\|v\|_{L^{10}_{t,x}(\{|t|>T_1\}\times\mathbb{R}^3)} < \delta, \tag{4-16}$$

$$\|v(t)-e^{it\Delta}v_\pm\|_{\dot H^1_x} < \delta \quad \text{for } \pm t > T_2, \tag{4-17}$$

where $v_\pm$ denote the asymptotic states for the solution v. Note that the existence of $v_\pm$ is a consequence of the global space-time bounds for v, as discussed in [Colliander et al. 2008].

We first estimate the contribution of the higher-order terms appearing in e_n on the space-time slab $[-T,T]\times\mathbb{R}^3$. Defining

$$I_n = \big\{|t-t_n| \le \lambda_n^2 T_1\big\} \cap [-T,T] \quad \text{and} \quad I_n^c = \big\{|t-t_n| > \lambda_n^2 T_1\big\} \cap [-T,T],$$

we use Strichartz, Hölder, (4-14) and (4-16) to obtain

$$
\begin{aligned}
&\Big\| \int_0^t e^{i(t-s)\Delta} \nabla \sum_{k=2}^{4} N_k(\tilde u_n)(s)\, ds \Big\|_{L_t^\infty L_x^2 \cap L_t^{10} L_x^{30/13}} \\
&\quad \lesssim \|\nabla N_2(\tilde u_n)\|_{L_t^{20/19} L_x^{30/16}} + \|\nabla N_3(\tilde u_n)\|_{L_t^{5/4} L_x^{30/19}} + \|\nabla N_4(\tilde u_n)\|_{L_t^{20/13} L_x^{30/22}} \\
&\quad \lesssim \|\nabla \tilde u_n\|_{L_t^{10} L_x^{30/13}} \big\{ \|\tilde u_n\|_{L_t^{20/17} L_x^{10}(I_n\times\mathbb{R}^3)} + \|\tilde u_n\|_{L_t^{20/17} L_x^{10}(I_n^c\times\mathbb{R}^3)} \big\} \\
&\qquad + \|\nabla \tilde u_n\|_{L_t^{10} L_x^{30/13}} \|\tilde u_n\|_{L_{t,x}^{10}} \big\{ \|\tilde u_n\|_{L_t^{5/3} L_x^{10}(I_n\times\mathbb{R}^3)} + \|\tilde u_n\|_{L_t^{5/3} L_x^{10}(I_n^c\times\mathbb{R}^3)} \big\} \\
&\qquad\quad + \|\nabla \tilde u_n\|_{L_t^{10} L_x^{30/13}} \|\tilde u_n\|_{L_{t,x}^{10}}^2 \big\{ \|\tilde u_n\|_{L_t^{20/7} L_x^{10}(I_n\times\mathbb{R}^3)} + \|\tilde u_n\|_{L_t^{20/7} L_x^{10}(I_n^c\times\mathbb{R}^3)} \big\} \\
&\quad \lesssim_{\|\phi\|_{\dot H_x^1}} \sum_{k=2}^{4} \big\{ (\lambda_n^2 T_1)^{\frac{5-k}{4}} + T^{\frac{5-k}{4}} \delta \big\}.
\end{aligned}
$$

Taking δ sufficiently small depending on T and n sufficiently large, we see this contribution is acceptable.

Next we consider the contribution of the linear term appearing in e_n, again on the space-time slab $[-T,T]\times\mathbb{R}^3$. First, we observe that by Strichartz and (4-14), we have

$$
\Big\| \int_0^t e^{i(t-s)\Delta} \overline{\tilde u_n(s)}\, ds \Big\|_{L_t^2 \dot H_x^{1,6}([-T,T]\times\mathbb{R}^3)} \lesssim \|\tilde u_n\|_{L_t^1 \dot H_x^1([-T,T]\times\mathbb{R}^3)} \lesssim_{\|\phi\|_{\dot H_x^1}} T. \tag{4-18}
$$

To continue, using (4-17) we cover $\mathbb{R}$ by three disjoint intervals I_n^0 and $I_n^\pm$ such that

$$
|I_n^0| \le 2\lambda_n^2 T_2 \quad \text{and} \quad \Big\| \tilde u_n - e^{-i\gamma t} e^{i(t-t_n)\Delta} \Big[\lambda_n^{-\frac12} v_\pm\Big(\frac{\cdot}{\lambda_n}\Big)\Big] \Big\|_{L_t^\infty \dot H_x^1(I_n^\pm\times\mathbb{R}^3)} < \delta. \tag{4-19}
$$

By Strichartz, Hölder, (4-14), and (4-19), we have

$$
\Big\| \int_0^t e^{i(t-s)\Delta} \chi_{I_n^0}(s) \overline{\tilde u_n(s)}\, ds \Big\|_{L_t^\infty \dot H_x^1([-T,T]\times\mathbb{R}^3)} \lesssim \|\tilde u_n\|_{L_t^1 \dot H_x^1(I_n^0\times\mathbb{R}^3)} \lesssim_{\|\phi\|_{\dot H_x^1}} \lambda_n^2 T_2. \tag{4-20}
$$

Using the triangle inequality, Strichartz, and (4-19),

$$
\begin{aligned}
&\Big\| \int_0^t e^{i(t-s)\Delta} \chi_{I_n^\pm}(s) \overline{\tilde u_n(s)}\, ds \Big\|_{L_t^\infty \dot H_x^1([-T,T]\times\mathbb{R}^3)} \\
&\qquad \lesssim T\delta + \Big\| \int_0^t e^{i(t+t_n-2s)\Delta} \chi_{I_n^\pm}(s) e^{i\gamma s} \lambda_n^{-\frac12} \overline{v_\pm}\Big(\frac{\cdot}{\lambda_n}\Big)\, ds \Big\|_{L_t^\infty \dot H_x^1([-T,T]\times\mathbb{R}^3)}.
\end{aligned} \tag{4-21}
$$

Now for any $-T \le a < b \le T$, an application of Plancherel gives

$$
\begin{aligned}
\Big\| \int_a^b e^{is(\gamma-2\Delta)} \lambda_n^{-\frac12} \overline{v_\pm}\Big(\frac{\cdot}{\lambda_n}\Big)\, ds \Big\|_{\dot H_x^1} &= \Big\| \int_a^b e^{is(\gamma+2|\xi|^2)} |\xi| \lambda_n^{\frac52} \widehat{\overline{v_\pm}}(\xi\lambda_n)\, ds \Big\|_{L_\xi^2} \\
&\lesssim \big\| (\gamma+2|\xi|^2)^{-1} |\xi| \lambda_n^{\frac52} \widehat{\overline{v_\pm}}(\xi\lambda_n) \big\|_{L_\xi^2} \\
&\lesssim \Big\| \frac{\lambda_n^2}{2|\xi|^2+\gamma\lambda_n^2} |\xi| \widehat{\overline{v_\pm}}(\xi) \Big\|_{L_\xi^2},
\end{aligned} \tag{4-22}
$$

which converges to zero as $n \to \infty$ by the dominated convergence theorem. Collecting (4-20), (4-21), and (4-22), we obtain that

$$\left\| \int_0^t e^{i(t-s)\Delta} \overline{\tilde{u}_n(s)}\, ds \right\|_{L_t^\infty \dot{H}_x^1([-T,T]\times\mathbb{R}^3)} \lesssim_{\|\phi\|_{\dot{H}_x^1}} \lambda_n^2 T_2 + T\delta + o(1) \quad \text{as } n \to \infty.$$

Interpolating with (4-18) and taking δ sufficiently small depending on T and taking n sufficiently large, we see that the contribution of the linear term in e_n is also acceptable. This completes the proof of (4-10).

Finally, we turn to (4-11) and (4-12). For $\varepsilon > 0$, we approximate v by $\phi_\varepsilon, \psi_\varepsilon \in C_c^\infty(\mathbb{R}\times\mathbb{R}^3)$ such that

$$\|v - \phi_\varepsilon\|_{L^{10}_{t,x}(\mathbb{R}\times\mathbb{R}^3)} < \tfrac{1}{2}\varepsilon \quad \text{and} \quad \|\nabla v - \psi_\varepsilon\|_{L^{10/3}_{t,x}(\mathbb{R}\times\mathbb{R}^3)} < \tfrac{1}{2}\varepsilon$$

and take n sufficiently large so that

$$\|u_n - \tilde{u}_n\|_{L^{10}_{t,x} \cap L_t^{10/3} \dot{H}_x^{1,10/3}([-T,T]\times\mathbb{R}^3)} < \tfrac{1}{2}\varepsilon.$$

The two claims now follow easily from the triangle inequality. □

Finally we turn to the proof of Theorem 4.1.

Proof of Theorem 4.1. As mentioned above, by Lemma 4.2 and Theorem 3.3 we have that u and all of the u_n are global-in-time solutions to (1-6).

Fix $T > 0$. We will show that for any subsequence in n there exists a further subsequence so that along that subsequence, $u_n(t) \rightharpoonup u(t)$ weakly in $\dot{H}_x^1$ for all $t \in [-T, T]$. As the limit is independent of the original subsequence, this will prove the theorem.

Given a subsequence in n, we apply Proposition 4.3 to $u_n(0) - u_0$ and pass to a further subsequence to obtain the decomposition

$$u_n(0) - u_0 = \sum_{j=1}^J \phi_n^j + w_n^J \quad \text{with } \phi_n^j(x) := e^{-it_n^j\Delta}\left[(\lambda_n^j)^{-\frac{1}{2}} \phi^j\left(\frac{x - x_n^j}{\lambda_n^j}\right)\right],$$

which satisfies the conclusions of that proposition. By hypothesis, $u_n(0) - u_0 \rightharpoonup 0$ weakly in $\dot{H}_x^1$; using also (4-6) and (4-7), this implies that for all $j \geq 1$ we must have

$$w_n^J \rightharpoonup 0 \quad \text{weakly in } \dot{H}_x^1 \qquad \text{and} \qquad (\lambda_n^j)^{-1} + |t_n^j| + |x_n^j| \to \infty \quad \text{as } n \to \infty. \tag{4-23}$$

Indeed, one can first prove the divergence of the parameters by a contradiction argument. Briefly, if some $(\lambda_n^j)^{-1} + |t_n^j| + |x_n^j|$ were to remain bounded as $n \to \infty$ then one could use (4-6) and (4-7) to deduce that $\phi^j = 0$, a contradiction. Once the divergence of the parameters is established, the weak convergence of w_n^J to zero then follows.

Throughout the proof we write

$$\phi_n^0(x) := e^{-it_n^0\Delta}\left[(\lambda_n^0)^{-\frac{1}{2}} u_0\left(\frac{x - x_n^0}{\lambda_n^0}\right)\right] \quad \text{with parameters } \lambda_n^0 \equiv 1,\ t_n^0 \equiv 0,\ x_n^0 \equiv 0.$$

In view of (4-23), the decomposition

$$u_n(0) = \sum_{j=0}^{J} \phi_n^j + w_n^J$$

satisfies the conclusions of Proposition 4.3.

We next construct nonlinear profiles associated to each ϕ_n^j. If j conforms to the first scenario described in Proposition 4.3, then (4-3) and Lemma 4.2 guarantee that $\phi_n^j \in \mathcal{E}$ and moreover, $\|\nabla \operatorname{Re} \phi_n^j\|_{L_x^2} \le \delta_\gamma$ and $E(\phi_n^j) \le \frac{1}{4}\delta_\gamma$ if $\gamma \in \left(0, \frac{2}{3}\right)$. Thus by Theorem 3.3 there exists a unique solution u_n^j to (1-6) with data $u_n^j(0) = \phi_n^j$; in particular, $\|u_n^j\|_{\dot{S}^1([-T,T])} < \infty$. Note that u_n^0 is simply the solution u from the statement of Theorem 4.1.

If j conforms to the second scenario described in Proposition 4.3, we let u_n^j denote the solution to (1-6) with data $u_n^j(0) = \phi_n^j$ constructed in Proposition 4.5.

In either scenario, for all $\varepsilon > 0$ there exists $\phi_\varepsilon^j, \psi_\varepsilon^j \in C_c^\infty([-T,T] \times \mathbb{R}^3)$ such that

$$\limsup_{n\to\infty} \left\| u_n^j(t,x) - e^{-i\gamma t} (\lambda_n^j)^{-\frac{1}{2}} \phi_\varepsilon^j \left(\frac{t - t_n^j}{(\lambda_n^j)^2}, \frac{x - x_n^j}{\lambda_n^j} \right) \right\|_{L_{t,x}^{10}([-T,T]\times\mathbb{R}^3)} \le \varepsilon, \tag{4-24}$$

$$\limsup_{n\to\infty} \left\| \nabla u_n^j(t,x) - e^{-i\gamma t} (\lambda_n^j)^{-\frac{3}{2}} \psi_\varepsilon^j \left(\frac{t - t_n^j}{(\lambda_n^j)^2}, \frac{x - x_n^j}{\lambda_n^j} \right) \right\|_{L_{t,x}^{10/3}([-T,T]\times\mathbb{R}^3)} \le \varepsilon. \tag{4-25}$$

Note that the phase $e^{-i\gamma t}$ has no significance for j conforming to the first scenario described in Proposition 4.3; we simply incorporate it so as to treat both cases uniformly. For these j, both ϕ_ε^j and ψ_n^j are chosen to approximate $e^{i\gamma t} u_n^j$.

As a consequence of (4-24), (4-25), and the asymptotic orthogonality of parameters given by (4-7), for all $j \ne l$ we have

$$\|u_n^j u_n^l\|_{L_{t,x}^5} + \|u_n^j \nabla u_n^l\|_{L_t^5 L_x^{15/8}} + \|\nabla u_n^j \nabla u_n^l\|_{L_t^5 L_x^{15/13}} \to 0, \tag{4-26}$$

where all space-time norms are over $[-T,T] \times \mathbb{R}^3$.

We next claim that for $j \ge 1$ we have

$$u_n^j(t) \rightharpoonup 0 \quad \text{weakly in } \dot{H}_x^1(\mathbb{R}^3) \text{ as } n \to \infty \text{ for every } t \in [-T,T]. \tag{4-27}$$

Indeed, if j conforms to the first scenario, then (4-23) implies that $|x_n^j| \to \infty$ and hence (4-27) follows from the space-translation invariance of (1-6) together with uniqueness. If j conforms to the second scenario, then we have $\lambda_n^j \to 0$; however, as (1-6) is not scale invariant, the argument just described does not apply directly. For this case, we recall that according to the construction in Proposition 4.5, u_n^j are asymptotically close in $L_t^\infty \dot{H}_x^1$ (up to a phase factor) to rescaled solutions to the defocusing energy-critical NLS as $n \to \infty$. Using the scaling symmetry and uniqueness for (4-13), we see that these rescaled solutions converge weakly to 0 in $\dot{H}_x^1$ at each time; by construction, u_n^j inherit this property.

To continue, we define

$$u_n^J(t) = \sum_{j=0}^{J} u_n^j(t) + V^{-1} e^{-itH} V w_n^J.$$

Note that $u_n^J(0)=u_n(0)$. In what follows we will prove that for n and J sufficiently large, u_n^J is an approximate solution to (1-6) with uniform space-time bounds on $[-T,T]\times\mathbb{R}^3$. An application of Proposition 3.5 will then yield that for any $\varepsilon>0$ there exist n and J sufficiently large so that

$$\|u_n-u_n^J\|_{L_t^\infty \dot H_x^1([-T,T]\times\mathbb{R}^3)}\le\varepsilon.$$

On the other hand, using (4-23) and (4-27) and recalling $u=u_n^0$, we see that for J fixed, $u_n^J(t)-u(t)\rightharpoonup 0$ weakly in $\dot H_x^1$ for all $t\in[-T,T]$. Thus by the triangle inequality, for any $\varphi\in C_c^\infty(\mathbb{R}^3)$, we have

$$\begin{aligned}\big|\langle u_n(t)-u(t),\varphi\rangle\big| &\le \big|\langle u_n(t)-u_n^J(t),\varphi\rangle\big|+\big|\langle u_n^J(t)-u(t),\varphi\rangle\big|\\ &\le \|u_n(t)-u_n^J(t)\|_{\dot H_x^1}\|\varphi\|_{\dot H_x^{-1}}+\big|\langle u_n^J(t)-u(t),\varphi\rangle\big|\\ &\lesssim_\varphi \varepsilon+o(1)\quad \text{as } n\to\infty,\end{aligned}$$

which proves the claim in Theorem 4.1.

Thus it remains to show that for n and J sufficiently large, u_n^J are approximate solutions to (1-6) with uniform space-time bounds on $[-T,T]\times\mathbb{R}^3$.

Our first step in this direction is the following lemma.

Lemma 4.6 (finite space-time bounds). *Given $T>0$, we have*

$$\sup_J\limsup_{n\to\infty}\big[\|u_n^J\|_{L_{t,x}^{10}([-T,T]\times\mathbb{R}^3)}+\|\nabla u_n^J\|_{L_t^{10}L_x^{30/10}([-T,T]\times\mathbb{R}^3)}\big]\lesssim 1. \tag{4-28}$$

Moreover, for any $\eta>0$ there exists $J'=J'(\eta)$ sufficiently large so that

$$\limsup_{n\to\infty}\bigg[\Big\|\sum_{j=J'}^J u_n^j\Big\|_{L_{t,x}^{10}([-T,T]\times\mathbb{R}^3)}+\Big\|\sum_{j=J'}^J \nabla u_n^j\Big\|_{L_t^{10}L_x^{30/10}([-T,T]\times\mathbb{R}^3)}\bigg]\le\eta \tag{4-29}$$

uniformly in $J\ge J'$.

Proof. By the asymptotic decoupling of the $\dot H_x^1$-norm in (4-5), there exists $J_0=J_0(T)$ such that for all $j\ge J_0$ we have $\|\phi^j\|_{\dot H_x^1}\le\eta(T)$, where $\eta(T)$ is as in Corollary 3.6. In particular,

$$\|u_n^j\|_{\dot S^1([-T,T]}\lesssim_T\|\phi^j\|_{\dot H_x^1}\quad\text{for all } j\ge J_0. \tag{4-30}$$

On the space-time slab $[-T,T]\times\mathbb{R}^3$ we use Lemma 2.6 to estimate

$$\begin{aligned}\|u_n^J\|_{L_{t,x}^{10}}^2 &\lesssim \|V^{-1}e^{-itH}Vw_n^J\|_{L_{t,x}^{10}}^2+\bigg\|\Big(\sum_{j=0}^J u_n^j\Big)^2\bigg\|_{L_{t,x}^5}\\ &\lesssim_T \|w_n^J\|_{\dot H_x^1}^2+\sum_{j=0}^J\|u_n^j\|_{L_{t,x}^{10}}^2+\sum_{j\ne l}\|u_n^ju_n^l\|_{L_{t,x}^5}.\end{aligned}$$

This suffices to show that the first term on the left-hand side of (4-28) is finite. Indeed, we use (4-5) and (4-30) to bound the first two summands and (4-26) to bound the last (double) sum. An analogous argument yields that the second term on the left-hand side of (4-28) is also bounded.

To prove (4-29) one argues as above, taking $J' \geq J_0$ large enough that

$$\sum_{j \geq J'} \|\phi^j\|_{\dot{H}_x^1}^2 \lesssim \eta.$$

Note that this is possible because of (4-5). □

We next prove that the u_n^J are indeed approximate solutions to (1-6).

Lemma 4.7 (asymptotic solution to (1-6)). *We have*

$$\lim_{J \to J^*} \limsup_{n \to \infty} \big\| \nabla \big[(i\partial_t + \Delta - 2\gamma \operatorname{Re}) u_n^J - N(u_n^J) \big] \big\|_{\dot{N}^0([-T,T])} = 0.$$

Proof. Throughout the proof of the lemma, all space-time norms will be over $[-T, T] \times \mathbb{R}^3$. Writing $\tilde{w}_n^J := V^{-1} e^{-itH} V w_n^J$, we have

$$\begin{aligned} e_n^J &:= (i\partial_t + \Delta - 2\gamma \operatorname{Re}) u_n^J - N(u_n^J) \\ &= \sum_{j=0}^{J} N(u_n^j) - N\Big(\sum_{j=0}^{J} u_n^j \Big) + N\big(u_n^J - \tilde{w}_n^J\big) - N(u_n^J). \end{aligned}$$

Computations similar to those employed in the proof of Proposition 3.5 yield

$$\bigg\| \nabla \bigg[\sum_{j=0}^{J} N(u_n^j) - N\Big(\sum_{j=0}^{J} u_n^j \Big) \bigg] \bigg\|_{\dot{N}^0} \lesssim \sum_{k=2}^{5} \sum_{j \neq l} \sum_{m=0}^{J} T^{\frac{5-k}{4}} \|u_n^l \nabla u_n^j\|_{L_t^{\frac{5}{2}} L_x^{15/8}} \|u_n^m\|_{L_{t,x}^{10}}^{k-2},$$

which converges to zero as $n \to \infty$ in view of (4-26) and (4-30).

Thus, it remains to show that

$$\lim_{J \to J^*} \limsup_{n \to \infty} \big\| \nabla \big[N(u_n^J - \tilde{w}_n^J) - N(u_n^J) \big] \big\|_{\dot{N}^0([-T,T])} = 0. \tag{4-31}$$

We argue as follows: First, we estimate

$$\begin{aligned} &\big\| \nabla \big[N(u_n^J - \tilde{w}_n^J) - N(u_n^J) \big] \big\|_{\dot{N}^0([-T,T])} \\ &\quad \lesssim \|\nabla u_n^J\|_{L_t^{10} L_x^{30/13}} \|\tilde{w}_n^J\|_{L_{t,x}^{10}} \sum_{k=2}^{5} T^{\frac{5-k}{4}} \big(\|u_n^J\|_{L_{t,x}^{10}}^{k-2} + \|\tilde{w}_n^J\|_{L_{t,x}^{10}}^{k-2} \big) \\ &\quad\quad + \|\nabla \tilde{w}_n^J\|_{L_t^{10} L_x^{30/13}} \sum_{k=2}^{5} T^{\frac{5-k}{4}} \|\tilde{w}_n^J\|_{L_{t,x}^{10}}^{k-1} + \|u_n^J \nabla \tilde{w}_n^J\|_{L_t^{\frac{5}{2}} L_x^{15/8}} \sum_{k=2}^{5} T^{\frac{5-k}{4}} \|u_n^J\|_{L_{t,x}^{10}}^{k-2}. \end{aligned}$$

That the first two summands above go to zero as $n \to \infty$ and then $J \to \infty$ follows from (4-4) and Lemma 4.6. Thus, (4-31) will follow from Lemma 4.6 once we establish

$$\lim_{J \to J^*} \limsup_{n \to \infty} \|u_n^J \nabla \tilde{w}_n^J\|_{L_t^{\frac{5}{2}} L_x^{15/8}([-T,T] \times \mathbb{R}^3)} = 0. \tag{4-32}$$

We will prove that the left-hand side of (4-32) is $\lesssim \eta$ for arbitrary $\eta > 0$. By the definition of u_n^J, the triangle inequality, and Hölder, we estimate

$$\begin{aligned}
&\|u_n^J \nabla \tilde{w}_n^J\|_{L_t^5 L_x^{15/8}} \\
&\qquad \lesssim \|\tilde{w}_n^J\|_{L_{t,x}^{10}} \|\nabla \tilde{w}_n^J\|_{L_t^{10} L_x^{30/13}} + \Big\| \sum_{j=J'}^{J} u_n^j \Big\|_{L_{t,x}^{10}} \|\nabla \tilde{w}_n^J\|_{L_t^{10} L_x^{30/13}} + \Big\| \sum_{j=0}^{J'-1} u_n^j \nabla \tilde{w}_n^J \Big\|_{L_t^5 L_x^{15/8}},
\end{aligned}$$

where $J' = J'(\eta)$ is as in the statement of Lemma 4.6. Using (4-4) and (4-29), we see that the contribution of the first two summands on the right-hand side of the formula above is acceptable.

It remains to prove that

$$\lim_{J \to J^*} \limsup_{n \to \infty} \|u_n^j \nabla \tilde{w}_n^J\|_{L_t^5 L_x^{15/8}([-T,T]\times\mathbb{R}^3)} = 0 \quad \text{for each } 0 \le j < J'. \tag{4-33}$$

Assume first that $0 \le j < J'$ conforms to the first scenario in Proposition 4.3. Fix $\varepsilon > 0$. Invoking (4-24) and using the triangle inequality, Hölder, interpolation, and Corollary 2.8, we estimate

$$\begin{aligned}
\|u_n^j \nabla \tilde{w}_n^J\|_{L_t^5 L_x^{15/8}} &\le \big\| u_n^j(t,x) - e^{-i\gamma t} \phi_\varepsilon^j(t, x - x_n^j) \big\|_{L_{t,x}^{10}} \|\nabla \tilde{w}_n^J\|_{L_t^{10} L_x^{30/13}} + \|\phi_\varepsilon^j \nabla \tilde{w}_n^J(x + x_n^j)\|_{L_t^5 L_x^{15/8}} \\
&\lesssim \varepsilon + \|\phi_\varepsilon^j\|_{L_t^\infty L_x^{12}} \|\nabla \tilde{w}_n^J(x + x_n^j)\|_{L_{t,x}^2(\operatorname{supp}\phi_\varepsilon^j)}^{\frac14} \|\nabla \tilde{w}_n^J\|_{L_t^{10} L_x^{30/13}}^{\frac34} \\
&\lesssim_{\phi_\varepsilon^j} \varepsilon + \|\tilde{w}_n^J\|_{L_{t,x}^{10}}^{\frac{1}{12}} \|w_n^J\|_{\dot{H}_x^1}^{\frac16} \|\nabla \tilde{w}_n^J\|_{L_t^{10} L_x^{30/13}}^{\frac34}.
\end{aligned}$$

By (4-4), we see that (4-33) follows in this case.

Now assume that $1 \le j < J'$ conforms to the second scenario in Proposition 4.3. We split $\tilde{w}_n^J$ into low and high frequencies and estimate them separately, starting with the low-frequency piece. Fix $\varepsilon > 0$. Arguing as before, using (4-24), Hölder, and Bernstein, we estimate

$$\begin{aligned}
\|u_n^j P_{\le(\lambda_n^j)^{-1}} \nabla \tilde{w}_n^J\|_{L_t^5 L_x^{15/8}} &\lesssim \varepsilon + \Big\| (\lambda_n^j)^{-\frac12} \phi_\varepsilon^j \Big(\frac{t - t_n^j}{(\lambda_n^j)^2}, \frac{x - x_n^j}{\lambda_n^j} \Big) \Big\|_{L_t^{10} L_x^{30/13}} \|P_{\le(\lambda_n^j)^{-1}} \nabla \tilde{w}_n^J\|_{L_{t,x}^{10}} \\
&\lesssim \varepsilon + \|\phi_\varepsilon^j\|_{L_t^{10} L_x^{30/13}} \|\tilde{w}_n^J\|_{L_{t,x}^{10}}.
\end{aligned}$$

In view of (4-4), this contribution is acceptable.

We now consider the high-frequency piece. Using (2-7) we can deduce

$$\big\| P_{\ge N} - P_{\ge N} e^{it(\gamma - \Delta)} V^{-1} e^{-itH} V \big\|_{\dot{H}_x^1 \to \dot{H}_x^1} \lesssim_T N^{-2}$$

uniformly for $N \ge 1$ and $t \in [-T, T]$. Thus

$$\begin{aligned}
&\|u_n^j \nabla P_{\ge(\lambda_n^j)^{-1}} \tilde{w}_n^J\|_{L_t^5 L_x^{15/8}([-T,T]\times\mathbb{R}^3)} \\
&\qquad \lesssim_T \|u_n^j P_{\ge(\lambda_n^j)^{-1}} \nabla e^{it\Delta} w_n^J\|_{L_t^5 L_x^{15/8}([-T,T]\times\mathbb{R}^3)} + (\lambda_n^j)^2 \|u_n^j\|_{L_{t,x}^{10}([-T,T]\times\mathbb{R}^3)} \|w_n^J\|_{\dot{H}_x^1} \\
&\qquad \lesssim_T \varepsilon + \|\phi_\varepsilon^j \nabla e^{it\Delta} f_n\|_{L_t^5 L_x^{15/8}(I_n \times \mathbb{R}^3)} + o(1) \quad \text{as } n \to \infty,
\end{aligned}$$

where

$$f_n(x) = P_{\ge 1} (\lambda_n^j)^{\frac12} w_n^J(\lambda_n^j x + x_n^j) \quad \text{and} \quad I_n = \{|t - t_n^j| \le (\lambda_n^j)^2 T\}.$$

To continue, we estimate in much the same manner as for j conforming to the first scenario:

$$\begin{aligned}\|\phi_\varepsilon^j \nabla e^{it\Delta} f_n\|_{L_t^5 L_x^{15/8}(I_n\times\mathbb{R}^3)} &\lesssim \|\phi_\varepsilon^j\|_{L_t^\infty L_x^{12}} \|\nabla e^{it\Delta} f_n\|_{L^2_{t,x}(\operatorname{supp}\phi_\varepsilon^j \cap I_n\times\mathbb{R}^3)}^{\frac14} \|\nabla e^{it\Delta} f_n\|_{L_t^{10} L_x^{30/13}(I_n\times\mathbb{R}^3)}^{\frac34} \\ &\lesssim_{\phi_\varepsilon^j} \|e^{it\Delta} f_n\|_{L^{10}_{t,x}(I_n\times\mathbb{R}^3)}^{\frac{1}{12}} \|f_n\|_{\dot H^1_x}^{\frac16} \|\nabla e^{it\Delta} f_n\|_{L_t^{10} L_x^{30/13}(I_n\times\mathbb{R}^3)}^{\frac34} \\ &\lesssim_{\phi_\varepsilon^j} \|e^{it\Delta} w_n^J\|_{L^{10}_{t,x}([-T,T]\times\mathbb{R}^3)}^{\frac{1}{12}} \|w_n^J\|_{\dot H^1_x}^{\frac16} \|\nabla e^{it\Delta} w_n^J\|_{L_t^{10} L_x^{30/13}([-T,T]\times\mathbb{R}^3)}^{\frac34},\end{aligned}$$

where we have again used Corollary 2.8. Recalling (4-4), we see that the contribution of the high-frequency piece is acceptable. This completes the proof of (4-33) and hence the proof of Lemma 4.7. □

The final step in checking the hypotheses of Proposition 3.5, which will finish the proof of Theorem 4.1, is to verify that

$$\limsup_{J\to J^*} \limsup_{n\to\infty} \|u_n^J\|_{L_t^\infty \dot H_x^1([-T,T]\times\mathbb{R}^3)} \lesssim_T 1.$$

In view of Lemma 2.6, we have

$$\|u_n^J\|_{L_t^\infty \dot H_x^1([-T,T]\times\mathbb{R}^3)} \lesssim_T \|u_n^J(0)\|_{\dot H^1_x} + \|\nabla e_n^J\|_{\dot N^0([-T,T])} + \|\nabla N(u_n^J)\|_{\dot N^0([-T,T])}.$$

The requisite bounds on the right-hand side now follow from Lemmas 4.6 and 4.7. □

5. Normal form transformation

In this section we discuss the normal form transformation that we use throughout the rest of the paper. The use of normal form transformations originates in work of Shatah [1985] and has since become a widely used technique in the setting of nonlinear dispersive equations. The transformation we use is similar to the one used by Gustafson et al. [2006; 2007; 2009] in the setting of the Gross–Pitaevskii equation.

Suppose u is a solution to (1-6). As mentioned in the Introduction, the quadratic terms in the nonlinearity are the most problematic when it comes to questions of long-time behavior; in particular, the worst terms are those containing $u_2 = \operatorname{Im} u$, since in the diagonal variables we have $u_2 = U^{-1} v_2$. We would like to find a normal form transformation that eliminates if not all, at least the worst quadratic terms.

To this end, we let $B_1[\cdot,\cdot]$ and $B_2[\cdot,\cdot]$ be arbitrary bilinear Fourier multiplier operators defined as in (2-1), with symmetric real-valued symbols $B_1(\xi_1,\xi_2)$ and $B_2(\xi_1,\xi_2)$. Then

$$\tilde u := u + B_1[u_1,u_1] + B_2[u_2,u_2]$$

satisfies the equation

$$\begin{aligned}(i\partial_t+\Delta)\tilde u - 2\gamma\tilde u_1 = {} & \\ & (3\gamma+4)u_1^2 - (2\gamma-\Delta)B_1[u_1,u_1] && \text{(5-1)}\\ & + \gamma u_2^2 - (2\gamma-\Delta)B_2[u_2,u_2] && \text{(5-2)}\\ & + 2i\big(\gamma u_1u_2 + B_1[u_1,-\Delta u_2] - B_2[u_2,(2\gamma-\Delta)u_1]\big) && \text{(5-3)}\\ & + \text{cubic and higher order terms.}\end{aligned}$$

While the symmetry of B_1 and B_2 makes it impossible to eliminate *all* of the quadratic terms, we see that if we choose

$$B_2(\xi_1,\xi_2)=\gamma(2\gamma+|\xi_1+\xi_2|^2)^{-1},\quad \text{i.e., } B_2[f,g]=\gamma\langle\nabla\rangle^{-2}(fg),$$

then (5-2) $=0$. This allows us to eliminate the worst quadratic term, namely, the one containing two copies of u_2. Moreover, choosing $B_1=B_2$ we get

$$\tilde{u}=u+\gamma\langle\nabla\rangle^{-2}|u|^2,$$

with

$$(5\text{-}1)=(2\gamma+4)u_1^2 \quad\text{and}\quad (5\text{-}3)=-4i\gamma\langle\nabla\rangle^{-2}\nabla\cdot[u_1\nabla u_2].$$

The derivative appearing in front of u_2 is a welcome addition in light of the problem at low frequencies.

Similarly one can compute the higher-order terms. In general, one finds that for $k\in\{3,4,5\}$ the terms of order k are given by

$$N_k(u)+2i\big\{B_1[u_1,\operatorname{Im}(N_{k-1}(u))]-B_2[u_2,\operatorname{Re}(N_{k-1}(u))]\big\},$$

where the N_k are as in (1-6). Notice that there are no sixth-order terms, since $B_1=B_2$ and $u_1\operatorname{Im}(N_5(u))=u_2\operatorname{Re}(N_5(u))$.

Finally, we employ the transformation $Vu=u_1+iUu_2$ to diagonalize the equation. Our normal form transformation is therefore given by

$$z:=M(u):=Vu+\gamma\langle\nabla\rangle^{-2}|u|^2, \tag{5-4}$$

and z satisfies the equation

$$(i\partial_t-H)z=N_z(u) \tag{5-5}$$

with

$$\begin{aligned}\operatorname{Re}[N_z(u)]&=U\operatorname{Re}[N(u)-\gamma|u|^2]\\&=U\big[(2\gamma+4)u_1^2+(\gamma+8)u_1^3+(\gamma+4)u_1u_2^2+(5u_1^4+6u_1^2u_2^2+u_2^4)+|u|^4u_1)\big],\\ \operatorname{Im}[N_z(u)]&=-\frac{\nabla}{\langle\nabla\rangle^2}\cdot\big[4\gamma u_1\nabla u_2+\nabla(\gamma|u|^2u_2+q^2u_2)\big]\\&=-\frac{\nabla}{\langle\nabla\rangle^2}\cdot\big[4\gamma u_1\nabla u_2\big]+U^2\big[(\gamma+4)u_1^2u_2+\gamma u_2^3+4u_1u_2|u|^2+|u|^4u_2)\big].\end{aligned}$$

We should briefly pause to point out the improvements present in (5-5) with respect to (1-6). Firstly, (5-5) does not contain a quadratic term involving two copies of u_2. Secondly, the remaining quadratic terms involving u_2 exhibit a derivative of this problematic term. Lastly, all the remaining terms appear with a derivative at low frequencies, which is helpful throughout.

We next discuss the invertibility of the transformation (5-4). Note that by using the definition of $\langle\nabla\rangle^{-2}$ we can rewrite the transformation as

$$M(u)=U^2u_1+\gamma\langle\nabla\rangle^{-2}q+iUu_2, \tag{5-6}$$

where $q=q(u)=2u_1+|u|^2$.

This normal form transformation is a homeomorphism from a neighborhood of zero in $\mathcal{E}$ onto a neighborhood of zero in H^1_x. To prove this, we make use of the neighborhoods

$$\mathcal{M}_{E_0,\varepsilon_0} := \{u \in \mathcal{E} : E(u) \le E_0^2,\ \|u\|_{L^6_x} \le \varepsilon_0\},$$
$$\mathcal{N}_{E_0',\varepsilon_0'} := \{f \in H^1_x : \|f\|_{H^1_x} \le CE_0',\ \|f\|_{L^6_x} \le C\varepsilon_0'\},$$

where C denotes an absolute constant depending on γ.

Proposition 5.1. *Fix $E_0 > 0$ and $\varepsilon_0 > 0$.*

(i) *If $\gamma \in \left(\frac{2}{3}, 1\right)$, then $M : \mathcal{M}_{E_0,\varepsilon_0} \to \mathcal{N}_{E_0,\varepsilon_0+\varepsilon_0^2}$ continuously.*

(ii) *If $\gamma = \frac{2}{3}$, then $M : \mathcal{M}_{E_0,\varepsilon_0} \to \mathcal{N}_{E_0+E_0^3,\varepsilon_0+\varepsilon_0^2}$ continuously.*

(iii) *If $\gamma \in \left(0, \frac{2}{3}\right)$, then $M : \mathcal{M}_{E_0,\varepsilon_0} \cap \{\|\nabla u_1\|_2^2 \le \delta_\gamma\} \to \mathcal{N}_{E_0,\varepsilon_0+\varepsilon_0^2}$ continuously, where δ_γ is as in Lemma 3.2.*

(iv) *Given $E_1 > 0$, there exists $\varepsilon_1 = \varepsilon_1(E_1)$ and a continuous mapping*

$$R : \mathcal{N}_{E_1,\varepsilon_1} \to \mathcal{E}$$

such that $M \circ R = \mathrm{Id}$ on $\mathcal{N}_{E_1,\varepsilon_1}$ and $\|R(f)\|_{\mathcal{E}} \lesssim E_1$ for $f \in \mathcal{N}_{E_1,\varepsilon_1}$.

(v) *Suppose $\gamma \ge \frac{2}{3}$. Given $E_2 > 0$, there exists $\varepsilon_2 = \varepsilon_2(E_2)$ so that M is a homeomorphism of $\mathcal{M}_{E_2,\varepsilon_2}$ onto a subset of $H^1_x(\mathbb{R}^3)$ and has inverse R. In particular, M is injective on $\mathcal{M}_{E_2,\varepsilon_2}$. If $\gamma < \frac{2}{3}$, then the analogous assertions hold on $\mathcal{M}_{E_2,\varepsilon_2} \cap \{\|\nabla u_1\|_2^2 \le \delta_\gamma\}$.*

Remark 5.2. We warn the reader that just because $M(u)$ is small in L^6_x, one cannot guarantee that $u = (R \circ M)(u)$. However, this would follow if u were sufficiently small in L^6_x. This subtlety contributes nontrivially to the complexity of the proof of Theorem 1.1.

Proof. The proofs of the first three claims parallel one another closely. We will only present the details when $\gamma \in \left(\frac{2}{3}, 1\right)$.

Let $u \in \mathcal{M}_{E_0,\varepsilon_0}$. Recall from Lemma 3.1 that

$$\|u\|_{\mathcal{E}}^2 \lesssim E(u) \lesssim E_0^2.$$

We first show that $M(u) \in \mathcal{N}_{E_0,\varepsilon_0+\varepsilon_0^2}$. Using the representation (5-6), we estimate

$$\|M(u)\|_{H^1_x} \lesssim \|U^2 u_1\|_{H^1_x} + \|\langle\nabla\rangle^{-2} q\|_{H^1_x} + \|U u_2\|_{H^1_x} \lesssim \|u\|_{\dot{H}^1_x} + \|q\|_{L^2_x} \lesssim E_0.$$

Using the representation (5-4) and Sobolev embedding, we estimate

$$\begin{aligned}
\|M(u)\|_{L^6_x} &\lesssim \|u_1\|_{L^6_x} + \|U u_2\|_{L^6_x} + \big\|\langle\nabla\rangle^{-2}|u|^2\big\|_{L^6_x} \\
&\lesssim \|u\|_{L^6_x} + \big\||\nabla|^{1/2}\langle\nabla\rangle^{-2}|u|^2\big\|_{L^3_x} \\
&\lesssim \|u\|_{L^6_x} + \|u\|_{L^6_x}^2 \\
&\lesssim \varepsilon_0 + \varepsilon_0^2.
\end{aligned}$$

Collecting these estimates, we conclude $M(u) \in \mathcal{N}_{E_0,\varepsilon_0+\varepsilon_0^2}$.

To prove the continuity of M, we note that for $u, v \in \mathcal{E}$ we may write

$$M(u) - M(v) = U^2(u_1 - v_1) + \gamma\langle\nabla\rangle^{-2}[q(u) - q(v)] + iU(u_2 - v_2).$$

Estimating as above we find

$$\|M(u) - M(v)\|_{H^1_x} \lesssim d_{\mathcal{E}}(u, v).$$

We turn now to the fourth claim in the statement of the proposition. Let $f \in \mathcal{N}_{E_1,\varepsilon_1}$. We aim to show that for $\varepsilon_1 = \varepsilon_1(E_1) > 0$ sufficiently small, we can find a unique $u \in \mathcal{E}$ such that $M(u) = f$, that is,

$$\begin{cases} u_2 = U^{-1} f_2, \\ u_1 = f_1 - \gamma\langle\nabla\rangle^{-2}[U^{-1} f_2]^2 - \gamma\langle\nabla\rangle^{-2}u_1^2. \end{cases}$$

To this end, we define

$$R_f(u_1) := f_1 - \gamma\langle\nabla\rangle^{-2}[U^{-1} f_2]^2 - \gamma\langle\nabla\rangle^{-2}u_1^2.$$

We will show that for $\varepsilon_1 = \varepsilon_1(E_1)$ sufficiently small, R_f is a contraction on

$$B := \bigl\{u_1 \in \dot H^1_x : \|u_1\|_{\dot H^1_x} \le CE_1,\ \|u_1\|_{L^6_x} \le C(\varepsilon_1 + \varepsilon_1^{\frac12} E_1^{\frac32})\bigr\}$$

with respect to the metric $d(u_1, v_1) = \|u_1 - v_1\|_{\dot H^1_x}$, where C denotes an absolute constant depending on γ.

We first show that $R_f : B \to B$. We have

$$\|R_f(u_1)\|_{L^6_x} \lesssim \|f_1\|_{L^6_x} + \bigl\|\langle\nabla\rangle^{-2}[U^{-1} f_2]^2\bigr\|_{L^6_x} + \|\langle\nabla\rangle^{-2}u_1^2\|_{L^6_x}. \tag{5-7}$$

The first term in (5-7) is controlled by ε_1 by assumption. For the second term in (5-7), we use Sobolev embedding, Bernstein, and interpolation to estimate

$$\begin{aligned}\|\langle\nabla\rangle^{-2}[U^{-1} f_2]^2\|_{L^6_x} &\lesssim \left\|\frac{\nabla}{\langle\nabla\rangle^2}[P_{\mathrm{lo}}U^{-1} f_2]^2\right\|_{L^2_x} + \left\|\frac{\nabla}{\langle\nabla\rangle^2}\bigl[(P_{\mathrm{hi}}U^{-1} f_2)\emptyset(U^{-1} f_2)\bigr]\right\|_{L^2_x} \\ &\lesssim \|\nabla P_{\mathrm{lo}}U^{-1} f_2\|_{L^3_x}\|U^{-1} f_2\|_{L^6_x} + \|P_{\mathrm{hi}}U^{-1} f_2\|_{L^3_x}\|U^{-1} f_2\|_{L^6_x} \\ &\lesssim \|f\|_{L^6_x}^{\frac12}\|f\|_{H^1_x}^{\frac32}.\end{aligned}$$

For the third term in (5-7), we have

$$\|\langle\nabla\rangle^{-2}u_1^2\|_{L^6_x} \lesssim \bigl\||\nabla|^{\frac12}\langle\nabla\rangle^{-2}u_1^2\bigr\|_{L^3_x} \lesssim \|u_1\|_{L^6_x}^2.$$

Thus, for $u_1 \in B$ and $\varepsilon_1 = \varepsilon_1(E_1)$ sufficiently small we obtain

$$\|R_f(u_1)\|_{L^6_x} \le C(\varepsilon_1 + \varepsilon_1^{\frac12} E_1^{\frac32}).$$

To continue, we estimate

$$\|R_f(u_1)\|_{\dot H^1_x} \lesssim \|f_1\|_{\dot H^1_x} + \bigl\|\langle\nabla\rangle^{-2}\bigl[U^{-1} f_2\bigr]^2\bigr\|_{\dot H^1_x} + \|\langle\nabla\rangle^{-2}u_1^2\|_{\dot H^1_x}. \tag{5-8}$$

The first term in (5-8) is controlled by E_1 by assumption. For the second term in (5-8), we argue as above to find

$$\|\langle\nabla\rangle^{-2}[U^{-1}f_2]^2\|_{\dot H^1_x}\lesssim\left\|\frac{\nabla}{\langle\nabla\rangle^2}[P_{\mathrm{lo}}U^{-1}f_2]^2\right\|_{L^2_x}+\left\|\frac{\nabla}{\langle\nabla\rangle^2}\big[(P_{\mathrm{hi}}U^{-1}f_2)\textrm{Ø}(U^{-1}f_2)\big]\right\|_{L^2_x}\lesssim\|f\|_{L^6_x}^{\frac12}\|f\|_{H^1_x}^{\frac32}.$$

For the third term in (5-8) we estimate

$$\|\langle\nabla\rangle^{-2}u_1^2\|_{\dot H^1_x}\lesssim\big\||\nabla|^{\frac32}\langle\nabla\rangle^{-2}u_1^2\big\|_{L^{3/2}_x}\lesssim\|u_1\|_{L^6_x}\|\nabla u_1\|_{L^2_x}.$$

Thus for $u_1\in B$ and $\varepsilon_1=\varepsilon_1(E_1)$ sufficiently small we have

$$\|R_f(u_1)\|_{\dot H^1_x}\le CE_1.$$

Collecting these estimates, we conclude that $R_f:B\to B$.

Next we show that R_f is a contraction with respect to the $\dot H^1_x$-norm. We first use Sobolev embedding, Bernstein, and interpolation to estimate

$$\begin{aligned}
&\left\|\frac{1}{\langle\nabla\rangle^2}\big[(u_1+v_1)(u_1-v_1)\big]\right\|_{\dot H^1_x}\\
&\qquad\lesssim\left\|\frac{|\nabla|^{\frac32}}{\langle\nabla\rangle^2}\big[(u_1+v_1)P_{\mathrm{hi}}(u_1-v_1)\big]\right\|_{L^{3/2}_x}+\left\|\frac{1}{\langle\nabla\rangle^2}\big[P_{\mathrm{lo}}(u_1+v_1)\,P_{\mathrm{lo}}(u_1-v_1)\big]\right\|_{\dot H^1_x}\\
&\qquad\qquad+\left\|\frac{1}{\langle\nabla\rangle^2}\big[P_{\mathrm{hi}}(u_1+v_1)\,P_{\mathrm{lo}}(u_1-v_1)\big]\right\|_{\dot H^1_x}\\
&\qquad\lesssim\|u_1+v_1\|_{L^6_x}\|P_{\mathrm{hi}}(u_1-v_1)\|_{L^2_x}+\|\nabla P_{\mathrm{lo}}(u_1+v_1)\|_{L^3_x}\|u_1-v_1\|_{L^6_x}\\
&\qquad\qquad+\|u_1+v_1\|_{L^6_x}\|\nabla P_{\mathrm{lo}}(u_1-v_1)\|_{L^3_x}+\|P_{\mathrm{hi}}(u_1+v_1)\|_{L^3_x}\|u_1-v_1\|_{L^6_x}\\
&\qquad\lesssim\big(\|u_1\|_{L^6_x}^{\frac12}\|u_1\|_{\dot H^1_x}^{\frac12}+\|v_1\|_{L^6_x}^{\frac12}\|v_1\|_{\dot H^1_x}^{\frac12}\big)\|u_1-v_1\|_{\dot H^1_x}.
\end{aligned}\tag{5-9}$$

In particular, for $\varepsilon_1=\varepsilon_1(E_1)$ sufficiently small we deduce that

$$\|R_f(u_1)-R_f(v_1)\|_{\dot H^1_x}\le\tfrac12\|u_1-v_1\|_{\dot H^1_x}.$$

Therefore, by the contraction mapping theorem there exists a unique $u_1\in B$ such that $R_f(u_1)=u_1$. We define $R(f):=u_1+iU^{-1}f_2$. By construction, we have $M(R(f))=f$.

It remains to see that $u:=R(f)\in\mathcal{E}$ with $\|u\|_{\mathcal{E}}\lesssim E_1$. As $u_1\in B$, we have

$$\|u\|_{\dot H^1_x}\lesssim\|u_1\|_{\dot H^1_x}+\|U^{-1}f_2\|_{\dot H^1_x}\lesssim E_1+\|f_2\|_{H^1_x}\lesssim E_1.$$

Moreover, by Hölder,

$$\begin{aligned}
\|q(u)\|_{L^2_x}=\|2f_1+U^2|u|^2\|_{L^2_x}&\lesssim\|f\|_{L^2_x}+\|U(|u|^2)\|_{L^2_x}\\
&\lesssim E_1+\|\nabla\textrm{Ø}[(P_{\mathrm{lo}}u)^2]\|_{L^2_x}+\|\textrm{Ø}(uP_{\mathrm{hi}}u)\|_{L^2_x}\\
&\lesssim E_1+\|\nabla P_{\mathrm{lo}}u\|_{L^2_x}\|P_{\mathrm{lo}}u\|_{L^\infty_x}+\|u\|_{L^6_x}\|P_{\mathrm{hi}}u\|_{L^3_x}\\
&\lesssim E_1+\|u\|_{\dot H^1_x}\big[\|P_{\mathrm{lo}}u\|_{L^\infty_x}+\|P_{\mathrm{hi}}u\|_{L^3_x}\big].
\end{aligned}$$

Using Bernstein, Hölder, and interpolation, we estimate

$$\|P_{\mathrm{lo}}u\|_{L_x^\infty} \lesssim \|u_1\|_{L_x^6} + \|P_{\mathrm{lo}}U^{-1}f_2\|_{L_x^{10}} \lesssim \|u_1\|_{L_x^6} + \|f_2\|_{L_x^{30/13}} \lesssim \|u_1\|_{L_x^6} + \|f_2\|_{L_x^6}^{\frac{1}{5}}\|f_2\|_{L_x^2}^{\frac{4}{5}}$$

and

$$\|P_{\mathrm{hi}}u\|_{L_x^3} \lesssim \|P_{\mathrm{hi}}u_1\|_{L_x^3} + \|P_{\mathrm{hi}}f_2\|_{L_x^3} \lesssim \|u_1\|_{L_x^6}^{\frac{1}{2}}\|u_1\|_{\dot H_x^1}^{\frac{1}{2}} + \|f_2\|_{L_x^6}^{\frac{1}{2}}\|f_2\|_{\dot H_x^1}^{\frac{1}{2}}.$$

Taking $\varepsilon_1 = \varepsilon_1(E_1)$ sufficiently small, this proves $\|q(u)\|_{L_x^2} \lesssim E_1$.

To complete the proof of the proposition, it remains to address part (v). From (5-4) and (5-9), we see that M is injective on $\mathcal{M}_{E_2,\varepsilon_2}$ provided ε_2 is sufficiently small depending on E_2. By shrinking ε_2, if necessary, we can further ensure that $M(\mathcal{M}_{E_2,\varepsilon_2})$ is contained in a region where R is defined (this relies on all the other parts of the proposition). It then follows that M is a homeomorphism on $M(\mathcal{M}_{E_2,\varepsilon_2})$ with inverse R. □

The last result of this section relates the energy and the inverse of the normal form transformation; this will be useful in the proof of Theorem 1.1.

Lemma 5.3. *Let $\{z_n\}_{n\geq 1} \subset H_x^1$ be uniformly bounded and assume that $z_n \to 0$ in L_x^6. Then*

$$E(R(z_n)) = \tfrac{1}{2}\|z_n\|_{H_x^1}^2 + o(1) \quad \text{as } n \to \infty.$$

Proof. By Proposition 5.1 (and its proof), we have that $R(z_n)$ exists for n large and

$$\limsup_n \|R(z_n)\|_{\mathcal{E}} \lesssim 1 \quad \text{and} \quad \lim_{n\to\infty} \|\operatorname{Re} R(z_n)\|_{L_x^6} = 0. \tag{5-10}$$

We first claim that

$$R(z_n) = V^{-1}z_n + o(1) \quad \text{in } \dot H_x^1 \text{ as } n \to \infty. \tag{5-11}$$

Indeed, from the construction of R via the fixed point argument in Proposition 5.1, this amounts to proving that

$$\|\langle\nabla\rangle^{-2}[U^{-1}\operatorname{Im} z_n]^2\|_{\dot H_x^1} + \|\langle\nabla\rangle^{-2}[\operatorname{Re} R(z_n)]^2\|_{\dot H_x^1} = o(1) \quad \text{as } n \to \infty.$$

To see this, we use the decomposition

$$[U^{-1}\operatorname{Im} z_n]^2 = [P_{\mathrm{lo}}U^{-1}\operatorname{Im} z_n]^2 + \emptyset\big[(U^{-1}\operatorname{Im} z_n)P_{\mathrm{hi}}U^{-1}\operatorname{Im} z_n\big] \tag{5-12}$$

together with Bernstein, Hölder, (5-10), and the hypotheses of the lemma to estimate

$$\begin{aligned}
\big\|\langle\nabla\rangle^{-2}[U^{-1}\operatorname{Im} z_n]^2\big\|_{\dot H_x^1} &\lesssim \big\||\nabla|[P_{\mathrm{lo}}U^{-1}\operatorname{Im} z_n]^2\big\|_{L_x^2} + \big\|(U^{-1}\operatorname{Im} z_n)P_{\mathrm{hi}}U^{-1}\operatorname{Im} z_n\big\|_{L_x^2} \\
&\lesssim \|\operatorname{Im} z_n\|_{L_x^3}\|U^{-1}\operatorname{Im} z_n\|_{L_x^6} \\
&\lesssim \|z_n\|_{L_x^6}^{\frac{1}{2}}\|z_n\|_{L_x^2}^{\frac{1}{2}}\|z_n\|_{H_x^1} = o(1) \quad \text{as } n \to \infty,
\end{aligned}$$

$$\begin{aligned}
\big\|\langle\nabla\rangle^{-2}[\operatorname{Re} R(z_n)]^2\big\|_{\dot H_x^1} &\lesssim \big\||\nabla|^{\frac{3}{2}}\langle\nabla\rangle^{-2}[\operatorname{Re} R(z_n)]^2\big\|_{L_x^{3/2}} \\
&\lesssim \|\nabla \operatorname{Re} R(z_n)\|_{L_x^2}\|\operatorname{Re} R(z_n)\|_{L_x^6} = o(1) \quad \text{as } n \to \infty.
\end{aligned}$$

This completes the proof of (5-11).

We now turn our attention to the terms in the formula for $E(R(z_n))$ containing $q(R(z_n))$. Using the representation (5-6), we observe that

$$M(u) = \tfrac{1}{2}q(u) - \tfrac{1}{2}U^2(|u|^2) + iU\operatorname{Im} u \quad \text{and so} \quad q(R(z_n)) = 2\operatorname{Re} z_n + U^2(|R(z_n)|^2).$$

We next claim that

$$q(R(z_n)) = 2\operatorname{Re} z_n + o(1) \quad \text{in } L_x^2 \text{ as } n \to \infty. \tag{5-13}$$

To prove this, we note that $\operatorname{Im} R(z_n) = U^{-1}\operatorname{Im} z_n$ and use the decomposition (5-12), as well as the analogous decomposition for $\operatorname{Re} R(z_n)$. Arguing as for (5-11), we estimate

$$\|U^2[U^{-1}\operatorname{Im} z_n]^2\|_{L_x^2} \lesssim \big\||\nabla|[P_{\mathrm{lo}}U^{-1}\operatorname{Im} z_n]^2\big\|_{L_x^2} + \big\|(U^{-1}\operatorname{Im} z_n)P_{\mathrm{hi}}U^{-1}\operatorname{Im} z_n\big\|_{L_x^2} = o(1),$$

$$\begin{aligned}\|U^2[\operatorname{Re} R(z_n)]^2\|_{L_x^2} &\lesssim \big\||\nabla|[P_{\mathrm{lo}}\operatorname{Re} R(z_n)]^2\big\|_{L_x^{3/2}} + \big\|(\operatorname{Re} R(z_n))P_{\mathrm{hi}}\operatorname{Re} R(z_n)\big\|_{L_x^2}\\ &\lesssim \|\nabla \operatorname{Re} R(z_n)\|_{L_x^2}\|\operatorname{Re} R(z_n)\|_{L_x^6} + \|\operatorname{Re} R(z_n)\|_{L_x^6}\|P_{\mathrm{hi}}\operatorname{Re} R(z_n)\|_{L_x^3}\\ &\lesssim \|\nabla R(z_n)\|_{L_x^2}\|\operatorname{Re} R(z_n)\|_{L_x^6} = o(1) \quad \text{as } n\to\infty.\end{aligned}$$

This completes the proof of (5-13).

Finally, we note that

$$q(R(z_n)) = o(1) \quad \text{in } L_x^3 \text{ as } n \to \infty. \tag{5-14}$$

Indeed, arguing as above we find

$$\|\operatorname{Re} z_n\|_{L_x^3} \lesssim \|z_n\|_{L_x^2}^{\frac12}\|z_n\|_{L_x^6}^{\frac12} = o(1) \quad \text{as } n\to\infty,$$

$$\|U^2[\operatorname{Re} R(z_n)]^2\|_{L_x^3} \lesssim \|\operatorname{Re} R(z_n)\|_{L_x^6}^2 = o(1) \quad \text{as } n\to\infty,$$

$$\begin{aligned}\|U^2[\operatorname{Im} R(z_n)]^2\|_{L_x^3} &\lesssim \big\||\nabla|[P_{\mathrm{lo}}U^{-1}\operatorname{Im} z_n]^2\big\|_{L_x^3} + \big\|(U^{-1}\operatorname{Im} z_n)P_{\mathrm{hi}}U^{-1}\operatorname{Im} z_n\big\|_{L_x^3}\\ &\lesssim \|\operatorname{Im} z_n\|_{L_x^6}\big\||\nabla|^{-1}\operatorname{Im} z_n\big\|_{L_x^6} + \|U^{-1}\operatorname{Im} z_n\|_{L_x^6}\|\operatorname{Im} z_n\|_{L_x^6}\\ &\lesssim \|z_n\|_{L_x^6}\|z_n\|_{H_x^1} = o(1) \quad \text{as } n\to\infty.\end{aligned}$$

Putting together (5-11), (5-13), and (5-14) completes the proof of the lemma. □

6. Proof of the main result

In this section we prove the main result, Theorem 1.1. As discussed in the Introduction, the proof is based off of a strategy of Nakanishi; see especially Theorem 1.3 and the sketch of proof thereafter. For the convenience of the reader, we restate the main theorem here.

Theorem 6.1. *Suppose $\gamma \in \big[\frac{2}{3}, 1\big)$. For any $u_+ \in H^1_{\mathrm{real}} + i\dot H^1_{\mathrm{real}}$, there exists a global solution $u \in C(\mathbb{R}; \mathcal{E})$ to (1-6) such that*

$$\lim_{t\to\infty}\|u(t) - u_{\mathrm{lin}}(t)\|_{\dot H_x^1} = 0, \tag{6-1}$$

where $u_{\mathrm{lin}}(t) := V^{-1}e^{-itH}Vu_+$. *Moreover,*

$$\lim_{t\to\infty} d_{\mathcal{E}}\big(u(t), u_{\mathrm{lin}}(t) - \gamma\langle\nabla\rangle^{-2}|u_{\mathrm{lin}}(t)|^2\big) = 0. \tag{6-2}$$

For $\gamma \in \big(0, \frac{2}{3}\big)$, *these conclusions hold if* $\|u_+\|_{H^1_{\mathrm{real}}+i\dot H^1_{\mathrm{real}}}$ *is sufficiently small.*

Proof of Theorem 6.1. Let $u_+ \in H^1_{\mathrm{real}} + i\dot H^1_{\mathrm{real}}$. We define $z_+ = Vu_+ \in H^1_x$ and we let $E_0 := \|z_+\|_{H^1_x}$.

We first claim that

$$\lim_{t\to\infty} \|e^{-itH}z_+\|_{L^6_x} = 0. \tag{6-3}$$

Indeed, given $\eta > 0$ we may find $\varphi \in \mathcal{S}(\mathbb{R}^3)$ such that $\|z_+ - \varphi\|_{\dot H^1_x} < \eta$. Using the dispersive estimate (2-3) and Sobolev embedding, we find

$$\|e^{-itH}z_+\|_{L^6_x} \lesssim \|e^{-itH}\varphi\|_{L^6_x} + \|z_+ - \varphi\|_{\dot H^1_x} \lesssim |t|^{-1}\|\varphi\|_{L^{6/5}_x} + \eta,$$

which yields (6-3).

Next, we choose ε_0 sufficiently small depending on E_0 as in Proposition 5.1. By (6-3), there exists $T_0 > 0$ such that $e^{-itH}z_+ \in \mathcal{N}_{E_0,\varepsilon_0}$ for $t \geq T_0$; thus for any $T \geq T_0$, we may define $R(e^{-iTH}z_+) \in \mathcal{E}$ so that $M(R(e^{-iTH}z_+)) = e^{-iTH}z_+$, with $\|R(e^{-iTH}z_+)\|_{\mathcal{E}} \lesssim E_0$.

By Theorem 3.3, there exists a global solution $u^T \in C(\mathbb{R}; \mathcal{E})$ to (1-6) with $u^T(T) = R(e^{-iTH}z_+)$. Note that when $\gamma \in \big(0, \frac{2}{3}\big)$, we require E_0 to be sufficiently small to guarantee that

$$\|\nabla \operatorname{Re}(u^T(0))\|^2_{L^2_x} \leq \delta_\gamma \quad \text{and} \quad E(u^T(0)) \leq \tfrac{1}{4}\delta_\gamma \tag{6-4}$$

uniformly in T, where δ_γ is as in Theorem 3.3. We define

$$q^T := q(u^T) = 2u_1^T + |u^T|^2 \quad \text{and} \quad z^T := M(u^T).$$

Note that (u^T, z^T) solves (5-4)–(5-5) with $z^T(T) = e^{-iTH}z_+$. Furthermore, we have

$$\|z^T(t)\|_{H^1_x} + \|u^T(t)\|_{\dot H^1_x} + \|q^T(t)\|_{L^2_x\cap L^3_x} + \|u_1^T(t)\|_{L^3_x\cap L^6_x} \lesssim_{E_0} 1 \tag{6-5}$$

uniformly in t and T.

As a consequence of (6-5), there exists a sequence $T_n \to \infty$ and a function $u_0 \in \dot H^1_x$ such that $u^{T_n}(0) \rightharpoonup u_0$ weakly in $\dot H^1_x$. As (6-5) and (6-4) imply that $\{u^{T_n}(0)\}$ satisfy the hypotheses of Theorem 4.1, we may apply this theorem to deduce that

$$u^{T_n}(t) \rightharpoonup u^\infty(t) \quad \text{weakly in } \dot H^1_x \text{ for all } t \in \mathbb{R}, \tag{6-6}$$

where $u^\infty \in C(\mathbb{R}; \mathcal{E})$ denotes the solution to (1-6) with initial data $u^\infty(0) = u_0 \in \mathcal{E}$.

We define $z^\infty := M(u^\infty)$ and note that (u^∞, z^∞) solves (5-4)–(5-5). We will prove that u^∞ is a solution to (1-6) that satisfies the conclusions of Theorem 1.1. A first step in this direction is the following weak convergence result.

Proposition 6.2. *We have*

$$e^{itH}z^\infty(t) \rightharpoonup z_+ \quad \textit{weakly in } H^1_x \textit{ as } t \to \infty.$$

Assuming Proposition 6.2 for now, we proceed with the proof of Theorem 1.1. We begin by upgrading the weak convergence from Proposition 6.2 to strong convergence, namely,

$$\lim_{t\to\infty} \|z^\infty(t) - e^{-itH} z_+\|_{H^1_x} = 0. \tag{6-7}$$

Using Lemma 4.2 combined with (6-6) and Lemma 5.3 combined with (6-3), we can first write

$$E(u^\infty) \le \liminf_{n\to\infty} E(u^{T_n}) = \liminf_{n\to\infty} E(R(e^{-iT_nH} z_+)) = \tfrac{1}{2}\|z_+\|^2_{H^1_x}. \tag{6-8}$$

At this moment, it is tempting to attempt a Radon–Riesz-style argument. Recall that the Radon–Riesz theorem says that if $x_n \rightharpoonup x$ weakly in some Banach space X and $\limsup F(x_n) \le F(x)$ for some uniformly convex function $F : X \to \mathbb{R}$, then $x_n \to x$ in norm. (This is most often quoted in the case of a uniformly convex Banach space with F being the norm.)

The ideas just sketched were adapted beautifully to the Gross–Pitaevskii setting treated in [Gustafson et al. 2009]. As discussed in the Introduction, those authors exploit

$$E_{GP}(u) = \tfrac{1}{2}\|M(u)\|^2_{H^1_x} + \tfrac{1}{4}\|U|u|^2\|^2_{L^2_x} \ge \tfrac{1}{2}\|M(u)\|^2_{H^1_x},$$

which holds under no additional hypotheses. As also discussed there (see (1-23), in particular) the energy functional for the cubic-quintic problem admits no such global inequality. Correspondingly, we need to keep track of the structure of $z^\infty(t_n)$ as $t_n \to \infty$ and then demonstrate the requisite coercivity is available in this particular limiting regime. To achieve this goal we will use the following lemma. Note that the result on $\tilde{E}$ plays a key role in controlling the kinetic energy of the real part when $\gamma < \frac{2}{3}$.

Lemma 6.3. *Let $\{u_n\}_{n\ge1} \subset \mathcal{E}$ be uniformly bounded. Assume that we may write $u_n = \xi_n + r_n$, where ξ_n satisfies*

$$\sup_n \|\xi_n\|_{H^1_{\mathrm{real}} + i\dot{H}^1_{\mathrm{real}}} \lesssim 1 \quad \text{and} \quad \lim_{n\to\infty} \|\xi_n\|_{L^3_x \cap L^6_x} = 0.$$

Then

$$E(u_n) = E(r_n) + \tfrac{1}{2}\|V\xi_n\|^2_{H^1_x} + \mathrm{Re}\langle M(r_n), V\xi_n\rangle_{H^1_x} + o(1) \quad \text{as } n\to\infty. \tag{6-9}$$

Furthermore, if $\tilde{E}$ denotes the reduced energy defined via

$$\tilde{E}(f) := \int \tfrac{1}{4}|\nabla f|^2 + \tfrac{1}{8}\gamma |q(f)|^2\,dx = \tfrac{1}{2}E(f) - \tfrac{1}{12}\int q(f)^3\,dx,$$

then

$$\tilde{E}(u_n) = \tilde{E}(r_n) + \tfrac{1}{4}\|V\xi_n\|^2_{H^1_x} + \tfrac{1}{2}\mathrm{Re}\langle M(r_n), V\xi_n\rangle_{H^1_x} + o(1) \quad \text{as } n\to\infty. \tag{6-10}$$

Proof. We will only prove (6-9). Claim (6-10) can be read off from the proof we give below.

To begin we observe that

$$q(u_n) = q(r_n) + 2\,\mathrm{Re}\,\xi_n + |\xi_n|^2 + 2\,\mathrm{Re}(\bar{\xi}_n r_n).$$

By hypothesis, $r_n = u_n - \xi_n$ is uniformly bounded in L^6_x. Using this and our assumptions on ξ_n, we see

$$\begin{aligned} q(u_n) &= q(r_n) + 2\,\mathrm{Re}\,\xi_n + o(1) && \text{in } L^2_x \text{ as } n\to\infty, \\ q(u_n) &= q(r_n) + o(1) && \text{in } L^3_x \text{ as } n\to\infty. \end{aligned}$$

Moreover, as u_n is bounded in $\mathcal{E}$ and $\operatorname{Re}\xi_n$ is bounded in L^2_x, we deduce that $q(r_n)$ is uniformly bounded in both L^2_x and L^3_x.

Therefore, we obtain

$$\begin{aligned} E(u_n) &= E(r_n) + \int \tfrac{1}{2}|\nabla\xi_n|^2 + \operatorname{Re}(\nabla\bar{\xi}_n\nabla r_n) + \gamma q(r_n)\operatorname{Re}\xi_n + \gamma(\operatorname{Re}\xi_n)^2\,dx + o(1) \\ &= E(r_n) + \tfrac{1}{2}\|V\xi_n\|^2_{H^1_x} + \operatorname{Re}\langle(2\gamma-\Delta)M(r_n), V\xi_n\rangle_{L^2_x} + o(1) \quad \text{as } n\to\infty. \qquad \square \end{aligned}$$

We return now to the proof of (6-7). Let us begin by showing that

$$E(u^\infty) \geq \tfrac{1}{2}\|z_+\|^2_{H^1_x}, \tag{6-11}$$

which combined with (6-8) fully identifies $E(u^\infty)$. While natural, this is not (in and of itself) essential to the argument; it does, however, force us to control the contributions of parts of the energy with the unhelpful sign. It will be this control that will ultimately allows us to complete the proof of (6-7).

Let $t_n \to \infty$ be an arbitrary sequence. We apply Lemma 6.3 with

$$u_n := u^\infty(t_n) \quad \text{and} \quad \xi_n := (\mathrm{Id}\oplus P_{\geq N_n})V^{-1}e^{-it_nH}z_+, \tag{6-12}$$

where $N_n \in 2^{\mathbb{Z}}$ converges to zero sufficiently slowly to guarantee that

$$\|\xi_n\|_{L^3_x\cap L^6_x} \to 0 \quad \text{as } n\to\infty. \tag{6-13}$$

Note that this is possible because of (6-3). In view of (6-9), we obtain

$$\begin{aligned} E(u^\infty) = E(r_n) + \tfrac{1}{2}\big\|(\mathrm{Id}\oplus P_{\geq N_n})e^{-it_nH}z_+\big\|^2_{H^1_x} \\ + \operatorname{Re}\langle M(r_n), (\mathrm{Id}\oplus P_{\geq N_n})e^{-it_nH}z_+\rangle_{H^1_x} + o(1) \quad \text{as } n\to\infty. \end{aligned} \tag{6-14}$$

By Proposition 6.2, $e^{it_nH}M(u^\infty(t_n)) = e^{it_nH}z^\infty(t_n) \rightharpoonup z_+$ weakly in H^1_x. On the other hand, by (6-13), we have

$$\begin{aligned} M(u^\infty(t_n)) &= e^{-it_nH}z_+ + M(r_n) - P_{\leq N_n}\operatorname{Im} e^{-it_nH}z_+ + \gamma\langle\nabla\rangle^{-2}[|\xi_n|^2 + 2\operatorname{Re}(\bar{\xi}_n r_n)] \\ &= e^{-it_nH}z_+ + M(r_n) + o(1) \quad \text{in } H^1_x \text{ as } n\to\infty. \end{aligned} \tag{6-15}$$

Thus, we may deduce that

$$e^{it_nH}M(r_n) \rightharpoonup 0 \quad \text{weakly in } H^1_x \text{ as } n\to\infty.$$

Combining this with the dominated convergence theorem (which allows us to replace $P_{\geq N_n}$ by Id), (6-14) becomes

$$E(u^\infty) = E(r_n) + \tfrac{1}{2}\|z_+\|^2_{H^1_x} + o(1) \quad \text{as } n\to\infty. \tag{6-16}$$

Arguing similarly and using (6-10) in place of (6-9), we obtain

$$\tilde{E}(u^\infty) = \tilde{E}(r_n) + \tfrac{1}{4}\|z_+\|^2_{H^1_x} + o(1) \quad \text{as } n\to\infty. \tag{6-17}$$

Note that (6-11) follows immediately from (6-16), provided that $E(r_n) \geq 0$. By Lemma 3.1, this is immediate if $\gamma \in \left[\frac{2}{3}, 1\right)$. In view of Lemma 3.2, if $\gamma \in \left(0, \frac{2}{3}\right)$ we simply have to verify that $\|\nabla \operatorname{Re} r_n\|_{L_x^2}^2 \leq \delta_\gamma$. This, however, follows from (6-8) and (6-17), provided E_0 is chosen sufficiently small depending on γ.

Combining (6-8) with (6-11) and (6-16), we deduce that

$$E(u^\infty) = \tfrac{1}{2}\|z_+\|_{H_x^1}^2 \qquad \text{and} \qquad E(r_n) \to 0 \quad \text{as } n \to \infty.$$

By the argument in the preceding paragraph, this implies

$$\|r_n\|_{\mathcal{E}} \to 0 \quad \text{as } n \to \infty. \tag{6-18}$$

Therefore, using the representation (5-6) for M, we see that

$$\begin{aligned}\|M(r_n)\|_{H_x^1} &\lesssim \|U \operatorname{Im} r_n\|_{H_x^1} + \|U^2 \operatorname{Re} r_n\|_{H_x^1} + \|\langle\nabla\rangle^{-2} q(r_n)\|_{H_x^1} \\ &\lesssim \|r_n\|_{\dot{H}_x^1} + \|q(r_n)\|_{L_x^2} \to 0 \quad \text{as } n \to \infty.\end{aligned}$$

Combining this with (6-15), we get

$$\|z^\infty(t_n) - e^{-it_n H} z_+\|_{H_x^1} \to 0 \quad \text{as } n \to \infty.$$

As the sequence $t_n \to \infty$ was arbitrary, this completes the proof of (6-7).

We next prove that (6-7) implies the conclusions of Theorem 6.1. We first show that (6-7) implies (6-1). Let $t_n \to \infty$ be an arbitrary sequence and define u_n and ξ_n as in (6-12). Using (6-13) and (6-18), we deduce that $u^\infty(t_n) \to 0$ in L_x^6. Furthermore, by (6-7) and (6-3), we have that $z^\infty(t_n) \to 0$ in L_x^6. Using Proposition 5.1(v), we find that $u^\infty(t_n) = R(z^\infty(t_n))$ for n sufficiently large. Arguing as in Lemma 5.3 and using (5-11), we may write $u^\infty(t_n) = V^{-1} z^\infty(t_n) + o(1)$ in $\dot{H}_x^1$, which together with (6-7) yields (6-1).

We now turn to (6-2). We begin with the following strengthening of (6-3):

$$\lim_{t\to\infty} \|U^{-1} e^{-itH} z_+\|_{L_x^6} = 0. \tag{6-19}$$

Given $0 < N < 1$, we have

$$\begin{aligned}\|U^{-1} P_{\leq N} e^{-itH} z_+\|_{L_x^6} &\lesssim \|\nabla U^{-1} P_{\leq N} e^{-itH} z_+\|_{L_x^2} \lesssim \|P_{\leq N} z_+\|_{L_x^2}, \\ \|U^{-1} P_{>N} e^{-itH} z_+\|_{L_x^6} &\lesssim N^{-1} \|e^{-itH} z_+\|_{L_x^6}.\end{aligned}$$

In view of (6-3), choosing N sufficiently small and then sending $t \to \infty$ yields (6-19).

Using (6-19), we now show that the modification $\gamma\langle\nabla\rangle^{-2}|u_{\text{lin}}|^2$ appearing in (6-2) is negligible in the $\dot{H}_x^1$-norm. Indeed, we have the stronger statement

$$\begin{aligned}\big\|\langle\nabla\rangle^{-1}|u_{\text{lin}}(t)|^2\big\|_{\dot{H}_x^1} &\lesssim \big\||\nabla|^{\frac{1}{2}} u_{\text{lin}}(t)\big\|_{L_x^3} \|u_{\text{lin}}(t)\|_{L_x^6} \\ &\lesssim \|\nabla u_{\text{lin}}(t)\|_{L_x^2} \|U^{-1} e^{-itH} z_+\|_{L_x^6} \\ &\lesssim \|z_+\|_{H_x^1} \|U^{-1} e^{-itH} z_+\|_{L_x^6} \to 0 \quad \text{as } t \to \infty.\end{aligned} \tag{6-20}$$

It remains to show

$$\lim_{t\to\infty}\big\|q\big(u^\infty(t)\big)-q\big(u_{\text{lin}}(t)-\gamma\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big)\big\|_{L^2_x}=0. \tag{6-21}$$

As demonstrated above, $u^\infty(t)=R(z^\infty(t))$ for t sufficiently large and $z^\infty(t)\to 0$ in L^6_x as $t\to\infty$. Thus, arguing as for (5-13) and using (6-7), we deduce that

$$q(u^\infty(t))=2\operatorname{Re}u_{\text{lin}}(t)+o(1)\quad\text{in } L^2_x \text{ as } t\to\infty.$$

On the other hand, a straightforward computation yields

$$\begin{aligned}q\big(u_{\text{lin}}(t)-\gamma\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big)&\\ =2\operatorname{Re}u_{\text{lin}}(t)+U^2|u_{\text{lin}}(t)|^2+\big[\gamma\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big]^2-2\gamma\big[\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big]\operatorname{Re}u_{\text{lin}}(t).&\end{aligned}$$

Thus, to prove (6-21) it suffices to show that the last three terms on the right-hand side above are $o(1)$ in L^2_x as $t\to\infty$. Indeed, we may estimate

$$\begin{aligned}\|U^2|u_{\text{lin}}(t)|^2\|_{L^2_x}&\lesssim\|\langle\nabla\rangle^{-1}|u_{\text{lin}}(t)|^2\|_{\dot H^1_x},\\ \big\|\big[\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big]^2\big\|_{L^2_x}&\lesssim\big\||\nabla|^{\frac14}\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big\|^2_{L^3_x}\lesssim\|U^{-1}e^{-itH}z_+\|^4_{L^6_x},\\ \big\|\big[\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big]\operatorname{Re}u_{\text{lin}}(t)\big\|_{L^2_x}&\lesssim\|U^{-1}e^{-itH}z_+\|^3_{L^6_x},\end{aligned}$$

and so by (6-19) and (6-20), we have

$$q\big(u_{\text{lin}}(t)-\gamma\langle\nabla\rangle^{-2}|u_{\text{lin}}(t)|^2\big)=2\operatorname{Re}u_{\text{lin}}(t)+o(1)\quad\text{in } L^2_x\text{ as } t\to\infty.$$

This completes the proof of (6-21) and hence that of Theorem 1.1. □

It remains to prove Proposition 6.2.

Proof of Proposition 6.2. We first claim that

$$z^{T_n}(t)\rightharpoonup z^\infty(t)\quad\text{weakly in } \dot H^1_x \text{ for all } t\in\mathbb{R}. \tag{6-22}$$

This relies in an essential way on Theorem 4.1 via (6-6). Henceforth, we let $t\in\mathbb{R}$ be fixed. Using (6-6) and Rellich–Kondrashov and passing to a subsequence, we have $u^{T_n}(t)\to u^\infty(t)$ strongly in $L^2_x(K)$ for any compact $K\subset\mathbb{R}^3$. Now fix $\varphi\in C^\infty_c(\mathbb{R}^3)$. Then $\langle\nabla\rangle^{-2}\varphi\in\mathcal{S}(\mathbb{R}^3)$; in particular, for any $\varepsilon>0$ there exists $\tilde\varphi_\varepsilon\in C^\infty_c(\mathbb{R}^3)$ such that

$$\|\langle\nabla\rangle^{-2}\varphi-\tilde\varphi_\varepsilon\|_{L^{3/2}_x}\le\varepsilon.$$

Using this, Hölder, (6-5), and (6-6), we obtain

$$\begin{aligned}\langle z^{T_n}(t),\varphi\rangle&=\langle u^{T_n}(t),V\varphi\rangle+\gamma\langle|u^{T_n}(t)|^2,\tilde\varphi_\varepsilon\rangle+\gamma\big\langle|u^{T_n}(t)|^2,\langle\nabla\rangle^{-2}\varphi-\tilde\varphi_\varepsilon\big\rangle\\ &=\langle u^{T_n}(t),V\varphi\rangle+\gamma\langle|u^{T_n}(t)|^2,\tilde\varphi_\varepsilon\rangle+O(\varepsilon)\\ &\to\langle u^\infty(t),V\varphi\rangle+\gamma\langle|u^\infty(t)|^2,\tilde\varphi_\varepsilon\rangle+O(\varepsilon)\\ &=\langle z^\infty(t),\varphi\rangle+O(\varepsilon).\end{aligned}$$

As $\varepsilon>0$ was arbitrary, this proves (6-22).

To continue, we write

$$e^{itH}z^{\infty}(t)-z_{+}$$
$$=\big[e^{itH}z^{\infty}(t)-e^{iT_0H}z^{\infty}(T_0)\big]+\big[e^{iT_0H}z^{\infty}(T_0)-e^{iT_0H}z^{T_n}(T_0)\big]+\big[e^{iT_0H}z^{T_n}(T_0)-e^{iT_nH}z^{T_n}(T_n)\big].$$

As the above is bounded in H^1_x, it suffices to prove weak convergence when testing against a dense set of functions in H^{-1}_x. In this role, we take $\varphi\in\mathcal{S}(\mathbb{R}^3)$ with $\hat{\varphi}\in C_c^{\infty}(\mathbb{R}^3\setminus\{0\})$. To continue, we choose $N_0\in 2^{\mathbb{Z}}$ so that $\operatorname{supp}\hat{\varphi}\subset\{|\xi|\geq 100N_0\}$ and fix $\varepsilon>0$. By (6-22), there exists n sufficiently large (depending on T_0) so that

$$\big|\big\langle e^{iT_0H}z^{\infty}(T_0)-e^{iT_0H}z^{T_n}(T_0),\varphi\big\rangle\big|\leq\varepsilon. \tag{6-23}$$

To handle the remaining two differences, we will prove the inequality

$$\big|\big\langle e^{it_2H}z(t_2)-e^{it_1H}z(t_1),\varphi\big\rangle\big|\lesssim_{\varphi}|t_1|^{-\frac14} \tag{6-24}$$

uniformly for $t_2>t_1$, where z denotes any of the functions z^{T_n}. In view of (6-22), we see that (6-24) also holds (with the same implicit constant) for $z=z^{\infty}$. Thus, taking T_0 large enough depending on ε and then n large enough so that $T_n>T_0$ and (6-23) holds, we get

$$\sup_{t>T_0}|\langle e^{itH}z^{\infty}(t)-z_{+},\varphi\rangle|\lesssim_{\varphi}\varepsilon.$$

As $\varepsilon>0$ was arbitrary, this proves Proposition 6.2.

It remains to verify (6-24). By Duhamel's formula, we have

$$\big|\big\langle e^{it_2H}z(t_2)-e^{it_1H}z(t_1),\varphi\big\rangle\big|\leq\int_{t_1}^{t_2}|\langle N_z(u(s)),e^{-isH}\varphi\rangle|\,ds. \tag{6-25}$$

To continue, we decompose the nonlinearity as

$$N_z(u)=N_z^1(u)-\gamma N_z^2(u),$$

where

$$N_z^1(u)=U\big[\tfrac12\gamma+2q^2+q^2u_1-\gamma u_1^3-\tfrac12\gamma|u|^4\big]-2\gamma i\,\langle\nabla\rangle^{-2}\nabla\cdot[q\nabla u_2-u_1^2\nabla u_2]+iU^2[\gamma u_1^2u_2+q^2u_2],$$
$$N_z^2(u)=U(u_1u_2^2)-\tfrac13 iU^2(u_2^3).$$

We first estimate the contribution of $N_z^1(u)$ to (6-25). By Hölder and the dispersive estimate (2-3), we can estimate

$$\begin{aligned}\int_{t_1}^{t_2}\big|\langle N_z^1(u(s)),e^{-isH}\varphi\rangle\big|\,ds&\lesssim\int_{t_1}^{t_2}\|N_z^1(u(s))\|_{L_x^{12/11}}\|e^{-isH}\varphi\|_{L_x^{12}}\,ds\\&\lesssim_{\varphi}\int_{t_1}^{t_2}|s|^{-\frac54}\|N_z^1(u(s))\|_{L_x^{12/11}}\,ds.\end{aligned}$$

Most of the terms appearing in $\operatorname{Re}(N_z^1)$ can be handled using Hölder and (6-5):

$$\big\|U\big[\tfrac12\gamma+2q^2+q^2u_1-\gamma u_1^3\big]\big\|_{L_x^{12/11}}\lesssim\|q\|^2_{L_x^{24/11}}+\|q\|^2_{L_x^{8/3}}\|u_1\|_{L_x^6}+\|u_1\|^3_{L_x^{36/11}}\lesssim 1.$$

To estimate the remaining term in $\mathrm{Re}(N_z^1)$ we also use the fractional chain rule and Sobolev embedding:

$$\|U(|u|^4)\|_{L_x^{12/11}} \lesssim \big\||\nabla|^{\frac{3}{4}}(|u|^4)\big\|_{L_x^{12/11}} \lesssim \big\||\nabla|^{\frac{3}{4}}u\big\|_{L_x^{12/5}}\|u\|_{L_x^6}^3 \lesssim \|\nabla u\|_{L_x^2}^4 \lesssim 1.$$

To estimate the terms in $\mathrm{Im}(N_z^1)$, we use Hölder and (6-5):

$$\big\|\langle\nabla\rangle^{-2}\nabla\cdot[q\nabla u_2 - u_1^2\nabla u_2]\big\|_{L_x^{12/11}} \lesssim \|\nabla u_2\|_{L_x^2}\|q\|_{L_x^{12/5}} + \|\nabla u_2\|_{L_x^2}\|u_1\|_{L_x^{24/5}}^2 \lesssim 1,$$

$$\begin{aligned}\big\|U^2[\gamma u_1^2 u_2 + q^2 u_2]\big\|_{L_x^{12/11}} &\lesssim \|\nabla(u_1^2 u_2)\|_{L_x^{12/11}} + \|q^2 u_2\|_{L_x^{12/11}} \\ &\lesssim \|u_1\|_{L_x^4}\|u\|_{L_x^6}\|\nabla u\|_{L_x^2} + \|q\|_{L_x^{8/3}}^2\|u_2\|_{L_x^6} \lesssim 1.\end{aligned}$$

Putting everything together, we find

$$\int_{t_1}^{t_2} \big|\langle N_z^1(u(s)), e^{-isH}\varphi\rangle\big|\, ds \lesssim_\varphi |t_1|^{-\frac{1}{4}}.$$

We turn now to estimating the contribution of $N_z^2(u)$ to (6-25). To complete the proof of (6-24) and so that of the proposition, we must show that

$$\int_{t_1}^{t_2} \big|\langle N_z^2(u(s)), e^{-isH}\varphi\rangle\big|\, ds \lesssim_\varphi |t_1|^{-\frac{1}{4}}. \tag{6-26}$$

Recalling that $\operatorname{supp}\hat{\varphi} \subset \{|\xi| \geq 100N_0\}$, we see that

$$\langle P_{\leq 20N_0} N_z^2(u(s)), e^{-isH}\varphi\rangle \equiv 0.$$

Writing $u_2 = P_{\leq N_0}u_2 + P_{>N_0}u_2$, we may decompose the remaining part of $N_z^2(u)$ as

$$\begin{aligned}P_{>20N_0}N_z^2(u) = P_{>20N_0}U\emptyset\big(u_1 u_2 P_{>N_0}u_2\big) + P_{>20N_0}U^2\emptyset\big(u_2[P_{>N_0}u_2]^2\big) \\ + P_{>20N_0}U\big\{[P_{>8N_0}u_1][P_{\leq N_0}u_2]^2 - iU\big([P_{>8N_0}u_2][P_{\leq N_0}u_2]^2\big)\big\}.\end{aligned}$$

Writing $u_1 = \bar{v} + iUu_2$ (with $v = Vu$) and $a := [P_{\leq N_0}u_2]^2$, we arrive at the decomposition

$$\begin{aligned}P_{>20N_0}N_z^2(u) = &\\ &P_{>20N_0}U\emptyset\big(u_1 u_2 P_{>N_0}u_2\big) + P_{>20N_0}U^2\emptyset\big(u_2[P_{>N_0}u_2]^2\big) &(6\text{-}27)\\ &+ iP_{>20N_0}U\big\{aU(P_{>8N_0}u_2) - U\big(aP_{>8N_0}u_2\big)\big\} &(6\text{-}28)\\ &+ P_{>20N_0}U\big\{aP_{>8N_0}\bar{v}\big\}. &(6\text{-}29)\end{aligned}$$

As we will see, the terms in (6-27) and (6-28) are small. However, there is no reason to believe that (6-29) is small pointwise in time; instead, we will show that this term is nonresonant.

We first consider (6-27). Using Hölder, Bernstein, and (6-5), we estimate

$$\begin{aligned}\|(6\text{-}27)\|_{L_x^{12/11}} &\lesssim \|u_1\|_{L_x^4}\|u_2\|_{L_x^6}\|P_{>N_0}u_2\|_{L_x^2} + \|u_2\|_{L_x^6}\|P_{>N_0}u_2\|_{L_x^{8/3}}^2 \\ &\lesssim_{N_0} \|u_1\|_{L_x^4}\|\nabla u_2\|_{L_x^2}^2 + \|\nabla u_2\|_{L_x^2}^3 \lesssim_{N_0} 1.\end{aligned}$$

Thus, the contribution of this term to (6-26) is acceptable.

We now turn to (6-28), which includes the commutator $[a, U]$. We regard this term as a bilinear operator $T(a, P_{>8N_0}u_2)$ with symbol given by

$$\left[1-\phi\left(\frac{\xi}{20N_0}\right)\right]U(\xi)[U(\xi_2)-U(\xi)]\phi\left(\frac{\xi_1}{4N_0}\right) = \left\{-2\gamma\frac{U(\xi)(\xi_1+2\xi_2)}{\langle\xi\rangle\langle\xi_2\rangle(|\xi_2|\langle\xi\rangle+|\xi|\langle\xi_2\rangle)}\left[1-\phi\left(\frac{\xi}{20N_0}\right)\right]\phi\left(\frac{\xi_1}{4N_0}\right)\right\}\cdot\xi_1,$$

where ϕ denotes the standard Littlewood–Paley multiplier. Observing that the multiplier inside the braces is amenable to Lemma 2.3, we may estimate

$$\|(6\text{-}28)\|_{L_x^{12/11}} \lesssim \|\nabla a\|_{L_x^{3/2}}\|P_{>8N_0}u_2\|_{L_x^4} \lesssim_{N_0} \|\nabla P_{\le N_0}u_2\|_{L_x^2}\|P_{\le N_0}u_2\|_{L_x^6}\|\nabla u_2\|_{L_x^2}.$$

In view of (6-5), the contribution of this term to (6-26) is acceptable.

Finally, we consider (6-29). Using (1-11), we find

$$i\partial_t\left\langle aP_{>8N_0}\bar{v}, e^{-itH}\frac{U}{2H}\varphi\right\rangle = \langle U(aP_{>8N_0}\bar{v}), e^{-itH}\varphi\rangle + \left\langle aP_{>8N_0}\overline{N_v(u)}, e^{-itH}\frac{U}{2H}\varphi\right\rangle + \left\langle [a,H]P_{>8N_0}\bar{v}, e^{-itH}\frac{U}{2H}\varphi\right\rangle + i\left\langle \dot{a}P_{>8N_0}\bar{v}, e^{-itH}\frac{U}{2H}\varphi\right\rangle.$$

By the fundamental theorem of calculus, we may thus estimate the contribution of (6-29) to (6-26) as

$$\begin{aligned}\int_{t_1}^{t_2} &\left|\langle U(aP_{>8N_0}\bar{v}), e^{-isH}\varphi\rangle\right| ds \\ &\lesssim \sup_{t\ge t_1}\left|\langle aP_{>8N_0}\bar{v}, e^{-itH}\langle\nabla\rangle^{-2}\varphi\rangle\right| + \int_{t_1}^{t_2}\left|\langle aP_{>8N_0}\overline{N_v(u)}, e^{-isH}\langle\nabla\rangle^{-2}\varphi\rangle\right| ds \\ &\quad + \int_{t_1}^{t_2}\left|\langle\langle\nabla\rangle^{-2}([a,H]P_{>8N_0}\bar{v}), e^{-isH}\varphi\rangle\right| ds + \int_{t_1}^{t_2}\left|\langle \dot{a}P_{>8N_0}\bar{v}, e^{-isH}\langle\nabla\rangle^{-2}\varphi\rangle\right| ds. \end{aligned} \tag{6-30}$$

To estimate the terms on the right-hand side of (6-30), we note that in view of (6-5),

$$\|\nabla a(t)\|_{L_x^{3/2}\cap L_x^\infty} + \|a(t)\|_{L_x^3\cap L_x^\infty} + \|\nabla v(t)\|_{L_x^2} + \|v(t)\|_{L_x^3\cap L_x^6} \lesssim_{N_0} 1 \tag{6-31}$$

uniformly for $t\in\mathbb{R}$. Using this, Hölder, and (2-3), we estimate the first term on the right-hand side of (6-30) as

$$\sup_{t\ge t_1}\left|\langle aP_{>8N_0}\bar{v}, e^{-itH}\langle\nabla\rangle^{-2}\varphi\rangle\right| \lesssim \|a\|_{L_t^\infty L_x^3}\|v\|_{L_t^\infty L_x^3}\sup_{t\ge t_1}\|e^{-itH}\langle\nabla\rangle^{-2}\varphi\|_{L_x^3} \lesssim_{\varphi,N_0} |t_1|^{-\frac12}.$$

Thus, the contribution of this term to (6-26) is acceptable.

Next we consider the second term on the right-hand side of (6-30). By Hölder and (2-3),

$$\int_{t_1}^{t_2}\left|\langle aP_{>8N_0}\overline{N_v(u)}, e^{-isH}\langle\nabla\rangle^{-2}\varphi\rangle\right| ds \lesssim_\varphi |t_1|^{-\frac12}\|aP_{>8N_0}N_v(u)\|_{L_t^\infty L_x^1}.$$

Note that

$$N_v(u) = \sum_{k=2}^5 U\mathcal{O}(u^k) + \mathcal{O}(u^k).$$

To estimate the contribution of the quadratic terms in $N_v(u)$, we use Bernstein, (6-5), and (6-31):

$$\begin{aligned}\big\|aP_{>8N_0}\big[U\mathcal{O}(u^2)+\mathcal{O}(u^2)\big]\big\|_{L_t^\infty L_x^1} &\lesssim_{N_0} \|a\|_{L_t^\infty L_x^3}\|\nabla\mathcal{O}(u^2)\|_{L_t^\infty L_x^{3/2}}\\ &\lesssim_{N_0} \|a\|_{L_t^\infty L_x^3}\|\nabla u\|_{L_t^\infty L_x^2}\|u\|_{L_t^\infty L_x^6} \lesssim_{N_0} 1.\end{aligned}$$

Similarly, we can estimate the cubic terms in $N_v(u)$ via

$$\begin{aligned}\big\|aP_{>8N_0}\big[U\mathcal{O}(u^3)+\mathcal{O}(u^3)\big]\big\|_{L_t^\infty L_x^1} &\lesssim_{N_0} \|a\|_{L_t^\infty L_x^6}\|\nabla\mathcal{O}(u^3)\|_{L_t^\infty L_x^{6/5}}\\ &\lesssim_{N_0} \|a\|_{L_t^\infty L_x^6}\|\nabla u\|_{L_t^\infty L_x^2}\|u\|^2_{L_t^\infty L_x^6} \lesssim_{N_0} 1.\end{aligned}$$

We estimate the quartic and quintic terms in $N_v(u)$ using Hölder, (6-5), and (6-31):

$$\begin{aligned}\big\|aP_{>8N_0}\big[U\mathcal{O}(u^4)+\mathcal{O}(u^4)\big]\big\|_{L_t^\infty L_x^1} &\lesssim \|a\|_{L_t^\infty L_x^3}\|u\|^4_{L_t^\infty L_x^6} \lesssim 1,\\ \big\|aP_{>8N_0}\big[U\mathcal{O}(u^5)+\mathcal{O}(u^5)\big]\big\|_{L_t^\infty L_x^1} &\lesssim \|a\|_{L_t^\infty L_x^6}\|u\|^5_{L_t^\infty L_x^6} \lesssim 1.\end{aligned}$$

Putting everything together, we see that the contribution of the second term on the right-hand side of (6-30) to (6-26) is acceptable.

We now turn to the third term on the right-hand side of (6-30). By Hölder and (2-3),

$$\int_{t_1}^{t_2} \big|\big\langle\langle\nabla\rangle^{-2}\big([a,H]P_{>8N_0}\bar v\big), e^{-isH}\varphi\big\rangle\big|\,ds \lesssim_\varphi |t_1|^{-\frac14}\big\|\langle\nabla\rangle^{-2}\big([a,H]P_{>8N_0}\bar v\big)\big\|_{L_t^\infty L_x^{12/11}}.$$

We regard the term on the right-hand side above as a bilinear operator $T(a,v)$ with symbol given by

$$\frac{H(\xi_2)-H(\xi)}{\langle\xi\rangle^2}\phi\Big(\frac{\xi_1}{4N_0}\Big)\Big[1-\phi\Big(\frac{\xi_2}{8N_0}\Big)\Big] = m(\xi_1,\xi_2)\cdot\xi_1,$$

where

$$m(\xi_1,\xi_2) = -\frac{(2\gamma+|\xi|^2+|\xi_2|^2)(\xi_1+2\xi_2)}{\langle\xi\rangle^2(|\xi_2|\langle\xi_2\rangle+|\xi|\langle\xi\rangle)}\phi\Big(\frac{\xi_1}{4N_0}\Big)\Big[1-\phi\Big(\frac{\xi_2}{8N_0}\Big)\Big]$$

is a bounded bilinear multiplier in view of Lemma 2.3. Using also (6-31), we get

$$\big\|\langle\nabla\rangle^{-2}\big([a,H]P_{>8N_0}\bar v\big)\big\|_{L_t^\infty L_x^{12/11}} \lesssim \|\nabla a\|_{L_t^\infty L_x^{3/2}}\|v\|_{L_t^\infty L_x^4} \lesssim_{N_0} 1.$$

Thus, the contribution of the third term on the right-hand side of (6-30) to (6-26) is acceptable.

We now turn to the fourth and last term on the right-hand side of (6-30). By Hölder, (2-3), and Bernstein,

$$\begin{aligned}\int_{t_1}^{t_2} \big|\big\langle\dot a P_{>8N_0}\bar v, e^{-isH}\langle\nabla\rangle^{-2}\varphi\big\rangle\big|\,ds &\lesssim_\varphi |t_1|^{-\frac12}\|\dot a\|_{L_t^\infty L_x^2}\|P_{>8N_0}v\|_{L_t^\infty L_x^2}\\ &\lesssim_{\varphi,N_0} |t_1|^{-\frac12}\|P_{\le N_0}\dot u_2\|_{L_t^\infty L_x^3}\|u_2\|_{L_t^\infty L_x^6}\|\nabla v\|_{L_t^\infty L_x^2}.\end{aligned}$$

In view of (6-5) and (6-31), we need only bound $P_{\leq N_0}\dot{u}_2$ in $L_t^\infty L_x^3$. To this end, we use (1-6), Bernstein, and (6-5):

$$\|P_{\leq N_0}\dot{u}_2\|_{L_t^\infty L_x^3} \lesssim \|P_{\leq N_0}(2\gamma-\Delta)u_1\|_{L_t^\infty L_x^3} + \sum_{k=2}^{5}\|P_{\leq N_0}\emptyset(u^k)\|_{L_t^\infty L_x^3}$$
$$\lesssim_{N_0} \|u_1\|_{L_t^\infty L_x^3} + \sum_{k=2}^{5}\|u\|^k_{L_t^\infty L_x^6} \lesssim_{N_0} 1.$$

Thus, the contribution of the fourth term on the right-hand side of (6-30) to (6-26) is acceptable. This completes the justification of (6-26) and so the proof of Proposition 6.2. □

7. Proof of Theorem 1.4

In this section we prove Theorem 1.4 and Corollary 1.7. We recall the norm

$$\|u\|_{X_T} := \sup_{t\geq T} t^{\frac{1}{2}}\|u(t)\|_{H_x^{1,3}(\mathbb{R}^3)}.$$

The proof of Theorem 1.4 will be effected by running a contraction mapping argument simultaneously for u and $z = M(u)$. The necessity of exploiting the normal form transformation can be seen when one endeavors to estimate the quadratic terms appearing in the nonlinearity.

Proof of Theorem 1.4.. We define maps

$$\begin{cases} [\Phi_1(u,z)](t) = V^{-1}z(t) - \gamma\langle\nabla\rangle^{-2}|u(t)|^2, \\ [\Phi_2(u)](t) = e^{-itH}Vu_+ + i\displaystyle\int_t^\infty e^{-i(t-s)H}N_z(u(s))\,ds, \end{cases}$$

where N_z is as in (5-5).

We will show that the map $(u,z)\mapsto \Phi(u,z) := (\Phi_1(u,z),\Phi_2(u))$ is a contraction on a suitable complete metric space, and so deduce that Φ has a unique fixed point (u,z) in this space, which then necessarily solves (5-4)–(5-5).

For $0<\eta<1$ and $T>1$ to be determined below, we define

$$B_1 = \{u : \|u\|_{L_t^\infty(H^1_{\mathrm{real}}+i\dot{H}^1_{\mathrm{real}})} \leq 4\|u_+\|_{H^1_{\mathrm{real}}+i\dot{H}^1_{\mathrm{real}}},\ \|u\|_{X_T}\leq 4\eta\},$$

and

$$B_2 = \{z : \|z\|_{L_t^\infty H_x^1} \leq 2\|u_+\|_{H^1_{\mathrm{real}}+i\dot{H}^1_{\mathrm{real}}},\ \|V^{-1}z\|_{X_T}\leq 2\eta\},$$

where here and in what follows all space-time norms are taken over $(T,\infty)\times\mathbb{R}^3$ unless stated otherwise. We define $B = B_1\times B_2$ and equip B with the metric

$$d((u,z),(\tilde{u},\tilde{z})) = \|u-\tilde{u}\|_{X_T} + 8\|V^{-1}(z-\tilde{z})\|_{X_T}.$$

We first show that $\Phi: B\to B$. By Sobolev embedding, for $(u,z)\in B$ and $t>T\geq 1$,

$$\big\|\langle\nabla\rangle^{-2}|u(t)|^2\big\|_{H_x^1} + \big\|\langle\nabla\rangle^{-2}|u(t)|^2\big\|_{H_x^{1,3}} \lesssim \big\||u(t)|^2\big\|_{L_x^{3/2}} \lesssim \|u(t)\|^2_{L_x^3} \lesssim \eta^2 t^{-1}.$$

Thus choosing $T = T(\|u_+\|_{H^1_{\text{real}}+i\dot H^1_{\text{real}}})$ large enough, we have

$$\big\|[\Phi_1(u,z)](t)\big\|_{H^1_{\text{real}}+i\dot H^1_{\text{real}}} \le \|V^{-1}z(t)\|_{H^1_{\text{real}}+i\dot H^1_{\text{real}}} + \gamma\big\|\langle\nabla\rangle^{-2}|u(t)|^2\big\|_{H^1_x} \le 2\|z\|_{L^\infty_t H^1_x}.$$

Similarly,

$$\big\|[\Phi_1(u,z)](t)\big\|_{H^{1,3}_x} \le \|V^{-1}z(t)\|_{H^{1,3}_x} + \gamma\big\|\langle\nabla\rangle^{-2}|u(t)|^2\big\|_{H^{1,3}_x} \le 4\eta t^{-\frac12},$$

provided η is chosen small enough. Thus $\Phi_1 : B \to B_1$.

We next show that $\Phi_2 : B_1 \to B_2$. We first estimate $N_z(u)$, which satisfies

$$U^{-1}N_z(u) = \emptyset(u^2+u^3+u^4+u^5) + U^{-1}\frac{\nabla}{\langle\nabla\rangle^2}\cdot\emptyset(u\nabla u) + U\emptyset(u^3+u^4+u^5). \tag{7-1}$$

We estimate the quadratic terms at fixed time $t > T \ge 1$, as

$$\Big\|\Big[\emptyset(u^2) + U^{-1}\frac{\nabla}{\langle\nabla\rangle^2}\cdot\emptyset(u\nabla u)\Big](t)\Big\|_{H^{1,3/2}_x} \lesssim t^{-1}\|u\|^2_{X_T} \lesssim t^{-1}\eta^2.$$

Similarly, for $k \in \{2,3,4\}$ we have

$$\begin{aligned}\big\|\big[\emptyset(u^{k+1}) + U\emptyset(u^{k+1})\big](t)\big\|_{H^{1,3/2}_x} &\lesssim \|u(t)\|^k_{L^{3k}_x}\|u(t)\|_{H^{1,3}_x}\\ &\lesssim \||\nabla|^{1-\frac1k}u(t)\|^k_{L^3_x}\|u(t)\|_{H^{1,3}_x} \lesssim t^{-\frac{k+1}{2}}\|u\|^{k+1}_{X_T} \lesssim t^{-\frac{k+1}{2}}\eta^{k+1}.\end{aligned} \tag{7-2}$$

Combining the above, we deduce that

$$\|U^{-1}N_z(u(t))\|_{H^{1,3/2}_x} \lesssim \sum_{k=1}^{4} t^{-\frac{k+1}{2}}\eta^{k+1} \quad \text{uniformly for } t > T \ge 1. \tag{7-3}$$

To continue, we use Strichartz and (7-3) to estimate

$$\begin{aligned}\|\Phi_2(u)\|_{L^\infty_t H^1_x} &\le \|Vu_+\|_{L^\infty_t H^1_x} + C\|N_z(u)\|_{L^{4/3}_t H^{1,3/2}_x}\\ &\le \|u_+\|_{H^1_{\text{real}}+i\dot H^1_{\text{real}}} + C\sum_{k=1}^{4} T^{-\frac{2k-1}{4}}\eta^{k+1} \le 2\|u_+\|_{H^1_{\text{real}}+i\dot H^1_{\text{real}}},\end{aligned}$$

provided $T = T(\|u_+\|_{H^1_{\text{real}}+i\dot H^1_{\text{real}}})$ is chosen sufficiently large.

We turn to estimating $V^{-1}\Phi_2(u)$ in the X-norm for $u \in B_1$. By hypothesis, the dispersive estimate (2-3), and (7-3), for $t > T \ge 1$ we have

$$\begin{aligned}\|[V^{-1}\Phi_2(u)](t)\|_{H^{1,3}_x} &\le \|V^{-1}e^{-itH}Vu_+\|_{H^{1,3}_x} + \Big\|\int_t^\infty V^{-1}[ie^{-i(t-s)H}N_z(u(s))]\,ds\Big\|_{H^{1,3}_x}\\ &\le \eta t^{-\frac12} + \int_t^\infty \|e^{-i(t-s)H}U^{-1}N_z(u(s))]\|_{H^{1,3}_x}\,ds\\ &\le \eta t^{-\frac12} + C\int_t^\infty |t-s|^{-\frac12}\sum_{k=1}^{4} s^{-\frac{k+1}{2}}\eta^{k+1}\,ds \le \eta t^{-\frac12} + C\sum_{k=1}^{4} t^{-\frac{k}{2}}\eta^{k+1} \le 2\eta t^{-\frac12},\end{aligned}$$

provided η is chosen sufficiently small. This completes the proof that $\Phi : B \to B$.

We next claim that Φ is a contraction with respect to the metric defined above. First, for $(u,z), (\tilde{u},\tilde{z}) \in B$, we estimate

$$\begin{aligned}\big\|\Phi_1(u,z)-\Phi_1(\tilde{u},\tilde{z})\big\|_{X_T} &\le \|V^{-1}(z-\tilde{z})\|_{X_T}+\gamma\big\|\langle\nabla\rangle^{-2}(|u|^2-|\tilde{u}|^2)\big\|_{X_T}\\ &\le \tfrac{1}{8}d((u,z),(\tilde{u},\tilde{z}))+C\sup_{t\ge T} t^{\frac{1}{2}}\big\|(u+\tilde{u})(t)(u-\tilde{u})(t)\big\|_{L_x^{3/2}}\\ &\le \tfrac{1}{8}d((u,z),(\tilde{u},\tilde{z}))+C\eta T^{-\frac{1}{2}}\|u-\tilde{u}\|_{X_T}\\ &\le \tfrac{1}{4}d((u,z),(\tilde{u},\tilde{z})),\end{aligned}$$

provided η is sufficiently small.

By (2-3), for $t>T\ge 1$ we estimate

$$\begin{aligned}\big\|V^{-1}[\Phi_2(u)-\Phi_2(\tilde{u})](t)\big\|_{H_x^{1,3}} &\le \bigg\|\int_t^\infty V^{-1}\big(ie^{-i(t-s)H}[N_z(u(s))-N_z(\tilde{u}(s))]\big)\,ds\bigg\|_{H_x^{1,3}}\\ &\le \int_t^\infty |t-s|^{-\frac{1}{2}}\big\|U^{-1}[N_z(u(s))-N_z(\tilde{u}(s))]\big\|_{H_x^{1,3/2}}\,ds.\end{aligned}$$

Writing w to indicate that either u or $\tilde{u}$ may appear, we have

$$U^{-1}[N_z(u)-N_z(\tilde{u})]=\sum_{k=1}^{4}\emptyset[w^k(u-\tilde{u})]+\frac{U^{-1}\nabla}{\langle\nabla\rangle^2}\cdot[(u-\tilde{u})\nabla w+w\nabla(u-\tilde{u})]+U\sum_{k=2}^{4}\emptyset[w^k(u-\tilde{u})].$$

We estimate the contribution of the quadratic terms via

$$\begin{aligned}&\int_t^\infty |t-s|^{-\frac{1}{2}}\bigg\|\emptyset[w(u-\tilde{u})](s)+\frac{U^{-1}\nabla}{\langle\nabla\rangle^2}\cdot[(u-\tilde{u})\nabla w+w\nabla(u-\tilde{u})](s)\bigg\|_{H_x^{1,3/2}}ds\\ &\qquad\lesssim \|w\|_{X_T}\|u-\tilde{u}\|_{X_T}\int_t^\infty |t-s|^{-\frac{1}{2}}s^{-1}\,ds \lesssim t^{-\frac{1}{2}}\eta\|u-\tilde{u}\|_{X_T}.\end{aligned}$$

Arguing as in (7-2), we obtain

$$\begin{aligned}&\int_t^\infty |t-s|^{-\frac{1}{2}}\bigg\|\sum_{k=2}^{4}\emptyset[w^k(u-\tilde{u})](s)+U\emptyset[w^k(u-\tilde{u})](s)\bigg\|_{H_x^{1,3/2}}ds\\ &\qquad\lesssim \|w\|_{X_T}^k\|u-\tilde{u}\|_{X_T}\int_t^\infty |t-s|^{-\frac{1}{2}}s^{-\frac{k+1}{2}}\,ds \lesssim t^{-\frac{k}{2}}\eta^k\|u-\tilde{u}\|_{X_T}.\end{aligned}$$

Thus for η sufficiently small we get

$$8\big\|V^{-1}[\Phi_2(u)-\Phi_2(\tilde{u})]\big\|_{X_T}\le \tfrac{1}{4}d((u,z),(\tilde{u},\tilde{z})).$$

This completes the proof that Φ is a contraction on B. Hence there exists a unique $(u,z)\in B$ such that $\Phi(u,z)=(u,z)$. In particular $z=M(u)$ and (u,z) solves (5-4)–(5-5) on $(T,\infty)\times\mathbb{R}^3$. We note that by construction we have $u_1\in H_x^1$ and $u\in L_x^3\cap L_x^6$. In particular, $q(u)=|u|^2+2u_1\in L_x^2$ and hence $u(t)\in\mathcal{E}$ for $t>T$.

For $\gamma \in [\frac{2}{3}, 1)$, Theorem 3.3 guarantees that the solution u can be extended (in a unique way) to be global in time. For $\gamma \in (0, \frac{2}{3})$, global existence follows from [Killip et al. 2012, Theorem 1.3], while uniqueness in the energy space follows from Theorem 3.3 (see also Remark 3.4).

Next we show that (1-25) holds; indeed, we prove the stronger claim (1-26). We first note that Strichartz combined with (7-3) gives

$$\|z(t) - e^{-itH} V u_+\|_{H_x^1} \lesssim t^{-\frac{1}{4}},$$

which in turn implies

$$\big\|V^{-1} z(t) - V^{-1} e^{-itH} V u_+\big\|_{H^1_{\text{real}} + i\dot{H}^1_{\text{real}}} \lesssim t^{-\frac{1}{4}}.$$

As $z = M(u)$, for $t > T$ we have

$$\big\|V^{-1} z(t) - u(t)\big\|_{H^1_{\text{real}} + i\dot{H}^1_{\text{real}}} \lesssim \big\|\langle\nabla\rangle^{-2} |u(t)|^2\big\|_{H_x^1} \lesssim \big\||u(t)|^2\big\|_{L_x^{3/2}} \lesssim t^{-1} \|u\|_{X_T}^2.$$

Therefore, by the triangle inequality we may conclude that

$$\big\|u(t) - V^{-1} e^{-itH} V u_+\big\|_{H^1_{\text{real}} + i\dot{H}^1_{\text{real}}} \lesssim t^{-\frac{1}{4}}.$$

By the arguments presented so far, it is clear that u is the unique solution in B_1 that obeys (1-25). This is slightly weaker than is claimed in Theorem 1.4, which places no restrictions on the $H^1_{\text{real}} + i\dot{H}^1_{\text{real}}$-norm of alternate solutions $v(t)$, nor any restriction on the value of T for which $\|v\|_{X_T} \leq 4\eta$; however, any solution $v(t)$ obeying (1-25) must have

$$\|v\|_{L_t^\infty([T,\infty); H^1_{\text{real}} + i\dot{H}^1_{\text{real}})} \leq 4\|u_+\|_{H^1_{\text{real}} + i\dot{H}^1_{\text{real}}}$$

for some T large enough. Thus the equality of $v(t)$ and $u(t)$ follows from the contraction mapping argument above with T large enough combined with uniqueness in the energy space. $\square$

Finally, we prove Corollary 1.7.

Proof of Corollary 1.7. The proof consists of showing that smallness of the weighted norms implies the smallness condition (1-24). In view of (2-4), it suffices to show

$$\|e^{\pm itH} u_+\|_{H_x^{1,3}} \lesssim |t|^{-\frac{1}{2}} \eta \quad \text{and} \quad \|e^{\pm itH} U^{-1} \operatorname{Re} u_+\|_{H_x^{1,3}} \lesssim |t|^{-\frac{1}{2}} \eta.$$

By the dispersive estimate (2-3) and Hölder,

$$\|e^{\pm itH} u_+\|_{H_x^{1,3}} \lesssim |t|^{-\frac{1}{2}} \|\langle\nabla\rangle u_+\|_{L_x^{3/2}} \lesssim |t|^{-\frac{1}{2}} \big\|\langle x\rangle^{\frac{1}{2}+} \langle\nabla\rangle u_+\big\|_{L_x^2}$$

and

$$\|e^{\pm itH} U^{-1} \operatorname{Re} u_+\|_{H_x^{1,3}} \lesssim |t|^{-\frac{1}{2}} \big\|U^{-\frac{5}{6}} \langle\nabla\rangle \operatorname{Re} u_+\big\|_{L_x^{3/2}}.$$

Using Hölder and Sobolev embedding, we obtain

$$\begin{aligned} \|\nabla U^{-\frac{5}{6}} \operatorname{Re} u_+\|_{L_x^{3/2}} &\lesssim \|\langle\nabla\rangle u_+\|_{L_x^{3/2}} \lesssim \|\langle x\rangle^{\frac{1}{2}+} \langle\nabla\rangle u_+\|_{L_x^2}, \\ \|U^{-\frac{5}{6}} \operatorname{Re} u_+\|_{L_x^{3/2}} &\lesssim \big\||\nabla|^{\frac{5}{6}} U^{-\frac{5}{6}} \operatorname{Re} u_+\big\|_{L_x^{18/17}} \lesssim \big\|\langle x\rangle^{\frac{4}{3}+} \langle\nabla\rangle^{\frac{5}{6}} \operatorname{Re} u_+\big\|_{L_x^2}. \end{aligned}$$

$\square$

8. Acknowledgements

Killip was supported by NSF grant DMS-1265868. Murphy was supported by DMS-1400706. Visan was supported by NSF grant DMS-1161396. We are indebted to the Hausdorff Institute of Mathematics, which hosted us during our work on this project.

References

[Bahouri and Gérard 1999] H. Bahouri and P. Gérard, "High frequency approximation of solutions to critical nonlinear wave equations", *Amer. J. Math.* **121**:1 (1999), 131–175. MR 1705001 Zbl 0919.35089

[Barashenkov et al. 1989] I. V. Barashenkov, A. D. Gocheva, V. G. Makhan'kov, and I. V. Puzynin, "Stability of the soliton-like 'bubbles' ", *Phys. D* **34**:1–2 (1989), 240–254. MR 982390 Zbl 0697.35127

[Bourgain 1999] J. Bourgain, "Global wellposedness of defocusing critical nonlinear Schrödinger equation in the radial case", *J. Amer. Math. Soc.* **12**:1 (1999), 145–171. MR 1626257 Zbl 0958.35126

[Cazenave 2003] T. Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics **10**, American Mathematical Society, Providence, RI, 2003. MR 2002047 Zbl 1055.35003

[Christ and Weinstein 1991] F. M. Christ and M. I. Weinstein, "Dispersion of small amplitude solutions of the generalized Korteweg–de Vries equation", *J. Funct. Anal.* **100**:1 (1991), 87–109. MR 1124294 Zbl 0743.35067

[Coifman and Meyer 1978] R. Coifman and Y. Meyer, "Commutateurs d'intégrales singulières et opérateurs multilinéaires", *Ann. Inst. Fourier* (*Grenoble*) **28**:3 (1978), 177–202. MR 511821 Zbl 0368.47031

[Colliander et al. 2008] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Global well-posedness and scattering for the energy-critical nonlinear Schrödinger equation in $\mathbb{R}^3$", *Ann. of Math.* (2) **167**:3 (2008), 767–865. MR 2415387 Zbl 1178.35345

[Gérard 2006] P. Gérard, "The Cauchy problem for the Gross–Pitaevskii equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23**:5 (2006), 765–779. MR 2259616 Zbl 1122.35133

[Ginibre and Velo 1992] J. Ginibre and G. Velo, "Smoothing properties and retarded estimates for some dispersive evolution equations", *Comm. Math. Phys.* **144**:1 (1992), 163–188. MR 1151250 Zbl 0762.35008

[Ginzburg and Pitaevskiĭ 1958] V. L. Ginzburg and L. P. Pitaevskiĭ, "On the theory of superfluidity", *Z. Eksper. Teoret. Fiz.* **32**:5 (1958), 1240–1245. In Russian; translated in *Sov. Phys. JETP* **34**:7 (1958), 858–861. MR 0105929

[Ginzburg and Sobyanin 1976] V. L. Ginzburg and A. A. Sobyanin, "Superfluidity of helium II near the λ-point", *Usp. Fiz. Nauk* **19**:10 (1976), 153–216. In Russian; translated in *Sov. Phys. Usp.* **19**:10 (1976), 773–812.

[Gustafson et al. 2006] S. Gustafson, K. Nakanishi, and T.-P. Tsai, "Scattering for the Gross–Pitaevskii equation", *Math. Res. Lett.* **13**:2–3 (2006), 273–285. MR 2231117 Zbl 1119.35084

[Gustafson et al. 2007] S. Gustafson, K. Nakanishi, and T.-P. Tsai, "Global dispersive solutions for the Gross–Pitaevskii equation in two and three dimensions", *Ann. Henri Poincaré* **8**:7 (2007), 1303–1331. MR 2360438 Zbl 05218113

[Gustafson et al. 2009] S. Gustafson, K. Nakanishi, and T.-P. Tsai, "Scattering theory for the Gross–Pitaevskii equation in three dimensions", *Commun. Contemp. Math.* **11**:4 (2009), 657–707. MR 2559713 Zbl 1180.35481

[Kartavenko 1984] V. G. Kartavenko, "Soliton-like solutions in nuclear hydrodynamics", *Yad. Fiz.* **40** (1984), 377–388. In Russian; translated in *Sov. J. Nucl. Phys.* **40** (1984), 240–246. Zbl 0596.76022

[Keel and Tao 1998] M. Keel and T. Tao, "Endpoint Strichartz estimates", *Amer. J. Math.* **120**:5 (1998), 955–980. MR 1646048 Zbl 0922.35028

[Keraani 2001] S. Keraani, "On the defect of compactness for the Strichartz estimates of the Schrödinger equations", *J. Differential Equations* **175**:2 (2001), 353–392. MR 1855973 Zbl 1038.35119

[Killip et al. 2012] R. Killip, T. Oh, O. Pocovnicu, and M. Vişan, "Global well-posedness of the Gross–Pitaevskii and cubic-quintic nonlinear Schrödinger equations with non-vanishing boundary conditions", *Math. Res. Lett.* **19**:5 (2012), 969–986. MR 3039823 Zbl 1291.35352

[Killip et al. 2014] R. Killip, T. Oh, O. Pocovnicu, and M. Visan, "Solitons and scattering for the cubic-quintic nonlinear Schrödinger equation on $\mathbb{R}^3$", preprint, 2014. arXiv 1409.6734

[Meyer and Coifman 1991] Y. Meyer and R. R. Coifman, *Ondelettes et opérateurs, III: Opérateurs multilinéaires*, Hermann, Paris, 1991. MR 1160989 Zbl 0745.42012

[Nakanishi 2001] K. Nakanishi, "Asymptotically-free solutions for the short-range nonlinear Schrödinger equation", *SIAM J. Math. Anal.* **32**:6 (2001), 1265–1271. MR 1856248 Zbl 0979.35132

[Pushkarov and Kojnov 1978] D. Pushkarov and Z. Kojnov, "Contribution to the theory of defectons in quantum crystals", *Z. Eksper. Teoret. Fiz.* **47**:5 (1978), 1845–1852. In Russian; translated in *Sov. Phys. JETP* **47**:5 (1978), 962–966.

[Pushkarov and Primatarova 1984] K. I. Pushkarov and M. T. Primatarova, "Solitary clusters of spin deviations and lattice deformation in an anharmonic ferromagnetic chain", *Phys. Stat. Sol.* (*b*) **123**:2 (1984), 573–584.

[Pushkarov and Primatarova 1986] K. I. Pushkarov and M. T. Primatarova, "Solitary excitations in molecular chains in the presence of vibronic states", *Phys. Stat. Sol.* (*b*) **133**:1 (1986), 253–265.

[Shatah 1985] J. Shatah, "Normal forms and quadratic nonlinear Klein–Gordon equations", *Comm. Pure Appl. Math.* **38**:5 (1985), 685–696. MR 803256 Zbl 0597.35101

[Strichartz 1977] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations", *Duke Math. J.* **44**:3 (1977), 705–714. MR 0512086 Zbl 0372.35001

[Tao et al. 2007] T. Tao, M. Visan, and X. Zhang, "The nonlinear Schrödinger equation with combined power-type nonlinearities", *Comm. Partial Differential Equations* **32**:7–9 (2007), 1281–1343. MR 2354495 Zbl 1187.35245

[Visan 2014] M. Visan, "Dispersive equations", pp. 223–312 in *Dispersive equations and nonlinear waves*, Oberwolfach Seminars **45**, Birkhäuser, Basel, 2014. Zbl 1304.35003

[Zhang 2006] X. Zhang, "On the Cauchy problem of 3-D energy-critical Schrödinger equations with subcritical perturbations", *J. Differential Equations* **230**:2 (2006), 422–445. MR 2271498 Zbl 1106.35108

Received 28 Jul 2015. Revised 4 May 2016. Accepted 9 Jul 2016.

ROWAN KILLIP: killip@math.ucla.edu
Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, United States

JASON MURPHY: murphy@math.berkeley.edu
Department of Mathematics, University of California, Berkeley, CA 94720-3840, United States

MONICA VISAN: visan@math.ucla.edu
Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, United States

ANALYSIS AND PDE
Vol. 9, No. 7, 2016
dx.doi.org/10.2140/apde.2016.9.1575

MAGNETIC WELLS IN DIMENSION THREE

BERNARD HELFFER, YURI KORDYUKOV, NICOLAS RAYMOND AND SAN VŨ NGỌC

This paper deals with semiclassical asymptotics of the three-dimensional magnetic Laplacian in the presence of magnetic confinement. Using generic assumptions on the geometry of the confinement, we exhibit three semiclassical scales and their corresponding effective quantum Hamiltonians, by means of three microlocal normal forms à la Birkhoff. As a consequence, when the magnetic field admits a unique and nondegenerate minimum, we are able to reduce the spectral analysis of the low-lying eigenvalues to a one-dimensional $\hbar$-pseudodifferential operator whose Weyl's symbol admits an asymptotic expansion in powers of $\hbar^{\frac{1}{2}}$.

1. Introduction

1A. ***Motivation and context.*** The analysis of the magnetic Laplacian $(-i\hbar\nabla - \boldsymbol{A})^2$ in the semiclassical limit $\hbar \to 0$ has been the object of many developments in the last twenty years. The existence of the discrete spectrum for this operator, together with the analysis of the eigenvalues, is related to the notion of a "magnetic bottle", or quantum confinement by a pure magnetic field, and has important applications in physics. Moreover, motivated by investigations of the third critical field in Ginzburg–Landau theory for superconductivity, there has been great attention focused on estimates of the lowest eigenvalue. In the last decade, it appears that the spectral analysis of the magnetic Laplacian has acquired a life of its own. For a story and discussions about the subject, the reader is referred to the recent reviews [Fournais and Helffer 2010; Helffer and Kordyukov 2014; Raymond 2016].

In contrast to the wealth of studies exploring the semiclassical approximations of the Schrödinger operator $-\hbar^2\Delta + V$, the *classical* picture associated with the Hamiltonian $\|p - \boldsymbol{A}(q)\|^2$ has almost never been investigated to describe the *semiclassical* bound states (i.e., the eigenfunctions of low energy) of the magnetic Laplacian. The paper [Raymond and Vũ Ngọc 2015] is to our knowledge the first rigorous work in this direction. In that paper, which deals with the two-dimensional case, the notion of magnetic drift, well known to physicists, is cast in a symplectic framework, and using a semiclassical Birkhoff normal form (see, for instance, [Vũ Ngọc 2006; 2009; Charles and Vũ Ngọc 2008]) it becomes possible to describe all the eigenvalues of order $\mathcal{O}(\hbar)$. Independently, the asymptotic expansion of a smaller set of eigenvalues was established in [Helffer and Kordyukov 2011; 2015] through different methods which act directly on the quantum side: explicit unitary transforms and a Grushin-like reduction are used to reduce the two-dimensional operator to an effective one-dimensional operator.

MSC2010: primary 81Q20, 35P15; secondary 37G05, 70H15.
Keywords: magnetic fields, Birkhoff normal forms, microlocal analysis.

The three-dimensional case happens to be much harder. The only known results in this case that provide a full asymptotic expansion of a given eigenvalue concern toy models where the confinement is obtained by a boundary carrying a Neumann condition on a half-space [Raymond 2012] or on a wedge in [Popoff and Raymond 2013]. In the case of smooth confinement without boundary, a construction of quasimodes by Helffer and Kordyukov [2013] suggests what the expansions of the low-lying eigenvalues could be. But, as was expected by Colin de Verdière [1996] in his list of open questions, extending the symplectic and microlocal techniques to the three-dimensional case contains an intrinsic difficulty in the fact that the symplectic form cannot be nondegenerate on the characteristic hypersurface. The goal of our paper is to answer this question by fully carrying out this strategy. After averaging the cyclotron motion, the effect of the degeneracy of the symplectic form can be observed on the fact that the reduced operator is only partially elliptic. Hence, the key ingredient will be a separation of scales via the introduction of a new semiclassical parameter for only one part of the variables. These semiclassical scales are reminiscent of the three scales that have been exhibited in the classical picture in the large field limit; see [Benettin and Sempio 1994; Cheverry 2015]. They are also related to the Born–Oppenheimer-type of approximation in quantum mechanics (see, for instance, [Born and Oppenheimer 1927; Martinez 2007]). In fact, in a partially semiclassical context and under generic assumptions, a full asymptotic expansion of the first magnetic eigenvalues (and the corresponding WKB expansions) has been recently established in any dimension in the paper by Bonnaillie-Noël, Hérau and Raymond [Bonnaillie-Noël et al. 2016].

1B. ***Magnetic geometry.*** Let us now describe the geometry of the problem. The configuration space is

$$\mathbb{R}^3 = \{q_1 \boldsymbol{e}_1 + q_2 \boldsymbol{e}_2 + q_3 \boldsymbol{e}_3 \mid q_j \in \mathbb{R},\ j = 1, 2, 3\},$$

where $(\boldsymbol{e}_j)_{j=1,2,3}$ is the canonical basis of $\mathbb{R}^3$. The phase space is

$$\mathbb{R}^6 = \{(q, p) \in \mathbb{R}^3 \times \mathbb{R}^3\}$$

and we endow it with the canonical 2-form

$$\omega_0 = \mathrm{d}p_1 \wedge \mathrm{d}q_1 + \mathrm{d}p_2 \wedge \mathrm{d}q_2 + \mathrm{d}p_3 \wedge \mathrm{d}q_3. \tag{1-1}$$

We will use the standard Euclidean scalar product $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^3$ and $\|\cdot\|$, the associated norm. In particular, we can rewrite ω_0 as

$$\omega_0((u_1, u_2), (v_1, v_2)) = \langle v_1, u_2\rangle - \langle v_2, u_1\rangle \quad \forall u_1, u_2, v_1, v_2 \in \mathbb{R}^3.$$

The main object of this paper is the magnetic Hamiltonian, defined for all $(q, p) \in \mathbb{R}^6$ by

$$H(q, p) = \|p - \boldsymbol{A}(q)\|^2, \tag{1-2}$$

where $\boldsymbol{A} \in \mathcal{C}^\infty(\mathbb{R}^3, \mathbb{R}^3)$.

Let us now introduce the magnetic field. The vector field $\boldsymbol{A} = (A_1, A_2, A_3)$ is associated (via the Euclidean structure) with the 1-form

$$\alpha = A_1 \mathrm{d}q_1 + A_2 \mathrm{d}q_2 + A_3 \mathrm{d}q_3$$

and its exterior derivative is a 2-form, called the magnetic 2-form and expressed as

$$\mathrm{d}\alpha = (\partial_1 A_2 - \partial_2 A_1)\mathrm{d}q_1 \wedge \mathrm{d}q_2 + (\partial_1 A_3 - \partial_3 A_1)\mathrm{d}q_1 \wedge \mathrm{d}q_3 + (\partial_2 A_3 - \partial_3 A_2)\mathrm{d}q_2 \wedge \mathrm{d}q_3.$$

The form $\mathrm{d}\alpha$ may be identified with a vector field. If we let

$$\boldsymbol{B} = \nabla\times\boldsymbol{A} = (\partial_2 A_3 - \partial_3 A_2,\ \partial_3 A_1 - \partial_1 A_3,\ \partial_1 A_2 - \partial_2 A_1) = (B_1, B_2, B_3),$$

then we can write

$$\mathrm{d}\alpha = B_3 \mathrm{d}q_1 \wedge \mathrm{d}q_2 - B_2 \mathrm{d}q_1 \wedge \mathrm{d}q_3 + B_1 \mathrm{d}q_2 \wedge \mathrm{d}q_3. \tag{1-3}$$

The vector field $\boldsymbol{B}$ is called the magnetic field. Notice that we can express the 2-form $\mathrm{d}\alpha$ thanks to the magnetic matrix

$$M_{\boldsymbol{B}} = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}.$$

Indeed we have

$$\mathrm{d}\alpha(U, V) = \langle U, M_{\boldsymbol{B}} V\rangle = \langle U, V \times \boldsymbol{B}\rangle = [U, V, \boldsymbol{B}] \quad \forall (U, V) \in \mathbb{R}^3 \times \mathbb{R}^3, \tag{1-4}$$

where $[\cdot,\cdot,\cdot]$ is the canonical mixed product on $\mathbb{R}^3$. We note that $\boldsymbol{B}$ belongs to the kernels of $M_{\boldsymbol{B}}$ and $\mathrm{d}\alpha$.

An important role will be played by the characteristic hypersurface

$$\Sigma = H^{-1}(0),$$

which is the submanifold defined by the parametrization

$$\mathbb{R}^3 \ni q \mapsto j(q) := (q, \boldsymbol{A}(q)) \in \mathbb{R}^3 \times \mathbb{R}^3.$$

We may notice the relation between Σ, the symplectic structure and the magnetic field given by

$$j^*\omega_0 = \mathrm{d}\alpha, \tag{1-5}$$

where $\mathrm{d}\alpha$ is defined in (1-3).

1C. ***Confinement assumptions and discrete spectrum.*** This paper is devoted to the semiclassical analysis of the discrete spectrum of the magnetic Laplacian $\mathcal{L}_{\hbar,\boldsymbol{A}} := (-i\hbar\nabla_q - \boldsymbol{A}(q))^2$, which is the semiclassical Weyl quantization of H (see (2-1)). This means that we will consider that $\hbar$ belongs to $(0, \hbar_0)$ with $\hbar_0$ small enough.

If $\mathcal{L}$ is a self-adjoint operator, we denote its spectrum by $\mathfrak{s}(\mathcal{L})$. The discrete spectrum of $\mathcal{L}$ consists of the isolated eigenvalues with finite multiplicity. The essential spectrum is by definition the complement in $\mathfrak{s}(\mathcal{L})$ of the discrete spectrum and is denoted by $\mathfrak{s}_{\mathrm{ess}}(\mathcal{L})$. It is empty when $\mathcal{L}$ has compact resolvent.

It is known (see, for example, [Avron et al. 1978]) that $\mathcal{L}_{\hbar,\boldsymbol{A}}$ is essentially self-adjoint and we always consider with the same notation its self-adjoint extension.

Let us recall the assumptions under which the discrete spectrum actually exists. In two dimensions, with a nonvanishing magnetic field, a standard estimate (see [Avron et al. 1978; Cycon et al. 1987]) gives

$$\hbar \int_{\mathbb{R}^2} |B(q)||u(q)|^2 \, \mathrm{d}q \leqslant \langle \mathcal{L}_{\hbar,\boldsymbol{A}} u \mid u \rangle \quad \forall u \in \mathcal{C}_0^\infty(\mathbb{R}^2). \tag{1-6}$$

This implies that, as $B(q) \to +\infty$, the magnetic Laplacian has compact resolvent. Except in special cases when some components of the magnetic field have constant sign, this doesn't hold anymore in higher dimensions (see [Dufresnoy 1983]). One can give examples where $|B(q)| \to +\infty$ and the operator doesn't have a compact resolvent. We should impose a control of the oscillations of $\boldsymbol{B}$ at infinity. Under this condition, we get an estimate similar to (1-6) at the price of a small loss. When there exists a constant $C > 0$ such that

$$\|\nabla \boldsymbol{B}(q)\| \leqslant C(1 + b(q)) \quad \forall q \in \mathbb{R}^3, \tag{1-7}$$

and $b(q) := \|\boldsymbol{B}(q)\|$ tends to $+\infty$, one can show again that the magnetic Laplacian has compact resolvent [Helffer and Mohamed 1996].

In the semiclassical context, we would like to consider the case of $\mathbb{R}^3$ and, in addition to (1-7), a confining assumption which allows the presence of the essential spectrum above a certain threshold. More precisely:

Assumption 1.1. We assume that (1-7) holds and

$$b(q) \geqslant b_0 := \inf_{q \in \mathbb{R}^3} b(q) > 0. \tag{1-8}$$

Under Assumption 1.1, it is proven in [Helffer and Mohamed 1996, Theorem 3.1] that there exist $h_0 > 0$ and $C_0 > 0$ such that, for all $\hbar \in (0, h_0)$,

$$\hbar(1 - C_0 \hbar^{\frac{1}{4}}) \int_{\mathbb{R}^3} b(q)|u(q)|^2 \, \mathrm{d}q \leqslant \langle \mathcal{L}_{\hbar,\boldsymbol{A}} u \mid u \rangle \quad \forall u \in \mathcal{C}_0^\infty(\mathbb{R}^3). \tag{1-9}$$

In this case, if we do not assume that $b(q) \to +\infty$, the spectrum is not necessarily discrete, but using this inequality and Persson's theorem [1960], we obtain that the bottom of the essential spectrum is asymptotically above $\hbar b_1$, where

$$b_1 := \liminf_{|q| \to +\infty} b(q).$$

More precisely, under Assumption 1.1, there exist $h_0 > 0$ and $C_0 > 0$ such that, for all $\hbar \in (0, h_0)$,

$$\mathfrak{s}_{\mathrm{ess}}(\mathcal{L}_{\hbar,\boldsymbol{A}}) \subset \big[\hbar b_1 (1 - C_0 \hbar^{\frac{1}{4}}), +\infty\big). \tag{1-10}$$

Assumption 1.2. We assume that

$$0 < b_0 < b_1. \tag{1-11}$$

Moreover, we will assume that there exists a point $q_0 \in \mathbb{R}^3$ and $\varepsilon > 0$, $\tilde{\beta}_0 \in (b_0, b_1)$ such that

$$\{b(q) \leqslant \tilde{\beta}_0\} \subset D(q_0, \varepsilon), \tag{1-12}$$

where $D(q_0,\varepsilon)$ is the Euclidean ball centered at q_0 and of radius ε. For the rest of the article we let $\beta_0 \in (b_0, \tilde{\beta}_0)$. Without loss of generality, we can assume that $q_0 = 0$ and that $\boldsymbol{A}(0) = 0$ (which can be obtained with a change of gauge).

Note that Assumption 1.2 implies that the minimal value of b is attained inside $D(q_0,\varepsilon)$.

Throughout this paper, we will strengthen the assumptions on the nature of the point q_0. At some stage of our investigation, q_0 will be the unique minimum of b. Note in particular that (1-12) is satisfied as soon as b admits a unique and nondegenerate minimum.

1D. ***Informal description of the results.*** Let us now informally walk through the main results of this paper. We will assume (as precisely formulated in (1-11)–(1-12)) that the magnetic field does not vanish and is confining.

Of course, for eigenvalues of order $\mathcal{O}(\hbar)$, the corresponding eigenfunctions are microlocalized in the semiclassical sense near the characteristic manifold Σ (see, for instance, [Robert 1987; Zworski 2012]). Moreover, the confinement assumption implies that the eigenfunctions of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ associated with eigenvalues less than $\beta_0\hbar$ enjoy localization estimates à la Agmon. Therefore we will be reduced to investigating the magnetic geometry locally in space near a point $q_0 = 0 \in \mathbb{R}^3$ belonging to the confinement region and which, for notational simplicity, we may assume to be the origin.

Then, in a neighborhood of $(0, \boldsymbol{A}(0)) \in \Sigma$, there exist symplectic coordinates $(x_1,\xi_1,x_2,\xi_2,x_3,\xi_3)$ such that $\Sigma = \{x_1 = \xi_1 = \xi_3 = 0\}$ and $(0,\boldsymbol{A}(0))$ has coordinates $0 \in \mathbb{R}^6$. Hence Σ is parametrized by (x_2,ξ_2,x_3).

1D1. *First Birkhoff form.* In these coordinates suited for the magnetic geometry, it is possible to perform a semiclassical Birkhoff normal form and microlocally unitarily conjugate $\mathcal{L}_{\hbar,\boldsymbol{A}}$ to a first normal form $\mathcal{N}_\hbar = \mathrm{Op}_\hbar^w(N_\hbar)$ with an operator-valued symbol $N_\hbar$ depending on (x_2,ξ_2,x_3,ξ_3) in the form

$$N_\hbar = \xi_3^2 + b(x_2,\xi_2,x_3)\mathcal{I}_\hbar + f^\star(\hbar,\mathcal{I}_\hbar,x_2,\xi_2,x_3,\xi_3) + \mathcal{O}(|\mathcal{I}_\hbar|^\infty, |\xi_3|^\infty),$$

where $\mathcal{I}_h = \hbar^2 D_{x_1}^2 + x_1^2$ is the first encountered harmonic oscillator and where $(\hbar, I, x_2,\xi_2,x_3,\xi_3) \mapsto f^\star(\hbar, I, x_2,\xi_2,x_3,\xi_3)$ satisfies, for $I \in (0, I_0)$,

$$|f^\star(\hbar,I,x_2,\xi_2,x_3,\xi_3)| \leqslant C(|I|^{\frac{3}{2}} + |\xi_3|^3 + \hbar^{\frac{3}{2}}).$$

Since we wish to describe the spectrum in a spectral window containing at least the lowest eigenvalues, we are led to replace $\mathcal{I}_\hbar$ by its lowest eigenvalue $\hbar$ and thus, we are reduced to the two-dimensional pseudodifferential operator $\mathcal{N}_\hbar^{[1]} = \mathrm{Op}_\hbar^w(N_\hbar^{[1]})$, where

$$N_\hbar^{[1]} = \xi_3^2 + b(x_2,\xi_2,x_3)\hbar + f^\star(\hbar,\hbar,x_2,\xi_2,x_3,\xi_3) + \mathcal{O}(\hbar^\infty, |\xi_3|^\infty).$$

1D2. *Second Birkhoff form.* If we want to continue the normalization, we shall assume a new nondegeneracy condition (the first one was the positivity of b).

Now we assume that, for any (x_2,ξ_2) in a neighborhood of $(0,0)$, the function $x_3 \mapsto b(x_2,\xi_2,x_3)$ admits a unique and nondegenerate minimum denoted by $s(x_2,\xi_2)$. Then, by using a new symplectic transformation in order to center the analysis at the partial minimum $s(x_2,\xi_2)$, we get a new operator $\underline{\mathcal{N}}_\hbar^{[1]}$ whose Weyl symbol is in the form

$$\underline{N}_\hbar^{[1]} = \nu^2(x_2,\xi_2)(\xi_3^2+\hbar x_3^2) + \hbar b(x_2,\xi_2,s(x_2,\xi_2)) + \mathsf{remainders},$$

with

$$\nu(x_2,\xi_2) = \left(\tfrac{1}{2}\partial_3^2 b(x_2,\xi_2,s(x_2,\xi_2))\right)^{\frac{1}{4}} \tag{1-13}$$

and where the remainders have been properly normalized to be at least formal perturbations of the second harmonic oscillator $\xi_3^2+\hbar x_3^2$. Since the frequency of this oscillator is $\hbar^{-\frac{1}{2}}$ in the classical picture, we are naturally led to introduce the new semiclassical parameter

$$h = \hbar^{\frac{1}{2}}$$

and the new impulsion

$$\xi = \hbar^{\frac{1}{2}}\tilde{\xi}$$

so that

$$\mathrm{Op}_\hbar^w(\xi_3^2+\hbar x_3^2) = h^2\,\mathrm{Op}_h^w(\tilde{\xi}_3^2+x_3^2).$$

We therefore get the h-symbol of $\underline{\mathcal{N}}_\hbar^{[1]}$,

$$\underline{\mathsf{N}}_h^{[1]} = h^2\nu^2(x_2,h\tilde{\xi}_2)(\tilde{\xi}_3^2+x_3^2) + h^2 b(x_2,h\tilde{\xi}_2,s(x_2,h\tilde{\xi}_2)) + \mathsf{remainders}.$$

We can again perform a Birkhoff analysis in the space of formal series given by $\mathscr{E} = \mathscr{F}[\![x_3,\tilde{\xi}_3,h]\!]$, where $\mathscr{F}$ is a space of symbols in the form $c(h,x_2,h\tilde{\xi}_2)$. We get the new operator $\mathfrak{M}_h = \mathrm{Op}_h^w(\mathsf{M}_h)$, with

$$\mathsf{M}_h = h^2 b(x_2,h\tilde{\xi}_2,s(x_2,h\tilde{\xi}_2)) + h^2\mathcal{J}_h\,\mathrm{Op}_h^w\,\nu^2(x_2,h\tilde{\xi}_2) + h^2 g^\star(h,\mathcal{J}_h,x_2,h\tilde{\xi}_2) + \mathsf{remainders},$$

where $\mathcal{J}_h = \mathrm{Op}_h^w(\tilde{\xi}_3^2+x_3^2)$ and $g^\star(h,J,x_2,\xi_2)$ is of order three with respect to $(J^{\frac{1}{2}},h^{\frac{1}{2}})$. Motivated again by the perspective of describing the low-lying eigenvalues, we replace $\mathcal{J}_h$ by h and rewrite the symbol with the old semiclassical parameter $\hbar$ to get the operator $\mathcal{M}_\hbar^{[1]} = \mathrm{Op}_h^w(\mathsf{M}_h^{[1]}) = \mathrm{Op}_\hbar^w(M_\hbar^{[1]})$, with

$$M_\hbar^{[1]} = \hbar b(x_2,\xi_2,s(x_2,\xi_2)) + \hbar^{\frac{3}{2}}\nu^2(x_2,\xi_2) + \hbar g^\star(\hbar^{\frac{1}{2}},\hbar^{\frac{1}{2}},x_2,\xi_2) + \mathsf{remainders}. \tag{1-14}$$

1D3. *Third Birkhoff form.* The last generic assumption is the uniqueness and nondegeneracy of the minimum of the new "principal" symbol

$$(x_2,\xi_2) \mapsto b(x_2,\xi_2,s(x_2,\xi_2))$$

that implies that b admits a unique and nondegenerate minimum at $(0,0,0)$. Up to an $\hbar^{\frac{1}{2}}$-dependent translation in the phase space and a rotation, we are essentially reduced to a standard Birkhoff normal form with respect to the third harmonic oscillator $\mathcal{K}_\hbar = \hbar^2 D_{x_2}^2 + x_2^2$.

Note that all our normal forms may be used to describe the classical dynamics of a charged particle in a confining magnetic field (see Figure 1).

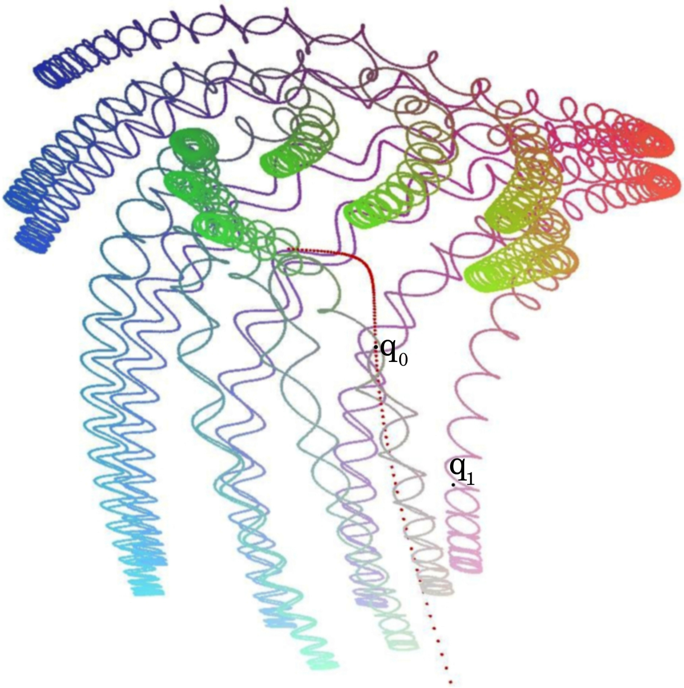


Figure 1. The dashed line represents the integral curve of the confining magnetic field $\boldsymbol{B} = \operatorname{curl} \boldsymbol{A}$ through $q_0 = (0.5, 0.6, 0.7)$ for $\boldsymbol{B}(x, y, z) = \left(\frac{1}{2}y, \frac{1}{2}z, \sqrt{1+x^2}\right)$ and the full line represents the projection in the q-space of the Hamiltonian trajectory with initial condition (q_0, p_0) (with $p_0 = (-0.6, 0.01, 0.2)$) ending at (q_1, p_1). The motion is easier to follow on a video: see http://tinyurl.com/3DMagneticFlow.

1D4. *Microlocalization.* Of course, at each step, we will have to provide accurate microlocal estimates of the eigenfunctions of the different operators to get a good control of the different remainders. In a first approximation, we will get localizations at the scales $x_1, \xi_1, \xi_3 \sim \hbar^\delta$ ($\delta > 0$ is small enough) and $x_2, \xi_2, x_3 \sim 1$. In a second approximation, we will get $x_3, \tilde{\xi}_3 \sim \hbar^\delta$. In the final step, we will refine the localization by $x_2, \xi_2 \sim \hbar^\delta$.

1E. *A semiclassical eigenvalue estimate.*

Let us already state one of the consequences of our investigation. It will follow from the third normal form that we have a complete description of the spectrum below the threshold $b_0\hbar + 3\nu^2(0,0)\hbar^{\frac{3}{2}}$. This description is reminiscent of the results à la Bohr–Sommerfeld of [Helffer and Robert 1984; Helffer and Sjöstrand 1989, Appendix B] (see also [Helffer and Kordyukov 2015, Remark 1.4]) obtained in the case of one-dimensional semiclassical operators.

Theorem 1.3. *Assume that b admits a unique and nondegenerate minimum at q_0. Define*

$$\sigma = \frac{\mathsf{Hess}_{q_0} b\,(\boldsymbol{B}(q_0), \boldsymbol{B}(q_0))}{2b_0^2}, \qquad \theta = \sqrt{\frac{\det \mathsf{Hess}_{q_0} b}{\mathsf{Hess}_{q_0} b\,(\boldsymbol{B}(q_0), \boldsymbol{B}(q_0))}}. \tag{1-15}$$

There exists a function $k^\star \in \mathcal{C}_0^\infty(\mathbb{R}^2)$ *with arbitrarily small compact support, and*

$$k^\star(\hbar^{\frac{1}{2}}, Z) = \mathcal{O}((\hbar + |Z|)^{\frac{3}{2}})$$

when $(\hbar, Z) \to (0,0)$, *such that the following holds:*

For all $c \in (0,3)$, *the spectrum of* $\mathcal{L}_{\hbar,\boldsymbol{A}}$ *below* $b_0\hbar + c\sigma^{\frac{1}{2}}\hbar^{\frac{3}{2}}$ *coincides modulo* $\mathcal{O}(\hbar^\infty)$ *with the spectrum of the operator* $\mathcal{F}_\hbar$ *acting on* $L^2(\mathbb{R}_x)$ *given by*

$$\mathcal{F}_\hbar = b_0\hbar + \sigma^{\frac{1}{2}}\hbar^{\frac{3}{2}} - \frac{\zeta}{2\theta}\hbar^2 + \hbar\big(\tfrac{1}{2}\theta\mathcal{K}_\hbar + k^\star(\hbar^{\frac{1}{2}}, \mathcal{K}_\hbar)\big), \qquad \mathcal{K}_\hbar = \hbar^2 D_x^2 + x^2,$$

with some constant ζ.

Remark 1.4. The constant ζ in Theorem 1.3 is given by the formula

$$\zeta = \|\nabla \nu^2(0,0)\|^2,$$

where the function ν is given in (1-13). Observe also that $\sigma = \nu^4(0,0)$.

Corollary 1.5. *Under the hypothesis of Theorem 1.3, let* $(\lambda_m(\hbar))_{m\geqslant 1}$ *be the nondecreasing sequence of the eigenvalues of* $\mathcal{L}_{\hbar,\boldsymbol{A}}$. *For any* $c \in (0,3)$, *let*

$$\mathsf{N}_{\hbar,c} := \{m \in \mathbb{N}^* \mid \lambda_m(\hbar) \leqslant \hbar b_0 + c\sigma^{\frac{1}{2}}\hbar^{\frac{3}{2}}\}.$$

Then the cardinal of $\mathsf{N}_{\hbar,c}$ *is of order* $\hbar^{-\frac{1}{2}}$, *and there exist* $\upsilon_1, \upsilon_2 \in \mathbb{R}$ *and* $\hbar_0 > 0$ *such that*

$$\lambda_m(\hbar) = \hbar b_0 + \sigma^{\frac{1}{2}}\hbar^{\frac{3}{2}} + \left[\theta\big(m - \tfrac{1}{2}\big) - \frac{\zeta}{2\theta}\right]\hbar^2 + \upsilon_1\big(m - \tfrac{1}{2}\big)\hbar^{\frac{5}{2}} + \upsilon_2\big(m - \tfrac{1}{2}\big)^2\hbar^3 + \mathcal{O}(\hbar^{\frac{5}{2}})$$

uniformly for $\hbar \in (0, \hbar_0)$ *and* $m \in \mathsf{N}_{\hbar,c}$.

In particular, the splitting between two consecutive eigenvalues satisfies

$$\lambda_{m+1}(\hbar) - \lambda_m(\hbar) = \theta\hbar^2 + \mathcal{O}(\hbar^{\frac{5}{2}}).$$

Proof. If the support of $k^\star$ is small enough, the hypothesis $k^\star(\hbar^{\frac{1}{2}}, Z) = \mathcal{O}((\hbar + |Z|)^{\frac{3}{2}})$ implies that, when $\hbar$ is small enough,

$$(1+\eta)\mathcal{K}_\hbar \geqslant \mathcal{K}_\hbar + \frac{2}{\theta}k^\star(\hbar^{\frac{1}{2}}, \mathcal{K}_\hbar) \geqslant (1-\eta)\mathcal{K}_\hbar$$

for some small $\eta > 0$. Therefore, since the eigenvalues of $\mathcal{K}_\hbar$ are $(2m-1)\hbar$, $m \in \mathbb{N}^*$, the variational principle implies that the number of eigenvalues of $\mathcal{K}_\hbar + (2/\theta)k^\star(\hbar^{\frac{1}{2}}, \mathcal{K}_\hbar)$ below a threshold $C_\hbar$ belongs to

$$\left[\frac{1}{2}\left(\frac{C_\hbar}{\hbar(1+\eta)} + 1\right), \frac{1}{2}\left(\frac{C_\hbar}{\hbar(1-\eta)} + 1\right)\right].$$

Taking $C_\hbar = (2/\theta)(c-1)\sigma^{\frac{1}{2}}\hbar^{\frac{1}{2}} + (\zeta/\theta^2)\hbar$, and applying the theorem, we obtain the estimate for the cardinal of $\mathsf{N}_{\hbar,c}$. The corresponding eigenvalues of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ are of the form

$$\lambda_m(\hbar) = \hbar b_0 + \sigma^{\frac{1}{2}}\hbar^{\frac{3}{2}} - \frac{\zeta}{2\theta}\hbar^2 + \hbar\big[\theta\big(m - \tfrac{1}{2}\big)\hbar + k^\star(\hbar^{\frac{1}{2}}, 2m-1)\big] + \mathcal{O}(\hbar^\infty),$$

with $(2m-1)\hbar \leqslant C_\hbar/(1-\eta)$. Therefore there exists a constant $\tilde{C}>0$, independent of $\hbar$, such that all $m \in \mathsf{N}_{\hbar,c}$ satisfy the inequality $(2m-1)\hbar \leqslant \tilde{C}\hbar^{\frac{1}{2}}$. Writing

$$k^\star(\hbar^{\frac{1}{2}}, Z) = c_0\hbar^{\frac{3}{2}} + \upsilon_1\hbar^{\frac{1}{2}}\left(\tfrac{1}{2}Z\right) + c_1\hbar^2 + \upsilon_2\left(\tfrac{1}{2}Z\right)^2 + \upsilon_3\hbar Z + \hbar^{\frac{1}{2}}\mathcal{O}(h+|Z|)^2 + \mathcal{O}(Z^3),$$

we see that, for $m \in \mathsf{N}_{\hbar,c}$,

$$k^\star(\hbar^{\frac{1}{2}}, (2m-1)\hbar) = \upsilon_1\hbar^{\frac{3}{2}}\left(m-\tfrac{1}{2}\right) + \upsilon_2\hbar^2\left(m-\tfrac{1}{2}\right)^2 + \mathcal{O}(\hbar^{\frac{3}{2}}),$$

which gives the result. □

Remark 1.6. An upper bound of $\lambda_m(\hbar)$ for fixed $\hbar$-independent m with remainder in $\mathcal{O}(\hbar^{\frac{9}{4}})$ was obtained in [Helffer and Kordyukov 2013] through a quasimodes construction involving powers of $\hbar^{\frac{1}{4}}$. To the authors' knowledge, Corollary 1.5 gives the most accurate description of magnetic eigenvalues in three dimensions, in such a large spectral window. Note also that the nondegeneracy assumption on the norm of $\boldsymbol{B}$ is not purely technical. Indeed, at the quantum level, it appears through microlocal reductions matching with the splitting of the Hamiltonian dynamics into three scales: the cyclotron motion around field lines, the center-guide oscillation along the field lines, and the oscillation within the space of field lines.

1F. *Organization of the paper.* The paper is organized as follows. In Section 2, we state our main results. Section 3 is devoted to the investigation of the first normal form (see Theorem 2.1 and Corollary 2.4). In Section 4 we analyze the second normal form (see Theorems 2.8 and 2.11 and Corollaries 2.9 and 2.13). Section 5 is devoted to the third normal form (see Theorem 2.15 and Corollary 2.16).

2. Statements of the main results

We recall (see [Dimassi and Sjöstrand 1999, Chapter 7]) that a function $m:\mathbb{R}^d \to [0,\infty)$ is an order function if there exist constants $N_0, C_0 > 0$ such that

$$m(X) \leqslant C_0\langle X-Y\rangle^{N_0} m(Y)$$

for any $X, Y \in \mathbb{R}^d$. The symbol class $S(m)$ is the space of smooth $\hbar$-dependent functions $a_\hbar : \mathbb{R}^d \to \mathbb{C}$ such that, for all $\alpha \in \mathbb{N}^d$,

$$|\partial_x^\alpha a_\hbar(x)| \leqslant C_\alpha m(x) \quad \forall h \in (0,1].$$

Throughout this paper, we assume that the components of the vector potential $\boldsymbol{A}$ belong to a symbol class $S(m)$. Note that this implies that $\boldsymbol{B} \in S(m)$, and conversely, if $\boldsymbol{B} \in S(m)$, then there exist a potential $\boldsymbol{A}$ and another order function m' such that $\boldsymbol{A} \in S(m')$. Moreover, the magnetic Hamiltonian $H(x,\xi) = \|\xi - \boldsymbol{A}(x)\|^2$ belongs to $S(m'')$ for an order function m'' on $\mathbb{R}^6$.

We will work with the Weyl quantization; for a classical symbol $a_\hbar = a(x,\xi;\hbar) \in S(m)$, it is defined as

$$\mathrm{Op}_\hbar^w\, a\, \psi(x) = \frac{1}{(2\pi\hbar)^d}\int_{\mathbb{R}^{2d}} e^{i\langle x-y,\xi\rangle/\hbar} a\left(\frac{x+y}{2},\xi\right)\psi(y)\,\mathrm{d}y\,\mathrm{d}\xi \quad \forall \psi \in \mathcal{S}(\mathbb{R}^d). \tag{2-1}$$

The Weyl quantization of H is the magnetic Laplacian $\mathcal{L}_{\hbar,\boldsymbol{A}} = (-i\hbar\nabla - \boldsymbol{A})^2$.

2A. ***Normal forms and spectral reductions.*** Let us introduce our first Birkhoff normal form $\mathcal{N}_\hbar$.

Theorem 2.1. *If $\boldsymbol{B}(0) \neq 0$, there exists a neighborhood of $(0, \boldsymbol{A}(0))$ endowed with symplectic coordinates $(x_1, \xi_1, x_2, \xi_2, x_3, \xi_3)$ in which $\Sigma = \{x_1 = \xi_1 = \xi_3 = 0\}$ and $(0, \boldsymbol{A}(0))$ has coordinates $0 \in \mathbb{R}^6$, and there exist an associated unitary Fourier integral operator $U_\hbar$ and a smooth function $f^\star(\hbar, Z, x_2, \xi_2, x_3, \xi_3)$ compactly supported with respect to Z and ξ_3, whose Taylor series with respect to $Z, \xi_3, \hbar$ is*

$$\sum_{k \geqslant 3} \sum_{2\ell+2m+\beta=k} \hbar^\ell c^\star_{\ell,m,\beta}(x_2, \xi_2, x_3) Z^m \xi_3^\beta,$$

such that

$$U_\hbar^* \mathcal{L}_{\hbar,\boldsymbol{A}} U_\hbar = \mathcal{N}_\hbar + \mathcal{R}_\hbar, \tag{2-2}$$

with

$$\mathcal{N}_\hbar = \hbar^2 D_{x_3}^2 + \mathcal{I}_\hbar \operatorname{Op}_\hbar^w b + \operatorname{Op}_\hbar^w f^\star(\hbar, \mathcal{I}_\hbar, x_2, \xi_2, x_3, \xi_3),$$

and where

(a) *we have $\mathcal{I}_\hbar = \hbar^2 D_{x_1}^2 + x_1^2$,*

(b) *the operator $\operatorname{Op}_\hbar^w f^\star(\hbar, \mathcal{I}_\hbar, x_2, \xi_2, x_3, \xi_3)$ has to be understood as the Weyl quantization of an operator-valued symbol,*

(c) *the remainder $\mathcal{R}_\hbar$ is a pseudodifferential operator such that, in a neighborhood of the origin, the Taylor series of its symbol with respect to $(x_1, \xi_1, \xi_3, \hbar)$ is* 0.

Remark 2.2. In Theorem 2.1, the direction of $\boldsymbol{B}$ considered as a vector field on Σ is $\partial/\partial x_3$ and the function $b \in \mathcal{C}^\infty(\mathbb{R}^6)$ stands for $b \circ j_{|\Sigma}^{-1} \circ \pi$, where $\pi : \mathbb{R}^6 \to \Sigma : \pi(x_1, \xi_1, x_2, \xi_2, x_3, \xi_3) = (0, 0, x_2, \xi_2, x_3, 0)$. In addition, note that the support of $f^\star$ in Z and ξ_3 may be chosen as small as we want.

Remark 2.3. In the context of Weyl's asymptotics, a close version of this theorem appears in [Ivrii 1998, Chapter 6].

In order to investigate the spectrum of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ near the low-lying energies, we introduce the pseudodifferential operator

$$\mathcal{N}_\hbar^{[1]} = \hbar^2 D_{x_3}^2 + \hbar \operatorname{Op}_\hbar^w b + \operatorname{Op}_\hbar^w f^\star(\hbar, \hbar, x_2, \xi_2, x_3, \xi_3),$$

obtained by replacing $\mathcal{I}_\hbar$ by $\hbar$.

Corollary 2.4. *We introduce*

$$\mathcal{N}_\hbar^\sharp = \operatorname{Op}_\hbar^w(N_\hbar^\sharp), \tag{2-3}$$

with

$$N_\hbar^\sharp = \xi_3^2 + \mathcal{I}_\hbar \underline{b}(x_2, \xi_2, x_3) + f^{\star,\sharp}(\hbar, \mathcal{I}_\hbar, x_2, \xi_2, x_3, \xi_3),$$

and where $\underline{b}$ is a smooth extension of b away from $D(0, \varepsilon)$ such that (1-12) still holds and where $f^{\star,\sharp} = \chi(x_2, \xi_2, x_3) f^\star$, with χ a smooth cutoff function that is 1 in a neighborhood of $D(0, \varepsilon)$. We also define the operator attached to the first eigenvalue of $\mathcal{I}_\hbar$,

$$\mathcal{N}_\hbar^{[1],\sharp} = \operatorname{Op}_\hbar^w(N_\hbar^{[1],\sharp}), \tag{2-4}$$

where $N_\hbar^{[1],\sharp} = \xi_3^2 + \hbar \underline{b}(x_2, \xi_2, x_3) + f^{\star,\sharp}(\hbar, \hbar, x_2, \xi_2, x_3, \xi_3)$.

If ε and the support of $f^\star$ are small enough, then we have:

(a) *The spectra of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ and $\mathcal{N}_\hbar^\sharp$ below $\beta_0\hbar$ coincide modulo $\mathcal{O}(\hbar^\infty)$.*

(b) *For all $c\in(0,\min(3b_0,\beta_0))$, the spectra of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ and $\mathcal{N}_\hbar^{[1],\sharp}$ below $c\hbar$ coincide modulo $\mathcal{O}(\hbar^\infty)$.*

Let us now state our results concerning the normal form of $\mathcal{N}_\hbar^{[1]}$ (or $\mathcal{N}_\hbar^{[1],\sharp}$) under the following assumption.

Notation 2.5. If $f=f(z)$ is a differentiable function, we denote by $T_{\boldsymbol{z}}f(\cdot)$ its tangent map at the point $\boldsymbol{z}$. Moreover, if f is twice differentiable, the second derivative of f is denoted by $T_{\boldsymbol{z}}^2 f(\cdot,\cdot)$.

Assumption 2.6. We assume that $T_0^2 b(\boldsymbol{B}(0),\boldsymbol{B}(0))>0$.

Remark 2.7. If the function b admits a unique and positive minimum at 0 and it is nondegenerate, then Assumption 2.6 is satisfied.

Under Assumption 2.6, we have $\partial_3 b(0,0,0)=0$ and, in the coordinates (x_2,ξ_2,x_3) given in Theorem 2.1,

$$\partial_3^2 b(0,0,0)>0. \tag{2-5}$$

It follows from (2-5) and the implicit function theorem that, for small x_2, there exists a smooth function $(x_2,\xi_2)\mapsto s(x_2,\xi_2)$, with $s(0,0)=0$, such that

$$\partial_3 b(x_2,\xi_2,s(x_2,\xi_2))=0. \tag{2-6}$$

The point $s(x_2,\xi_2)$ is the unique (in a neighborhood of $(0,0,0)$) minimum of $x_3\mapsto b(x_2,\xi_2,x_3)$. We define

$$\nu(x_2,\xi_2):=\left(\tfrac{1}{2}\partial_3^2 b(x_2,\xi_2,s(x_2,\xi_2))\right)^{\frac{1}{4}}.$$

Theorem 2.8. *Under Assumption 2.6, there exist a neighborhood $\mathcal{V}_0$ of 0 and a Fourier integral operator $V_\hbar$ which is microlocally unitary near $\mathcal{V}_0$ and such that*

$$V_\hbar^*\mathcal{N}_\hbar^{[1]}V_\hbar=:\underline{\mathcal{N}}_\hbar^{[1]}=\mathrm{Op}_\hbar^w(\underline{N}_\hbar^{[1]}),$$

where $\underline{N}_\hbar^{[1]}=\nu^2(x_2,\xi_2)(\xi_3^2+\hbar x_3^2)+\hbar b(x_2,\xi_2,s(x_2,\xi_2))+\underline{r}_\hbar$ and $\underline{r}_\hbar$ is a semiclassical symbol such that $\underline{r}_\hbar=\mathcal{O}(\hbar x_3^3)+\mathcal{O}(\hbar\xi_3^2)+\mathcal{O}(\xi_3^3)+\mathcal{O}(\hbar^2)$.

Corollary 2.9. *Let us introduce*

$$\underline{\mathcal{N}}_\hbar^{[1],\sharp}=\mathrm{Op}_\hbar^w(\underline{N}_\hbar^{[1],\sharp}),$$

where $\underline{N}_\hbar^{[1],\sharp}=\underline{\nu}^2(x_2,\xi_2)(\xi_3^2+\hbar x_3^2)+\hbar\underline{b}(x_2,\xi_2,s(x_2,\xi_2))+\underline{r}_\hbar^\sharp$, with $\underline{r}_\hbar^\sharp=\chi(x_2,\xi_2,x_3,\xi_3)\underline{r}_\hbar$, and where $\underline{\nu}$ denotes a smooth and constant (with a positive constant) extension of the function ν.

There exists a constant $\tilde{c}>0$ such that, for any cut-off function χ equal to 1 on $D(0,\varepsilon)$ with support in $D(0,2\varepsilon)$, we have:

(a) *The spectra of $\underline{\mathcal{N}}_\hbar^{[1],\sharp}$ and $\mathcal{N}_\hbar^{[1],\sharp}$ below $(b_0+\tilde{c}\varepsilon^2)\hbar$ coincide modulo $\mathcal{O}(\hbar^\infty)$.*

(b) *For all $c\in(0,\min(3b_0,b_0+\tilde{c}\varepsilon^2))$, the spectra of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ and $\underline{\mathcal{N}}_\hbar^{[1],\sharp}$ below $c\hbar$ coincide modulo $\mathcal{O}(\hbar^\infty)$.*

Notation 2.10 (change of semiclassical parameter). We let $h = \hbar^{\frac{1}{2}}$ and, if $A_\hbar$ is a semiclassical symbol on $T^*\mathbb{R}^2$, admitting a semiclassical expansion in $\hbar^{\frac{1}{2}}$, we write

$$\mathcal{A}_\hbar := \mathrm{Op}_\hbar^w A_\hbar = \mathrm{Op}_h^w \mathsf{A}_h =: \mathfrak{A}_h,$$

with

$$\mathsf{A}_h(x_2, \tilde{\xi}_2, x_3, \tilde{\xi}_3) = A_{h^2}(x_2, h\tilde{\xi}_2, x_3, h\tilde{\xi}_3).$$

Thus, $\mathcal{A}_\hbar$ and $\mathfrak{A}_h$ represent the same operator when $h = \hbar^{\frac{1}{2}}$, but the former is viewed as an $\hbar$-quantization of the symbol $A_\hbar$, while the latter is an h-pseudodifferential operator with symbol A_h. Notice that, if $A_\hbar$ belongs to some class $S(m)$, then $\mathsf{A}_h \in S(m)$ as well. This is of course not true the other way around.

Theorem 2.11. *Under Assumption 2.6, there exists a unitary operator W_h as well as a smooth function $g^\star(h, Z, x_2, \xi_2)$, with compact support as small as we want with respect to Z and with compact support in (x_2, ξ_2), whose Taylor series with respect to Z, h is*

$$\sum_{2m+2\ell \geqslant 3} c_{m,\ell}(x_2, \xi_2) Z^m h^\ell,$$

such that

$$W_h^* \underline{\mathfrak{N}}_h^{[1],\sharp} W_h =: \mathfrak{M}_h = \mathrm{Op}_h^w(\mathsf{M}_h),$$

with

$$\mathsf{M}_h = h^2 \underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2)) + h^2 \mathcal{J}_h \,\mathrm{Op}_h^w\, \underline{\nu}^2(x_2, h\tilde{\xi}_2) + h^2 g^\star(h, \mathcal{J}_h, x_2, h\tilde{\xi}_2) + h^2 \mathsf{R}_h + h^\infty S(1),$$

where

(a) *the operator $\underline{\mathfrak{N}}_h^{[1],\sharp}$ is $\underline{\mathcal{N}}_\hbar^{[1],\sharp}$ (but written in the h-quantization),*

(b) *we have let $\mathcal{J}_h = \mathrm{Op}_h^w(\tilde{\xi}_3^2 + x_3^2)$,*

(c) *the function R_h satisfies $\mathsf{R}_h(x_2, h\tilde{\xi}_2, x_3, \tilde{\xi}_3) = \mathcal{O}((x_3, \tilde{\xi}_3)^\infty)$.*

Remark 2.12. Note that the support of $g^\star$ with respect to Z may be chosen as small as we want. Note also that we have used $\underline{\mathfrak{N}}_h^{[1],\sharp}$ instead of $\underline{\mathfrak{N}}_h^{[1]}$: since W_h is exactly unitary, we get a direct comparison of the spectra.

Corollary 2.13. *We introduce*

$$\mathfrak{M}_h^\sharp = \mathrm{Op}_h^w(\mathsf{M}_h^\sharp),$$

with

$$\mathsf{M}_h^\sharp = h^2 \underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2)) + h^2 \mathcal{J}_h \underline{\nu}^2(x_2, h\tilde{\xi}_2) + h^2 g^\star(h, \mathcal{J}_h, x_2, h\tilde{\xi}_2).$$

We also define

$$\mathfrak{M}_h^{[1],\sharp} = \mathrm{Op}_h^w(\mathsf{M}_h^{[1],\sharp}),$$

with

$$\mathsf{M}_h^{[1],\sharp} = h^2 \underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2)) + h^3 \underline{\nu}^2(x_2, h\tilde{\xi}_2) + h^2 g^\star(h, h, x_2, h\tilde{\xi}_2).$$

If ε and the support of $g^\star$ are small enough, we have:

(a) *For all $\eta > 0$, the spectra of $\underline{\mathfrak{N}}_h^{[1],\sharp}$ and $\mathfrak{M}_h^\sharp$ below $b_0 h^2 + \mathscr{O}(h^{2+\eta})$ coincide modulo $\mathscr{O}(h^\infty)$.*

(b) *For $c \in (0,3)$, the spectra of $\mathfrak{M}_h^\sharp$ and $\mathfrak{M}_h^{[1],\sharp}$ below $b_0 h^2 + c\sigma^{\frac{1}{2}} h^3$ coincide modulo $\mathscr{O}(h^\infty)$.*

(c) *If $c \in (0,3)$, the spectra of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ and $\mathcal{M}_\hbar^{[1],\sharp} = \mathfrak{M}_h^{[1],\sharp}$ below $b_0\hbar + c\sigma^{\frac{1}{2}}\hbar^{\frac{3}{2}}$ coincide modulo $\mathscr{O}(\hbar^\infty)$.*

Finally, we can perform a last Birkhoff normal form for the operator $\mathcal{M}_\hbar^{[1],\sharp}$ as soon as $(x_2,\xi_2) \mapsto \underline{b}(x_2,\xi_2,s(x_2,\xi_2))$ admits a unique and nondegenerate minimum at $(0,0)$. Under this additional assumption, b admits a unique and nondegenerate minimum at $(0,0,0)$.

Therefore we will use the following stronger assumption.

Assumption 2.14. The function b admits a unique and positive minimum at 0 and it is nondegenerate.

Theorem 2.15. *Under Assumption 2.14, there exists a unitary $\hbar$-Fourier integral operator $Q_{\hbar^{1/2}}$ whose phase admits an expansion in powers of $\hbar^{\frac{1}{2}}$ such that*

$$Q_{\hbar^{1/2}}^* \mathcal{M}_\hbar^{[1],\sharp} Q_{\hbar^{1/2}} = \mathcal{F}_\hbar + \mathcal{G}_\hbar,$$

where

(a) *$\mathcal{F}_\hbar$ is defined in Theorem 1.3,*

(b) *the remainder is in the form $\mathcal{G}_\hbar = \mathrm{Op}_\hbar^w(G_\hbar)$, with $G_\hbar = \hbar\mathscr{O}(|z_2|^\infty)$.*

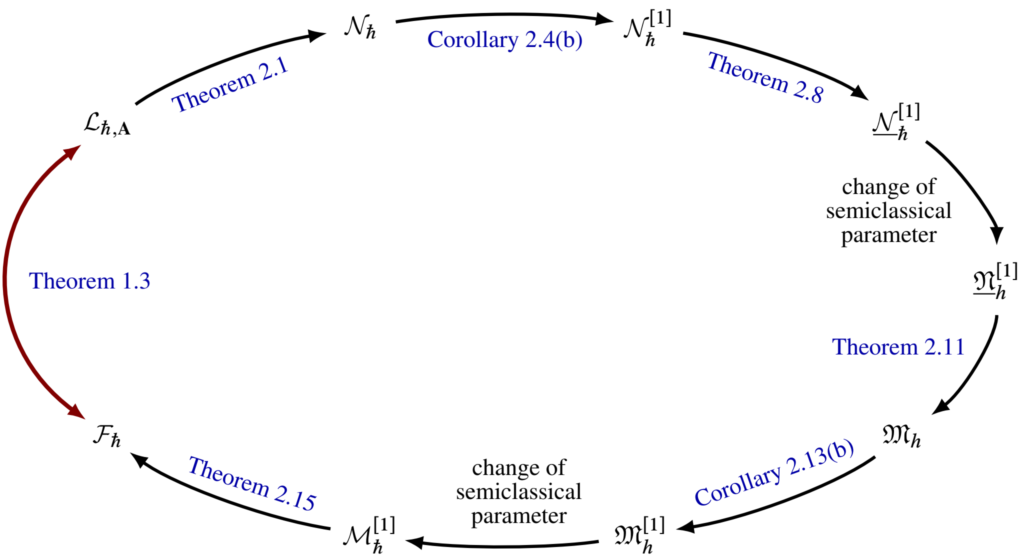
Corollary 2.16. *If ε and the support of $k^\star$ are small enough, we have*:

(a) *For all $\eta \in \left(0,\frac{1}{2}\right)$, the spectra of $\mathcal{M}_\hbar^{[1],\sharp}$ and $\mathcal{F}_\hbar$ below $b_0\hbar + \mathscr{O}(\hbar^{1+\eta})$ coincide modulo $\mathscr{O}(\hbar^\infty)$.*

(b) *For all $c \in (0,3)$, the spectra of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ and $\mathcal{F}_\hbar$ below $b_0\hbar + c\sigma^{\frac{1}{2}}\hbar^{\frac{3}{2}}$ coincide modulo $\mathscr{O}(\hbar^\infty)$.*

Remark 2.17. Since the spectral analysis of $\mathcal{F}_\hbar$ is straightforward, Corollary 2.16(b) implies Theorem 1.3.

The next sections are devoted to the proofs of our main results.



3. First Birkhoff normal form

We assume that $\boldsymbol{B}(0) \neq 0$ so that in some neighborhood Ω of 0 the magnetic field does not vanish. Up to a rotation in $\mathbb{R}^3$ (extended to a symplectic transformation in $\mathbb{R}^6$) we may assume that $\boldsymbol{B}(0) = \|\boldsymbol{B}(0)\|\boldsymbol{e}_3$. In this neighborhood, we may define the unit vector

$$\boldsymbol{b} = \frac{\boldsymbol{B}}{\|\boldsymbol{B}\|} \tag{3-1}$$

and find vectors $\boldsymbol{c}$ and $\boldsymbol{d}$ depending smoothly on q such that $(\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d})$ is a direct orthonormal basis.

3A. *Symplectic coordinates.*

3A1. *Straightening the magnetic vector field.* Let $\widehat{\Omega}$ be a small neighborhood of $0 \in \mathbb{R}^2$. We consider the form $\mathrm{d}\alpha$ and we would like to find a diffeomorphism χ, defined on $\widehat{\Omega}$, such that $\chi^*(\mathrm{d}\alpha) = \mathrm{d}\hat{q}_1 \wedge \mathrm{d}\hat{q}_2$, where we use the notation $\chi(\hat{q}) = q$. First, it is easy to find a local diffeomorphism φ such that

$$\partial_3\varphi(\tilde{q}) = \boldsymbol{b}(\varphi(\tilde{q}))$$

and $\varphi(\tilde{q}_1, \tilde{q}_2, 0) = (\tilde{q}_1, \tilde{q}_2, 0)$. This is just the standard straightening lemma for the nonvanishing vector field $\boldsymbol{b}$.

The vector $\boldsymbol{e}_3$ is in the kernel of $\varphi^*(\mathrm{d}\alpha)$, which implies that we have $\varphi^*(\mathrm{d}\alpha) = f(\tilde{q})\,\mathrm{d}\tilde{q}_1 \wedge \mathrm{d}\tilde{q}_2$ for some smooth function f.

But since the form $\varphi^*(\mathrm{d}\alpha)$ is closed, f does not depend on $\tilde{q}_3$. It is then easy to find another diffeomorphism ψ, corresponding to the change of variables

$$\hat{q} = \psi(\tilde{q}) = \big(\psi_1(\tilde{q}_1, \tilde{q}_2),\ \psi_2(\tilde{q}_1, \tilde{q}_2),\ \tilde{q}_3\big),$$

such that

$$\psi^*(\varphi^*(\mathrm{d}\alpha)) = \mathrm{d}\hat{q}_1 \wedge \mathrm{d}\hat{q}_2.$$

We let $\chi = \varphi \circ \psi$ and we notice that

$$\chi^*(\mathrm{d}\alpha) = \mathrm{d}\hat{q}_1 \wedge \mathrm{d}\hat{q}_2, \quad \partial_3\chi(\hat{q}) = \boldsymbol{b}(\chi(\hat{q})). \tag{3-2}$$

Remark 3.1. It follows from (3-2) and (1-4) that $\det T\chi = \|\boldsymbol{B}\|^{-1}$.

3A2. *Symplectic coordinates.* Let us consider the new parametrization of Σ given by

$$\begin{aligned} \iota : \widehat{\Omega} &\to \Sigma, \\ \hat{q} &\mapsto \big(\chi(\hat{q}),\ A_1(\chi(\hat{q})),\ A_2(\chi(\hat{q})),\ A_3(\chi(\hat{q}))\big), \end{aligned}$$

which gives a basis $(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3)$ of $T\Sigma$,

$$\boldsymbol{f}_j = \big(T\chi(\boldsymbol{e}_j), T\boldsymbol{A} \circ T\chi(\boldsymbol{e}_j)\big), \quad j = 1, 2, 3.$$

Using (1-5), and the fact that $\boldsymbol{f}_3$ is in the kernel of $\mathrm{d}\alpha$, we find $\omega_0(\boldsymbol{f}_j, \boldsymbol{f}_3) = 0$, $j = 1, 2$. Finally, $\omega_0(f_1, f_2) = \mathrm{d}\alpha(T\chi\boldsymbol{e}_1, T\chi\boldsymbol{e}_2) = \chi^*(\mathrm{d}\alpha)(\boldsymbol{e}_1, \boldsymbol{e}_2) = 1$.

The following vectors of $\mathbb{R}^3 \times \mathbb{R}^3$ form a basis of the symplectic orthogonal of $T_{\iota(\hat{q})}\Sigma$:

$$\begin{aligned} \boldsymbol{f}_4 &= \|\boldsymbol{B}\|^{-\frac{1}{2}}(\boldsymbol{c}, ({}^tT_{\chi(\hat{q})}\boldsymbol{A})\boldsymbol{c}), \\ \boldsymbol{f}_5 &= \|\boldsymbol{B}\|^{-\frac{1}{2}}(\boldsymbol{d}, ({}^tT_{\chi(\hat{q})}\boldsymbol{A})\boldsymbol{d}), \end{aligned} \tag{3-3}$$

so that

$$\omega_0(\boldsymbol{f}_4, \boldsymbol{f}_5) = -1.$$

We let $\boldsymbol{f}_6 = (0, \boldsymbol{b}) + \rho_1 \boldsymbol{f}_1 + \rho_2 \boldsymbol{f}_2$, where ρ_1 and ρ_2 are determined so that $\omega_0(\boldsymbol{f}_j, \boldsymbol{f}_6) = 0$ for $j = 1, 2$. We notice that $\omega_0(\boldsymbol{f}_j, \boldsymbol{f}_6) = 0$ for $j = 4, 5$ and $\omega_0(\boldsymbol{f}_3, \boldsymbol{f}_6) = -1$.

3A3. *Diagonalizing the Hessian.* We recall that

$$H(q, p) = \|p - \boldsymbol{A}(q)\|^2$$

so that, at a critical point $p = \boldsymbol{A}(q)$, the Hessian is

$$T^2H\big((U_1, V_1), (U_2, V_2)\big) = 2\big\langle V_1 - T_q\boldsymbol{A}(U_1), V_2 - T_q\boldsymbol{A}(U_2)\big\rangle.$$

Let us notice that

$$\begin{aligned} T^2H(\boldsymbol{f}_4, \boldsymbol{f}_5) &= 2\|\boldsymbol{B}\|^{-1}\langle \boldsymbol{B} \times \boldsymbol{c}, \boldsymbol{B} \times \boldsymbol{d}\rangle = 0, \\ T^2H(\boldsymbol{f}_4, \boldsymbol{f}_6) &= 2\langle \boldsymbol{B} \times \boldsymbol{c}, \boldsymbol{b}\rangle = 0, \\ T^2H(\boldsymbol{f}_5, \boldsymbol{f}_6) &= 2\langle \boldsymbol{B} \times \boldsymbol{d}, \boldsymbol{b}\rangle = 0. \end{aligned}$$

The Hessian, restricted to the symplectic orthogonal of $T_{\iota(\hat{q})}\Sigma$, is diagonal in the basis $(\boldsymbol{f}_4, \boldsymbol{f}_5, \boldsymbol{f}_6)$. Moreover we have

$$T^2H(\boldsymbol{f}_4, \boldsymbol{f}_4) = \mathrm{d}^2H(\boldsymbol{f}_5, \boldsymbol{f}_5) = 2\|\boldsymbol{B}\|^{-1}\|\boldsymbol{B} \times \boldsymbol{c}\|^2 = 2\|\boldsymbol{B}\|^{-1}\|\boldsymbol{B} \times \boldsymbol{d}\|^2 = 2\|\boldsymbol{B}\|.$$

Finally we have

$$T^2H(\boldsymbol{f}_6, \boldsymbol{f}_6) = 2.$$

Now we consider the local diffeomorphism

$$(x, \xi) \mapsto \iota(x_2, \xi_2, x_3) + x_1 \boldsymbol{f}_4(x_2, \xi_2, x_3) + \xi_1 \boldsymbol{f}_5(x_2, \xi_2, x_3) + \xi_3 \boldsymbol{f}_6(x_2, \xi_2, x_3).$$

The Jacobian of this map is a symplectic matrix on Σ. We may apply the Moser–Weinstein argument (see [Weinstein 1971]) to make this map locally symplectic near Σ modulo a change of variable which is tangent to the identity.

Near Σ, in these new coordinates, the Hamiltonian H admits the expansion

$$\widehat{H} = H^0 + \mathcal{O}(|x_1|^3 + |\xi_1|^3 + |\xi_3|^3), \tag{3-4}$$

where $\widehat{H}$ denotes H in the coordinates $(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3)$, and with

$$H^0 = \xi_3^2 + b(x_2, \xi_2, x_3)(x_1^2 + \xi_1^2), \quad b = \|\boldsymbol{B}(x_2, \xi_2, x_3)\|. \tag{3-5}$$

3B. *Semiclassical Birkhoff normal form.*

3B1. *Birkhoff procedure in formal series.* Let us consider the space $\mathcal{E}$ of formal power series in $(x_1,\xi_1,\xi_3,\hbar)$ with coefficients smoothly depending on $\tilde{x}=(x_2,\xi_2,x_3)$:

$$\mathcal{E}=C^\infty_{x_2,\xi_2,x_3}[\![x_1,\xi_1,\xi_3,\hbar]\!].$$

We endow $\mathcal{E}$ with the semiclassical Moyal product (with respect to all variables $(x_1,x_2,x_3,\xi_1,\xi_2,\xi_3)$) denoted by $\star$ and the commutator of two series κ_1 and κ_2 is defined as

$$[\kappa_1,\kappa_2]=\kappa_1\star\kappa_2-\kappa_2\star\kappa_1.$$

The degree of $x_1^{\alpha_1}\xi_1^{\alpha_2}\xi_3^{\beta}\hbar^\ell=z_1^\alpha\xi_3^\beta\hbar^\ell$ is $\alpha_1+\alpha_2+\beta+2\ell=|\alpha|+\beta+2\ell$. The space of monomials of degree N is denoted $\mathcal{D}_N$, and $\mathcal{O}_N$ is the space of formal series with valuation at least N. For any $\tau,\gamma\in\mathcal{E}$, we define $\operatorname{ad}_\tau\gamma=[\tau,\gamma]$.

Proposition 3.2. *Given $\gamma\in\mathcal{O}_3$, there exist formal power series $\tau,\kappa\in\mathcal{O}_3$ such that*

$$e^{i\hbar^{-1}\operatorname{ad}_\tau}(H^0+\gamma)=H^0+\kappa,$$

with $[\kappa,|z_1|^2]=0$.

Proof. Let $N\geqslant 1$. Assume that we have, for $\tau_N\in\mathcal{O}_3$,

$$e^{i\hbar^{-1}\operatorname{ad}_{\tau_N}}(H^0+\gamma)=H^0+K_3+\cdots+K_{N+1}+R_{N+2}+\mathcal{O}_{N+3},$$

with $K_i\in\mathcal{D}_i$, $[K_i,|z_1|^2]=0$ and $R_{N+2}\in\mathcal{D}_{N+2}$.

Let $\tau'\in\mathcal{D}_{N+2}$. Then we have

$$e^{i\hbar^{-1}\operatorname{ad}_{\tau_N+\tau'}}(H^0+\gamma)=H^0+K_3+\cdots+K_{N+1}+K_{N+2}+\mathcal{O}_{N+3},$$

with $K_{N+2}\in\mathcal{D}_{N+2}$ such that

$$K_{N+2}=R_{N+2}+i\hbar^{-1}\operatorname{ad}_{\tau'}H^0+\mathcal{O}_{N+3}.$$

Let us temporarily admit that (see Lemma 3.3 below)

$$i\hbar^{-1}\operatorname{ad}_{\tau'}H^0=i\hbar^{-1}b\operatorname{ad}_{\tau'}|z_1|^2+\mathcal{O}_{N+3}.$$

We obtain

$$K_{N+2}=R_{N+2}+b\operatorname{ad}_{\tau'}|z_1|^2,$$

which we rewrite as

$$R_{N+2}=K_{N+2}+i\hbar^{-1}b\operatorname{ad}_{|z_1|^2}\tau'=K_{N+2}+b\{|z_1|^2,\tau'\}.$$

Since $b(\tilde{x})\neq 0$, we deduce the existence of τ' and K_{N+2} such that K_{N+2} commutes with $|z_1|^2$. $\square$

Lemma 3.3. *For $\tau'\in\mathcal{D}_{N+2}$, we have*

$$i\hbar^{-1}\operatorname{ad}_{\tau'}H^0=i\hbar^{-1}b\operatorname{ad}_{\tau'}|z_1|^2+\mathcal{O}_{N+3}.$$

Proof. We observe that

$$i\hbar^{-1}\,\mathrm{ad}_{\tau'} H^0 = i\hbar^{-1}\,\mathrm{ad}_{\tau'}\xi_3^2 + i\hbar^{-1}\,\mathrm{ad}_{\tau'}(b(\tilde{x})|z_1|^2).$$

Let us write

$$\tau' = \sum_{|\alpha|+\beta+2l=N+2} a_{\alpha,\beta,l}(\tilde{x}) z_1^{\alpha}\xi_3^{\beta}\hbar^{l}.$$

Then, for the first term, we have

$$i\hbar^{-1}\,\mathrm{ad}_{\tau'}\xi_3^2 = \{\tau',\xi_3^2\} = -2\xi_3\frac{\partial\tau'}{\partial x_3} = -2\sum_{|\alpha|+\beta+2\ell=N+2}\frac{\partial a_{\alpha,\beta,\ell}}{\partial x_3}(\tilde{x}) z_1^{\alpha}\xi_3^{\beta+1}\hbar^{\ell} \in \mathcal{O}_{N+3}.$$

We also have

$$\begin{aligned} i\hbar^{-1}(\mathrm{ad}_{\tau'} b(\tilde{x})) &= \{\tau', b\} + \hbar^2\mathcal{O}_N = \frac{\partial\tau'}{\partial\xi_3}\frac{\partial b}{\partial x_3} + \frac{\partial\tau'}{\partial\xi_2}\frac{\partial b}{\partial x_2} - \frac{\partial\tau'}{\partial x_2}\frac{\partial b}{\partial\xi_2} + \mathcal{O}_{N+1} \\ &= \sum_{|\alpha|+\beta+2\ell=N+2}\beta a(\tilde{x})\frac{\partial b}{\partial x_3} z_1^{\alpha}|z_1|^2\xi_3^{\beta-1}\hbar^{\ell} + \mathcal{O}_{N+1} \in \mathcal{O}_{N+1}. \end{aligned}$$

Therefore, for the second term, we get

$$\begin{aligned} i\hbar^{-1}\,\mathrm{ad}_{\tau'}(b(\tilde{x})|z_1|^2) &= i\hbar^{-1}(\mathrm{ad}_{\tau'} b(\tilde{x}))|z_1|^2 + i\hbar^{-1} b(\tilde{x})\,\mathrm{ad}_{\tau'}|z_1|^2 \\ &= i\hbar^{-1} b(\tilde{x})\,\mathrm{ad}_{\tau'}|z_1|^2 + \mathcal{O}_{N+3}, \end{aligned}$$

which completes the proof of the lemma. □

3B2. *Quantizing the formal procedure.* Let us now prove Theorem 2.1. Using (3-4) and applying the Egorov theorem (see [Robert 1987; Zworski 2012] or Theorem A.2), we can find a unitary Fourier integral operator $U_\hbar$ such that

$$U_\hbar^* \mathcal{L}_{\hbar,\mathbf{A}} U_\hbar = C_0\hbar + \mathrm{Op}_\hbar^w(H^0) + \mathrm{Op}_\hbar^w(r_\hbar),$$

where the Taylor series (with respect to x_1, ξ_1, ξ_3, $\hbar$) of $r_\hbar$ satisfies $r_\hbar^T = \gamma \in \mathcal{O}_3$ and C_0 is the value at the origin of the subprincipal symbol of $U_\hbar^*\mathcal{L}_{\hbar,\mathbf{A}}U_\hbar$. One can choose $U_\hbar$ such that the subprincipal symbol is preserved by conjugation,[1] which implies $C_0 = 0$. Applying Proposition 3.2, we obtain τ and κ in $\mathcal{O}_3$ such that

$$e^{i\hbar^{-1}\,\mathrm{ad}_\tau}(H^0 + \gamma) = H^0 + \kappa,$$

with $[\kappa, |z_1|^2] = 0$.

We can introduce a smooth symbol $a_\hbar$ with compact support such that we have $a_\hbar^T = \tau$ in a neighborhood of the origin. By Proposition 3.2 and Theorem A.4, we obtain that the operator

$$e^{i\hbar^{-1}\,\mathrm{Op}_\hbar^w(a_\hbar)}\big(\mathrm{Op}_\hbar^w(H^0) + \mathrm{Op}_\hbar^w(r_\hbar)\big)e^{-i\hbar^{-1}\,\mathrm{Op}_\hbar^w(a_\hbar)}$$

[1] This is sometimes called the *improved Egorov theorem*. It was first discovered by Weinstein [1975] in the homogeneous setting. For the semiclassical case, see, for instance, [Helffer and Sjöstrand 1989, Appendix A].

is a pseudodifferential operator such that the formal Taylor series of its symbol is $H^0+\kappa$. In this application of Theorem A.4, we have used the filtration $\mathcal{O}_j$ defined in Section 3B1. Since κ commutes with $|z_1|^2$, we can write it as a formal series in $|z_1|^2$:

$$\kappa=\sum_{k\geqslant 3}\sum_{2\ell+2m+\beta=k}\hbar^\ell c_{\ell,m}(x_2,\xi_2,x_3)|z_1|^{2m}\xi_3^\beta.$$

This formal series can be reordered by using monomials $(|z_1|^2)^{\star m}$:

$$\kappa=\sum_{k\geqslant 3}\sum_{2\ell+2m+\beta=k}\hbar^\ell c^\star_{\ell,m}(x_2,\xi_2,x_3)(|z_1|^2)^{\star m}\xi_3^\beta.$$

Thanks to the Borel lemma, we may find a smooth function $f^\star(\hbar,I,x_2,\xi_2,x_3,\xi_3)$, with compact support as small as we want with respect to $\hbar$, I and ξ_3, such that its Taylor series with respect to $\hbar,I,\xi_3$ is

$$\sum_{k\geqslant 3}\sum_{2\ell+2m+\beta=k}\hbar^\ell c^\star_{\ell,m}(x_2,\xi_2,x_3)I^m\xi_3^\beta.$$

This achieves the proof of Theorem 2.1.

3C. *Spectral reduction to the first normal form.* This section is devoted to the proof of Corollary 2.4.

3C1. *Numbers of eigenvalues.*

Lemma 3.4. *Under Assumption 1.2, there exists $h_0>0$ and $\varepsilon_0>0$ such that, for all $\hbar\in(0,h_0)$, the essential spectrum of $\mathcal{N}_\hbar^\sharp$ admits the lower bound*

$$\inf\mathfrak{s}_{\mathrm{ess}}(\mathcal{N}_\hbar^\sharp)\geqslant(\beta_0+\varepsilon_0)\hbar.$$

Proof. By using the assumption, we may consider a smooth function χ with compact support and $\varepsilon_0>0$ such that

$$\xi_3^2+b(x_2,\xi_2,x_3)+\chi(x_2,x_3,\xi_2,\xi_3)\geqslant\beta_0+2\varepsilon_0.$$

Then, given $\eta\in(0,1)$ and estimating the second term in (2-3) by using that the support of $f^\star$ is chosen small enough and the semiclassical Calderon–Vaillancourt theorem, we notice that, for $\hbar$ small enough,

$$\mathcal{N}_\hbar^\sharp\geqslant(1-\eta)\operatorname{Op}_\hbar^w\big(\xi_3^2+|z_1|^2b(x_2,\xi_2,x_3)\big).\tag{3-6}$$

Since the essential spectrum is invariant by (relatively) compact perturbations, we have

$$\mathfrak{s}_{\mathrm{ess}}\big(\mathcal{N}_\hbar^\sharp+(1-\eta)\hbar\operatorname{Op}_\hbar^w\chi(x_2,x_3,\xi_2,\xi_3)\big)=\mathfrak{s}_{\mathrm{ess}}(\mathcal{N}_\hbar^\sharp).$$

Hence

$$\inf\mathfrak{s}_{\mathrm{ess}}(\mathcal{N}_\hbar^\sharp)\geqslant\inf\mathfrak{s}\big(\mathcal{N}_\hbar^\sharp+(1-\eta)\hbar\operatorname{Op}_\hbar^w\chi(x_2,x_3,\xi_2,\xi_3)\big).$$

In order to bound the right-hand side from below, we write

$$\begin{aligned}\mathcal{N}_\hbar^\sharp+(1-\eta)\hbar\operatorname{Op}_\hbar^w\chi(x_2,x_3,\xi_2,\xi_3)&\geqslant(1-\eta)\operatorname{Op}_\hbar^w\big(\xi_3^2+|z_1|^2b(x_2,\xi_2,x_3)\big)+(1-\eta)\hbar\operatorname{Op}_\hbar^w\chi(x_2,x_3,\xi_2,\xi_3)\\&\geqslant\hbar(1-\eta)\operatorname{Op}_\hbar^w\big(\xi_3^2+b(x_2,\xi_2,x_3)+\chi(x_2,x_3,\xi_2,\xi_3)\big)\\&\geqslant\hbar(1-\eta)(\beta_0+2\varepsilon_0-C\hbar),\end{aligned}$$

where we have used the semiclassical Gårding inequality. Taking η and then $\hbar$ small enough, this concludes the proof. □

By using the Hilbertian decomposition given by the Hermite functions $(e_{k,\hbar})_{k\geqslant 1}$ associated with $\mathcal{I}_\hbar$, we notice that

$$\mathcal{N}_\hbar^\sharp = \bigoplus_{k\geqslant 1} \mathcal{N}_\hbar^{[k],\sharp},$$

where

$$\mathcal{N}_\hbar^{[k],\sharp} = \hbar^2 D_{x_3}^2 + (2k-1)\hbar \operatorname{Op}_\hbar^w b + \operatorname{Op}_\hbar^w f^{\star,\sharp}\big(\hbar, (2k-1)\hbar, x_2, \xi_2, x_3, \xi_3\big) \tag{3-7}$$

acting on $L^2(\mathbb{R}^2)$.

Lemma 3.5. *For all $\eta \in (0,1)$, there exist $C > 0$ and $h_0 > 0$ such that, for all $k \geqslant 1$ and $\hbar \in (0, h_0)$, we have* $\inf \mathfrak{s}(\mathcal{N}_\hbar^{[k],\sharp}) \geqslant (1-2\eta) b_0 (2k-1)\hbar$.

Proof. Applying (3-6) to $\psi(x_1, x_2, x_3) = \varphi(x_2, x_3) e_{k,\hbar}(x_1)$, we infer that

$$\langle \mathcal{N}_\hbar^{[k],\sharp} \varphi, \varphi\rangle \geqslant (2k-1)\hbar(1-\eta)\langle \operatorname{Op}_\hbar^w(b)\varphi, \varphi\rangle.$$

With the Gårding inequality, we get

$$\langle \operatorname{Op}_\hbar^w(b)\varphi, \varphi\rangle \geqslant (b_0 - C\hbar)\|\varphi\|^2,$$

and the conclusion follows by the min-max principle. □

We immediately deduce the following proposition.

Proposition 3.6. *We have the following descriptions of the low-lying spectrum of $\mathcal{N}_\hbar^\sharp$.*

(a) *There exist $\hbar_0 > 0$ and $K \in \mathbb{N}$ such that, for $\hbar \in (0, \hbar_0)$, the spectrum of $\mathcal{N}_\hbar^\sharp$ lying below $\beta_0\hbar$ is contained in the union $\bigcup_{k=1}^K \mathsf{sp}(\mathcal{N}_\hbar^{[k],\sharp})$.*

(b) *If $c \in (0, \min(3b_0, \beta_0))$, then there exists $\hbar_0 > 0$ such that for all $\hbar \in (0, \hbar_0)$ the eigenvalues of $\mathcal{N}_\hbar^\sharp$ lying below $c\hbar$ coincide with the eigenvalues of $\mathcal{N}_\hbar^{[1],\sharp}$ below $c\hbar$.*

Notation 3.7. If $\mathcal{L}$ is a self-adjoint operator and $E < \inf \mathfrak{s}_{\mathrm{ess}}(\mathcal{L})$, we denote by $\mathsf{N}(\mathcal{L}, E)$ the number of eigenvalues of $\mathcal{L}$ lying in $(-\infty, E)$.

We deduce the following proposition.

Corollary 3.8. *Under assumption (1-11), we have*

$$\mathsf{N}(\mathcal{L}_{\hbar,\boldsymbol{A}}, \beta_0\hbar) = \mathcal{O}(\hbar^{-\frac{3}{2}}), \quad \mathsf{N}(\mathcal{N}_\hbar^\sharp, \beta_0\hbar) = \mathcal{O}(\hbar^{-2}).$$

Proof. To get the first estimate, we use the Lieb–Thirring inequalities (which provide an upper bound on the number of eigenvalues in dimension three) and the diamagnetic inequality (see [Raymond and Vũ Ngọc 2015] and (1-9)). To get the second estimate, we use the first point in Proposition 3.6. Moreover, given $\eta \in (0,1)$, by using $\hbar \in (0,1)$ we infer

$$\langle \mathcal{N}_\hbar^{[k],\sharp}\psi, \psi\rangle \geqslant (1-\eta)\hbar\big\langle \operatorname{Op}_\hbar^w\big(\xi_3^2 + b(x_2, \xi_2, x_3)\big)\psi, \psi\big\rangle.$$

Note that the last inequality is very rough. By the min-max principle, we deduce that

$$\mathsf{N}(\mathcal{N}_\hbar^{[k],\sharp}, \beta_0\hbar) \leqslant \mathsf{N}\big(\mathrm{Op}_\hbar^w(\xi_3^2 + b(x_2,\xi_2,x_3)), (1-\eta)^{-1}\beta_0\big).$$

Then, we conclude by using the Weyl asymptotics and our confinement assumption:

$$\mathsf{N}\big(\mathrm{Op}_\hbar^w(\xi_3^2 + b(x_2,\xi_2,x_3)), (1-\eta)^{-1}\beta_0\big) = \mathcal{O}(\hbar^{-2}). \qquad \square$$

Since $\mathcal{N}_\hbar^\sharp$ commutes with $\mathcal{I}_\hbar$, we also deduce the following corollary.

Corollary 3.9. *For any eigenvalue λ of $\mathcal{N}_\hbar^\sharp$ such that $\lambda \leqslant \beta_0\hbar$, we may consider an orthonormal eigenbasis of the space $\ker(\mathcal{N}_\hbar^\sharp - \lambda)$ formed with functions in the form $e_{k,\hbar}(x_1)\varphi_\hbar(x_2,x_3)$ with $k \in \{1,\dots,K\}$. Moreover, we have $\mathbb{1}_{(-\infty,\beta_0\hbar)}(\mathcal{N}_\hbar^\sharp) = \mathcal{O}(\hbar^{-2})$ and each eigenfunction associated with $\lambda \leqslant \beta_0\hbar$ is a linear combination of at most $\mathcal{O}(\hbar^{-2})$ such tensor products.*

3C2. *Microlocalization estimates.* The following proposition follows from the same lines as in dimension two (see [Helffer and Mohamed 1996, Theorem 2.1]).

Proposition 3.10. *Under Assumptions 1.1 and 1.2, for any $\varepsilon > 0$, there exist $C(\varepsilon) > 0$ and $h_0(\varepsilon) > 0$ such that, for any eigenpair (λ,ψ) of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ with $\lambda \leqslant \beta_0\hbar$, we have for $\hbar \in (0, h_0(\varepsilon))$,*

$$\int_{\mathbb{R}^3} e^{2(1-\varepsilon)\phi(q)/\hbar^{\frac12}} |\psi|^2 \,\mathrm{d}q \leqslant C(\varepsilon)\exp(\varepsilon\hbar^{-\frac12})\|\psi\|^2,$$

$$\mathcal{Q}_{\hbar,\boldsymbol{A}}(e^{(1-\varepsilon)\phi(q)/\hbar^{\frac12}}\psi) \leqslant C(\varepsilon)\exp(\varepsilon\hbar^{-\frac12})\|\psi\|^2,$$

where ϕ is the distance to the bounded set $\{\|\boldsymbol{B}(q)\| \leqslant \beta_0\}$ for the Agmon metric $(\|(\boldsymbol{B}(q)\| - \beta_0)_+ g$, with g the standard metric.

Proposition 3.11. *Under Assumptions 1.1 and 1.2, we consider $0 < b_0 < \beta_0 < b_1$ and there exist $C > 0$ and $\hbar_0 > 0$ such that, for any eigenpair (λ,ψ) of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ with $\lambda \leqslant \beta_0\hbar$, we have for $\hbar \in (0,\hbar_0)$ and $\delta \in \big(0,\frac12\big)$,*

$$\psi = \chi_0(\hbar^{-2\delta}\mathcal{L}_{\hbar,\boldsymbol{A}})\chi_1(q)\psi + \mathcal{O}(\hbar^\infty)\|\psi\|,$$

where χ_0 is a cutoff function compactly supported in the ball of center 0 and radius 1 and where χ_1 is a compactly supported smooth cutoff function that is 1 in an open neighborhood of $\{\|\boldsymbol{B}(q)\| \leqslant \beta_0\}$.

Let us now investigate the microlocalization of the eigenfunctions of $\mathcal{N}_\hbar^\sharp$.

Proposition 3.12. *Let χ be a smooth cutoff function that is 0 on $\{b \leqslant \beta_0\}$ and 1 on the set $\{b \geqslant \beta_0 + \varepsilon\}$. If λ is an eigenvalue of $\mathcal{N}_\hbar^\sharp$ such that $\lambda \leqslant \beta_0\hbar$ and if ψ is an associated eigenfunction, then we have*

$$\mathrm{Op}_\hbar^w(\chi(x_2,\xi_2,x_3))\psi = \mathcal{O}(\hbar^\infty)\|\psi\|.$$

Proof. Due to Corollary 3.9, it is sufficient to prove the estimate for a function in the form $\psi(x_1,x_2,x_3) = e_{k,\hbar}(x_1)\varphi(x_2,x_3)$, where k lies in $\{1,\dots,K\}$ and we have

$$\mathcal{N}_\hbar^\sharp\psi = \lambda\psi, \quad \text{or equivalently} \quad \mathcal{N}_\hbar^{[k],\sharp}\varphi = \lambda\varphi,$$

where we recall (3-7). Then, we write

$$\mathcal{N}_\hbar^{[k],\sharp}\,\mathrm{Op}_\hbar^w(\chi)\varphi = \lambda\,\mathrm{Op}_\hbar^w(\chi)\varphi + [\mathcal{N}_\hbar^{[k],\sharp}, \mathrm{Op}_\hbar^w(\chi)]\varphi$$

and it follows that

$$\langle \mathcal{N}_\hbar^{[k],\sharp}\,\mathrm{Op}_\hbar^w(\chi)\varphi, \mathrm{Op}_\hbar^w(\chi)\varphi\rangle = \lambda\|\,\mathrm{Op}_\hbar^w(\chi)\varphi\|^2 + \langle [\mathcal{N}_\hbar^{[k],\sharp}, \mathrm{Op}_\hbar^w(\chi)]\varphi, \mathrm{Op}_\hbar^w(\chi)\varphi\rangle. \tag{3-8}$$

Rough pseudodifferential estimates imply that there exist $C>0$, $\hbar_0>0$ such that, for all $\hbar\in(0,\hbar_0)$,

$$\begin{aligned}&\bigl|\langle [\mathcal{N}_\hbar^{[k],\sharp}, \mathrm{Op}_\hbar^w(\chi)]\varphi, \mathrm{Op}_\hbar^w(\chi)\varphi\rangle\bigr| \\ &\qquad\leqslant C\hbar^2\bigl\|\mathrm{Op}_\hbar^w(\underline{\chi})\varphi\bigr\|^2 + C\hbar\bigl\|\mathrm{Op}_\hbar^w(\underline{\chi})\varphi\bigr\|^2 + C\hbar\langle \mathrm{Op}_\hbar^w(\partial_3\chi)\varphi, \mathrm{Op}_\hbar^w(\xi_3)\,\mathrm{Op}_\hbar^w(\chi)\varphi\rangle.\end{aligned} \tag{3-9}$$

Combining (3-9) and (3-8), we get

$$\bigl\|\mathrm{Op}_\hbar^w(\xi_3)\,\mathrm{Op}_\hbar^w(\chi)\varphi\bigr\| \leqslant C\hbar^{\frac12}\bigl\|\mathrm{Op}_\hbar^w(\underline{\chi})\varphi\bigr\|, \tag{3-10}$$

where $\underline{\chi}$ is a smooth cutoff function living on a slightly larger support than χ. By using (3-10), we can improve the commutator estimate

$$\bigl|\langle [\mathcal{N}_\hbar^{[k],\sharp}, \mathrm{Op}_\hbar^w(\chi)]\varphi, \mathrm{Op}_\hbar^w(\chi)\varphi\rangle\bigr| \leqslant C\hbar^{\frac32}\bigl\|\mathrm{Op}_\hbar^w(\underline{\chi})\varphi\bigr\|^2.$$

We infer that, there exist $C>0$, $\hbar_0>0$ such that for $\hbar\in(0,\hbar_0)$,

$$\langle \mathcal{N}_\hbar^{[k],\sharp}\,\mathrm{Op}_\hbar^w(\chi)\varphi, \mathrm{Op}_\hbar^w(\chi)\varphi\rangle \leqslant \beta_0\hbar\bigl\|\mathrm{Op}_\hbar^w(\chi)\varphi\bigr\|^2 + C\hbar^{\frac32}\bigl\|\mathrm{Op}_\hbar^w(\underline{\chi})\varphi\bigr\|^2.$$

By using the semiclassical Gårding inequality and the support of χ, we get

$$\langle \mathcal{N}_\hbar^{[k],\sharp}\,\mathrm{Op}_\hbar^w(\chi)\varphi, \mathrm{Op}_\hbar^w(\chi)\varphi\rangle \geqslant (\beta_0+\varepsilon_0)\hbar\bigl\|\mathrm{Op}_\hbar^w(\chi)\varphi\bigr\|^2$$

and we deduce

$$\bigl\|\mathrm{Op}_\hbar^w(\chi)\varphi\bigr\|^2 \leqslant C\hbar^{\frac12}\bigl\|\mathrm{Op}_\hbar^w(\underline{\chi})\varphi\bigr\|^2.$$

The conclusion follows by a standard iteration argument. □

The following proposition is concerned with the microlocalization with respect to ξ_3.

Proposition 3.13. *Let χ_0 be a smooth cutoff function that is 0 in a neighborhood of 0 and let $\delta\in\bigl(0,\frac12\bigr)$. If λ is an eigenvalue of $\mathcal{N}_\hbar^\sharp$ such that $\lambda\leqslant\beta_0\hbar$ and if ψ is an associated eigenfunction, then we have*

$$\mathrm{Op}_\hbar^w(\chi_0(\hbar^{-\delta}\xi_3))\psi = \mathcal{O}(\hbar^\infty)\|\psi\|.$$

Proof. We write again $\psi(x_1,x_2,x_3) = e_{k,\hbar}(x_1)\varphi(x_2,x_3)$ with $k\in\{1,\dots,K\}$ and we have $\mathcal{N}_\hbar^{[k],\sharp}\varphi=\lambda\varphi$. We use again the formula (3-8) with $\chi_0(\hbar^{-\delta}\xi_3)$. We get the commutator estimate

$$\bigl|\langle [\mathcal{N}_\hbar^{[k],\sharp}, \mathrm{Op}_\hbar^w(\chi_0(\hbar^{-\delta}\xi_3))]\varphi, \mathrm{Op}_\hbar^w(\chi_0(\hbar^{-\delta}\xi_3))\varphi\rangle\bigr| \leqslant C\hbar^{\frac32-\delta}\bigl\|\mathrm{Op}_\hbar^w(\underline{\chi}_0(\hbar^{-\delta}\xi_3))\varphi\bigr\|^2.$$

We have

$$\mathrm{Op}_\hbar^w\bigl((\hbar^{-\delta}\xi_3)^2\chi_0^2(\hbar^{-\delta}\xi_3)\bigr) = \mathrm{Op}_{\hbar^{1-\delta}}^w(\xi_3^2\chi_0^2(\xi_3)),$$

so that, with the Gårding inequality,

$$\left\langle \mathrm{Op}_\hbar^w\left((\hbar^{-\delta}\xi_3)^2\chi_0^2(\hbar^{-\delta}\xi_3)\right)\varphi,\varphi\right\rangle \geqslant (1-C\hbar^{1-\delta})\|\varphi\|^2.$$

We infer

$$\left(\hbar^{2\delta}(1-Ch^{1-\delta})-\beta_0\hbar\right)\left\|\mathrm{Op}_\hbar^w(\chi_0(\hbar^{-\delta}\xi_3))\varphi\right\|^2 \leqslant C\hbar^{\frac{3}{2}-\delta}\left\|\mathrm{Op}_\hbar^w(\underline{\chi}_0(\hbar^{-\delta}\xi_3))\varphi\right\|^2. \qquad \square$$

Using $\mathrm{Op}_\hbar^w\, f^\star(\hbar,\mathcal{I}_\hbar,x_2,\xi_2,x_3,\xi_3) = \mathrm{Op}_\hbar^w\, f(\hbar,|z_1|^2,x_2,\xi_2,x_3,\xi_3)$, we deduce the following in the same way.

Proposition 3.14. *Let χ_1 be a smooth cutoff function that is 0 in a neighborhood of 0 and let $\delta \in \left(0,\frac{1}{2}\right)$. If λ is an eigenvalue of $\mathcal{N}_\hbar^\sharp$ such that $\lambda \leqslant \beta_0\hbar$ and if ψ is an associated eigenfunction, then we have*

$$\mathrm{Op}_\hbar^w\left(\chi_1(\hbar^{-\delta}(x_1,\xi_1))\right)\psi = \mathcal{O}(\hbar^\infty)\|\psi\|.$$

Proposition 3.15. *The spectra of $\mathcal{L}_{\hbar,\boldsymbol{A}}$ and $\mathcal{N}_\hbar^\sharp$ below $\beta_0\hbar$ coincide modulo $\mathcal{O}(\hbar^\infty)$.*

Proof. We refer to [Raymond and Vũ Ngọc 2015, Section 4.3], which contains similar arguments. $\square$

This proposition provides Corollary 2.4(a). With Proposition 3.6, we deduce point (b).

4. Second Birkhoff normal form

4A. *Birkhoff analysis of the first level.* This section is devoted to the proofs of Theorems 2.8 and 2.11.

The goal now is to normalize an $\hbar$-pseudodifferential operator $\mathcal{N}_\hbar^{[1]}$ on $\mathbb{R}^2$ whose Weyl symbol has the form

$$N_\hbar^{[1]} = \xi_3^2 + \hbar b(x_2,\xi_2,x_3) + r_\hbar(x_2,\xi_2,x_3,\xi_3),$$

where $r_\hbar$ is a classical symbol with the asymptotic expansion

$$r_\hbar = r_0 + \hbar r_1 + \hbar^2 r_2 + \cdots$$

(in the symbol class topology), where each r_ℓ has a formal expansion in ξ_3 of the form

$$r_\ell(x_2,\xi_2,x_3,\xi_3) \sim \sum_{2\ell+\beta\geqslant 3} c_{\ell,\beta}(x_2,\xi_2,x_3)\xi_3^\beta. \tag{4-1}$$

The leading terms of $N_\hbar^{[1]}$ are

$$N_\hbar^{[1]} = \xi_3^2 + \hbar b(x_2,\xi_2,x_3) + c_{1,1}(x_2,\xi_2,x_3)\hbar\xi_3 + \mathcal{O}(\hbar\xi_3^2) + \mathcal{O}(\xi_3^3) + \mathcal{O}(\hbar^2). \tag{4-2}$$

4A1. *First normalization of the symbol.* We consider the local change of variables $\hat\varphi(x_2,\xi_2,x_3,\xi_3) = (\hat x_2,\hat\xi_2,\hat x_3,\hat\xi_3)$, where

$$\begin{aligned} \hat x_2 &:= x_2 + \xi_3\partial_2 s(x_2,\xi_2), & \hat x_3 &:= x_3 - s(x_2,\xi_2),\\ \hat\xi_2 &:= \xi_2 + \xi_3\partial_1 s(x_2,\xi_2), & \hat\xi_3 &:= \xi_3. \end{aligned} \tag{4-3}$$

It is easy to check that the differential of $\hat\varphi$ is invertible as soon as ξ_3 is small enough. Moreover, we have

$$\hat\varphi^*\omega_0 - \omega_0 = \mathcal{O}(|\xi_3|).$$

By the Darboux–Weinstein theorem (see, for instance, [Raymond and Vũ Ngọc 2015, Lemma 2.4]), there exists a local diffeomorphism ψ such that

$$\psi = \mathrm{Id} + \mathcal{O}(\xi_3^2) \quad \text{and} \quad \psi^*\hat{\varphi}^*\omega_0 = \omega_0. \tag{4-4}$$

Using the improved Egorov theorem, one can find a unitary Fourier integral operator $V_\hbar$ such that the Weyl symbol of $V_\hbar^* \mathcal{N}_\hbar^{[1]} V_\hbar$ is $\widehat{N}_\hbar := N_\hbar^{[1]} \circ \hat{\varphi} \circ \psi + \mathcal{O}(\hbar^2)$. From (4-4), and (4-3), we see that $\hat{r}_\hbar := r_\hbar \circ \hat{\varphi} \circ \psi$ is still of the form (4-1), with modified coefficients $c_{\ell,\beta}$. Thus, using the new variables and a Taylor expansion in ξ_3, we get

$$\widehat{N}_\hbar = \hat{\xi}_3^2 + \hbar b\big(\hat{x}_2 + \mathcal{O}(\hat{\xi}_3),\, \hat{\xi}_2 + \mathcal{O}(\xi_3),\, \hat{x}_3 + s\big(\hat{x}_2 + \mathcal{O}(\hat{\xi}_3),\, \hat{\xi}_2 + \mathcal{O}(\hat{\xi}_3)\big) + \mathcal{O}(\hat{\xi}_3^2)\big) + \mathcal{O}(\hat{\xi}_3^3) + \hat{r}_\hbar + \mathcal{O}(\hbar^2)$$

and thus

$$\widehat{N}_\hbar = \hat{\xi}_3^2 + \hbar b\big(\hat{x}_2, \hat{\xi}_2, \hat{x}_3 + s(\hat{x}_2, \hat{\xi}_2)\big) + \hbar\hat{\xi}_3 g(\hat{x}_2, \hat{\xi}_2, \hat{x}_3) + \mathcal{O}(\hbar\hat{\xi}_3^2) + \hat{r}_\hbar + \mathcal{O}(\hat{\xi}_3^3) + \mathcal{O}(\hbar^2) \tag{4-5}$$

for some smooth function $g(\hat{x}_2, \hat{\xi}_2, \hat{x}_3)$.

Therefore $\widehat{N}_\hbar$ has the form

$$\widehat{N}_\hbar = \hat{\xi}_3^2 + \hbar b\big(\hat{x}_2, \hat{\xi}_2, \hat{x}_3 + s(\hat{x}_2, \hat{\xi}_2)\big) + \hat{c}_{1,1}(x_2, \hat{\xi}_2, \hat{x}_3)\hbar\hat{\xi}_3 + \mathcal{O}(\hbar\hat{\xi}_3^2) + \mathcal{O}(\hat{\xi}_3^3) + \mathcal{O}(\hbar^2).$$

4A2. *Where the second harmonic oscillator appears.* We now drop all the hats off the variables. We use a Taylor expansion with respect to x_3, which, in view of (2-6), yields

$$b(x_2, \xi_2, x_3 + s(x_2, \xi_2)) = b(x_2, \xi_2, s(x_2, \xi_2)) + \tfrac{1}{2}x_3^2\partial_3^2 b(x_2, \xi_2, s(x_2, \xi_2)) + \mathcal{O}(x_3^3).$$

We let

$$\nu = \big(\tfrac{1}{2}\partial_3^2 b(x_2, \xi_2, s(x_2, \xi_2))\big)^{\frac{1}{4}} \quad \text{and} \quad \gamma = \ln \nu. \tag{4-6}$$

We introduce the change of coordinates $(\check{x}_2, \check{x}_3, \check{\xi}_2, \check{\xi}_3) = C(x_2, x_3, \xi_2, \xi_3)$ defined by

$$\begin{aligned} \check{x}_2 &= x_2 + \frac{\partial\gamma}{\partial\xi_2}x_3\xi_3, & \check{\xi}_2 &= \xi_2 - \frac{\partial\gamma}{\partial x_2}x_3\xi_3, \\ \check{x}_3 &= \nu x_3, & \check{\xi}_3 &= \nu^{-1}\xi_3, \end{aligned} \tag{4-7}$$

for which one can check that $C^*\omega_0 - \omega_0 = \mathcal{O}(x_3\xi_3) = \mathcal{O}(\xi_3)$. As before, we can make this local diffeomorphism symplectic by the Darboux–Weinstein theorem, which modifies (4-7) by $\mathcal{O}(\xi_3^2)$. In the new variables (which we call (x_2, x_3, ξ_2, ξ_3) again), the symbol $\check{N}_\hbar$ has the form

$$\begin{aligned} \check{N}_\hbar = \nu^2(x_2, \xi_2)(\xi_3^2 + \hbar x_3^2) + \hbar b(x_2, \xi_2, s(x_2, \xi_2)) + \check{c}_{1,1}(x_2, \xi_2, x_3)\hbar\xi_3 \\ + \mathcal{O}(\hbar x_3^3) + \mathcal{O}(\hbar\xi_3^2) + \mathcal{O}(\xi_3^3) + \mathcal{O}(\hbar^2) \end{aligned}$$

for some smooth function $\check{c}_{1,1}(x_2, \xi_2, x_3)$.

4A3. *Normalizing the remainder.* The next step is to get rid of the term $\check{c}_{1,1}(x_2,\xi_2,x_3)\hbar\xi_3$. Let

$$a(x_2,\xi_2,x_3) := -\frac{1}{2}\int_0^{x_3} \check{c}_{1,1}(x_2,\xi_2,t)\,dt.$$

Since $\check{c}_{1,1}$ is compactly supported, a is bounded, and one can form the unitary pseudodifferential operator $\exp(iA)$, where $A = \mathrm{Op}_\hbar^w(a)$. We have

$$\exp(-iA)\,\mathrm{Op}_\hbar^w(\check{N}_\hbar)\exp(iA) = \mathrm{Op}_\hbar^w(\check{N}_\hbar) + \exp(-iA)[\mathrm{Op}_\hbar^w(\check{N}_\hbar),\exp(iA)].$$

The symbol of $[\exp(-iA)\,\mathrm{Op}_\hbar^w(\check{N}_\hbar),\exp(iA)]$ is

$$\frac{\hbar}{i}e^{-ia}\{N,e^{ia}\} + \mathcal{O}(\hbar^2) = \hbar\{\check{N}_\hbar,a\} + \mathcal{O}(\hbar^2) = \hbar\{\check{N}_0,a\} + \mathcal{O}(\hbar^2),$$

where $\check{N}_0$ is the principal symbol of $\check{N}_\hbar$, which satisfies

$$\check{N}_0 = \xi_3^2 + \mathcal{O}(\xi_3^3).$$

Therefore $\{\check{N}_\hbar,a\} = \{\xi_3^2,a\} + \mathcal{O}(\xi_3^2)$. Since

$$\{\xi_3^2,a\} = 2\xi_3\frac{\partial a}{\partial x_3} = -\xi_3\check{c}_{1,1},$$

we get

$$\exp(-iA)\,\mathrm{Op}_\hbar^w(\check{N}_\hbar)\exp(iA) = \mathrm{Op}_\hbar^w\big(\check{N}_\hbar - \hbar\xi_3\check{c}_{1,1} + \mathcal{O}(\hbar\xi_3^2) + \mathcal{O}(\hbar^2)\big),$$

which shows that we can remove the coefficient of $\hbar\xi_3$. The new operator given by the conjugation formula $\underline{\mathcal{N}}_\hbar^{[1]} = \exp(-iA)\,\mathrm{Op}_\hbar^w(\check{N}_\hbar)\exp(iA)$ has a symbol of the form

$$\underline{N}_\hbar^{[1]} = \nu^2(x_2,\xi_2)(\xi_3^2 + \hbar x_3^2) + \hbar b(x_2,\xi_2,s(x_2,\xi_2)) + \underline{r}_\hbar, \tag{4-8}$$

where $\underline{r}_\hbar = \mathcal{O}(\hbar x_3^3) + \mathcal{O}(\hbar\xi_3^2) + \mathcal{O}(\xi_3^3) + \mathcal{O}(\hbar^2)$.

This proves Theorem 2.8.

4A4. *The second Birkhoff normal form.* We now want to perform a Birkhoff normal form for $\underline{\mathcal{N}}_\hbar^{[1],\sharp}$ relative to the "second harmonic oscillator"

$$\underline{\nu}^2(x_2,\xi_2)(\xi_3^2 + \hbar x_3^2).$$

Using Notation 2.10, we introduce the new semiclassical parameter $h = \hbar^{\frac{1}{2}}$, and use the relation

$$\mathrm{Op}_\hbar^w(\underline{N}_\hbar^{[1],\sharp}) = \mathrm{Op}_h^w(\underline{\mathsf{N}}_h^{[1],\sharp}).$$

Thus, let $\tilde{\xi}_j := \hbar^{-\frac{1}{2}}\xi_j$. The new symbol $\underline{\mathsf{N}}_h^{[1],\sharp}$ has the form

$$\underline{\mathsf{N}}_h^{[1],\sharp}(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3) = h^2\big(\underline{\nu}^2(x_2,h\tilde{\xi}_2)(\tilde{\xi}_3^2 + x_3^2) + \underline{b}(x_2,h\tilde{\xi}_2,s(x_2,h\tilde{\xi}_2)) + h^{-2}\underline{r}_{h^2}^\sharp(x_2,h\tilde{\xi}_2,x_3,h\tilde{\xi}_3)\big).$$

We introduce momentarily a new parameter μ and define

$$\underline{\mathsf{N}}_h^{[1],\sharp}(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3;\mu) := \underline{\nu}^2(x_2,\mu\tilde{\xi}_2)(\tilde{\xi}_3^2 + x_3^2) + \underline{b}(x_2,\mu\tilde{\xi}_2,s(x_2,\mu\tilde{\xi}_2)) + h^{-2}\underline{r}_{h^2}^\sharp(x_2,\mu\tilde{\xi}_2,x_3,h\tilde{\xi}_3).$$

Notice that $\underline{\mathsf{N}}_h^{[1],\sharp}(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3;h) = h^{-2}\underline{\mathsf{N}}_h^{[1],\sharp}(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3)$. We define now a space of functions suitable for the Birkhoff normal form in $(x_3,\tilde{\xi}_3,h)$. Let us now use the notation of the Appendix introduced in (A-4) in the case when the family of smooth linear maps $\mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$\varphi_{\mu,\mathbb{R}^2}(x_2,\tilde{\xi}_2) = (x_2,\mu\tilde{\xi}_2).$$

Let

$$\mathscr{F} := \mathcal{C}(1)_{\mathbb{R}^2},$$

where the index $\mathbb{R}^2$ means that we consider symbols on $\mathbb{R}^2$. More explicitly, we have

$$\mathscr{F} = \bigl\{d \text{ s.t. } \exists c \in S(1;[0,1]\times(0,1])_{\mathbb{R}^2} \bigm| d(x_2,\tilde{\xi}_2;\mu,h) = c(\varphi_{\mu,\mathbb{R}^2}(x_2,\tilde{\xi}_2);\mu,h)\bigr\}.$$

Then we define

$$\mathscr{E} := \mathscr{F}[\![x_3,\tilde{\xi}_3,h]\!],$$

endowed with the full Poisson bracket

$$\mathscr{E}\times\mathscr{E} \ni (f,g) \mapsto \{f,g\} = \sum_{j=2,3} \frac{\partial f}{\partial\tilde{\xi}_j}\frac{\partial g}{\partial x_j} - \frac{\partial g}{\partial\tilde{\xi}_j}\frac{\partial f}{\partial x_j} \in \mathscr{E},$$

and the corresponding Moyal bracket $[f,g]$. We remark that the formal Taylor series of the symbol $\underline{\mathsf{N}}_h^{[1],\sharp}(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3;\mu)$ with respect to $(x_3,\tilde{\xi}_3,h)$ belongs to $\mathscr{E}$. We may apply the semiclassical Birkhoff normal form relative to the main term $\underline{\nu}^2(x_2,\mu\tilde{\xi}_2)(\tilde{\xi}_3^2+x_3^2)$ exactly as in Section 3B1 (and also [Raymond and Vũ Ngọc 2015, Proposition 2.7]), where we use the fact that the function

$$(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3;\mu,h) \mapsto (\underline{\nu}^2(x_2,\mu\tilde{\xi}_2))^{-1}$$

belongs to $\mathscr{E}$ because $\underline{\nu}^2 > C > 0$ uniformly with respect to μ. Let us consider $\gamma\in\mathscr{E}$, the formal Taylor expansion of $h^{-2}\underline{r}^\sharp_{h^2}(x_2,\mu\tilde{\xi}_2,x_3,h\tilde{\xi}_3)$ with respect to $(x_3,\tilde{\xi}_3,h)$. The series γ is of valuation 3 and we obtain two formal series $\kappa,\tau\in\mathscr{E}$ of valuation at least 3 such that

$$[\kappa, x_3^2+\tilde{\xi}_3^2] = 0$$

and

$$e^{ih^{-1}\mathrm{ad}_\tau}(\underline{\nu}^2(x_2,\mu\tilde{\xi}_2)(\tilde{\xi}_3^2+x_3^2)+\gamma) = \underline{\nu}^2(x_2,\mu\tilde{\xi}_2)(\tilde{\xi}_3^2+x_3^2)+\kappa.$$

The coefficients of τ are in $S(1)$ and one can find a smooth function $\tau_h\in S(1)$ with compact support with respect to $(x_3,\tilde{\xi}_3,h)$ and whose Taylor series in $(x_3,\tilde{\xi}_3,h)$ is τ. By the Borel summation, τ_h will actually lie in $S(m')$ with $m'(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3) = \langle(x_3,\tilde{\xi}_3)\rangle^{-k}$ for any $k>0$, uniformly for small $h>0$ and $\mu\in[0,1]$. Notice that $\underline{\mathsf{N}}_h^{[1],\sharp}\in\mathcal{C}(m)$ with $m = \langle(x_3,\tilde{\xi}_3)\rangle^2 \geqslant 1$, and that $mm' = \mathscr{O}(1)$.

Then, we can apply Theorem A.3 with the family of endomorphisms of $\mathbb{R}^4$ defined by

$$\varphi_{\mu,\mathbb{R}^4}(x_2,\tilde{\xi}_2,x_3,\tilde{\xi}_3) = (x_2,\mu\tilde{\xi}_2,x_3,\tilde{\xi}_3).$$

Thus, the new operator

$$\mathfrak{M}_h = e^{ih^{-1}\mathrm{Op}_h^w\tau_h}\underline{\mathfrak{N}}_h^{[1],\sharp}e^{-ih^{-1}\mathrm{Op}_h^w\tau_h}$$

is a pseudodifferential operator whose Weyl symbol belongs to the class $\mathcal{C}(m)$ modulo $h^\infty S(1)$ (see the notation of Theorem 2.11). Moreover, thanks to Theorem A.4, its symbol M_h admits the following Taylor expansion (with respect to $(x_3, \tilde{\xi}_3, h)$)

$$\tilde{b}(x_2, \mu\tilde{\xi}_2, s(x_2, \mu\tilde{\xi}_2)) + \underline{\nu}^2(x_2, \mu\tilde{\xi}_2)(\tilde{\xi}_3^2 + x_3^2) + \kappa.$$

We write $\kappa = \sum_{m+2\ell\geqslant 3} c_{m,\ell}(x_2, \mu\tilde{\xi}_2)|\tilde{z}_3|^{\star 2m} h^\ell$ and we may find a smooth function $g^\star(x_2, \mu\tilde{\xi}_2, Z, h)$ such that its Taylor series with respect to Z, h is

$$\sum_{2m+2\ell\geqslant 3} c_{m,\ell}(x_2, \mu\tilde{\xi}_2) Z^m h^\ell.$$

We may now replace μ by h, which achieves the proof of Theorem 2.11.

4B. ***Spectral reduction to the second normal form.*** This section is devoted to the proof of Corollary 2.13.

4B1. *From $\mathcal{N}_\hbar^{[1],\sharp}$ to $\underline{\mathcal{N}}_\hbar^{[1],\sharp}$.* In this section, we prove Corollary 2.9.

Lemma 4.1. *We have*

$$\mathsf{N}(\mathcal{N}_\hbar^{[1],\sharp}, \beta_0\hbar) = \mathcal{O}(\hbar^{-2}), \quad \mathsf{N}(\underline{\mathcal{N}}_\hbar^{[1],\sharp}, \beta_0\hbar) = \mathcal{O}(\hbar^{-2}).$$

Proof. The first estimate comes from Proposition 3.6 and Corollary 3.8. The second estimate can be obtained by the same method as in the proof of Corollary 3.8. □

Let us now summarize the microlocalization properties of the eigenfunctions of $\underline{\mathcal{N}}_\hbar^{[1],\sharp}$ in the following proposition.

Proposition 4.2. *Let χ_0 be a smooth cutoff function on $\mathbb{R}$ that is 0 in a neighborhood of 0 and let $\delta \in \left(0, \frac{1}{2}\right)$. Let χ be a smooth cutoff function that is 0 on the bounded set $\{x_3^2 + \underline{b}(x_2, \xi_2, s(x_2, \xi_2)) \leqslant \beta_0\}$ and 1 on the set $\{x_3^2 + \underline{b}(x_2, \xi_2, s(x_2, \xi_2)) \geqslant \beta_0 + \tilde{\varepsilon}\}$, with $\tilde{\varepsilon} > 0$. If λ is an eigenvalue of $\underline{\mathcal{N}}_\hbar^{[1],\sharp}$ such that $\lambda \leqslant \beta_0\hbar$ and if ψ is an associated eigenfunction, then we have*

$$\mathrm{Op}_\hbar^w(\chi(x_2, \xi_2, x_3))\psi = \mathcal{O}(\hbar^\infty)\|\psi\|,$$

and

$$\mathrm{Op}_\hbar^w(\chi_0(\hbar^{-\delta}\xi_3))\psi = \mathcal{O}(\hbar^\infty)\|\psi\|.$$

Proof. The proof follows exactly the same lines as for Propositions 3.12 and 3.13. □

Lemma 4.1 and Proposition 4.2 on the one hand and Propositions 3.12 and 3.13 on the other hand are enough to deduce Corollary 2.9(a) from Theorem 2.8. Part (b) easily follows from Corollary 2.4.

4B2. *From $\underline{\mathfrak{N}}_h^{[1],\sharp}$ to $\mathfrak{M}_h^\sharp$.* Let us now prove Corollary 2.13(a). We get the following rough estimate of the number of eigenvalues.

Lemma 4.3. *We have*

$$\mathsf{N}(\underline{\mathfrak{N}}_h^{[1],\sharp}, \beta_0 h^2) = \mathsf{N}(\mathfrak{M}_h, \beta_0 h^2) = \mathcal{O}(h^{-4}), \tag{4-9}$$

$$\mathsf{N}(\mathfrak{M}_h^\sharp, \beta_0 h^2) = \mathcal{O}(h^{-4}). \tag{4-10}$$

Proof. First, we notice that $\underline{\mathfrak{N}}_h^{[1],\sharp}$ and $\mathfrak{M}_h$ are unitarily equivalent so that (4-9) holds. Then, given $\eta > 0$ and h small enough and up to shrinking the support of $g^\star$ and by using the Calderon–Vaillancourt theorem (as in the proof of Lemma 3.4), $\mathfrak{M}_h^\sharp \geqslant \widetilde{\mathfrak{M}}_h^\sharp$ in the sense of quadratic forms, with

$$\widetilde{\mathfrak{M}}_h^\sharp = \mathrm{Op}_h^w\big(h^2\underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2))\big) + h^2\mathcal{J}_h \,\mathrm{Op}_h^w\big((\underline{\nu}^2(x_2, h\tilde{\xi}_2)) - \eta\big).$$

Since $\underline{\nu}^2 \geqslant c > 0$, we get

$$\begin{aligned}\mathrm{Op}_h^w\big(h^2\underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2))\big) + h^2\mathcal{J}_h \,\mathrm{Op}_h^w\big((\underline{\nu}^2(x_2, h\tilde{\xi}_2)) - \eta\big)&\\ \geqslant \mathrm{Op}_h^w\big(h^2\underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2))\big) + \tfrac{1}{2}ch^2\mathcal{J}_h.&\end{aligned}$$

We deduce the upper bound (4-10) by separation of variables and the min-max principle. □

The following proposition deals with the microlocal properties of the eigenfunctions of $\underline{\mathfrak{N}}_h^{[1],\sharp}$.

Proposition 4.4. *Let $\eta \in (0, 1)$, $\delta \in \big(0, \frac{1}{2}\eta\big)$, and $C > 0$. Let χ be a smooth cutoff function that is 0 on $\{\underline{b}(x_2, \xi_2, s(x_2, \xi_2)) \leqslant \beta_0\}$ and 1 on the set $\{\underline{b}(x_2, \xi_2, s(x_2, \xi_2)) \geqslant \beta_0 + \tilde{\varepsilon}\}$, with $\tilde{\varepsilon} > 0$. Let also χ_1 be a smooth cutoff function on $\mathbb{R}^2$ that is 0 in a neighborhood of 0.*

If λ is an eigenvalue of $\underline{\mathfrak{N}}_h^{[1],\sharp}$ such that $\lambda \leqslant \beta_0 h^2$ and if ψ is an associated eigenfunction, we have

$$\mathrm{Op}_h^w\big(\chi(x_2, h\tilde{\xi}_2)\big)\psi = \mathcal{O}(h^\infty)\|\psi\| \tag{4-11}$$

and if λ is an eigenvalue of $\underline{\mathfrak{N}}_h^{[1],\sharp}$ such that $\lambda \leqslant b_0 h^2 + Ch^{2+\eta}$ and if ψ is an associated eigenfunction, we have

$$\mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi = \mathcal{O}(h^\infty)\|\psi\|. \tag{4-12}$$

Proof. The estimate (4-11) is a consequence of Proposition 4.2. Then, let us write the symbol of $\underline{\mathfrak{N}}_h^{[1],\sharp}$,

$$\underline{\mathsf{N}}_h^{[1],\sharp} = h^2\underline{\nu}^2(x_2, h\tilde{\xi}_2)(\tilde{\xi}_3^2 + x_3^2) + h^2\underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2)) + \underline{r}^\sharp_{h^2}(x_2, h\tilde{\xi}_2, x_3, h\tilde{\xi}_3).$$

We write

$$\begin{aligned}&\big\langle \underline{\mathfrak{N}}_h^{[1],\sharp}\,\mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi, \mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\big\rangle\\ &\quad= \lambda\big\|\mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi\big\|^2 + \big\langle\big[\underline{\mathfrak{N}}_h^{[1],\sharp}, \mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\big], \mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi\big\rangle.\end{aligned}$$

We get

$$\big\langle\big[\underline{\mathfrak{N}}_h^{[1],\sharp}, \mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\big], \mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi\big\rangle \leqslant Ch^3\big\|\mathrm{Op}_h^w\big(\underline{\chi}_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi\big\|^2,$$

where we have used (4-11). Then, we use that

$$\underline{b}(x_2, h\tilde{\xi}_2, s(x_2, h\tilde{\xi}_2)) \geqslant b_0, \quad \underline{\nu}^2(x_2, h\tilde{\xi}_2) \geqslant c_0 > 0, \quad \lambda \leqslant b_0 h^2 + Ch^{2+\eta},$$

and the Gårding inequality to deduce

$$h^2(Ch^{2\delta} - Ch^\eta)\big\|\mathrm{Op}_h^w\big(\chi_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi\big\|^2 \leqslant Ch^3\big\|\mathrm{Op}_h^w\big(\underline{\chi}_1(h^{-\delta}(x_3, \tilde{\xi}_3))\big)\psi\big\|^2.$$

The desired estimate follows by an iteration argument. □

In the same way we can deal with $\mathfrak{M}_h^\sharp$.

Proposition 4.5. *Let $\eta \in (0,1)$, $\delta \in \left(0, \frac{1}{2}\eta\right)$, and $C > 0$. Let χ be a smooth cutoff function that is 0 on $\{\underline{b}(x_2,\xi_2,s(x_2,\xi_2)) \leqslant \beta_0\}$ and 1 on the set $\{\underline{b}(x_2,\xi_2,s(x_2,\xi_2)) \geqslant \beta_0+\tilde{\varepsilon}\}$, with $\tilde{\varepsilon} > 0$. If λ is an eigenvalue of $\mathfrak{M}_h^\sharp$ such that $\lambda \leqslant \beta_0 h^2$ and if ψ is an associated eigenfunction, we have*

$$\mathrm{Op}_h^w(\chi(x_2,h\tilde{\xi}_2))\psi = \mathcal{O}(h^\infty)\|\psi\| \tag{4-13}$$

and if λ is an eigenvalue of $\mathfrak{M}_h^\sharp$ such that $\lambda \leqslant b_0 h^2 + Ch^{2+\eta}$ and if ψ is an associated eigenfunction, we have

$$\mathrm{Op}_h^w\left(\chi_1(h^{-\delta}(x_3,\tilde{\xi}_3))\right)\psi = \mathcal{O}(h^\infty)\|\psi\|. \tag{4-14}$$

Proof. In order to get (4-13), it is enough to go back to the representation with semiclassical $\hbar$, that is, $\mathfrak{M}_h^\sharp = \mathcal{M}_\hbar^\sharp$. Indeed the microlocal estimate follows by the same arguments as in Propositions 3.12 and 3.13. Then, (4-14) follows as in Proposition 4.4. □

Propositions 4.4 and 4.5 and Theorem 2.11 standardly imply Corollary 2.13(a).

4B3. *From $\mathfrak{M}_h^\sharp$ to $\mathfrak{M}_h^{[1],\sharp}$.* Let us now prove Corollary 2.13(b). Note that part (c) is just a reformulation of (b).

Let us consider the Hilbertian decomposition $\mathfrak{M}_h^\sharp = \bigoplus_{k\geqslant 1}\mathfrak{M}_h^{[k],\sharp}$, where the symbol $\mathsf{M}_h^{[k],\sharp}$ of $\mathfrak{M}_h^{[k],\sharp}$ is

$$h^2\underline{b}\left(x_2,h\tilde{\xi}_2,s(x_2,h\tilde{\xi}_2)\right) + (2k-1)h^3\underline{\nu}^2(x_2,h\tilde{\xi}_2) + h^2 g^\star\left(h,(2k-1)h,x_2,h\tilde{\xi}_2\right).$$

There exists $h_0 > 0$ such that for all $k \geqslant 1$ and $h \in (0,h_0)$,

$$\langle \mathfrak{M}_h^{[k],\sharp}\psi,\psi\rangle \geqslant \left\langle \mathrm{Op}_h^w\left(h^2\underline{b}\left(x_2,h\tilde{\xi}_2,s(x_2,h\tilde{\xi}_2)\right) + (2k-1)h^3(\underline{\nu}^2(x_2,h\tilde{\xi}_2)-\varepsilon)\right)\psi,\psi\right\rangle.$$

Since each eigenfunction of $\mathfrak{M}_h^{[k],\sharp}$ associated with an eigenvalue less than $\beta_0 h^2$ provides an eigenfunction of $\mathfrak{M}_h^\sharp$, we infer that the eigenfunctions of $\mathfrak{M}_h^{[k],\sharp}$ are uniformly microlocalized in an (x_2,ξ_2)-neighborhood of $(0,0)$ as small as we want. Therefore, on the range of $\mathbb{1}_{(-\infty,b_0h^2)}(\mathfrak{M}_h^{[k],\sharp})$, we have

$$\langle \mathfrak{M}_h^{[k],\sharp}\psi,\psi\rangle \geqslant \left\langle \mathrm{Op}_h^w\left(h^2\underline{b}\left(x_2,h\tilde{\xi}_2,s(x_2,h\tilde{\xi}_2)\right) + (2k-1)h^3(\nu^2(0,0)-2\varepsilon)\right)\psi,\psi\right\rangle,$$

and, with the Gårding inequality in the $\hbar$-quantization, we get

$$\langle \mathfrak{M}_h^{[k],\sharp}\psi,\psi\rangle \geqslant \left\langle \mathrm{Op}_h^w\left(h^2 b_0 + (2k-1)h^3(\nu^2(0,0)-\varepsilon) - Ch^4\right)\psi,\psi\right\rangle.$$

This implies Corollary 2.13(b).

5. Third Birkhoff normal form

5A. ***Birkhoff analysis of the first level.*** In this section we prove Theorem 2.15.

We consider $\mathcal{M}_\hbar^{[1],\sharp} = \mathrm{Op}_\hbar^w(M_\hbar^{[1],\sharp})$, with

$$M_\hbar^{[1],\sharp} = \hbar\underline{b}(x_2,\xi_2,s(x_2,\xi_2)) + \hbar^{\frac{3}{2}}\underline{\nu}^2(x_2,\xi_2) + \hbar g^\star(\hbar^{\frac{1}{2}},\hbar^{\frac{1}{2}},x_2,\xi_2).$$

By using a Taylor expansion, we get,

$$M_\hbar^{[1],\sharp} = \hbar b_0 + \tfrac{1}{2}\hbar\,\mathsf{Hess}_{(0,0)}\underline{b}\left(x_2,\xi_2,s(x_2,\xi_2)\right) + \hbar^{\frac{3}{2}}\nu^2(0,0) + cx_2\hbar^{\frac{3}{2}} + d\xi_2\hbar^{\frac{3}{2}} + \hbar\mathcal{O}((\hbar^{\frac{1}{2}},z_2)^3), \tag{5-1}$$

where $c = \partial_{x_2}\nu^2(0,0)$ and $d = \partial_{\xi_2}\nu^2(0,0)$, and we have identified the Hessian with its quadratic form in (x_2, ξ_2).

Then, there exists a linear symplectic change of variables that diagonalizes the Hessian, so that, if $L_\hbar$ is the associated unitary transform,

$$L_\hbar^* \mathcal{M}_\hbar^{[1],\sharp} L_\hbar = \mathrm{Op}_\hbar^w(\widehat{M}_\hbar^{[1],\sharp}),$$

with

$$\widehat{M}_\hbar^{[1],\sharp} = \hbar b_0 + \tfrac{1}{2}\hbar\theta(x_2^2+\xi_2^2) + \hbar^{\frac{3}{2}}\nu^2(0,0) + \hat{c}x_2\hbar^{\frac{3}{2}} + \hat{d}\xi_2\hbar^{\frac{3}{2}} + \hbar\mathcal{O}((\hbar^{\frac{1}{2}}, z_2)^3),$$

where

$$\theta = \sqrt{\det \mathsf{Hess}_{(0,0)} b(x_2,\xi_2,s(x_2,\xi_2))}.$$

Since $(\partial_{x_3}b)(x_2,\xi_2,s(x_2,\xi_2)) = 0$ and $(0,0)$ is a critical point of s, we notice that $\partial^2_{x_2x_3}b(0,0,0) = \partial^2_{\xi_2x_3}b(0,0,0) = 0$. Thus

$$\det \mathsf{Hess}_{(0,0,0)} b(0,0,0) = \theta^2\partial^2_{x_3}b(0,0,0).$$

Using that b is identified with $b\circ\chi$ (see Remarks 2.2 and 3.1), this provides the expression given in (1-15).

Note that $\hat{c}^2+\hat{d}^2 = \|(\nabla_{x_2,\xi_2}\nu^2)(0,0)\|^2$ since the symplectic transform is in fact a rotation. Moreover, we have

$$\theta(x_2^2+\xi_2^2) + \hat{c}x_2\hbar^{\frac{1}{2}} + \hat{d}\xi_2\hbar^{\frac{1}{2}} = \theta\left(\left(x_2 - \frac{\hat{c}\hbar^{\frac{1}{2}}}{\theta}\right)^2 + \left(\xi_2 - \frac{\hat{d}\hbar^{\frac{1}{2}}}{\theta}\right)^2\right) - \hbar\frac{\hat{c}^2+\hat{d}^2}{\theta}.$$

Thus, there exists a unitary transform $\widehat{U}_{\hbar^{1/2}}$, which is in fact an $\hbar$-Fourier integral operator whose phase admits a Taylor expansion in powers of $\hbar^{\frac{1}{2}}$, such that

$$\widehat{U}^*_{\hbar^{1/2}} L_\hbar^* \mathcal{M}_\hbar^{[1],\sharp} L_\hbar \widehat{U}_{\hbar^{1/2}} =: \underline{\mathcal{F}}_\hbar = \mathrm{Op}_\hbar^w(\underline{F}_\hbar),$$

where

$$\underline{F}_\hbar = \hbar b_0 + \hbar^{\frac{3}{2}}\nu^2(0,0) - \frac{\|(\nabla_{x_2,\xi_2}\nu^2)(0,0)\|^2}{2\theta}\hbar^2 + \hbar\big(\tfrac{1}{2}\theta|z_2|^2 + \mathcal{O}((\hbar^{\frac{1}{2}}, z_2)^3)\big).$$

Now we perform a semiclassical Birkhoff normal form in the space of formal series $\mathbb{R}[\![x_2,\xi_2,\hbar^{\frac{1}{2}}]\!]$ equipped with the degree such that $x_2^\ell\xi_2^m\hbar^{\frac{n}{2}}$ is $\ell+m+n$ and endowed with the Moyal product. Let $\underline{F}_\hbar^T$ be the full Taylor series of $\underline{F}_\hbar$. We find a formal series $\tau(x_2,\xi_2,\hbar^{\frac{1}{2}})$ with a valuation at least 3 such that

$$e^{i\hbar^{-1}\mathrm{ad}_\tau}\underline{F}_\hbar^T = F_\hbar^T,$$

where $F_\hbar^T$ is a formal series of the form

$$F_\hbar^T = \hbar b_0 + \hbar^{\frac{3}{2}}\nu^2(0,0) - \frac{\|(\nabla_{x_2,\xi_2}\nu^2)(0,0)\|^2}{2\theta}\hbar^2 + \tfrac{1}{2}\theta\hbar|z_2|^2 + \hbar k^T(\hbar^{\frac{1}{2}}, |z_2|^2),$$

and k^T is a formal series in $\mathbb{R}[\![\hbar^{\frac{1}{2}}, |z_2|^2]\!]$ (and that can be also written as a formal series in Moyal power of $|z_2|^2$, say $(k^T)^\star$).

Let $\tilde{\tau}(x_2,\xi_2,\mu)$ be a compactly supported function whose Taylor expansion at $(0,0,0)$ is equal to $\tau(x_2,\xi_2,\mu)$. By the Egorov theorem (Theorem A.2), uniformly with respect to the parameter μ, we obtain

$$e^{-i\hbar^{-1}\mathrm{Op}_\hbar^w(\tilde{\tau})}\,\mathrm{Op}_\hbar^w(\underline{F}_{\mu^2})e^{i\hbar^{-1}\mathrm{Op}_\hbar^w(\tilde{\tau})} =: \mathrm{Op}_\hbar^w(\widetilde{F}_\mu)$$

is an $\hbar$-pseudodifferential operator depending smoothly on μ. Expanding $\widetilde{F}_\mu$ in powers of μ in the $S(1)$ topology, and letting $\mu=\sqrt{\hbar}$, we see that $\widetilde{F}_{\sqrt{\hbar}}=F_\hbar+\widetilde{G}_\hbar$, where

$$F_\hbar=\hbar b_0+\hbar^{\frac{3}{2}}\nu^2(0,0)-\frac{\|(\nabla_{x_2,\xi_2}\nu^2)(0,0)\|^2}{2\theta}\hbar^2+\tfrac{1}{2}\theta\hbar|z_2|^2+\hbar k(\hbar^{\frac{1}{2}},|z_2|^2),$$

with k a smooth function with support as small as desired with respect to its second variable, and $\widetilde{G}_h=\hbar\mathcal{O}(|z_2|^\infty)$. It remains to notice that $\mathrm{Op}_\hbar^w\big(k(\hbar^{\frac{1}{2}},|z_2|^2)\big)$ can be written as $k^\star(\hbar^{\frac{1}{2}},\mathcal{K}_\hbar)$ modulo $\mathrm{Op}_\hbar^w(\mathcal{O}(|z_2|^\infty))$. This achieves the proof of Theorem 2.15.

5B. ***Spectral reduction to the third normal form.*** Corollary 2.16 is a consequence of the following lemma and proposition.

Lemma 5.1. *We have*

$$\mathsf{N}(\mathcal{M}_\hbar^{[1],\sharp},\beta_0\hbar)=\mathcal{O}(\hbar^{-2}),\quad \mathsf{N}(\mathcal{F}_\hbar,b_0\hbar+C\hbar^{1+\eta})=\mathcal{O}(\hbar^{-1+\eta}).$$

Proof. The first estimate follows from Lemma 4.3 and the second one from a comparison with the harmonic oscillator in x_2. □

The last proposition concerns the microlocalization of the eigenfunctions.

Proposition 5.2. *Let $\eta\in(0,1)$, $\delta\in\big(0,\frac{1}{2}\eta\big)$, and $C>0$. Let χ be a smooth cutoff function that is 0 in a bounded neighborhood of $(0,0)$ and 1 outside a bounded neighborhood of $(0,0)$. If λ is an eigenvalue of $\mathcal{M}_\hbar^{[1],\sharp}$ or of $\mathcal{F}_\hbar$ such that $\lambda\leqslant b_0\hbar+C\hbar^{1+\eta}$ and if ψ is an associated eigenfunction, we have*

$$\mathrm{Op}_\hbar^w\big(\chi(\hbar^{-\delta}(x_2,\xi_2))\big)\psi=\mathcal{O}(\hbar^\infty).$$

Proof. The proof is similar to that of Proposition 4.4. □

Appendix: Egorov theorems

We start with the classical result (see, for instance, [Zworski 2012, Theorem 11.1; Robert 1987, Théorème IV.10]).

Theorem A.1 [Zworski 2012, Theorem 11.1, Remark (ii) on p. 251]. *Let P and Q be h-pseudodifferential operators on $\mathbb{R}^d$, with $P\in\mathrm{Op}_h^w(S(1))$ and $Q\in\mathrm{Op}_h^w(S(1))$. Then the operator $e^{\frac{i}{h}Q}Pe^{-\frac{i}{h}Q}$ is a pseudodifferential operator in $\mathrm{Op}_h^w(S(1))$, and*

$$e^{\frac{i}{h}Q}Pe^{-\frac{i}{h}Q}-\mathrm{Op}_h^w(p\circ\kappa)\in h\,\mathrm{Op}_h^w(S(1)).$$

Here p is the Weyl symbol of P, and the canonical transformation κ is the time-1 Hamiltonian flow associated with principal symbol of Q.

From this classical version of Egorov's theorem, one can deduce the following refinement that is useful when p does not belong to $S(1)$ (as is the case in this paper).

Theorem A.2. *Let P and Q be h-pseudodifferential operators on $\mathbb{R}^d$, with $P \in \mathrm{Op}_h^w(S(m))$ and $Q \in \mathrm{Op}_h^w(S(m'))$, where m and m' are order functions such that*

$$m' = \mathcal{O}(1), \quad mm' = \mathcal{O}(1). \tag{A-1}$$

Then the operator $e^{\frac{i}{h}Q}Pe^{-\frac{i}{h}Q}$ is a pseudodifferential whose symbol is in $S(m)$, and $e^{\frac{i}{h}Q}Pe^{-\frac{i}{h}Q} - \mathrm{Op}_h^w(p\circ\kappa) \in h\,\mathrm{Op}_h^w(S(1))$.

Proof. The proof is based on the following observation. In order to compare $\mathrm{Op}_h^w(p\circ\kappa^t)$ and $e^{\frac{it}{h}Q}Pe^{-\frac{it}{h}Q}$, we consider the derivative

$$\frac{d}{d\tau}\big(e^{\frac{i\tau}{h}Q}\,\mathrm{Op}_h^w(p\circ\kappa^{t-\tau})e^{-\frac{i\tau}{h}Q}\big) = e^{\frac{i\tau}{h}Q}\Big(\frac{i}{h}[Q,\mathrm{Op}_h^w(p\circ\kappa^{t-\tau})] + \frac{d}{d\tau}\mathrm{Op}_h^w(p\circ\kappa^{t-\tau})\Big)e^{-\frac{i\tau}{h}Q}.$$

From hypothesis (A-1), the term $[Q,\mathrm{Op}_h^w(p\circ\kappa^{t-\tau})]$ belongs to $\mathrm{Op}_h^w(S(1))$; moreover, if we denote by q_0 the principal symbol of Q, we have

$$\frac{d}{d\tau}\mathrm{Op}_h^w(p\circ\kappa^{t-\tau}) = -\mathrm{Op}_h^w(\{q_0, p\circ\kappa^{t-\tau}\}),$$

which implies that this term is also in $\mathrm{Op}_h^w(S(1))$. By symbolic calculus, we see that

$$\frac{i}{h}[Q,\mathrm{Op}_h^w(p\circ\kappa^{t-\tau})] + \frac{d}{d\tau}\mathrm{Op}_h^w(p\circ\kappa^{t-\tau}) \in h\,\mathrm{Op}_h^w(S(1)) \tag{A-2}$$

uniformly for t,τ in compact sets. It follows by integration from 0 to t that

$$e^{\frac{it}{h}Q}Pe^{-\frac{it}{h}Q} = \mathrm{Op}_h^w(p\circ\kappa^t) + h\int_0^t e^{\frac{is}{h}Q}P_1(s)e^{-\frac{is}{h}Q}\,ds \tag{A-3}$$

for some $P_1(s) \in \mathrm{Op}_h^w(S(1))$, uniformly for $s\in[0,t]$. Applying Theorem A.1 to the integrand, we see that $e^{\frac{it}{h}Q}Pe^{-\frac{it}{h}Q} - \mathrm{Op}_h^w(p\circ\kappa^t) \in h\,\mathrm{Op}_h^w(S(1))$. □

In order to quantize the formal Birkhoff procedure of Section 4A4, one needs to consider symbols in a class $\mathcal{C}$ stable under the Moyal product. For that purpose we first define the families of symbols $S(m;[0,1]\times(0,1])$, that is, of smooth functions $a:\mathbb{R}^{2d}\times[0,1]\times(0,1]\to\mathbb{C}$ such that, for any $\alpha\in\mathbb{N}^{2d}$, there exists C_α such that, for all $(z;\mu,h)\in\mathbb{R}^{2d}\times[0,1]\times(0,1]$,

$$|\partial_z^\alpha a(z;\mu,h)| \leqslant C_\alpha m(z)$$

and where m is an order function on $\mathbb{R}^{2d}$. The pair (μ,h) is considered as a parameter.

Then, let $(\varphi_\mu)_{\mu\in[0,1]}$ be a smooth family of linear maps $\mathbb{R}^{2d}\to\mathbb{R}^{2d}$ and define the following families of symbols on $\mathbb{R}^{2d}$ by

$$\mathcal{C}(m) = \big\{a\in S(m;[0,1]\times(0,1]) \;\big|\; a(z;\mu,h) = \tilde a(\varphi_\mu(z);\mu,h) \text{ with } \tilde a\in S(m;[0,1]\times(0,1])\big\}. \tag{A-4}$$

Theorem A.3. *Let P and Q be h-pseudodifferential operators on $\mathbb{R}^d$, with $P\in\mathrm{Op}_h^w(\mathcal{C}(m))$ and $Q\in\mathrm{Op}_h^w(\mathcal{C}(m'))$, where m and m' are order functions such that*

$$m\geqslant 1, \quad m' = \mathcal{O}(1), \quad mm' = \mathcal{O}(1).$$

Then $e^{\frac{i}{h}Q}Pe^{-\frac{i}{h}Q} = \widetilde{P} + R$, where $\widetilde{P}\in\mathrm{Op}_h^w(\mathcal{C}(m))$, $R\in h^\infty\,\mathrm{Op}_h^w(S(1))$, and with $\widetilde{P} - \mathrm{Op}_h^w(p\circ\kappa)\in h\,\mathrm{Op}_h^w(\mathcal{C}(1))$.

Proof. Since φ_μ is linear, one can see (using, for instance, [Zworski 2012, Theorem 4.17]) that $\mathcal{C}$ is stable under the formal Moyal product, i.e., for all order functions m_1 and m_2, we have

$$(\mathcal{C}(m_1)) \star (\mathcal{C}(m_2)) \subset \mathcal{C}(m_1 m_2) + h^\infty S(1).$$

Let κ be the canonical transformation associated with Q. Then, since $m \geqslant 1$, we have $p \circ \kappa \in \mathcal{C}(m)$; indeed, if we write the Hamiltonian flow of Q in terms of the variable $\tilde{z} = \varphi_\mu(z)$, we see from the linearity of φ_μ that the components of the transformed vector field belong to $\mathcal{C}(m')$. Therefore $\varphi_\mu \circ \kappa$ is of the form $\tilde{\kappa}_\mu \circ \varphi_\mu$ for some diffeomorphism $\tilde{\kappa}_\mu$ depending smoothly on μ.

Therefore, both terms in (A-2) belong to $\mathrm{Op}_h^w(\mathcal{C}(1))$. Applying this argument inductively in (A-3), we may write, for any $k > 0$,

$$e^{\frac{i}{h}Q} P e^{-\frac{i}{h}Q} - \mathrm{Op}_h^w(p \circ \kappa) - (h\widetilde{P}_1 + h^2 \widetilde{P}_2 + \cdots + h^k \widetilde{P}_k) \in h^{k+1}\,\mathrm{Op}_h^w(S(1)),$$

with $\widetilde{P}_j \in \mathrm{Op}_h^w(\mathcal{C}(1))$. By a Borel summation in h, parametrized by $\tilde{z} = \varphi_\mu(z)$, we can find a symbol $\widehat{P} \in \mathrm{Op}_h^w(\mathcal{C}(1))$ such that we have the asymptotic expansion in $\mathrm{Op}_h^w(S(1))$

$$\widehat{P} \sim h\widetilde{P}_1 + h^2 \widetilde{P}_2 + \cdots .$$

We conclude by letting $\widetilde{P} = \mathrm{Op}_h^w(p \circ \kappa) + \widehat{P}$. □

We will also need to examine how the Egorov theorem behaves with respect to taking formal power series of symbols. For this, it is convenient to introduce a filtration of $S(m)$.

Theorem A.4. *Let m be an order function on $\mathbb{R}^{2d}$, and let $(\mathcal{O}_j)_{j\in\mathbb{N}}$ be a filtration of $S(m)$, i.e.,*

$$\mathcal{O}_0 = S(m), \quad \mathcal{O}_{j+1} \subset \mathcal{O}_j.$$

Let $P = \mathrm{Op}_h^w\, p$ and $Q = \mathrm{Op}_h^w\, q$ be h-pseudodifferential operators on $\mathbb{R}^d$, with $p \in S(m)$ and $q \in S(m')$, where m' is an order function such that m' and $m m'$ are bounded.

Assume that

$$\frac{i}{h}\,\mathrm{ad}_q(\mathcal{O}_j) \subset \mathcal{O}_{j+1} \quad \forall j \geqslant 0. \tag{A-5}$$

Then for any $k \geqslant 0$, the Weyl symbol of the pseudodifferential operator

$$e^{\frac{i}{h}Q} P e^{-\frac{i}{h}Q} - \sum_{j=0}^{k} \frac{1}{j!}\big(\tfrac{i}{h}\,\mathrm{ad}_Q\big)^j P$$

belongs to $\mathrm{Op}_h^w(\mathcal{O}_{k+1})$. In other words, the series of $\exp\big(\frac{i}{h}\,\mathrm{ad}_Q\big)P$ converges to $e^{\frac{i}{h}Q} P e^{-\frac{i}{h}Q}$ for the filtration $(\mathcal{O}_j)_{j\in\mathbb{N}}$.

Proof. By the Taylor formula, we can write

$$e^{\frac{i}{h}Q} P e^{-\frac{i}{h}Q} = \sum_{j=0}^{k} \frac{1}{j!}(\mathrm{ad}_{ih^{-1}Q})^j P + \frac{1}{k!}(\mathrm{ad}_{ih^{-1}Q})^{k+1} \int_0^1 (1-t)^k e^{\frac{it}{h}Q} P e^{-\frac{it}{h}Q}\, dt.$$

By Theorem A.2, we see that the integral belongs to $\mathrm{Op}_h^w(S(m)) = \mathrm{Op}_h^w(\mathcal{O}_0)$. Therefore, by assumption (A-5), the remainder in the Taylor formula lies in $\mathrm{Op}_h^w(\mathcal{O}_{k+1})$. □

Acknowledgments

The authors would like to thank Yves Colin de Verdière for stimulating discussions. This work was partially supported by the ANR (Agence Nationale de la Recherche), project NOSEVOL n° ANR-11-BS01-0019 and by the Centre Henri Lebesgue (program "Investissements d'avenir" — n° ANR-11-LABX-0020-01). Kordyukov was partially supported by the Russian Foundation of Basic Research, project 13-01-91052-NCNI-a. During the completion of this work, Helffer was Simons foundation visiting fellow at the Isaac Newton Institute in Cambridge.

References

[Avron et al. 1978] J. Avron, I. Herbst, and B. Simon, "Schrödinger operators with magnetic fields, I: General interactions", *Duke Math. J.* **45**:4 (1978), 847–883. MR 518109 Zbl 0399.35029

[Benettin and Sempio 1994] G. Benettin and P. Sempio, "Adiabatic invariants and trapping of a point charge in a strong nonuniform magnetic field", *Nonlinearity* **7**:1 (1994), 281–303. MR 1260143 Zbl 0856.70010

[Bonnaillie-Noël et al. 2016] V. Bonnaillie-Noël, F. Hérau, and N. Raymond, "Magnetic WKB constructions", *Arch. Ration. Mech. Anal.* **221**:2 (2016), 817–891. MR 3488538 Zbl 1338.35379

[Born and Oppenheimer 1927] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln", *Ann. Phys.* **84**:20 (1927), 457–484. JFM 53.0845.04

[Charles and Vũ Ngọc 2008] L. Charles and S. Vũ Ngọc, "Spectral asymptotics via the semiclassical Birkhoff normal form", *Duke Math. J.* **143**:3 (2008), 463–511. MR 2423760 Zbl 1154.58015

[Cheverry 2015] C. Cheverry, "Can one hear whistler waves?", *Comm. Math. Phys.* **338**:2 (2015), 641–703. MR 3351054 Zbl 1333.35290

[Colin de Verdière 1996] Y. Colin de Verdière, "La méthode de moyennisation en mécanique semi-classique", pp. exposé no. 5 in *Journées "Équations aux Dérivées Partielles"* (Saint-Jean-de-Monts, 1996), École Polytech., Palaiseau, 1996. MR 1417729 Zbl 0898.35080

[Cycon et al. 1987] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger operators with application to quantum mechanics and global geometry*, Springer, 1987. MR 883643 Zbl 0619.47005

[Dimassi and Sjöstrand 1999] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit*, London Mathematical Society Lecture Note Series **268**, Cambridge University Press, 1999. MR 1735654 Zbl 0926.35002

[Dufresnoy 1983] A. Dufresnoy, "Un exemple de champ magnétique dans $\mathbb{R}^{\nu}$", *Duke Math. J.* **50**:3 (1983), 729–734. MR 714827 Zbl 0532.35021

[Fournais and Helffer 2010] S. Fournais and B. Helffer, *Spectral methods in surface superconductivity*, Progress in Nonlinear Differential Equations and their Applications **77**, Birkhäuser, Boston, 2010. MR 2662319 Zbl 1256.35001

[Helffer and Kordyukov 2011] B. Helffer and Y. A. Kordyukov, "Semiclassical spectral asymptotics for a two-dimensional magnetic Schrödinger operator: the case of discrete wells", pp. 55–78 in *Spectral theory and geometric analysis*, Contemp. Math. **535**, Amer. Math. Soc., Providence, RI, 2011. MR 2560751 Zbl 1218.58017

[Helffer and Kordyukov 2013] B. Helffer and Y. A. Kordyukov, "Eigenvalue estimates for a three-dimensional magnetic Schrödinger operator", *Asymptot. Anal.* **82**:1–2 (2013), 65–89. MR 3088341 Zbl 1325.35123

[Helffer and Kordyukov 2014] B. Helffer and Y. A. Kordyukov, "Semiclassical spectral asymptotics for a magnetic Schrödinger operator with non-vanishing magnetic field", pp. 259–278 in *Geometric Methods in Physics* (Bialowieza, Poland, 2013), edited by P. Kielanowski et al., Birkhäuser, Basel, 2014. Zbl 1317.81101

[Helffer and Kordyukov 2015] B. Helffer and Y. A. Kordyukov, "Accurate semiclassical spectral asymptotics for a two-dimensional magnetic Schrödinger operator", *Ann. Henri Poincaré* **16**:7 (2015), 1651–1688. MR 3356100 Zbl 1338.81187

[Helffer and Mohamed 1996] B. Helffer and A. Mohamed, "Semiclassical analysis for the ground state energy of a Schrödinger operator with magnetic wells", *J. Funct. Anal.* **138**:1 (1996), 40–81. MR 1391630 Zbl 0851.58046

[Helffer and Robert 1984] B. Helffer and D. Robert, "Puits de potentiel généralisés et asymptotique semi-classique", *Ann. Inst. H. Poincaré Phys. Théor.* **41**:3 (1984), 291–331. MR 776281 Zbl 0565.35082

[Helffer and Sjöstrand 1989] B. Helffer and J. Sjöstrand, *Semiclassical analysis for Harper's equation, III: Cantor structure of the spectrum*, Mém. Soc. Math. France (N.S.) **39**, Soc. Math. France, Paris, 1989. MR 1041490

[Ivrii 1998] V. Ivrii, *Microlocal analysis and precise spectral asymptotics*, Springer, 1998. MR 1631419 Zbl 0906.35003

[Martinez 2007] A. Martinez, "A general effective Hamiltonian method", *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **18**:3 (2007), 269–277. MR 2318820 Zbl 1139.35383

[Persson 1960] A. Persson, "Bounds for the discrete part of the spectrum of a semi-bounded Schrödinger operator", *Math. Scand.* **8** (1960), 143–153. MR 0133586 Zbl 0145.14901

[Popoff and Raymond 2013] N. Popoff and N. Raymond, "When the 3D magnetic Laplacian meets a curved edge in the semiclassical limit", *SIAM J. Math. Anal.* **45**:4 (2013), 2354–2395. MR 3085117 Zbl 06219425

[Raymond 2012] N. Raymond, "Semiclassical 3D Neumann Laplacian with variable magnetic field: a toy model", *Comm. Partial Differential Equations* **37**:9 (2012), 1528–1552. MR 2969489 Zbl 1254.35167

[Raymond 2016] N. Raymond, "Bound states of the magnetic Schrödinger operator", preprint, 2016, available at https://blogperso.univ-rennes1.fr/nicolas.raymond/public/Magnetic-Book.pdf.

[Raymond and Vũ Ngọc 2015] N. Raymond and S. Vũ Ngọc, "Geometry and spectrum in 2D magnetic wells", *Ann. Inst. Fourier* (*Grenoble*) **65**:1 (2015), 137–169. MR 3449150 Zbl 1327.81207

[Robert 1987] D. Robert, *Autour de l'approximation semi-classique*, Progress in Mathematics **68**, Birkhäuser, Boston, 1987. MR 897108 Zbl 0621.35001

[Vũ Ngọc 2006] S. Vũ Ngọc, *Systèmes intégrables semi-classiques: du local au global*, Panoramas et Synthèses **22**, Soc. Math. France, Paris, 2006. MR 2331010 Zbl 1118.37001

[Vũ Ngọc 2009] S. Vũ Ngọc, "Quantum Birkhoff normal forms and semiclassical analysis", pp. 99–116 in *Noncommutativity and singularities*, Adv. Stud. Pure Math. **55**, Math. Soc. Japan, Tokyo, 2009. MR 2463493 Zbl 1185.58012

[Weinstein 1971] A. Weinstein, "Symplectic manifolds and their Lagrangian submanifolds", *Advances in Math.* **6** (1971), 329–346. MR 0286137 Zbl 0213.48203

[Weinstein 1975] A. Weinstein, "Fourier integral operators, quantization, and the spectra of Riemannian manifolds", pp. 289–298 in *Géométrie symplectique et physique mathématique* (Aix-en-Provence, 1974), Éditions Centre Nat. Recherche Sci., Paris, 1975. MR 0650990 Zbl 0327.58013

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, American Mathematical Society, Providence, RI, 2012. MR 2952218 Zbl 1252.58001

Received 29 Sep 2015. Revised 6 Jun 2016. Accepted 9 Jul 2016.

BERNARD HELFFER: bernard.helffer@math.u-psud.fr
Département de Mathématiques, Bât. 425, Université Paris-Sud et CNRS, 91405 Orsay Cedex, France

and

Laboratoire de Mathématiques Jean Leray, Université de Nantes, 2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3, France

YURI KORDYUKOV: yurikor@matem.anrb.ru
Institute of Mathematics, Russian Academy of Sciences, 112 Chernyshevsky str., 450008 Ufa, Russia

NICOLAS RAYMOND: nicolas.raymond@univ-rennes1.fr
Institut de Recherches Mathématiques de Rennes (UMR 6625), Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France

SAN VŨ NGỌC: san.vu-ngoc@univ-rennes1.fr
Institut de Recherches Mathématiques de Rennes (UMR 6625), Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France

and

Institut Universitaire de France, 1 rue Descartes, 75231 Paris Cedex 5, France

ANALYSIS AND PDE
Vol. 9, No. 7, 2016
dx.doi.org/10.2140/apde.2016.9.1609

AN ANALYTICAL AND NUMERICAL STUDY OF STEADY PATCHES IN THE DISC

FRANCISCO DE LA HOZ, ZINEB HASSAINIA, TAOUFIK HMIDI AND JOAN MATEU

We prove the existence of m-fold rotating patches for the Euler equations in the disc, for the simply connected and doubly connected cases. Compared to the planar case, the rigid boundary introduces rich dynamics for the lowest symmetries $m = 1$ and $m = 2$. We also discuss some numerical experiments highlighting the interaction between the boundary of the patch and the rigid one.

1. Introduction

In this paper, we shall discuss some aspects of the vortex motion for the Euler system in the unit disc $\mathbb{D}$ of the Euclidean space $\mathbb{R}^2$. That system is described by the equations

$$\begin{cases} \partial_t v + v \cdot \nabla v + \nabla p = 0, & (t,x) \in \mathbb{R}_+ \times \mathbb{D}, \\ \operatorname{div} v = 0, & \\ v \cdot \nu = 0 & \text{on } \partial\mathbb{D}, \\ v|_{t=0} = v_0. & \end{cases} \tag{1}$$

MSC2010: primary 35Q35, 37G40, 35Q31; secondary 76B47.
Keywords: Euler equations, V-states, bifurcation.

Here, $v=(v^1,v^2)$ is the velocity field, and the pressure p is a scalar potential that can be related to the velocity using the incompressibility condition. The boundary equation means that there is no matter flow through the rigid boundary $\partial\mathbb{D}=\mathbb{T}$; the vector ν is the outer unitary vector orthogonal to the boundary. The main feature of two-dimensional flows is that they can be illustrated through their vorticity structure; this can be identified with the scalar function $\omega=\partial_1 v_2-\partial_2 v_1$, and its evolution is governed by the nonlinear transport equation

$$\partial_t\omega+v\cdot\nabla\omega=0. \tag{2}$$

To recover the velocity from the vorticity, we use the stream function Ψ, which is defined as the unique solution of the Dirichlet problem on the unit disc:

$$\begin{cases} \Delta\Psi=\omega,\\ \psi|_{\partial\mathbb{D}}=0.\end{cases}$$

Therefore, the velocity is given by

$$v=\nabla^\perp\Psi,\quad \nabla^\perp=(-\partial_2,\partial_1).$$

By using the Green function of the unit disc, we get the expression

$$\Psi(z)=\frac{1}{4\pi}\int_{\mathbb{D}}\log\left|\frac{z-\xi}{1-z\bar{\xi}}\right|^2\omega(\xi)\,dA(\xi), \tag{3}$$

with dA being the planar Lebesgue measure. In what follows, we shall identify the Euclidean and the complex planes, so the velocity field is identified with the complex function

$$v(z)=v_1(x_1,x_2)+iv_2(x_1,x_2),\quad z=x_1+ix_2.$$

Therefore, we get the compact formula

$$\begin{aligned} v(t,z)&=2i\partial_{\bar z}\Psi(t,z)\\ &=\frac{i}{2\pi}\int_{\mathbb{D}}\frac{|\xi|^2-1}{(\bar z-\bar\xi)(\xi\bar z-1)}\omega(t,\xi)\,dA(\xi)\\ &=\frac{i}{2\pi}\int_{\mathbb{D}}\frac{\omega(t,\xi)}{\bar z-\bar\xi}\,dA(\xi)+\frac{i}{2\pi}\int_{\mathbb{D}}\frac{\xi}{1-\xi\bar z}\omega(t,\xi)\,dA(\xi). \end{aligned} \tag{4}$$

We recognize in the first part of the last formula the structure of the Biot–Savart law in the plane $\mathbb{R}^2$, which is given by

$$v(t,z)=\frac{i}{2\pi}\int_{\mathbb{C}}\frac{\omega(t,\xi)}{\bar z-\bar\xi}\,dA(\xi),\quad z\in\mathbb{C}. \tag{5}$$

The second term of (4) is absent in the planar case. It describes the contribution of the rigid boundary $\mathbb{T}$, and our main task is to investigate the boundary effects on the dynamics of special long-lived vortex structures. Before going further into details, we recall first that, from the equivalent formulation (2)–(4) of the Euler system (1), Yudovich [1963] was able to construct a unique global solution in the weak sense, provided that the initial vorticity ω_0 is compactly supported and bounded. This result is very important because it allows one to deal rigorously with vortex patches, which are vortices uniformly distributed in a bounded region D, i.e., $\omega_0 = \chi_D$. These structures are preserved by the evolution, and at each time t, the vorticity is given by χ_{D_t}, with $D_t = \psi(t, D)$ being the image of D by the flow. As we shall see later in (16), the contour dynamics equation of the boundary ∂D_t is described by the following nonlinear integral equation. Let $\gamma_t : \mathbb{T} \to \partial D_t$ be the Lagrangian parametrization of the boundary; then

$$\partial_t \gamma_t = -\frac{1}{2\pi} \int_{\partial D_t} \log|\gamma_t - \xi| \, d\xi + \frac{1}{4\pi} \int_{\partial D_t} \frac{|\xi|^2}{1 - \overline{\gamma_t}\xi} \, d\xi.$$

We point out that, when the initial boundary is smooth enough, roughly speaking more regular than C^1, then the regularity is propagated for long times without any loss. This was first achieved by Chemin [1998] in the plane and extended in bounded domains by Depauw [1999]. Note also that we can find in [Bertozzi and Constantin 1993] another proof of Chemin's result. It appears that the boundary dynamics of the patch is very complicate to tackle and, to our knowledge, the only known explicit example is the stationary one given by a small disc centered at the origin. Even though explicit solutions form a poor class, one can try to find implicit patches with prescribed dynamics, such as rotating patches, also known as V-states. These patches are subject to perpetual rotation around some fixed point that we can assume to be the origin and with uniform angular velocity Ω; this means that $D_t = e^{it\Omega} D$. We shall see in Section 2.3 that the V-states equation, when D is symmetric with respect to the real axis, is given by

$$\operatorname{Re}\left\{\left(2\Omega\bar{z} + \fint_{\Gamma} \frac{\bar{z} - \bar{\xi}}{z - \xi} \, d\xi - \fint_{\Gamma} \frac{|\xi|^2}{1 - z\xi} \, d\xi\right) z'\right\} = 0, \quad z \in \Gamma \triangleq \partial D, \tag{6}$$

with z' being a tangent vector to the boundary ∂D_0 at the point z; note that we have used the notation $\fint_{\Gamma} \equiv (1/2i\pi) \int_{\Gamma}$. In the flat case, the boundary equation (6) becomes

$$\operatorname{Re}\left\{\left(2\Omega\bar{z} + \fint_{\Gamma} \frac{\bar{z} - \bar{\xi}}{z - \xi} \, d\xi\right) z'\right\} = 0, \quad z \in \Gamma. \tag{7}$$

Note that circular patches are stationary solutions for (7); however, elliptical vortex patches perform a steady rotation about their centers without changing shape. This latter fact was discovered by Kirchhoff [1876], who proved that, when D is an ellipse centered at zero, $D_t = e^{it\Omega} D$, where the angular velocity Ω is determined by the semiaxes a and b through the formula $\Omega = ab/(a+b)^2$. These ellipses are often referred to in the literature as the Kirchhoff elliptic vortices; see for instance [Majda and Bertozzi 2002, p. 304] or [Lamb 1945, p. 232].

One century later, several examples of rotating patches were obtained by Deem and Zabusky [1978], using contour dynamics simulations. Burbea [1982] gave an analytical proof and showed the existence of V-states with m-fold symmetry for each integer $m \geq 2$. In this countable family, the case $m = 2$ corresponds to the Kirchhoff elliptic vortices. Burbea's approach consists of using complex analysis tools, combined with bifurcation theory. It should be noted that, from this standpoint, the rotating patches are arranged in a collection of countable curves bifurcating from Rankine vortices (trivial disc solution) at the discrete angular velocities set $\{(m-1)/2m : m \geq 2\}$. The numerical analysis of limiting V-states which are the ends of each branch is done in [Overman 1986; Wu et al. 1984] and reveals interesting behavior: the boundary develops corners at right angles. Recently, the C^∞ regularity and the convexity of the patches near the trivial solutions have been investigated in [Hmidi et al. 2013]. More recently, this result has been improved by Castro, Córdoba and Gómez-Serrano [Castro et al. 2016b], who showed the analyticity of the V-states close to the disc. We point out that similar research has been carried out in the past few years for more singular nonlinear transport equations arising in geophysical flows, such as the surface quasigeostrophic equations or the quasigeostrophic shallow-water equations; see for instance [Castro et al. 2016a; 2016b; Hassainia and Hmidi 2015; Płotka and Dritschel 2012]. It should be noted that the angular velocities of the bifurcating V-states for (7) are contained in the interval $]0, \frac{1}{2}[$. However, it is not clear whether we can find a V-state when Ω does not lie in this range. Fraenkel [2000] proved, always in the simply connected case, that the solutions associated with $\Omega = 0$ are trivial and reduced to Rankine patches. This was established by using the moving plane method, which seems to be flexible and has been recently adapted in [Hmidi 2015] to $\Omega < 0$ but with a convexity restriction. The case $\Omega = \frac{1}{2}$ was also solved in that paper, using the maximum principle for harmonic functions.

Another related subject is to see whether a second bifurcation occurs at the branches discovered by Deem and Zabusky. This has been explored for the branch of the ellipses corresponding to $m = 2$. Kamm [1987] gave numerical evidence of the existence of some branches bifurcating from the ellipses; see also [Saffman 1992]. In [Luzzatto-Fegiz and Williamson 2010], one can find more details about the diagram for the first bifurcations and some illustrations of the limiting V-states. The proof of the existence and analyticity of the boundary has been recently investigated in [Castro et al. 2016b; Hmidi and Mateu 2016]. Another interesting topic which has been studied since the pioneering work of Love [1893] is the linear and nonlinear stability of the m-folds. For the ellipses, we mention [Guo et al. 2004; Tang 1987], and for the general case of the m-fold symmetric V-states, we refer to [Burbea and Landau 1982; Wan 1986]. For further numerical discussions, see also [Cerretelli and Williamson 2003; Dritschel 1986; Mitchell and Rossi 2008]. Recently [Hmidi et al. 2015; de la Hoz et al. 2016b] have shown a special interest in the study of doubly connected V-states which are bounded patches and delimited by two disjoint Jordan curves. For example, an annulus is doubly connected, and by rotation invariance, it is a stationary V-state. No other explicit doubly connected V-state is known in the literature. In [Hmidi et al. 2015], a full characterization of the V-states (with nonzero magnitude in the interior domain) with at least one elliptical interface has been achieved, complementing the results of Polvani and Flierl [1986]. As a byproduct, it is shown that the domain between two ellipses is a V-state only if it is an annulus. The proof of existence of nonradial doubly connected V-states has been achieved very recently in [de la Hoz et al. 2016b] by using bifurcation theory. More precisely, we get the following result. Let $0 < b < 1$ and $m \geq 3$, such that

$$1 + b^m - \frac{1-b^2}{2} m < 0.$$

Then there exist two curves of m-fold symmetric doubly connected V-states bifurcating from the annulus $\{z \in \mathbb{C} : b < |z| < 1\}$ at each of the angular velocities

$$\Omega_m^{\pm} = \frac{1-b^2}{4} \pm \frac{1}{2m} \sqrt{\left(\frac{m(1-b^2)}{2} - 1\right)^2 - b^{2m}}. \tag{8}$$

The main goal of the current paper is to explore the existence of rotating patches (6) for Euler equations posed on the unit disc $\mathbb{D}$. We shall focus on the simply connected and doubly connected cases and study the influence of the rigid boundary on these structures. Before stating our main results, we define the set $\mathbb{D}_b = \{z \in \mathbb{C} : |z| < b\}$. Our first result dealing with the simply connected V-states is:

Theorem 1. *Let $b \in]0, 1[$ and m be a positive integer. Then there exists a family of m-fold symmetric V-states $(V_m)_{m\geq 1}$ for (6) bifurcating from the trivial solution $\omega_0 = \chi_{\mathbb{D}_b}$ at the angular velocity*

$$\Omega_m \triangleq \frac{m-1+b^{2m}}{2m}.$$

The proof of this theorem is done in the spirit of [Burbea 1980; de la Hoz et al. 2016b], using the conformal mapping parametrization $\phi : \mathbb{T} \to \partial D$ of the V-states, combined with bifurcation theory. As we shall see later in (17), the function ϕ satisfies the following nonlinear equation, for all $w \in \mathbb{T}$:

$$\operatorname{Im}\left\{\left[2\Omega\overline{\phi(w)} + \fint_{\mathbb{T}} \frac{\overline{\phi(w)} - \overline{\phi(\tau)}}{\phi(w) - \phi(\tau)} \phi'(\tau)\, d\tau - \fint_{\mathbb{T}} \frac{|\phi(\tau)|^2 \phi'(\tau)}{1 - \phi(w)\phi(\tau)}\, d\tau\right] w\phi'(w)\right\} = 0.$$

Denote by $F(\Omega, \phi)$ the term in the left-hand side of the preceding equality. Then the linearized operator around the trivial solution $\phi = b\,\mathrm{Id}$ can be explicitly computed and is given by the following Fourier multiplier: for $h(w) = \sum_{n\in\mathbb{N}} a_n \overline{w}^n$,

$$\partial_\phi F(\Omega, b\,\mathrm{Id})h(w) = b \sum_{n\geq 1} n\left(\frac{n-1+b^{2n}}{n} - 2\Omega\right) a_{n-1} e_n, \quad e_n = \frac{1}{2i}(\overline{w}^n - w^n).$$

Therefore, the *nonlinear eigenvalues* leading to nontrivial kernels of dimension 1 are explicitly described by the quantity Ω_m appearing in Theorem 1. Later on, we check that all the assumptions of the Crandall–Rabinowitz theorem stated in Section 2.2 are satisfied, and our result follows easily. In Section 5.1, we implement some numerical experiments concerning the limiting V-states. We observe two regimes depending on the size of b: b small and b close to 1. In the first case, as expected, corners do appear as in the planar case. However, for b close to 1, the effect of the rigid boundary is not negligible. We observe that the limiting V-states are tangentially touching the unit circle; see Figure 5. Some remarks are in order.

Remark 2. For the Euler equations in the plane, there are no curves of 1-fold V-states close to Rankine vortices. However, we deduce from our main theorem that this mode appears for spherical bounded domains. Its existence is the result of the interaction between the patch and the rigid boundary $\mathbb{T}$. Moreover, according to the numerical experiments, these V-states are not necessarily centered at the origin, and this fact is completely new. For the symmetry $m \geq 2$, all the discovered V-states are necessarily centered at zero because they have at least two axes of symmetry passing through zero.

Remark 3. By a scaling argument, when the domain of the fluid is the ball $B(0, R)$, with $R > 1$, then from the preceding theorem, the bifurcation from the unit disc occurs at the angular velocities

$$\Omega_{m,R} \triangleq \frac{m-1+R^{-2m}}{2m}.$$

Therefore, we obtain Burbea's result [1980] by letting R tend to $+\infty$.

Remark 4. From the numerical experiments done in [de la Hoz et al. 2016b], we note that, in the plane, the bifurcation is pitchfork and occurs to the left of Ω_m. Furthermore, the branches of bifurcation are "monotonic" with respect to the angular velocity. In particular, this means that, for each value of Ω, we have at most only one V-state with that angular velocity. This behavior is no longer true in the disc as will be discussed later in the numerical experiments; see Figure 3.

Remark 5. Due to the boundary effects, the ellipses are no longer solutions for the rotating patch equation (6). Whether explicit solutions can be found for this model is an interesting problem. However, we believe that the conformal mapping of any nontrivial V-state has a necessary infinite expansion. Note that Burbea [1982] proved that, in the planar case when the conformal mapping associated to the V-state has a finite expansion, it is necessarily an ellipse. His approach is based on Faber polynomials, and this could give insight to solving the same problem in the disc.

The second part of this paper deals with the existence of doubly connected V-states for the system (1), governed by (6). Note that the annular patches centered at zero, which are given by

$$\mathbb{A}_{b_1,b_2} = \{z \in \mathbb{C} : b_1 < |z| < b_2\}, \quad b_1 < b_2 < 1,$$

are indeed stationary solutions. Our main task is to study the bifurcation of the V-states from these trivial solutions in the spirit of the recent works [de la Hoz et al. 2016a; 2016b]. We shall first start by studying the existence with the symmetry $m \geq 2$, followed by the special case $m = 1$.

Theorem 6. *Let $0 < b_2 < b_1 < 1$, and set $b \triangleq b_2/b_1$. Let $m \geq 2$, such that*

$$m > \frac{2 + 2b^m - (b_1^m + b_2^m)^2}{1 - b^2}.$$

Then there exist two curves of m-fold symmetric doubly connected V-states bifurcating from the annulus $\mathbb{A}_{b_1,b_2}$ at the angular velocities

$$\Omega_m^{\pm} = \frac{1-b^2}{4} + \frac{b_1^{2m} - b_2^{2m}}{4m} \pm \tfrac{1}{2}\sqrt{\Delta_m},$$

with

$$\Delta_m = \left(\frac{1-b^2}{2} - \frac{2 - b_1^{2m} - b_2^{2m}}{2m}\right)^2 - b^{2m}\left(\frac{1-b_1^{2m}}{m}\right)^2.$$

Before outlining the ideas of the proof, a few remarks are necessary.

Remark 7. As was discussed in Remark 3, one can use a scaling argument and obtain the result previously established in [de la Hoz et al. 2016b] for the planar case. Indeed, when the domain of the fluid is the ball $B(0, R)$, with $R > 1$, then the bifurcation from the annulus $\mathbb{A}_{b,1}$ amounts to making the changes $b_1 = 1/R$ and $b_2 = b/R$ in Theorem 6. Thus, by letting R tend to infinity, we get exactly the nonlinear eigenvalues of the Euler equations in the plane (8).

Remark 8. Unlike in the plane, where the frequency m is assumed to be larger than 3, we can reach $m = 2$ in the case of the disc. This can be checked for b_2 small with respect to b_1. This illustrates once again the interaction between the rigid boundary and the V-states.

Now we shall sketch the proof of Theorem 6, which follows along the lines of [de la Hoz et al. 2016b] and stems from bifurcation theory. The first step is to write down the analytical equations of the boundaries of the V-states. This can be done for example through the conformal parametrization of the domains D_1 and D_2, which are close to the discs $b_1\mathbb{D}$ and $b_2\mathbb{D}$, respectively. Set $\phi_j : \mathbb{D}^c \to D_j^c$, the conformal mappings which have the expansions,

$$\text{for all } |w| \geq 1, \quad \phi_1(w) = b_1 w + \sum_{n\in\mathbb{N}} \frac{a_{1,n}}{w^n}, \quad \phi_2(w) = b_2 w + \sum_{n\in\mathbb{N}} \frac{a_{2,n}}{w^n}.$$

In addition, we assume that the Fourier coefficients are real, which means that we are looking only for V-states that are symmetric with respect to the real axis. As we shall see later in Section 4.1, the conformal mappings are subject to two coupled nonlinear equations defined as follows: for $j \in \{1, 2\}$ and $w \in \mathbb{T}$,

$$F_j(\lambda, \phi_1, \phi_2)(w) \triangleq \operatorname{Im}\big\{\big((1-\lambda)\overline{\phi_j(w)} + I(\phi_j(w)) - J(\phi_j(w))\big) w \phi_j'(w)\big\} = 0,$$

with

$$\begin{aligned} I(z) &= \fint_{\mathbb{T}} \frac{\bar{z} - \overline{\phi_1(\xi)}}{z - \phi_1(\xi)} \phi_1'(\xi)\, d\xi - \fint_{\mathbb{T}} \frac{\bar{z} - \overline{\phi_2(\xi)}}{z - \phi_2(\xi)} \phi_2'(\xi)\, d\xi, \\ J(z) &= \fint_{\mathbb{T}} \frac{|\phi_1(\xi)|^2}{1 - z\phi_1(\xi)} \phi_1'(\xi)\, d\xi - \fint_{\mathbb{T}} \frac{|\phi_2(\xi)|^2}{1 - z\phi_2(\xi)} \phi_2'(\xi)\, d\xi, \end{aligned} \qquad \lambda \triangleq 1 - 2\Omega.$$

In order to apply bifurcation theory, we should understand the structure of the linearized operator around the trivial solution $(\phi_1, \phi_2) = (b_1 \operatorname{Id}, b_2 \operatorname{Id})$ corresponding to the annulus with radii b_1 and b_2 and identify the range of Ω where this operator has a one-dimensional kernel. The computations of the linear operator $DF(\Omega, b_1 \operatorname{Id}, b_2 \operatorname{Id})$ with $F = (F_1, F_2)$ in terms of the Fourier coefficients are fairly lengthy, and we find that it acts as a Fourier multiplier matrix. More precisely, for

$$h_1(w) = \sum_{n\geq 1} \frac{a_{1,n}}{w^n}, \qquad h_2(w) = \sum_{n\geq 1} \frac{a_{2,n}}{w^n},$$

we obtain the formula

$$DF(\lambda, b_1 \operatorname{Id}, b_2 \operatorname{Id})(h_1, h_2) = \sum_{n\geq 1} M_n(\lambda) \begin{pmatrix} a_{1,n-1} \\ a_{2,n-1} \end{pmatrix} e_n, \quad e_n(w) \triangleq \frac{1}{2i}(\overline{w}^n - w^n),$$

where the matrix M_n is given by

$$M_n(\lambda) = \begin{pmatrix} b_1[n\lambda - 1 + b_1^{2n} - n(b_2/b_1)^2] & b_2[(b_2/b_1)^n - (b_1 b_2)^n] \\ -b_1[(b_2/b_1)^n - (b_1 b_2)^n] & b_2[n\lambda - n + 1 - b_2^{2n}] \end{pmatrix}.$$

Therefore, the values of Ω associated with nontrivial kernels are the solutions of a second-degree polynomial in λ,

$$P_n(\lambda) \triangleq \det M_n(\lambda) = 0. \tag{9}$$

The polynomial P_n has real roots when the discriminant $\Delta_n(\alpha, b)$ introduced in Theorem 6 is positive. The calculation of the dimension of the kernel is significantly more complicated than the cases considered before in [Burbea 1980; de la Hoz et al. 2016b]. The matter reduces to counting, for a given λ, the discrete set

$$\{n \geq 2 : P_n(\lambda) = 0\}.$$

Note that, in [Burbea 1980; de la Hoz et al. 2016b], this set has only one element, and therefore, the kernel is one-dimensional. This follows from the monotonicity of the roots of P_n with respect to n. In the current situation, we get similar results but with a more refined analysis.

Now we shall move on to the existence of 1-fold symmetries, which are completely absent in the plane. The study in the general case is slightly subtler, and we have only carried out partial results, so some other cases are left open and deserve to be explored. Before stating our main result, we need to do some preparation. As we shall see in Section 4.4.3, the equation $P_1(\lambda)=0$ admits exactly two solutions

$$\lambda_1^- = (b_2/b_1)^2 \quad \text{or} \quad \lambda_1^+ = 1+b_2^2-b_1^2.$$

Similarly to the planar case [de la Hoz et al. 2016b], there is no hope of bifurcating from the first eigenvalue λ_1^- because the range of the linearized operator around the trivial solution has an infinite codimension, and thus, the Crandall–Rabinowitz theorem stated in Section 2.2 is useless. However, for the second eigenvalue λ_1^+, the range is at most of codimension 2, and in order to bifurcate, we should avoid a special set of b_1 and b_2 that we shall describe now. Fix b_1 in $]0,1[$, and set

$$\mathscr{E}_{b_1} \triangleq \{b_2 \in]0, b_1[: \text{there exists } n \geq 2 \text{ such that } P_n(1+b_2^2-b_1^2)=0\},$$

where P_n is defined in (9). As we shall see in Proposition 20, this set is countable and composed of a strictly increasing sequence $(x_m)_{m\geq 1}$ converging to b_1. Now we state our result.

Theorem 9. *Given $b_1 \in]0,1[$, then for any $b_2 \notin \mathscr{E}_{b_1}$, there exists a curve of nontrivial 1-fold doubly connected V-states bifurcating from the annulus $\mathbb{A}_{b_1,b_2}$ at the angular velocity*

$$\Omega_1 = \frac{b_1^2-b_2^2}{2}.$$

The proof is done in the spirit of Theorem 6. When $b_2 \notin \mathscr{E}_{b_1}$, then all the conditions of the Crandall–Rabinowitz theorem are satisfied. However, when $b_2 \in \mathscr{E}_{b_1}$, then the range of the linearized operator has codimension 2. Whether the bifurcation occurs in this special case is an interesting problem which is left open here.

Notation. We need to collect some useful notation that will be frequently used along this paper. We shall use the symbol $\triangleq$ to define an object. The unit disc is denoted by $\mathbb{D}$ and its boundary, the unit circle, by $\mathbb{T}$. For a given continuous complex function $f:\mathbb{T}\to\mathbb{C}$, we set

$$\fint_{\mathbb{T}} f(\tau)\,d\tau \triangleq \frac{1}{2i\pi}\int_{\mathbb{T}} f(\tau)\,d\tau,$$

where $d\tau$ stands for complex integration.

Let X and Y be two normed spaces. We denote by $\mathscr{L}(X,Y)$ the space of all continuous linear maps $T:X\to Y$ endowed with its usual strong topology. We denote by $\operatorname{Ker} T$ and $R(T)$ the null space and the range of T, respectively. Finally, if F is a subspace of Y, then Y/F denotes the quotient space.

2. Preliminaries and background

In this introductory section, we shall collect some basic facts on Hölder spaces and bifurcation theory and shall recall how to use conformal mappings to obtain the equations of V-states.

2.1. ***Function spaces.*** In this paper as well as in the preceding ones [Hmidi et al. 2013; de la Hoz et al. 2016b], we find it more convenient to think of a 2π-periodic function $g:\mathbb{R}\to\mathbb{C}$ as a function of the complex variable $w=e^{i\theta}$. To be more precise, let $f:\mathbb{T}\to\mathbb{R}^2$ be a smooth function; then it can be assimilated to a 2π-periodic function $g:\mathbb{R}\to\mathbb{R}^2$ via the relation

$$f(w)=g(\eta),\quad w=e^{i\eta}.$$

By Fourier expansion, there exist complex numbers $(c_n)_{n\in\mathbb{Z}}$ such that

$$f(w)=\sum_{n\in\mathbb{Z}}c_n w^n,$$

and the differentiation with respect to w is understood in the complex sense. Now we shall introduce Hölder spaces on the unit circle $\mathbb{T}$.

Definition. Let $0<\gamma<1$. We denote by $C^\gamma(\mathbb{T})$ the space of continuous functions f such that

$$\|f\|_{C^\gamma(\mathbb{T})}\triangleq\|f\|_{L^\infty(\mathbb{T})}+\sup_{\tau\neq w\in\mathbb{T}}\frac{|f(\tau)-f(w)|}{|\tau-w|^\gamma}<\infty.$$

For any nonnegative integer n, the space $C^{n+\gamma}(\mathbb{T})$ stands for the set of functions f of class C^n whose n-th order derivatives are Hölder continuous with exponent γ. It is equipped with the usual norm

$$\|f\|_{C^{n+\gamma}(\mathbb{T})}\triangleq\sum_{k=0}^{n}\left\|\frac{d^kf}{d^kw}\right\|_{L^\infty(\mathbb{T})}+\left\|\frac{d^nf}{d^nw}\right\|_{C^\gamma(\mathbb{T})}.$$

Recall that the Lipschitz seminorm is defined by

$$\|f\|_{\mathrm{Lip}(\mathbb{T})}=\sup_{\tau\neq w\in\mathbb{T}}\frac{|f(\tau)-f(w)|}{|\tau-w|}.$$

Now we list some classical properties that will be useful later.

(i) For $n\in\mathbb{N}$ and $\gamma\in\,]0,1[$, the space $C^{n+\gamma}(\mathbb{T})$ is an algebra.

(ii) For $K\in L^1(\mathbb{T})$ and $f\in C^{n+\gamma}(\mathbb{T})$, we have the convolution inequality

$$\|K\star f\|_{C^{n+\gamma}(\mathbb{T})}\le\|K\|_{L^1(\mathbb{T})}\|f\|_{C^{n+\gamma}(\mathbb{T})}.$$

2.2. ***Elements of bifurcation theory.*** We shall now recall an important theorem of bifurcation theory which plays a central role in the proofs of our main results. This theorem was established by Crandall and Rabinowitz [1971]. Consider a continuous function $F:\mathbb{R}\times X\to Y$ with X and Y being two Banach spaces. Assume that $F(\lambda,0)=0$ for any λ belonging to nontrivial interval I. The Crandall–Rabinowitz theorem gives sufficient conditions for the existence of branches of nontrivial solutions to the equation $F(\lambda,x)=0$ bifurcating at some point $(\lambda_0,0)$. For more general results, we refer the reader to [Kielhöfer 2012].

Theorem 10. *Let X and Y be two Banach spaces and V a neighborhood of 0 in X, and let $F : \mathbb{R} \times V \to Y$. Set $\mathscr{L}_0 \triangleq \partial_x F(0,0)$; then the following properties are satisfied.*

(i) *$F(\lambda, 0) = 0$ for any $\lambda \in \mathbb{R}$.*

(ii) *The partial derivatives F_λ, F_x and $F_{\lambda x}$ exist and are continuous.*

(iii) *The spaces $N(\mathscr{L}_0)$ and $Y/R(\mathscr{L}_0)$ are one-dimensional.*

(iv) *The* transversality assumption $\partial_\lambda \partial_x F(0,0)x_0 \notin R(\mathscr{L}_0)$ *holds, where*

$$N(\mathscr{L}_0) = \operatorname{span}\{x_0\}.$$

If Z is any complement of $N(\mathscr{L}_0)$ in X, then there is a neighborhood U of $(0,0)$ in $\mathbb{R} \times X$, an interval $]-a, a[$ and continuous functions $\varphi :]-a, a[\to \mathbb{R}$ and $\psi :]-a, a[\to Z$ such that $\varphi(0) = 0$, $\psi(0) = 0$ and

$$F^{-1}(0) \cap U = \{(\varphi(\xi), \xi x_0 + \xi \psi(\xi)) : |\xi| < a\} \cup \{(\lambda, 0) : (\lambda, 0) \in U\}.$$

Before proceeding further with the consideration of the V-states, we shall recall the Riemann mapping theorem, a central result in complex analysis. To restate this result, we need to recall the definition of *simply connected* domains. Let $\widehat{\mathbb{C}} \triangleq \mathbb{C} \cup \{\infty\}$ denote the Riemann sphere. We say that a domain $\Omega \subset \widehat{\mathbb{C}}$ is *simply connected* if the set $\widehat{\mathbb{C}} \setminus \Omega$ is connected. Let $\mathbb{D}$ denote the unit open disc and $\Omega \subset \mathbb{C}$ be a simply connected bounded domain. Then according to the Riemann mapping theorem, there is a unique biholomorphic map $\Phi : \mathbb{C} \setminus \overline{\mathbb{D}} \to \mathbb{C} \setminus \overline{\Omega}$ taking the form

$$\Phi(z) = az + \sum_{n \in \mathbb{N}} \frac{a_n}{z^n}, \quad a > 0.$$

In this theorem, the regularity of the boundary has no effect on the existence of the conformal mapping, but it plays a role in determining the boundary behavior of the conformal mapping. See for instance [Pommerenke 1992; Warschawski 1935]. Here, we shall recall the following result.

Kellogg and Warschawski's theorem ([Warschawski 1935] or [Pommerenke 1992, Theorem 3.6]). *If the conformal map $\Phi : \mathbb{C} \setminus \overline{\mathbb{D}} \to \mathbb{C} \setminus \overline{\Omega}$ has a continuous extension to $\mathbb{C} \setminus \mathbb{D}$ which is of class $C^{n+\beta}$, with $n \in \mathbb{N}$ and $0 < \beta < 1$, then the boundary $\Phi(\mathbb{T})$ is of class $C^{n+\beta}$.*

2.3. ***Boundary equations.*** Our next task is to write down the equations of V-states using the conformal parametrization. First recall that the vorticity $\omega = \partial_1 v_2 - \partial_2 v_1$ satisfies the transport equation

$$\partial_t \omega + v \cdot \nabla \omega = 0$$

and the associated velocity is related to the vorticity through the stream function Ψ as

$$v = 2i\partial_{\bar{z}} \Psi,$$

with

$$\Psi(z) = \frac{1}{4\pi} \int_{\mathbb{D}} \log \left| \frac{z - \xi}{1 - z\bar{\xi}} \right|^2 \omega(\xi)\, dA(\xi).$$

When the vorticity is a patch of the form $\omega = \chi_D$ with D a bounded domain strictly contained in $\mathbb{D}$, then

$$\Psi(z) = \frac{1}{4\pi}\int_D \log\left|\frac{z-\xi}{1-z\bar{\xi}}\right|^2 dA(\xi).$$

For a complex function $\varphi : \mathbb{C} \to \mathbb{C}$ of class C^1 in the Euclidean variables (as a function of $\mathbb{R}^2$), we define

$$\partial_z\varphi = \frac{1}{2}\left(\frac{\partial\varphi}{\partial x} - i\frac{\partial\varphi}{\partial y}\right), \qquad \partial_{\bar{z}}\varphi = \frac{1}{2}\left(\frac{\partial\varphi}{\partial x} + i\frac{\partial\varphi}{\partial y}\right).$$

As we have seen in the introduction, a rotating patch or V-state is a special solution of the vorticity equation (2) with initial data $\omega_0 = \chi_D$ and such that

$$\omega(t) = \chi_{D_t}, \quad D_t = e^{it\Omega}D.$$

In this definition and for simplicity, we have only considered patches rotating around zero. According to [Burbea 1980; Hmidi et al. 2013; de la Hoz et al. 2016b], the boundary equation of the rotating patches is

$$\mathrm{Re}\{(\Omega\bar{z} - 2\partial_z\Psi)z'\} = 0, \quad z \in \Gamma \triangleq \partial D, \tag{10}$$

where z' denotes a tangent vector to the boundary at the point z. We point out that the existence of a rigid boundary does not alter this equation which in fact was established in the planar case. The purpose now is to transform (10) into an equation involving only the boundary ∂D of the V-state. To do so, we need to write $\partial_z\Psi$ as an integral on the boundary ∂D based on the use of the Cauchy–Pompeiu formula. Consider a finitely connected domain D bounded by finitely many smooth Jordan curves, and let Γ be the boundary ∂D endowed with the positive orientation; then

$$\text{for all } z \in \mathbb{C}, \quad \fint_\Gamma \frac{\varphi(z)-\varphi(\xi)}{z-\xi}\,d\xi = -\frac{1}{\pi}\int_D \partial_{\bar{\xi}}\varphi(\xi)\frac{dA(\xi)}{z-\xi}. \tag{11}$$

Differentiating (3) with respect to the variable z yields

$$\partial_z\Psi(z) = \frac{1}{4\pi}\int_D \frac{\bar{\xi}}{1-z\bar{\xi}}\,dA(\xi) + \frac{1}{4\pi}\int_D \frac{1}{z-\xi}\,dA(\xi). \tag{12}$$

Applying the Cauchy–Pompeiu formula with $\varphi(z) = \bar{z}$, we find

$$\frac{1}{\pi}\int_D \frac{1}{z-\xi}\,dA(\xi) = -\fint_\Gamma \frac{\bar{z}-\bar{\xi}}{z-\xi}\,d\xi \quad \text{for all } z \in \overline{D}.$$

Using the change of variable $\xi \to \bar{\xi}$ which keeps the Lebesgue measure invariant,

$$\frac{1}{\pi}\int_D \frac{\bar{\xi}}{1-z\bar{\xi}}\,dA(\xi) = \frac{1}{\pi z}\int_{\widetilde{D}} \frac{\xi}{1/z-\xi}\,dA(\xi)$$

with $\widetilde{D}$ being the image of D by complex conjugation. A second application of the Cauchy–Pompeiu formula, using that $1/z \notin \mathbb{D}$ for $z \in D$, yields

$$\frac{1}{\pi z}\int_{\widetilde{D}} \frac{\xi}{1/z-\xi}\,dA(\xi) = \fint_{\widetilde{\Gamma}} \frac{|\xi|^2}{1-z\xi}\,d\xi \quad \text{for all } z \in \overline{D}, \quad \widetilde{\Gamma} = \partial\widetilde{D}.$$

Using once again the change of variable $\xi \to \bar{\xi}$ which reverses the orientation,

$$\fint_{\widetilde{\Gamma}} \frac{|\xi|^2}{1-z\xi}\,d\xi = -\fint_{\Gamma} \frac{|\xi|^2}{1-z\bar{\xi}}\,d\bar{\xi} \quad \text{for all } z \in \overline{D}.$$

Therefore, we obtain

$$4\partial_z \Psi(z) = -\fint_{\Gamma} \frac{|\xi|^2}{1-z\bar{\xi}}\,d\bar{\xi} - \fint_{\Gamma} \frac{\bar{z}-\bar{\xi}}{z-\xi}\,d\xi. \tag{13}$$

Inserting the last identity in (10), we get an equation involving only the boundary

$$\mathrm{Re}\left\{\left(2\Omega\bar{z} + \fint_{\Gamma} \frac{\bar{z}-\bar{\xi}}{z-\xi}\,d\xi + \fint_{\Gamma} \frac{|\xi|^2}{1-z\bar{\xi}}\,d\bar{\xi}\right)z'\right\} = 0 \quad \text{for all } z \in \Gamma.$$

It is more convenient in the formulas to replace the angular velocity Ω in the preceding equation by the parameter $\lambda = 1 - 2\Omega$, leading to the V-states equation

$$\mathrm{Re}\left\{\left((1-\lambda)\bar{z} + \fint_{\Gamma} \frac{\bar{z}-\bar{\xi}}{z-\xi}\,d\xi + \fint_{\Gamma} \frac{|\xi|^2}{1-z\bar{\xi}}\,d\bar{\xi}\right)z'\right\} = 0 \quad \text{for all } z \in \Gamma. \tag{14}$$

It is worth pointing out that (14) characterizes V-states among domains with C^1 boundary, regardless of the number of boundary components. If the domain is simply connected, then there is only one boundary component and so only one equation. However, if the domain is doubly connected, then (14) gives rise to two coupled equations, one for each boundary component. We note that all the V-states that we shall consider admit at least one axis of symmetry passing through zero and without loss of generality it can be supposed to be the real axis. This implies that the boundary ∂D is invariant by the reflection symmetry $\xi \to \bar{\xi}$. Therefore, using this change of variables, which reverses orientation, in the last integral term of the equation (14), we obtain

$$\mathrm{Re}\left\{\left((1-\lambda)\bar{z} + \fint_{\Gamma} \frac{\bar{z}-\bar{\xi}}{z-\xi}\,d\xi - \fint_{\Gamma} \frac{|\xi|^2}{1-z\xi}\,d\xi\right)z'\right\} = 0 \quad \text{for all } z \in \Gamma. \tag{15}$$

To end this section, we mention that in the general framework the dynamics of any vortex patch can be described by its Lagrangian parametrization $\gamma_t : \mathbb{T} \to \partial D_t \triangleq \Gamma_t$ as

$$\partial_t \gamma_t = v(t, \gamma_t).$$

Since Ψ is a real-valued function,

$$\partial_{\bar{z}}\Psi = \overline{\partial_z \Psi},$$

which implies according to (13)

$$\begin{aligned} v(t,z) &= 2i\,\partial_{\bar{z}}\Psi(t,z) \\ &= -\frac{1}{4\pi}\int_{\Gamma_t} \log|z-\xi|^2\,d\xi + \frac{1}{4\pi}\int_{\Gamma_t} \frac{|\xi|^2}{1-\bar{z}\xi}\,d\xi. \end{aligned}$$

Consequently, we find that the Lagrangian parametrization satisfies the nonlinear ODE

$$\partial_t \gamma_t = -\frac{1}{4\pi}\int_{\Gamma_t} \log|\gamma_t - \xi|^2\,d\xi + \frac{1}{4\pi}\int_{\Gamma_t} \frac{|\xi|^2}{1-\bar{\gamma}_t\xi}\,d\xi. \tag{16}$$

The ultimate goal of this section is to relate the V-states described above to stationary solutions for Euler equations when the rigid boundary rotates at some specific angular velocity. To do so, suppose that the disc $\mathbb{D}$ rotates with a constant angular velocity Ω; then the equations (1) written in the frame of the rotating disc take the form

$$\partial_t u + u\cdot\nabla u - \Omega y^{\perp}\cdot\nabla u + \Omega u^{\perp} + \nabla q = 0$$

with

$$y = e^{-it\Omega}x, \qquad v(t,x) = e^{-it\Omega}u(t,y), \qquad q(t,y) = p(t,x).$$

For more details about the derivation of this equation, we refer the reader for instance to [Farwig and Hishida 2011]. Here the variable in the rotating frame is denoted by y. Applying the *curl* operator to the equation of u, we find that the vorticity of u, which still denoted by ω, is governed by the transport equation

$$\partial_t\omega + (u - \Omega y^{\perp})\cdot\nabla\omega = 0.$$

Consequently, any stationary solution in the patch form is actually a V-state rotating with the angular velocity Ω. Relating this observation to Theorems 1 and 6, we deduce that rotating the disc at some suitable angular velocities creates stationary patches with m-fold symmetry.

3. Simply connected V-states

In this section, we shall gather all the pieces needed for the proof of Theorem 1. The strategy is analogous to [Burbea 1980; Hmidi et al. 2013; de la Hoz et al. 2016b]. It consists of first writing down the V-states equation through the conformal parametrization and second applying the Crandall–Rabinowitz theorem. As can be noted from Theorem 1, the result is local meaning that we are looking for V-states which are smooth and cause a small perturbation of the Rankine patch $\chi_{\mathbb{D}_b}$ with $\mathbb{D}_b = b\mathbb{D}$. We also assume that the patch is symmetric with respect to the real axis, and this fact has been crucial in deriving (15). Note that as $D \Subset \mathbb{D}$ the exterior conformal mapping $\phi : \mathbb{D}^c \to D^c$ has the expansion

$$\phi(w) = bw + \sum_{n\geq 0}\frac{b_n}{w^n}, \qquad b_n \in \mathbb{R},$$

and satisfies $0 < b < 1$. This latter fact follows from the Schwarz lemma. Indeed, let

$$\psi(z) \triangleq \frac{1}{\phi(1/z)};$$

then $\psi : \mathbb{D} \to \widehat{D}$ is conformal, with $\widehat{D}$ the image of D by the map $z \mapsto 1/z$. Clearly $\mathbb{D} \subset \widehat{D}$, and therefore, the restriction $\psi^{-1} : \mathbb{D} \to \mathbb{D}$ is well defined and holomorphic and satisfies $\psi(0) = 0$. From the Schwarz lemma, we deduce that $|(\psi^{-1})'(0)| < 1$; otherwise D will coincide with $\mathbb{D}$. It suffices now to use that $(\psi^{-1})'(0) = b$.

Now we shall transform (15) into an equation on the unit circle $\mathbb{T}$. For this purpose, we make the change of variables $z = \phi(w)$ and $\xi = \phi(\tau)$. Note that for $w \in \mathbb{T}$ a tangent vector at the point $z = \phi(w)$

is given by

$$z' = iw\phi'(w)$$

and thus (15) becomes

$$\operatorname{Im}\left\{\left[(1-\lambda)\overline{\phi(w)} + \fint_{\mathbb{T}} \frac{\overline{\phi(w)} - \overline{\phi(\tau)}}{\phi(w) - \phi(\tau)} \phi'(\tau)\, d\tau - \fint_{\mathbb{T}} \frac{|\phi(\tau)|^2 \phi'(\tau)}{1 - \phi(w)\phi(\tau)}\, d\tau\right] w\phi'(w)\right\} = 0. \tag{17}$$

Set $\phi \triangleq b\,\mathrm{Id} + f$; then the foregoing functional can be split into three parts

$$\begin{aligned} F_1(f)(w) &\triangleq \operatorname{Im}\{\overline{\phi(w)} w\phi'(w)\}, \\ F_2(f)(w) &\triangleq \operatorname{Im}\left\{\fint_{\mathbb{T}} \frac{\overline{\phi(w)} - \overline{\phi(\tau)}}{\phi(w) - \phi(\tau)} \phi'(\tau)\, d\tau\, w\phi'(w)\right\}, \\ F_3(f)(w) &\triangleq \operatorname{Im}\left\{\fint_{\mathbb{T}} \frac{|\phi(\tau)|^2 \phi'(\tau)}{1 - \phi(w)\phi(\tau)}\, d\tau\, w\phi'(w)\right\}, \end{aligned} \tag{18}$$

and consequently, (17) becomes

$$F(\lambda, f) = 0, \quad F(\lambda, f) \triangleq (1-\lambda)F_1(f) + F_2(f) - F_3(f). \tag{19}$$

Observe that we can decompose F into two parts $F(\lambda, f) = G(\lambda, f) - F_3(f)$ where $G(\lambda, f)$ is the functional appearing in the flat space $\mathbb{R}^2$ and the new term F_3 describes the interaction between the patch and the rigid boundary $\mathbb{T}$. Now it is easy from the complex formulation to check that the disc $\mathbb{D}_b$ is a rotating patch for any $\Omega \in \mathbb{R}$. Indeed, as the disc is a trivial solution for the full space $\mathbb{R}^2$, $G(\lambda, 0) = 0$. Moreover,

$$F_3(0)(w) \triangleq \operatorname{Im}\left\{b^4 w \fint_{\mathbb{T}} \frac{d\tau}{1 - b^2 w\tau}\right\} = 0$$

because the integrand is analytic in the open disc $(1/b^2)\mathbb{D}$ and therefore we apply residue theorem.

3.1. ***Regularity of the functional F.*** This section is devoted to the study of the regularity assumptions stated in the Crandall–Rabinowitz theorem for the functional F introduced in (19). The application of this theorem at this stage of the presentation requires one to fix the function spaces X and Y. We should look for Banach spaces X and Y of Hölder type in the spirit of [Hmidi et al. 2013; de la Hoz et al. 2016b], and they are given by

$$X = \left\{f \in C^{1+\alpha}(\mathbb{T}) : f(w) = \sum_{n \geq 0} a_n \overline{w}^n,\ a_n \in \mathbb{R},\ w \in \mathbb{T}\right\},$$

$$Y = \left\{g \in C^{\alpha}(\mathbb{T}) : g(w) = \sum_{n \geq 1} b_n e_n,\ b_n \in \mathbb{R},\ w \in \mathbb{T}\right\}, \quad e_n \triangleq \frac{1}{2i}(\overline{w}^n - w^n),$$

with $\alpha \in]0, 1[$. For $r \in]0, 1[$, we denote by B_r the open ball of X with center 0 and radius r

$$B_r = \{f \in X : \|f\|_{C^{1+\alpha}} \leq r\}.$$

It is straightforward to see that for any $f \in B_r$ the function $w \mapsto \phi(w) = bw + f(w)$ is conformal on

$\mathbb{C} \setminus \overline{\mathbb{D}}$ provided that $r < b$. Moreover, according to the Kellogg–Warshawski result [Warschawski 1935], the boundary of $\phi(\mathbb{C} \setminus \overline{\mathbb{D}})$ is a Jordan curve of class $C^{1+\alpha}$. We propose to prove the following result concerning the regularity of F.

Proposition 11. *Let* $b \in]0, 1[$ *and* $0 < r < \min(b, 1-b)$; *then the following hold true*:

(i) $F : \mathbb{R} \times B_r \to Y$ *is* C^1 (*it is in fact* C^∞).

(ii) *The partial derivative* $\partial_\lambda \partial_f F : \mathbb{R} \times B_r \to \mathscr{L}(X, Y)$ *exists and is continuous* (*it is in fact* C^∞).

Proof. (i) We shall only sketch the proof because most of the details are done in [Hmidi et al. 2013; de la Hoz et al. 2016b]. First recall from (19) the decomposition

$$F(\lambda, f) = (1-\lambda)F_1(f) + F_2(f) - F_3(f).$$

The part $(1-\lambda)F_1(f) + F_2(f)$ coincides with the nonlinear functional appearing in the plane, and its regularity was studied in [Hmidi et al. 2013; de la Hoz et al. 2016b]. Therefore, it remains to check the regularity assumptions for the term F_3 given in (18). Since $C^\alpha(\mathbb{T})$ is an algebra, it suffices to prove that the mapping $F_4 : \phi \in b\,\mathrm{Id} + B_r \to C^\alpha$ defined by

$$F_4(\phi(w)) = \fint_{\mathbb{T}} \frac{|\phi(\tau)|^2 \phi'(\tau)}{1 - \phi(w)\phi(\tau)}\, d\tau \tag{20}$$

is C^1 and admits real Fourier coefficients. Observe that this functional is well defined and is given by the series expansion

$$F_4(\phi(w)) = \sum_{n\in\mathbb{N}} \phi^n(w) \fint_{\mathbb{T}} \phi^n(\tau) |\phi(\tau)|^2 \phi'(\tau)\, d\tau.$$

This sum is defined pointwisely because $\|\phi\|_{L^\infty} \le b + r < 1$. This series converges absolutely in $C^\alpha(\mathbb{T})$. To get this, we use the law product which can be proved by induction

$$\|\phi^n\|_{C^\alpha} \le n \|\phi\|_{L^\infty}^{n-1} \|\phi\|_{C^\alpha},$$

and therefore, we obtain

$$\begin{aligned} \|F_4(\phi)\|_{C^\alpha} &\le \sum_{n\in\mathbb{N}} n\|\phi\|_{L^\infty}^{n-1} \|\phi\|_{C^\alpha} \left| \fint_{\mathbb{T}} \phi^n(\tau)|\phi(\tau)|^2 \phi'(\tau)\, d\tau \right| \\ &\le \|\phi'\|_{L^\infty} \|\phi\|_{C^\alpha} \sum_{n\in\mathbb{N}} n\|\phi\|_{L^\infty}^{2n+1} \\ &\le \|\phi'\|_{L^\infty} \|\phi\|_{C^\alpha} \sum_{n\in\mathbb{N}} n(b+r)^{2n+1} < \infty. \end{aligned}$$

From the completeness of $C^\alpha(\mathbb{T})$, we obtain that $F_4(\phi)$ belongs to this space. Again from the series expansion, we can check that $\phi \mapsto F_4(\phi)$ is not only C^1 but also C^∞. To end the proof, we need to check that all the Fourier coefficients of $F_4(\phi)$ are real, and this fact is equivalent to showing that

$$\overline{F_4(\phi(w))} = F_4(\phi(\overline{w})) \quad \text{for all } w \in \mathbb{T}.$$

As $\overline{\phi(w)} = \phi(\overline{w})$ and $\overline{\phi'(w)} = \phi'(\overline{w})$, we may write successively

$$\begin{aligned}\overline{F_4(\phi(w))} &= -\fint_{\mathbb{T}} \frac{|\phi(\bar\tau)|^2\phi'(\bar\tau)}{1-\phi(\overline{w})\phi(\bar\tau)}\,d\bar\tau \\ &= \fint_{\mathbb{T}} \frac{|\phi(\tau)|^2\phi'(\tau)}{1-\phi(w)\phi(\tau)}\,d\tau\end{aligned}$$

where in the last equality we have used the change of variable $\tau \mapsto \bar\tau$.

(ii) Following the arguments developed in [Hmidi et al. 2013; de la Hoz et al. 2016b], we get what is expected formally, that is

$$\begin{aligned}\partial_\lambda \partial_f F(\lambda, f)h &= -\partial_f F_1(f) \\ &= \operatorname{Im}\{\overline{\phi(w)}wh'(w) + \overline{h(w)}w\phi'(w)\},\end{aligned}$$

from which we deduce that $\partial_\lambda\partial_f F(\lambda, f) \in \mathscr{L}(X, Y)$ and the mapping $f \mapsto \partial_\lambda\partial_f F(\lambda, f)$ is in fact C^∞, which is clearly better than the statement of the proposition. □

3.2. *Spectral study.* This part is crucial for implementing the Crandall–Rabinowitz theorem. We shall in particular compute the linearized operator $\partial_f F(\lambda, 0)$ around the trivial solution and look for the values of λ associated with the nontrivial kernel. For these values of λ, we shall see that the linearized operator has a one-dimensional kernel and is in fact of Fredholm type with zero index. Before giving the main result of this subsection, we recall the notation $e_n = (\overline{w}^n - w^n)/2i$.

Proposition 12. *Let $h \in X$ take the form $h(w) = \sum_{n\geq 0} a_n/w^n$. Then the following hold true*:

(i) *The structure of $\partial_f F(\lambda, 0)$ is given by*

$$\partial_f F(\lambda, 0)h(w) = b\sum_{n\geq 1} n\left(\lambda - \frac{1-b^{2n}}{n}\right)a_{n-1}e_n.$$

(ii) *The kernel of $\partial_f F(\lambda, 0)$ is nontrivial if and only if there exists $m \in \mathbb{N}^\star$ such that*

$$\lambda = \lambda_m \triangleq \frac{1-b^{2m}}{m}, \qquad m \in \mathbb{N}^\star,$$

and in this case, the kernel is one-dimensional and generated by $v_m(w) = \overline{w}^{m-1}$.

(iii) *The range of $\partial_f F(\lambda_m, 0)$ is of codimension* 1.

(iv) *The transversality condition holds*: *for $m \in \mathbb{N}^\star$,*

$$\partial_\lambda\partial_f F(\lambda_m, 0)v_m \notin R\partial_f F(\lambda_m, 0).$$

Proof. (i) The computations of the terms $\partial_f F_i(\lambda,0)h$ were almost done in [de la Hoz et al. 2016b], and we shall only give some details. By straightforward computations, we obtain

$$\begin{aligned}\partial_f F_1(0,0)h(w) &= \operatorname{Im}\{b\overline{h(w)}w + bh'(w)\}\\ &= b\operatorname{Im}\Big\{\sum_{n\geq 0} a_n w^{n+1} - \sum_{n\geq 1} na_n\overline{w}^{n+1}\Big\}\\ &= -\frac{b}{2i}\sum_{n\geq 0}(n+1)a_n(\overline{w}^{n+1}-w^{n+1})\\ &= -b\sum_{n\geq 0}(n+1)a_n e_{n+1}.\end{aligned}\tag{21}$$

Concerning $\partial_f F_2(0,0)$, one may write

$$\partial_f F_2(0,0)h(w) = \operatorname{Im}\Big\{bw \fint_{\mathbb{T}} \frac{\overline{h(\tau)}-\overline{h(w)}}{\tau - w}\,d\tau + b\fint_{\mathbb{T}}\frac{h(\tau)-h(w)}{\tau-w}\overline{\tau}\,d\tau - b\fint_{\mathbb{T}} h'(\tau)\overline{\tau}\,d\tau - bh'(w)\Big\}.$$

Therefore, using the residue theorem at infinity,

$$\begin{aligned}\partial_f F_2(0,0)h(w) &= \operatorname{Im}\Big\{bw \fint_{\mathbb{T}} \frac{\overline{h(\tau)}-\overline{h(w)}}{\tau - w}\,d\tau - bh'(w)\Big\}\\ &= -\operatorname{Im}\{bh'(w)\},\end{aligned}$$

where we have used in the last line the fact

$$\begin{aligned}\fint_{\mathbb{T}} \frac{\overline{h(\tau)}-\overline{h(w)}}{\tau - w}\,d\tau &= \sum_{n\in\mathbb{N}} a_n \fint_{\mathbb{T}} \frac{w^n-\tau^n}{\tau-w}\,d\tau\\ &= 0.\end{aligned}$$

Consequently, we obtain

$$\partial_f F_2(0,0)h(w) = b\sum_{n\geq 1} na_n e_{n+1}.\tag{22}$$

As for the third term $\partial_f F_3(0,0)h$, we get by plain computation

$$\begin{aligned}\partial_f F_3(0,0)h(w) = \operatorname{Im}\Big\{& b^3 w\fint_{\mathbb{T}}\frac{d\tau}{1-b^2 w\tau}h'(w) + b^3 w\fint_{\mathbb{T}}\frac{h'(\tau)\,d\tau}{1-b^2w\tau}\\ &+ 2b^3 w\fint_{\mathbb{T}}\frac{\operatorname{Re}\{h(\tau)\overline{\tau}\}}{1-b^2w\tau}\,d\tau + b^5 w\fint_{\mathbb{T}}\frac{wh(\tau)+\tau h(w)}{(1-b^2w\tau)^2}\,d\tau\Big\}\\ \triangleq \operatorname{Im}\{& I_1(w)+I_2(w)+I_3(w)+I_4(w)\}.\end{aligned}\tag{23}$$

By once again invoking the residue theorem,

$$I_1(w) = 0.\tag{24}$$

To compute the second term $I_2(w)$, we use the Taylor series of $1/(1-\zeta)$, leading to

$$\begin{aligned} I_2(w) &= b^3 w \fint_{\mathbb{T}} \frac{h'(\tau)\,d\tau}{1-b^2 w\tau} \\ &= \sum_{n\geq 0} b^{2n+3} w^{n+1} \fint_{\mathbb{T}} \tau^n h'(\tau)\,d\tau. \end{aligned}$$

From the Fourier expansions of h, we infer that

$$\fint_{\mathbb{T}} \tau^n h'(\tau)\,d\tau = -na_n,$$

which implies that

$$I_2(w) = -\sum_{n\geq 1} n a_n b^{2n+3} w^{n+1}. \tag{25}$$

In regard to the third term $I_3(w)$, it may be written in the form

$$I_3(w) = b^3 w \fint_{\mathbb{T}} \frac{\tau\overline{h(\tau)}}{1-b^2 w\tau}\,d\tau + b^3 w \fint_{\mathbb{T}} \frac{\bar{\tau} h(\tau)}{1-b^2 w\tau}\,d\tau.$$

The first integral term is zero due to the fact that the integrand is analytic in the open unit disc and continuous up to the boundary. Therefore, we get similarly to $I_2(w)$

$$\begin{aligned} I_3(w) &= b^3 w \fint_{\mathbb{T}} \frac{\bar{\tau} h(\tau)}{1-b^2 w\tau}\,d\tau \\ &= \sum_{n\geq 0} b^{2n+3} w^{n+1} \fint_{\mathbb{T}} \tau^{n-1} h(\tau)\,d\tau. \end{aligned}$$

Note that

$$\fint_{\mathbb{T}} \tau^{n-1} h(\tau)\,d\tau = a_n,$$

which implies in turn that

$$I_3(w) = \sum_{n\geq 0} a_n b^{2n+3} w^{n+1}. \tag{26}$$

Now we come back to the last term $I_4(w)$, and one may write using again the residue theorem

$$\begin{aligned} I_4(w) &= b^5 w^2 \fint_{\mathbb{T}} \frac{h(\tau)\,d\tau}{(1-b^2 w\tau)^2} + b^5 w h(w) \fint_{\mathbb{T}} \frac{\tau\,d\tau}{(1-b^2 w\tau)^2} \\ &= b^5 w^2 \fint_{\mathbb{T}} \frac{h(\tau)\,d\tau}{(1-b^2 w\tau)^2} + 0. \end{aligned}$$

Using the Taylor expansion

$$\frac{1}{(1-\zeta)^2} = \sum_{n\geq 1} n\zeta^{n-1}, \quad |\zeta| < 1, \tag{27}$$

we deduce that

$$\begin{aligned} I_4(w) &= \sum_{n\geq 1} n b^{2n+3} w^{n+1} \fint_{\mathbb{T}} \tau^{n-1} h(\tau)\, d\tau \\ &= \sum_{n\geq 1} n a_n b^{2n+3} w^{n+1}. \end{aligned} \tag{28}$$

Inserting the identities (24), (25), (26) and (28) into (23), we find

$$\begin{aligned} \partial_f F_3(0,0) h(w) &= \operatorname{Im}\Big\{ \sum_{n\geq 0} a_n b^{2n+3} w^{n+1} \Big\} \\ &= -\sum_{n\geq 0} a_n b^{2n+3} e_{n+1}. \end{aligned} \tag{29}$$

Hence, by plugging (21), (22) and (29) into (19), we obtain

$$\begin{aligned} \partial_f F(\lambda, 0) h(w) &= b \sum_{n\geq 0} (n+1) \Big(\lambda - \frac{1-b^{2n+2}}{n+1} \Big) a_n e_{n+1} \\ &= b \sum_{n\geq 1} n \Big(\lambda - \frac{1-b^{2n}}{n} \Big) a_{n-1} e_n. \end{aligned} \tag{30}$$

This finishes the proof of the first part (i).

(ii) From (30), we immediately deduce that the kernel of $\partial_f F(\lambda, 0)$ is nontrivial if and only if there exists $m \geq 1$ such that

$$\lambda = \lambda_m \triangleq \frac{1-b^{2m}}{m}.$$

We shall prove that the sequence $n \mapsto \lambda_n$ is strictly decreasing, from which we conclude immediately that the kernel is one-dimensional. Assume that for two integers $n > m \geq 1$ one has

$$\frac{1-b^{2m}}{m} = \frac{1-b^{2n}}{n}.$$

This implies that

$$\frac{1-b^{2n}}{1-b^{2m}} = \frac{n}{m}.$$

Set $\alpha = n/m$ and $x = b^{2m}$; then the preceding equality becomes

$$f(x) \triangleq \frac{1-x^\alpha}{1-x} = \alpha.$$

If we prove that this equation has no solution $x \in]0,1[$ for any $\alpha > 1$, then the result follows without difficulty. To do so, we get after differentiating f

$$f'(x) = \frac{(\alpha-1)x^\alpha - \alpha x^{\alpha-1} + 1}{(1-x)^2} \triangleq \frac{g(x)}{(1-x)^2}.$$

Now we note that

$$g'(x) = \alpha(\alpha-1)x^{\alpha-2}(x-1) < 0.$$

As $g(1)=0$, then we deduce

$$g(x)>0 \quad \text{for all } x\in\,]0,1[.$$

Thus, f is strictly increasing. Furthermore,

$$\lim_{x\to 1} f(x)=\alpha.$$

This implies that,

$$\text{for all } x\in\,]0,1[,\qquad f(x)<\alpha.$$

Therefore, we get the strict monotonicity of the "eigenvalues", and consequently, the kernel of $\partial_f F(\lambda_m,0)$ is a one-dimensional vector space generated by the function $v_m(w)=\overline{w}^{m-1}$.

(iii) We shall prove that the range of $\partial_f F(\lambda_m,0)$ is described by

$$R\partial_f F(\lambda_m,0)=\Big\{g\in Y : g(w)=\sum_{\substack{n\geq 1\\ n\neq m}} b_n e_n\Big\}\triangleq\mathscr{Z}.$$

Combining Propositions 11 and 12(i), we conclude that the range is contained in the right space. So what is left is to prove the converse. Let $g\in\mathscr{Z}$; we will solve in X the equation

$$\partial_f F(\lambda_m,0)h=g,\qquad h=\sum_{n\geq 0} a_n\overline{w}^n.$$

By virtue of (30), this equation is equivalent to

$$a_{n-1}=\frac{b_n}{bn(\lambda_m-\lambda_n)},\qquad n\geq 1,\ n\neq m.$$

Thus, the problem reduces to showing that

$$h: w\mapsto \sum_{\substack{n\geq 1\\ n\neq m}}\frac{b_n}{bn(\lambda_m-\lambda_n)}\overline{w}^{n-1}\in C^{1+\alpha}(\mathbb{T}).$$

Observe that

$$\inf_{n\neq m}|\lambda_n-\lambda_m|\triangleq c_0>0,$$

and thus, we deduce by Cauchy–Schwarz

$$\begin{aligned}\|h\|_{L^\infty}&\leq\frac{1}{b}\sum_{\substack{n\geq 1\\ n\neq m}}\frac{|b_n|}{n|\lambda_m-\lambda_n|}\\&\leq\frac{1}{c_0 b}\sum_{\substack{n\geq 1\\ n\neq m}}\frac{|b_n|}{n}\\&\lesssim\|g\|_{L^2}\lesssim\|g\|_{C^\alpha}.\end{aligned}$$

To finish the proof, we shall check that $h' \in C^\alpha(\mathbb{T})$ or equivalently $(\overline{w}h)' \in C^\alpha(\mathbb{T})$. It is obvious that

$$
\begin{aligned}
(\overline{w}h(w))' &= -\sum_{\substack{n\geq 1\\ n\neq m}} \frac{b_n}{b(\lambda_m-\lambda_n)}\overline{w}^{n+1}\\
&= -\frac{1}{b\lambda_m}\sum_{\substack{n\geq 1\\ n\neq m}} b_n\overline{w}^{n+1} + \frac{1}{b\lambda_m}\sum_{\substack{n\geq 1\\ n\neq m}} \frac{\lambda_n}{\lambda_n-\lambda_m} b_n\overline{w}^{n+1}.
\end{aligned}
$$

We shall write the preceding expression with the Szegő projection

$$
\Pi : \sum_{n\in\mathbb{Z}} a_n w^n \mapsto \sum_{n\in -\mathbb{N}} a_n w^n, \quad (\overline{w}h(w))' = -\frac{\overline{w}}{2ib\lambda_m}\Pi g(w) + \frac{\overline{w}}{2ib\lambda_m}(K\star\Pi g)(w),
$$

with

$$
K(w) \triangleq \sum_{\substack{n\geq 1\\ n\neq m}} \frac{\lambda_n}{\lambda_n-\lambda_m}\overline{w}^n.
$$

Notice that

$$
\frac{\lambda_n}{|\lambda_n-\lambda_m|} \leq c_0^{-1}\frac{1}{n},
$$

and therefore, $K \in L^2(\mathbb{T})$ which implies in particular that $K \in L^1(\mathbb{T})$. Now to complete the proof of $(\overline{w}h)' \in C^\alpha(\mathbb{T})$, it suffices to use the continuity of the Szegő projection on $C^\alpha(\mathbb{T})$ combined with $L^1 \star C^\alpha(\mathbb{T}) \subset C^\alpha(\mathbb{T})$.

(iv) To check the transversality assumption, we differentiate (30) with respect to λ:

$$
\partial_\lambda\partial_f F(\lambda_m, 0)h = b\sum_{n\geq 1} n a_{n-1} e_n.
$$

Therefore,

$$
\partial_\lambda\partial_f F(\lambda_m, 0)v_m = bme_m \notin R(\partial_f F(\lambda_m, 0)).
$$

This completes the proof of the proposition. □

3.3. ***Proof of Theorem 1.*** According to Propositions 14 and 11, all the assumptions of the Crandall–Rabinowitz theorem are satisfied, and therefore, we conclude for each $m \geq 1$ the existence of only one nontrivial curve bifurcating from the trivial one at the angular velocity

$$
\Omega_m = \frac{1-\lambda_m}{2} = \frac{m-1+b^{2m}}{2m}.
$$

To complete the proof, it remains to check the m-fold symmetry of the V-states. This can be done by including the required symmetry in the function spaces. More precisely, instead of dealing with X and Y, we should work with the spaces

$$X_m = \Big\{ f \in C^{1+\alpha}(\mathbb{T}) : f(w) = \sum_{n=1}^{\infty} a_n \overline{w}^{nm-1},\ a_n \in \mathbb{R} \Big\},$$
$$Y_m = \Big\{ g \in C^{\alpha}(\mathbb{T}) : g(w) = \sum_{n\geq 1} b_n e_{nm},\ b_n \in \mathbb{R} \Big\}, \quad e_n = \frac{1}{2i}(\overline{w}^n - w^n).$$

The conformal mapping describing the V-state takes the form

$$\phi(w) = bw + \sum_{n=1}^{\infty} a_n \overline{w}^{nm-1},$$

and the m-fold symmetry of the V-state means that

$$\phi(e^{2i\pi/m} w) = e^{2i\pi/m} \phi(w) \quad \text{for all } w \in \mathbb{T}.$$

The ball B_r is changed to $B_r^m = \{ f \in X_m : \| f \|_{C^{1+\alpha}} < r \}$. Then Proposition 11 holds true according to this adaptation, and the only point that one must check is the stability of the spaces; that is, for $f \in B_r^m$, we have $F(\lambda, f) \in Y_m$. This result was checked in [de la Hoz et al. 2016b] for the terms F_1 and F_2, and it remains to check that $F_3(f)$ belongs to Y_m. Recall that

$$F_3(f(w)) = \operatorname{Im}\{ F_4(\phi(w)) w \phi'(w) \}, \quad \phi(w) = bw + f(w),$$

where F_4 is defined in (20). By change of variables and using the symmetry of ϕ,

$$\begin{aligned} F_4(\phi(e^{i2\pi/m} w)) &= \fint_{\mathbb{T}} \frac{|\phi(\xi)|^2 \phi'(\xi)}{1 - \phi(e^{i2\pi/m} w)\phi(\xi)}\, d\xi \\ &= e^{-i2\pi/m} \fint_{\mathbb{T}} \frac{|\phi(e^{-i2\pi/m}\zeta)|^2 \phi'(e^{-i2\pi/m}\zeta)}{1 - \phi(e^{i2\pi/m} w)\phi(e^{-i2\pi/m}\zeta)}\, d\zeta \\ &= e^{-i2\pi/m} \fint_{\mathbb{T}} \frac{|\phi(\tau)|^2 \phi'(\tau)}{1 - \phi(w)\phi(\tau)}\, d\tau \\ &= e^{-i2\pi/m} F_4(\phi(w)). \end{aligned}$$

Consequently, we obtain

$$F_3(f(e^{i2\pi/m} w)) = F_3(f(w)),$$

and this shows the stability result.

4. Doubly connected V-states

In this section, we shall establish all the ingredients required for the proofs of Theorems 6 and 9, and this will be carried out in several steps. First we shall write the equations governing the doubly connected V-states which are described by two coupled nonlinear equations. Second we briefly discuss the regularity of the functionals and compute the linearized operator around the trivial solution. The delicate part to which we will pay careful attention is the computation of the kernel dimension. This will be implemented through the study of the monotonicity of the nonlinear eigenvalues. As we shall see, the fact that we have multiple parameters introduces many more complications to this study compared to the result of [de la Hoz et al. 2016b]. Finally, we shall prove Theorem 6 in Section 4.5.2.

4.1. ***Boundary equations.*** Let D be a doubly connected domain of the form $D = D_1 \setminus D_2$ with $D_2 \subset D_1$ two simply connected domains. Denote by Γ_j the boundary of the domain D_j. In this case, the V-states equation (15) reduces to two coupled equations, one for each boundary component Γ_j. More precisely,

$$\operatorname{Re}\{((1-\lambda)\bar{z} + I(z) - J(z))z'\} = 0 \quad \text{for all } z \in \Gamma_1 \cup \Gamma_2, \tag{31}$$

with

$$\begin{aligned} I(z) &= \fint_{\Gamma_1} \frac{\bar{z}-\bar{\xi}}{z-\xi}\, d\xi - \fint_{\Gamma_2} \frac{\bar{z}-\bar{\xi}}{z-\xi}\, d\xi, \\ J(z) &= \fint_{\Gamma_1} \frac{|\xi|^2}{1-z\xi}\, d\xi - \fint_{\Gamma_2} \frac{|\xi|^2}{1-z\xi}\, d\xi. \end{aligned}$$

As for the simply connected case, we prefer using the conformal parametrization of the boundaries. Let $\phi_j : \mathbb{D}^c \to D_j^c$ satisfy

$$\phi_j(w) = b_j w + \sum_{n \geq 0} \frac{a_{j,n}}{w^n}$$

with $0 < b_j < 1$, $j = 1, 2$ and $b_2 < b_1$. We assume moreover that all the Fourier coefficients are real because we shall look for V-states which are symmetric with respect to the real axis. Then by change of variables, we obtain

$$\begin{aligned} I(z) &= \fint_{\mathbb{T}} \frac{\bar{z}-\overline{\phi_1(\xi)}}{z-\phi_1(\xi)} \phi_1'(\xi)\, d\xi - \fint_{\mathbb{T}} \frac{\bar{z}-\overline{\phi_2(\xi)}}{z-\phi_2(\xi)} \phi_2'(\xi)\, d\xi, \\ J(z) &= \fint_{\mathbb{T}} \frac{|\phi_1(\xi)|^2}{1-z\phi_1(\xi)} \phi_1'(\xi)\, d\xi - \fint_{\mathbb{T}} \frac{|\phi_2(\xi)|^2}{1-z\phi_2(\xi)} \phi_2'(\xi)\, d\xi. \end{aligned}$$

Setting $\phi_j = b_j \operatorname{Id} + f_j$, (31) becomes,

$$\text{for all } w \in \mathbb{T}, \quad G_j(\lambda, f_1, f_2)(w) = 0, \quad j = 1, 2,$$

where

$$G_j(\lambda, f_1, f_2)(w) \triangleq \operatorname{Im}\big\{\big((1-\lambda)\overline{\phi_j(w)} + I(\phi_j(w)) - J(\phi_j(w))\big) w \phi_j'(w)\big\}.$$

Note that one can easily check that

$$G(\lambda, 0, 0) = 0 \quad \text{for all } \lambda \in \mathbb{R}.$$

This is consistent with the fact that the annulus is a stationary solution and therefore rotates with any angular velocity since the shape is rotational invariant.

4.2. ***Regularity of the functional G.*** In this short subsection, we shall quickly state the regularity result of the functional $G \triangleq (G_1, G_2)$ needed in the Crandall–Rabinowitz theorem. Following the simply connected case, the spaces X and Y involved in the bifurcation will be chosen in a similar way: set

$$X = \Big\{ f \in (C^{1+\alpha}(\mathbb{T}))^2 : f(w) = \sum_{n\geq 0} A_n \overline{w}^n,\ A_n \in \mathbb{R}^2,\ w \in \mathbb{T} \Big\},$$

$$Y = \Big\{ g \in (C^{\alpha}(\mathbb{T}))^2 : g(w) = \sum_{n\geq 1} B_n e_n,\ B_n \in \mathbb{R}^2,\ w \in \mathbb{T} \Big\}, \quad e_n \triangleq \frac{1}{2i}(\overline{w}^n - w^n),$$

with $\alpha \in]0, 1[$. For $r \in (0, 1)$, we denote by B_r the open ball of X with center 0 and radius r,

$$B_r = \{ f \in X : \|f\|_{C^{1+\alpha}} \leq r \}.$$

Similarly to Proposition 11, one can establish the regularity assumptions needed for the Crandall–Rabinowitz theorem. Compared to the simply connected case, the only terms that one should care about are those describing the interaction between the boundaries of the patches which are supposed to be disjoint. Therefore, the involved kernels are sufficiently smooth and actually do not cause significant difficulties in their treatment. For this reason, we prefer skip the details and restrict ourselves to the following statement.

Proposition 13. *Let $b \in]0, 1[$ and $0 < r < \min(b, 1-b)$; then the following hold true*:

(i) *$G : \mathbb{R} \times B_r \to Y$ is C^1 (it is in fact C^∞).*

(ii) *The partial derivative $\partial_\lambda \partial_f G : \mathbb{R} \times B_r \to \mathscr{L}(X, Y)$ exists and is continuous (it is in fact C^∞).*

4.3. ***Structure of the linearized operator.*** In this subsection, we shall compute the linearized operator $\partial_f G(\lambda, 0)$ around the annulus $\mathbb{A}_{b_1,b_2}$ of radii b_1 and b_2. The study of the eigenvalues is postponed to the next subsections. From the regularity assumptions of G, we assert that the Fréchet derivative and Gâteaux derivatives coincide and

$$DG(\lambda, 0, 0)(h_1, h_2) = \frac{d}{dt} G(\lambda, th_1, th_2)|_{t=0}.$$

Note that $DG(\lambda, 0, 0)$ is nothing but the partial derivative $\partial_f G(\lambda, 0, 0)$. Our main result reads as follows.

Proposition 14. *Let $h = (h_1, h_2) \in X$ take the form $h_j(w) = \sum_{n\geq 0} a_{j,n}/w^n$. Then*

$$DG(\lambda, 0, 0)(h_1, h_2) = \sum_{n\geq 1} M_n(\lambda) \begin{pmatrix} a_{1,n-1} \\ a_{2,n-1} \end{pmatrix} e_n,$$

where the matrix M_n is given by

$$M_n(\lambda) = \begin{pmatrix} b_1[n\lambda - 1 + b_1^{2n} - n(b_2/b_1)^2] & b_2[(b_2/b_1)^n - (b_1 b_2)^n] \\ -b_1[(b_2/b_1)^n - (b_1 b_2)^n] & b_2[n\lambda - n + 1 - b_2^{2n}] \end{pmatrix} \quad \text{and} \quad e_n(w) = \frac{1}{2i}(\overline{w}^n - w^n).$$

Proof. Since $G = (G_1, G_2)$, for a given couple of functions $(h_1, h_2) \in X$,

$$DG(\lambda, 0, 0)(h_1, h_2) = \begin{pmatrix} \partial_{f_1} G_1(\lambda, 0, 0)h_1 + \partial_{f_2} G_1(\lambda, 0, 0)h_2 \\ \partial_{f_1} G_2(\lambda, 0, 0)h_1 + \partial_{f_2} G_2(\lambda, 0, 0)h_2 \end{pmatrix}.$$

We shall split G_j into three terms

$$G_j(\lambda, f_1, f_2) = G_j^1(\lambda, f_j) + G_j^2(f_1, f_2) + G_j^3(f_1, f_2),$$

where

$$G_j^1(\lambda, f_j)(w) \triangleq \mathrm{Im}\bigg\{\bigg[(1-\lambda)\overline{\phi_j(w)} + (-1)^{j+1} \fint_{\mathbb{T}} \frac{\overline{\phi_j(w)} - \overline{\phi_j(\tau)}}{\phi_j(w) - \phi_j(\tau)} \phi_j'(\tau)\, d\tau + (-1)^j \fint_{\mathbb{T}} \frac{|\phi_j(\tau)|^2 \phi_j'(\tau)}{1 - \phi_j(w)\phi_j(\tau)}\, d\tau\bigg] w\phi_j'(w)\bigg\},$$

$$G_j^2(f_1, f_2) \triangleq (-1)^j\, \mathrm{Im}\bigg\{\fint_{\mathbb{T}} \frac{\overline{\phi_j(w)} - \overline{\phi_i(\tau)}}{\phi_j(w) - \phi_i(\tau)} \phi_i'(\tau)\, d\tau\, w\phi_j'(w)\bigg\}, \quad i \neq j,$$

$$G_j^3(f_1, f_2) \triangleq (-1)^{j+1}\, \mathrm{Im}\bigg\{\fint_{\mathbb{T}} \frac{|\phi_i(\tau)|^2 \phi_i'(\tau)}{1 - \phi_j(w)\phi_i(\tau)}\, d\tau\, w\phi_j'(w)\bigg\}, \quad i \neq j,$$

with $\phi_j = b_j\, \mathrm{Id} + f_j$, $j = 1, 2$.

• *Computation of* $\partial_{f_j} G_j^1(\lambda, 0, 0)h_j$. First observe that

$$G_1^1(\lambda, f_1)(w) = \mathrm{Im}\bigg\{\bigg[(1-\lambda)\overline{\phi_1(w)} + \fint_{\mathbb{T}} \frac{\overline{\phi_1(w)} - \overline{\phi_1(\tau)}}{\phi_1(w) - \phi_1(\tau)} \phi_1'(\tau)\, d\tau - \fint_{\mathbb{T}} \frac{|\phi_1(\tau)|^2 \phi_1'(\tau)}{1 - \phi_1(w)\phi_1(\tau)}\, d\tau\bigg] w\phi_1'(w)\bigg\}.$$

This functional is exactly the defining function in the simply connected case, and thus, using merely (30),

$$\partial_{f_1} G_1^1(\lambda, 0)h_1 = b_1 \sum_{n\geq 0} (\lambda(n+1) - 1 + b_1^{2n+2}) a_{1,n} e_{n+1}. \tag{32}$$

In regard to $G_2^1(\lambda, f_2)$, we get from the definition

$$G_2^1(\lambda, f_2)(w) = \mathrm{Im}\bigg\{\bigg[(1-\lambda)\overline{\phi_2(w)} - \fint_{\mathbb{T}} \frac{\overline{\phi_2(w)} - \overline{\phi_2(\tau)}}{\phi_2(w) - \phi_2(\tau)} \phi_2'(\tau)\, d\tau + \fint_{\mathbb{T}} \frac{|\phi_2(\tau)|^2 \phi_2'(\tau)}{1 - \phi_2(w)\phi_2(\tau)}\, d\tau\bigg] w\phi_2'(w)\bigg\}.$$

It is easy to check the algebraic relation $G_2^1(\lambda, f_2) = -G_1^1(2-\lambda, f_2)$, and thus, by applying (32),

$$\partial_{f_2} G_2^1(\lambda, 0)h_2 = b_2 \sum_{n\geq 0} (\lambda(n+1) - 2n - 1 - b_2^{2n+2}) a_{2,n} e_{n+1}. \tag{33}$$

• *Computation of* $\partial_{f_j} G_j^2(\lambda, 0, 0)h_j$. This quantity is given by

$$\partial_{f_j} G_j^2(0,0)h_j = (-1)^j \frac{d}{dt} \mathrm{Im}\bigg\{b_i w \fint_{\mathbb{T}} \frac{b_j\overline{w} - b_i\overline{\tau} + t\overline{h_j(w)}}{b_j w - b_i\tau + th_j(w)}\, d\tau\, (b_j + th_j'(w))\bigg\}\bigg|_{t=0}.$$

Straightforward computations yield

$$\partial_{f_j} G_j^2(0,0)h_j = (-1)^j b_i\, \mathrm{Im}\bigg\{h_j'(w) w \fint_{\mathbb{T}} \frac{b_j\overline{w} - b_i\overline{\tau}}{b_j w - b_i\tau}\, d\tau + b_j w\overline{h_j(w)} \fint_{\mathbb{T}} \frac{d\tau}{b_j w - b_i\tau} - b_j w h_j(w) \fint_{\mathbb{T}} \frac{b_j\overline{w} - b_i\overline{\tau}}{(b_j w - b_i\tau)^2}\, d\tau\bigg\}.$$

According to the residue theorem,

$$\int_{\mathbb{T}} \frac{d\tau}{b_1 w - b_2 \tau} = 0, \quad \int_{\mathbb{T}} \frac{d\tau}{(b_1 w - b_2 \tau)^2} = 0 \quad \text{for all } w \in \mathbb{T},$$

and therefore,

$$\begin{aligned}
\partial_{f_1} G_1^2(0,0) h_1(w) &= -b_2^2 \operatorname{Im}\left\{ -\fint_{\mathbb{T}} \frac{w h_1'(w)}{b_1 w - b_2 \tau} \frac{d\tau}{\tau} + b_1 \fint_{\mathbb{T}} \frac{w h_1(w)}{(b_1 w - b_2 \tau)^2} \frac{d\tau}{\tau} \right\} \\
&= -b_2^2 \operatorname{Im}\left\{ -\frac{1}{b_1} h_1'(w) + \frac{1}{b_1} \overline{w} h_1(w) \right\} \\
&= -\frac{b_2^2}{b_1} \sum_{n \geq 0} (n+1) a_{1,n} e_{n+1}.
\end{aligned} \tag{34}$$

Now using the vanishing integrals

$$\int_{\mathbb{T}} \frac{\bar{\tau}\, d\tau}{b_2 w - b_1 \tau} = 0, \qquad \int_{\mathbb{T}} \frac{\bar{\tau}\, d\tau}{(b_2 w - b_1 \tau)^2} = 0, \qquad \int_{\mathbb{T}} \frac{d\tau}{(b_2 w - b_1 \tau)^2} = 0,$$

we may obtain

$$\begin{aligned}
\partial_{f_2} G_2^2(0,0) h_2(w) &= b_1 \operatorname{Im}\left\{ b_2 h_2'(w) \fint_{\mathbb{T}} \frac{d\tau}{b_2 w - b_1 \tau} + b_2 w \overline{h_2(w)} \fint_{\mathbb{T}} \frac{d\tau}{b_2 w - b_1 \tau} \right\} \\
&= b_1 \operatorname{Im}\left\{ -\frac{b_2}{b_1} h_2'(w) - \frac{b_2}{b_1} w \overline{h_2(w)} \right\} \\
&= b_2 \sum_{n \geq 0} (n+1) a_{2,n} e_{n+1}.
\end{aligned} \tag{35}$$

• *Computation of* $\partial_{f_i} G_j^2(\lambda, 0, 0) h_i$, $i \neq j$. By straightforward computations, we obtain

$$\begin{aligned}
\partial_{f_i} G_j^2(0,0) h_i(w) = (-1)^j b_j \operatorname{Im}\Bigg\{ & w \fint_{\mathbb{T}} \frac{(b_j \overline{w} - b_i \bar{\tau})}{b_j w - b_i \tau} h_i'(\tau) d\tau - b_i w \fint_{\mathbb{T}} \frac{\overline{h_i(\tau)}}{b_j w - b_i \tau} d\tau \\
& + b_i w \fint_{\mathbb{T}} \frac{(b_j \overline{w} - b_i \bar{\tau}) h_i(\tau)\, d\tau}{(b_j w - b_i \tau)^2} \Bigg\}.
\end{aligned} \tag{36}$$

As $\overline{h_i}$ is holomorphic inside the open unit disc, by the residue theorem, we deduce that

$$\fint_{\mathbb{T}} \frac{\overline{h_i(\tau)}}{b_1 w - b_2 \tau} d\tau = 0, \quad w \in \mathbb{T}.$$

It follows that

$$\begin{aligned}
\partial_{f_2} G_1^2(0,0) h_2(w) = -b_1 \operatorname{Im}\Bigg\{ & b_1 \fint_{\mathbb{T}} \frac{h_2'(\tau)}{b_1 w - b_2 \tau} d\tau - b_2 w \fint_{\mathbb{T}} \frac{\bar{\tau} h_2'(\tau)}{b_1 w - b_2 \tau} d\tau \\
& + b_1 b_2 \fint_{\mathbb{T}} \frac{h_2(\tau)\, d\tau}{(b_1 w - b_2 \tau)^2} - b_2^2 w \fint_{\mathbb{T}} \frac{\bar{\tau} h_2(\tau)\, d\tau}{(b_1 w - b_2 \tau)^2} \Bigg\} \\
\triangleq -b_1 \operatorname{Im}\{J_1 + J_2 + J_3 + J_4\}. &
\end{aligned} \tag{37}$$

To compute the first term $J_1(w)$, we write after using the series expansion of $1/(1-(b_2/b_1)\overline{w}\tau)$

$$\begin{aligned}
J_1 &= \overline{w}\,\fint_{\mathbb{T}} \frac{h_2'(\tau)}{1-(b_2/b_1)\overline{w}\tau}\,d\tau \\
&= \sum_{n\geq 0}\left(\frac{b_2}{b_1}\right)^n \overline{w}^{n+1} \fint_{\mathbb{T}} \tau^n h_2'(\tau)\,d\tau.
\end{aligned}$$

Note that

$$\fint_{\mathbb{T}} \tau^n h_2'(\tau)\,d\tau = -na_{2,n},$$

which us enables to get

$$J_1 = -\sum_{n\geq 1} na_{2,n}\left(\frac{b_2}{b_1}\right)^n \overline{w}^{n+1}. \tag{38}$$

As for the term $J_2(w)$, we write in a similar way

$$\begin{aligned}
J_2 &= -\frac{b_2}{b_1}\fint_{\mathbb{T}} \frac{\bar{\tau}h_2'(\tau)}{1-(b_2/b_1)\overline{w}\tau}\,d\tau \\
&= -\sum_{n\geq 0}\left(\frac{b_2}{b_1}\right)^{n+1} \overline{w}^{n} \fint_{\mathbb{T}} \tau^{n-1} h_2'(\tau)\,d\tau.
\end{aligned}$$

Since $\fint_{\mathbb{T}} \tau^{-k} h_2'(\tau)\,d\tau = 0$ for $k\in\{0, 1\}$, the preceding sum starts at $n=2$ and by shifting the summation index

$$\begin{aligned}
J_2 &= -\sum_{n\geq 1}\left(\frac{b_2}{b_1}\right)^{n+2} \overline{w}^{n+1} \fint_{\mathbb{T}} \tau^{n} h_2'(\tau)\,d\tau \\
&= \sum_{n\geq 1} na_{2,n}\left(\frac{b_2}{b_1}\right)^{n+2} \overline{w}^{n+1}.
\end{aligned} \tag{39}$$

Concerning the third term J_3, we write by virtue of (27)

$$\begin{aligned}
J_3 &= \frac{b_2}{b_1}\overline{w}^2 \fint_{\mathbb{T}} \frac{h_2(\tau)}{(1-(b_2/b_1)\overline{w}\tau)^2}\,d\tau \\
&= \sum_{n\geq 1} n\left(\frac{b_2}{b_1}\right)^{n} \overline{w}^{n+1} \fint_{\mathbb{T}} \tau^{n-1} h_2(\tau)\,d\tau.
\end{aligned}$$

Therefore, we find

$$J_3 = \sum_{n\geq 1} na_{2,n}\left(\frac{b_2}{b_1}\right)^n \overline{w}^{n+1}. \tag{40}$$

Similarly, we get

$$\begin{aligned} J_4 &= -\left(\frac{b_2}{b_1}\right)^2 \overline{w} \fint_{\mathbb{T}} \frac{\overline{\tau} h_2(\tau)}{(1-(b_2/b_1)\overline{w}\tau)^2}\, d\tau \\ &= -\sum_{n\geq 1} n \left(\frac{b_2}{b_1}\right)^{n+1} \overline{w}^n \fint_{\mathbb{T}} \tau^{n-2} h_2(\tau)\, d\tau \\ &= -\sum_{n\geq 0} (n+1) a_{2,n} \left(\frac{b_2}{b_1}\right)^{n+2} \overline{w}^{n+1}. \end{aligned} \tag{41}$$

Inserting the identities (38), (39), (40) and (41) into (37), we find

$$\begin{aligned} \partial_{f_2} G_1^2(0,0) h_2(w) &= b_1 \operatorname{Im}\left\{ \sum_{n\geq 0} a_{2,n} \left(\frac{b_2}{b_1}\right)^{n+2} \overline{w}^{n+1} \right\} \\ &= b_1 \sum_{n\geq 0} a_{2,n} \left(\frac{b_2}{b_1}\right)^{n+2} e_{n+1}(w). \end{aligned} \tag{42}$$

Next, we shall move to the computation of $\partial_{f_1} G_2^2(0,0)h_1$. In view of (36),

$$\begin{aligned} \partial_{f_1} G_2^2(0,0) h_1(w) = b_2 \operatorname{Im}\bigg\{ w \fint_{\mathbb{T}} \frac{(b_2\overline{w} - b_1\overline{\tau})}{b_2 w - b_1 \tau} h_1'(\tau)\, d\tau - b_1 w \fint_{\mathbb{T}} \frac{\overline{h_1(\tau)}}{b_2 w - b_1\tau}\, d\tau \\ + b_1 w \fint_{\mathbb{T}} \frac{(b_2 \overline{w} - b_1 \overline{\tau}) h_1(\tau)\, d\tau}{(b_2 w - b_1 \tau)^2} \bigg\}. \end{aligned}$$

The residue theorem at infinity enables us to get rid of the first and third integrals in the right-hand side, and thus,

$$\partial_{f_1} G_2^2(0,0) h_1(w) = -b_1 b_2 \operatorname{Im}\left\{ w \fint_{\mathbb{T}} \frac{\overline{h_1(\tau)}}{b_2 w - b_1 \tau}\, d\tau \right\}.$$

A second application of the residue theorem in the disc yields

$$\begin{aligned} \partial_{f_1} G_2^2(0,0) h_1(w) &= b_2 \operatorname{Im}\left\{ w \overline{h_1}\left(\frac{b_2 w}{b_1}\right) \right\} \\ &= -b_2 \sum_{n\geq 0} a_{1,n} \left(\frac{b_2}{b_1}\right)^n e_{n+1}(w). \end{aligned} \tag{43}$$

• *Computation of* $\partial_{f_i} G_j^3(\lambda, 0, 0) h_i$. The diagonal terms $i = j$ can be easily computed:

$$\begin{aligned} \partial_{f_i} G_i^3(0,0) h_j(w) &= (-1)^{i+1} b_i^3 \operatorname{Im}\left\{ w \fint_{\mathbb{T}} \frac{h_i'(w)\, d\tau}{1 - b_i^2 w \tau} + w h_i(w) \fint_{\mathbb{T}} \frac{\tau\, d\tau}{(1 - b_i^2 w\tau)^2} \right\} \\ &= 0. \end{aligned} \tag{44}$$

Let us now calculate $\partial_{f_i} G_j^3(\lambda, 0, 0)h_i$ for $i \neq j$. One can check with difficulty that

$$\partial_{f_i} G_j^3(0,0)h_i(w) = (-1)^{j+1} b_j b_i^2 \operatorname{Im}\left\{ w \fint_{\mathbb{T}} \frac{h_i'(\tau)}{1 - b_i b_j w\tau}\, d\tau + 2w \fint_{\mathbb{T}} \frac{\operatorname{Re}\{\tau \overline{h_i(\tau)}\}}{1 - b_i b_j w\tau}\, d\tau \right.$$
$$\left. + b_i b_j w^2 \fint_{\mathbb{T}} \frac{h_i(\tau)\, d\tau}{(1 - b_i b_j w\tau)^2} \right\}.$$

Invoking once again the residue theorem, we find

$$\begin{aligned}\partial_{f_i} G_j^3(0,0)h_i(w) &= (-1)^{j+1} b_j b_i^2 \operatorname{Im}\Big\{ -\sum_{n\geq 0} n a_{i,n}(b_j b_i)^n w^{n+1} + \sum_{n\geq 0} a_{i,n}(b_j b_i)^n w^{n+1} \\ &\qquad + \sum_{n\geq 0} n a_{i,n}(b_j b_i)^n w^{n+1} \Big\} \\ &= (-1)^j b_i \sum_{n\geq 0} a_{i,n}(b_j b_i)^{n+1} e_{n+1}. \end{aligned} \tag{45}$$

The details are left to the reader because most of them were done previously. Now putting together the identities (32), (34) and (44),

$$\partial_{f_1} G_1(\lambda, 0, 0)h_1 = \sum_{n\geq 0} b_1 \left[(n+1)\lambda - 1 + b_1^{2n+2} - (n+1)\left(\frac{b_2}{b_1}\right)^2 \right] a_{1,n} e_{n+1}. \tag{46}$$

From (33), (35) and (44), one obtains

$$\partial_{f_2} G_2(\lambda, 0, 0)h_2 = \sum_{n\geq 0} b_2((n+1)\lambda - n - b_2^{2n+2}) a_{2,n} e_{n+1}. \tag{47}$$

On the other hand, we observe that for $i \neq j$

$$\partial_{f_i} G_j^1(\lambda, 0)h_i(w) = 0. \tag{48}$$

Gathering the identities (48), (42) and (45) yields

$$\partial_{f_2} G_1(\lambda, 0, 0)h_2 = \sum_{n\geq 0} b_2 \left[\left(\frac{b_2}{b_1}\right)^{n+1} - (b_1 b_2)^{n+1} \right] a_{2,n} e_{n+1}.$$

Furthermore, combining (48), (43) and (45), we can assert that

$$\partial_{f_1} G_2(\lambda, 0, 0)h_1 = \sum_{n\geq 0} b_1 \left[(b_1 b_2)^{n+1} - \left(\frac{b_2}{b_1}\right)^{n+1} \right] a_{1,n} e_{n+1}.$$

Consequently, we get in view of the last two expressions combined with (47) and (48)

$$DG(\lambda, 0, 0)(h_1, h_2) = \sum_{n\geq 0} M_{n+1} \begin{pmatrix} a_{1,n} \\ a_{2,n} \end{pmatrix} e_{n+1}, \tag{49}$$

where the matrix M_n is given for each $n \geq 1$ by

$$M_n \triangleq \begin{pmatrix} b_1[n\lambda - 1 + b_1^{2n} - n(b_2/b_1)^2] & b_2[(b_2/b_1)^n - (b_1 b_2)^n] \\ -b_1[(b_2/b_1)^n - (b_1 b_2)^n] & b_2[n\lambda - n + 1 - b_2^{2n}] \end{pmatrix}. \tag{50}$$

This completes the proof of Proposition 14. □

4.4. *Eigenvalues study.* The current subsection will be devoted to the study of the structure of the *nonlinear eigenvalues* which are the values λ such that the linearized operator $DG(\lambda, 0, 0)$ given by (49) has a nontrivial kernel. Note that these eigenvalues correspond exactly to matrices M_n which are not invertible for some integer $n \geq 1$. In other words, λ is an eigenvalue if and only if there exists $n \geq 1$ such that $\det M_n = 0$, that is,

$$\begin{aligned}\det M_n(\lambda) = b_1 b_2 \bigg[n^2\lambda^2 - n\Big(n + b_2^{2n} - b_1^{2n} + n\Big(\frac{b_2}{b_1}\Big)^2\Big)\lambda + (n-1)\Big(1 - b_1^{2n} + n\Big(\frac{b_2}{b_1}\Big)^2\Big) \\ + \Big(\frac{b_2}{b_1}\Big)^{2n} + n b_2^{2n}\Big(\frac{b_2}{b_1}\Big)^2 - b_2^{2n}\bigg] \\ = 0.\end{aligned}$$

This is equivalent to

$$\begin{aligned}P_n(\lambda) \triangleq \lambda^2 - \bigg[1 + \Big(\frac{b_2}{b_1}\Big)^2 - \Big(\frac{b_1^{2n} - b_2^{2n}}{n}\Big)\bigg]\lambda \\ + \Big(\frac{b_2}{b_1}\Big)^2 - \frac{1-(b_2/b_1)^{2n}}{n^2} + \frac{1-(b_2/b_1)^2}{n} - \frac{b_1^{2n} - b_2^{2n}(b_2/b_1)^2}{n} + \frac{b_1^{2n} - b_2^{2n}}{n^2} \\ = 0. \end{aligned} \tag{51}$$

The reduced discriminant of this second-degree polynomial in λ is given by

$$\Delta_n = \bigg(\frac{1-(b_2/b_1)^2}{2} - \frac{2 - b_2^{2n} - b_1^{2n}}{2n}\bigg)^2 - \Big(\frac{b_2}{b_1}\Big)^{2n}\Big(\frac{1-b_1^{2n}}{n}\Big)^2. \tag{52}$$

Thereby P_n admits two real roots if and only if $\Delta_n \geq 0$, and they are given by

$$\lambda_n^{\pm} = \frac{1+(b_2/b_1)^2}{2} - \Big(\frac{b_1^{2n} - b_2^{2n}}{2n}\Big) \pm \sqrt{\Delta_n}.$$

To understand the structure of the eigenvalues and their dependence on the involved parameters, it would be better to fix the radius b_1 and to vary n and $b_2 \in]0, b_1[$. We shall distinguish the cases $n \geq 2$ from $n = 1$, which is very special. For given $n \geq 2$, we wish to draw the curves $b_2 \mapsto \lambda_n^{\pm}(b_2)$. As we shall see in Proposition 19, the maximal domains of existence of these curves are a common connected set of the form $[0, b_n^\star]$ and $b_n^\star$ is defined as the unique $b_2 \in]0, b_1[$ such that $\Delta_n = 0$. We introduce the graphs $\mathscr{C}_n^{\pm}$ of $\lambda_n^{\pm}(b_2)$:

$$\mathscr{C}_n^{\pm} \triangleq \{(b_2, \lambda_n^{\pm}(b_2)) : b_2 \in [0, b_n^\star]\}, \quad \mathscr{C}_n = \mathscr{C}_n^- \cup \mathscr{C}_n^+, \quad n \geq 2. \tag{53}$$

It is not hard to check that $\mathscr{C}_n^+$ intersects $\mathscr{C}_n^-$ at only one point whose abscissa is $b_n^\star$, that is, when the discriminant vanishes. Furthermore, and this is not trivial, we shall see that the domain enclosed by the curve $\mathscr{C}_n$ and located in the first quadrant of the plane is a strictly increasing set on n. This will give in particular the monotonicity of the eigenvalues with respect to n. Nevertheless, the dynamics of the first eigenvalues corresponding to $n = 1$ is completely different from the preceding ones. Indeed, according to

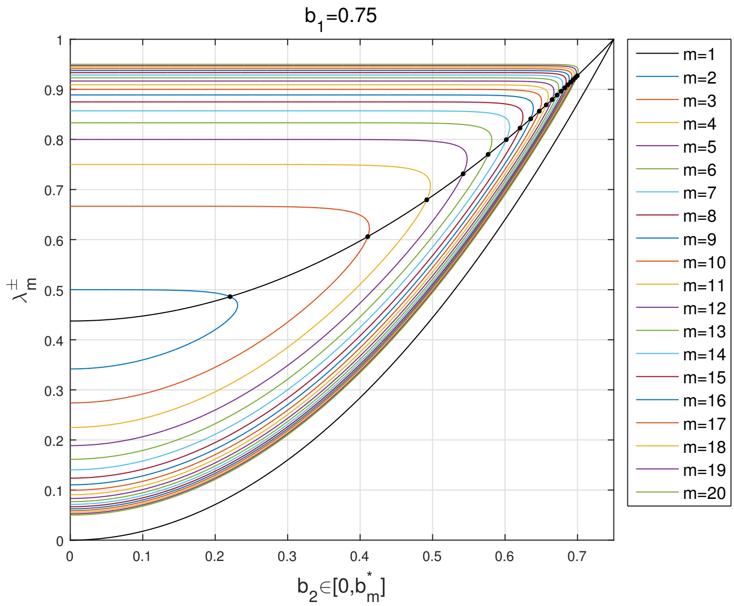


Figure 1. $\lambda_m^{\pm}$ as a function of $b_2 \in [0, b_m^\star]$, for $m = 2, \dots, 20$, together with the case $m = 1$ (black), for $b_1 = 0.75$.

Section 4.4.3, we find for $n = 1$ two eigenvalues given explicitly by

$$\lambda_1^- = (b_2/b_1)^2 \quad \text{or} \quad \lambda_1^+ = 1 + b_2^2 - b_1^2.$$

It turns out that for the first one the range of the linearized operator has an infinite codimension, and therefore, there is no hope to bifurcate using only the classical results of bifurcation theory. However, for the second eigenvalue, the range is "almost everywhere" of codimension 1 and the bifurcation is likely to happen. As for the structure of this eigenvalue, it is strictly increasing with respect to b_2, and by working more, we prove that the curve $\mathscr{C}_1^+$ of $b_2 \in]0, b_1[\mapsto \lambda_1^+$ intersects $\mathscr{C}_n$ if and only if $n \geq b_1^{-2}$. We can now make precise statements of these results, and for the complete ones, we refer the reader to Lemma 18 and Propositions 19 and 20.

Proposition 15. *Let* $b_1 \in]0, 1[$; *then the following hold true*:

(i) *The sequence* $n \geq 2 \mapsto b_n^\star$ *is strictly increasing.*

(ii) *Let* $2 \leq n < m$ *and* $b_2 \in [0, b_n^\star[$; *then*

$$\lambda_m^- < \lambda_n^- < \lambda_n^+ < \lambda_m^+.$$

(iii) *The curve* $\mathscr{C}_1^+$ *intersects* $\mathscr{C}_n$ *if and only if* $n \geq 1/b_1^2$. *In this case, we have a single point* $(x_n, \lambda_1^+(x_n))$, *with* $x_n \in]0, b_n^\star]$ *being the only solution* b_2 *of the equation*

$$P_n(1 + b_2^2 - b_1^2) = 0,$$

where P_n *is defined in* (51).

The properties mentioned in the preceding proposition can be illustrated by Figure 1. Further illustrations will be given in Figure 7.

For the proof of Proposition 15, it appears to be more convenient to work with a continuous variable instead of the discrete one n. This is advantageous especially in the study of the variations of the eigenvalues with respect to n and the radius b_2 for b_1 fixed. To do so, we extend in a natural way $(\Delta_n)_{n\geq 1}$ to a smooth function defined on $[1, +\infty[$ as

$$\Delta_x = \left(\frac{1-(b_2/b_1)^2}{2} - \frac{2-b_2^{2x}-b_1^{2x}}{2x}\right)^2 - \left(\frac{b_2}{b_1}\right)^{2x}\left(\frac{1-b_1^{2x}}{x}\right)^2, \quad x \in [1, +\infty[.$$

It is easy to see that Δ_x is positive if and only if

$$\left(1-\left(\frac{b_2}{b_1}\right)^2\right)x - (2-b_2^{2x}-b_1^{2x}) - 2\left(\frac{b_2}{b_1}\right)^x(1-b_1^{2x}) \geq 0 \tag{54}$$

or

$$E_x \triangleq \left(1-\left(\frac{b_2}{b_1}\right)^2\right)x - (2-b_2^{2x}-b_1^{2x}) + 2\left(\frac{b_2}{b_1}\right)^x(1-b_1^{2x}) < 0.$$

We shall prove that the last possibility $E_x < 0$ is excluded for $x \geq 2$. Indeed,

$$\begin{aligned}
E_x &= (1-(b_2/b_1)^2)x - 2(1-(b_2/b_1)^x) + (b_2^x - b_1^x)^2 \\
&= 2(1-(b_2/b_1)^2)\left[\frac{x}{2} - \frac{1-((b_2/b_1)^2)^{x/2}}{1-(b_2/b_1)^2}\right] + (b_2^x - b_1^x)^2 \\
&\geq (b_2^x - b_1^x)^2 > 0,
\end{aligned}$$

where we have used the classical inequality,

$$\text{for all } b \in (0, 1) \text{ and } x \geq 1, \quad \frac{1-b^x}{1-b} \leq x.$$

Thus, for $x \geq 2$, the condition $\Delta_x \geq 0$ is equivalent to the first one of (54) or, in other words,

$$x \geq \frac{2+2(b_2/b_1)^x - (b_1^x + b_2^x)^2}{1-(b_2/b_1)^2} \triangleq g_x(b_1, b_2). \tag{55}$$

In this case, the roots of the polynomial P_n can also be continuously extended as

$$\begin{aligned}
\lambda_x^+ &= \frac{1+(b_2/b_1)^2}{2} - \left(\frac{b_1^{2x}-b_2^{2x}}{2x}\right) + \sqrt{\Delta_x}, \\
\lambda_x^- &= \frac{1+(b_2/b_1)^2}{2} - \left(\frac{b_1^{2x}-b_2^{2x}}{2x}\right) - \sqrt{\Delta_x}.
\end{aligned}$$

4.4.1. *Monotonicity for $n \geq 2$.* To settle the proof of the second point (ii) of Proposition 15, we should look for the variations of the eigenvalues with respect to x but with fixed radii b_1 and b_2. For this purpose, we need to first understand the topological structure of the domain of definition of $x \mapsto \lambda_x^\pm$

$$\mathcal{I}_{b_1,b_2} \triangleq \{x \geq 2 : \Delta_x > 0\}$$

and see in particular whether this set is connected. We shall establish the following:

Lemma 16. *Let* $0 < b_2 < b_1 < 1$ *be two fixed numbers*; *then the following hold true*:

(i) *The set* $\mathcal{I}_{b_1,b_2}$ *is connected and of the form* $]\mu_{b_1,b_2}, \infty[$.

(ii) *The map* $x \in \mathcal{I}_{b_1,b_2} \mapsto \Delta_x$ *is strictly increasing.*

Remark 17. If the discriminant Δ_x admits a zero, then it is unique and coincides with the value μ_{b_1,b_2}. Otherwise, μ_{b_1,b_2} will be equal to 2.

Proof. To get this result, it suffices to check the following: for any $a \in \mathcal{I}_{b_1,b_2}$,

$$[a, +\infty[\subset \mathcal{I}_{b_1,b_2}.$$

By the continuity of the discriminant, there exists $\eta > a$ such that $[a, \eta[\subset \mathcal{I}_{b_1,b_2}$, and let $[a, \eta^\star[$ be the maximal interval contained in $\mathcal{I}_{b_1,b_2}$. If $\eta^\star$ is finite, then necessarily $\Delta_{\eta^\star} = 0$. If we could show that the discriminant is strictly increasing in this interval, then this will contradict the preceding assumption. To see this, observe that Δ_x can be rewritten in the form

$$\Delta_x = \frac{1}{4}\left(f_1\left(\frac{b_2}{b_1}\right) - f_x(b_1) - f_x(b_2)\right)^2 - \left(\frac{b_2}{b_1}\right)^{2x} f_x^2(b_1) \tag{56}$$

with the notation

$$f_x(t) \triangleq \frac{1-t^{2x}}{x}.$$

Differentiating Δ_x with respect to x,

$$\begin{aligned}\partial_x \Delta_x = -\tfrac{1}{2}(\partial_x f_x(b_1) + \partial_x f_x(b_2))\left(f_1\left(\frac{b_2}{b_1}\right) - f_x(b_1) - f_x(b_2)\right)\\ - 2 f_x(b_1)\left(\frac{b_2}{b_1}\right)^{2x}\left(f_x(b_1)\log\left(\frac{b_2}{b_1}\right) + \partial_x f_x(b_1)\right).\end{aligned} \tag{57}$$

We shall prove that, for all $t \in]0, 1[$, the mapping $x \in [2, \infty[\mapsto f_x(t)$ is strictly decreasing. It is clear that

$$\partial_x f_x(t) = \frac{t^{2x}(1 - 2x\log t) - 1}{x^2} \triangleq \frac{g_x(t)}{x^2}. \tag{58}$$

To study the variation of $t \mapsto g_x(t)$, note that

$$g_x'(t) = -4x^2 t^{2x-1}\log t > 0 \quad \text{for all } t \in]0, 1[$$

and therefore g_x is strictly increasing, which implies that

$$\partial_x f_x(t) < \frac{g_x(1)}{x^2} = 0.$$

Using this fact, we deduce that the last term of (57) is positive and consequently

$$\partial_x \Delta_x \geq -\tfrac{1}{2}(\partial_x f_x(b_1) + \partial_x f_x(b_2))\left(f_1\left(\frac{b_2}{b_1}\right) - f_x(b_1) - f_x(b_2)\right).$$

Hence, to get $\partial_x \Delta_x > 0$ it suffices to establish that

$$f_1\left(\frac{b_2}{b_1}\right) - f_x(b_1) - f_x(b_2) > 0, \tag{59}$$

which is equivalent to

$$x > \frac{2 - b_1^{2x} - b_2^{2x}}{1 - b^2}, \quad b = \frac{b_2}{b_1}.$$

Note that we have already seen that the positivity of Δ_x for $x \geq 2$ is equivalent to the condition (55) which actually implies the preceding one owing to the strict inequality

$$b^x - (b_1 b_2)^x > 0.$$

This shows that (59) is true and consequently,

$$\text{for all } x \in [a, \eta^\star[, \quad \partial_x \Delta_x > 0.$$

This shows that the discriminant, which is positive, is strictly increasing in $[a, \eta^\star[$, and this excludes the fact that $\Delta_{\eta^\star}$ vanishes. Therefore, $\eta^\star = \infty$, and thus, (i) and (ii) are simultaneously proved. ☐

The next goal is to establish the monotonicity of the eigenvalues.

Lemma 18. *Let* $0 < b_2 < b_1 < 1$. *Then*:

(i) *The mapping* $x \in \mathcal{I}_{b_1,b_2} \mapsto \lambda_x^+$ *is strictly increasing.*

(ii) *The mapping* $x \in \mathcal{I}_{b_1,b_2} \mapsto \lambda_x^-$ *is strictly decreasing.*

(iii) *For any* $x < y \in \mathcal{I}_{b_1,b_2}$,

$$\lambda_y^- < \lambda_x^- < \lambda_x^+ < \lambda_y^+.$$

Proof. (i) Note that

$$\lambda_x^+ = \frac{1+b^2}{2} - \frac{b_1^{2x}}{2} f_x(b) + \sqrt{\Delta_x}, \quad b = \frac{b_2}{b_1}.$$

We have already seen in the proof of Lemma 16 that for any $t \in]0, 1[$ the mapping $x \in [2, \infty[\mapsto f_x(t)$ is strictly decreasing, and therefore, $x \mapsto b_1^{2x} f_x(b_2/b_1)$ is also strictly decreasing. To get the strict increasing of $x \mapsto \lambda_x^+$, it suffices to combine this last fact with the increasing property of $x \mapsto \Delta_x$.

(ii) It is clear that

$$\lambda_x^- = \frac{1+b^2}{2} + \frac{f_x(b_1) - f_x(b_2)}{2} - \sqrt{\Delta_x}.$$

The derivative of λ_x^- with respect to x is given by

$$\partial_x \lambda_x^- = \tfrac{1}{2}\partial_x f_x(b_1) - \tfrac{1}{2}\partial_x f_x(b_2) - \frac{\partial_x \Delta_x}{2\sqrt{\Delta_x}}.$$

By virtue of (57), we can split the preceding function into three parts:

$$\partial_x \lambda_x^- = \mathrm{I} + \mathrm{II} + \mathrm{III},$$

where

$$\mathrm{I} \triangleq \tfrac{1}{2}\partial_x f_x(b_1)\left(1+\frac{f_1(b)-f_x(b_1)-f_x(b_2)}{2\sqrt{\Delta_x}}\right),$$

$$\mathrm{II} \triangleq \tfrac{1}{2}\partial_x f_x(b_2)\left(-1+\frac{f_1(b)-f_x(b_1)-f_x(b_2)}{2\sqrt{\Delta_x}}\right),$$

$$\mathrm{III} \triangleq \frac{b^{2x} f_x(b_1)(f_x(b_1)\log(b)+\partial_x f_x(b_1))}{\sqrt{\Delta_x}}.$$

Keeping in mind the inequality (59) and $\partial_x f_x(t) < 0$ for any $t \in]0, 1[$, we can see that I is negative. To prove that the term II is also negative, it suffices to check that

$$\frac{f_1(b)-f_x(b_1)-f_x(b_2)}{2\sqrt{\Delta_x}} > 1.$$

From (59), we can deduce by squaring that the last expression is actually equivalent to

$$\frac{1}{4}\left(f_1\left(\frac{b_2}{b_1}\right)-f_x(b_1)-f_x(b_2)\right)^2 > \Delta_x.$$

From (56), we immediately conclude that the last inequality is always verified.

In regard to the negativity of the third term III, we just use the fact that $0 < b < 1$ and the decreasing of the function $x \mapsto f_x(t)$.

(iii) This follows easily from (i), (ii) and the obvious fact,

$$\text{for all } x \in \mathscr{I}_{b_1,b_2}, \quad \lambda_x^- < \lambda_x^+. \qquad \square$$

4.4.2. *Lifespan of the eigenvalues with respect to b_2.* We shall study in this section some properties of the eigenvalue functions $b_2 \mapsto \lambda_n^\pm$ for $n \geq 2$ and b_1 fixed. This will be crucial for studying the dynamics of the first eigenvalue λ_1^+ and especially in counting the intersections between the curves $\mathscr{C}_1^+$ and $\mathscr{C}_n$ which has been the subject of the part (iii) of Proposition 15. Note that in this paragraph we shall give up using the continuous version $\lambda_x^\pm$ of the roots $\lambda_n^\pm$ as it has been done in the preceding section. The results that we shall state can actually be proved with the continuous parameter; however, this does not matter a lot for our final purpose. We define the following set: for $n \geq 2$ and $b_1 \in]0, 1[$,

$$\mathscr{I}_{n,b_1} \triangleq \left\{ b_2 \in [0, b_1[\, : n \geq \frac{2+2(b_2/b_1)^n-(b_1^n+b_2^n)^2}{1-(b_2/b_1)^2} \right\}.$$

We shall prove the following:

Proposition 19. *Let $b_1 \in]0, 1[$ fixed and $n \geq 2$; then the following hold true*:

(i) *The set $\mathscr{I}_{n,b_1}$ is an interval of the form $[0, b_n^\star]$, with $b_n^\star \in]0, b_1[$.*

(ii) *The eigenvalues $b_2 \mapsto \lambda_n^\pm$ are defined together in $[0, b_n^\star]$.*

(iii) *The sequence $n \mapsto b_n^\star$ is strictly increasing, and we have the asymptotics*

$$b_n^\star = b_1(1-\alpha/n)+o(1/n), \quad e^{-\alpha}+1 = \alpha, \ \alpha \approx 1.27846.$$

(iv) *The function* $b_2 \in [0, b_n^\star] \mapsto \lambda_n^-(b_2) - b_2^2$ *is strictly increasing.*

(v) *The function* $b_2 \in [0, b_n^\star] \mapsto \lambda_n^+(b_2) - b_2^2$ *is strictly decreasing.*

Proof. (i) This follows from studying the function $h : [0, b_1] \to \mathbb{R}$, defined by

$$h(x) = n(1-(x/b_1)^2) - 2 - 2(x/b_1)^n + (b_1^n + x^n)^2.$$

We claim that h is strictly decreasing. Indeed, by differentiating,

$$\begin{aligned} h'(x) &= \frac{2nx}{b_1^2}(-1+b_1^2 x^{2n-2}) + \frac{2nx^{n-1}}{b_1^n}(-1+b_1^{2n}) \\ &< 0. \end{aligned}$$

As $h(0) = n-2+b_1^{2n} > 0$ and $h(b_1) = 4(-1+b_1^{2n}) < 0$, we deduce from the intermediate value theorem that the set $\mathcal{J}_{n,b_1}$ is in fact an interval of the form $[0, b_n^\star]$. The number $b_n^\star \in [0, b_1[$ is defined by the unique solution of the equation

$$h(b_n^\star) = 0. \tag{60}$$

(ii) Observe that the domain of definition of the eigenvalues $\lambda_n^\pm$ coincides with the domain of the discriminant Δ_n, which is in turn given by $\mathcal{J}_{n,b_1}$ according to (55). Therefore, (60) implies the vanishing of Δ_n at the point $b_n^\star$, and consequently both eigenvalues coincide.

(iii) Recall from (53) the definitions of the curves $\mathscr{C}_n^\pm$ and $\mathscr{C}_n = \mathscr{C}_n^- \cup \mathscr{C}_n^+$. Since the eigenvalues $\lambda_n^+(b_n^\star)$ and $\lambda_n^-(b_n^\star)$ coincide, curves $\mathscr{C}_n^+$ and $\mathscr{C}_n^-$ end at the same point which is a turning point for $\mathscr{C}_n$. Furthermore, we can see that $\mathscr{C}_n$ lies on the left side of the vertical axis $x = b_n^\star$. Now let $m > n \geq 2$, and we intend to check by some elementary geometric considerations that $b_m^\star > b_n^\star$. From the monotonicity of the eigenvalues $n \mapsto \lambda_n^\pm$,

$$\lambda_m^-(0) < \lambda_n^-(0), \qquad \lambda_m^+(0) > \lambda_n^+(0).$$

If $b_m^\star \leq b_n^\star$, then the curve $\mathscr{C}_m$ will intersect $\mathscr{C}_n$ at some point and this contradicts the strict monotonicity of the eigenvalues with respect to n. Thus, we deduce that $n \mapsto b_n^\star$ is strictly increasing and therefore should converge to some value $b^\star \leq b_1$. Assume that $b^\star < b_1$; then from (60) and the continuity of h, we find by letting $n \to +\infty$ that

$$\lim_{n\to+\infty} h(b_n^\star) = 0.$$

On the other hand,

$$\begin{aligned} \lim_{n\to+\infty} h(b_n^\star) &= \lim_{n\to+\infty} n(1-(b_n^\star/b_1)^2) - 2 \\ &= +\infty, \end{aligned}$$

which is clearly a contradiction, and thus, $b^\star = b_1$. For the asymptotic behavior of $b_n^\star$, which is a marginal part here, we shall settle for a formal reasoning by taking a first-order Taylor expansion of $1/n$. We shall look for α such that

$$b_n^\star = b_1(1-\alpha/n) + o(1/n).$$

At the first order of h,

$$h(b_n^\star) = \alpha(2-\alpha/n) - 2 - 2(1-\alpha/n)^n + o(1).$$

By taking the limit as $n \to \infty$, we find that α must satisfy

$$e^{-\alpha} + 1 = \alpha.$$

This equation admits a unique solution lying in the interval $]1, 2[$ and can be given explicitly by the Lambert W function:

$$\alpha = W(e^{-1}) + 1 \approx 1.27846.$$

(iv) Set $x = (b_2/b_1)^2$ and define the functions

$$f_\pm(x) = \lambda_n^\pm(b_2) = \frac{1+x}{2} + \frac{b_1^{2n}}{2n}(x^n - 1) \pm \sqrt{\Delta_n(x)}, \quad x \in \left[0, \frac{b_n^{\star 2}}{b_1^2}\right],$$

with

$$\Delta_n(x) = \left(\frac{1-x}{2} - \frac{2-b_1^{2n}(1+x^n)}{2n}\right)^2 - x^n\left(\frac{1-b_1^{2n}}{n}\right)^2.$$

Differentiating with respect to x yields

$$\Delta_n'(x) = -\left(\frac{1-x}{2} - \frac{2-b_1^{2n}(1+x^n)}{2n}\right)(1-b_1^{2n}x^{n-1}) - nx^{n-1}\left(\frac{1-b_1^{2n}}{n}\right)^2.$$

Note from the assumption (55), by switching the parameters n and x, that

$$\frac{1-x}{2} - \frac{2-b_1^{2n}(1+x^n)}{2n} > 0,$$

and therefore,

$$\Delta_n'(x) < 0 \quad \text{for all } x \in \left[0, \frac{b_n^{\star 2}}{b_1^2}\right] \subset [0, 1[.$$

Coming back to the function $f_\pm$ and taking the derivative, we find

$$f_\pm'(x) = \frac{1}{2} + \frac{b_1^{2n}}{2}x^{n-1} \pm \frac{\Delta_n'(x)}{2\sqrt{\Delta_n(x)}}.$$

Using the definition of Δ_n and (54), one has

$$\left(\frac{1-x}{2} - \frac{2-b_1^{2n}(1+x^n)}{2n}\right) > \sqrt{\Delta_n(x)}$$

and consequently

$$\begin{aligned}\frac{\Delta_n'(x)}{\sqrt{\Delta_n(x)}} &\le -\frac{(1-x)/2 - (2-b_1^{2n}(1+x^n))/2n}{\sqrt{\Delta_n(x)}}(1-b_1^{2n}x^{n-1}) \\ &< -(1-b_1^{2n}x^{n-1}).\end{aligned}$$

Therefore, we obtain that for all $x \in [0, b_n^{\star 2}/b_1^2]$

$$\begin{aligned} &f_-'(x) > 1, \\ &f_+'(x) \le b_1^{2n} x^{n-1} < b_1^2. \end{aligned}$$

This shows that the function $g_- : x \mapsto f_-(x) - b_1^2 x$ is strictly increasing; however, $g_+ : x \mapsto f_+(x) - b_1^2 x$ is strictly decreasing. This finishes the proof of the desired result. □

4.4.3. *Dynamics of the first eigenvalue.* We shall in this paragraph discuss the behavior of the first eigenvalues corresponding to $n = 1$. Note from (51) that these eigenvalues are in fact the solutions of the polynomial

$$P_1(\lambda) = \lambda^2 - (1 + b_2^2 - b_1^2 + (b_2/b_1)^2)\lambda + (b_2/b_1)^2 + b_2^2(b_2/b_1)^2 - b_2^2,$$

which vanishes exactly at the points

$$\lambda_1^- = (b_2/b_1)^2 \quad \text{or} \quad \lambda_1^+ = 1 + b_2^2 - b_1^2.$$

Recall from the preceding sections the definition

$$\mathscr{C}_n^\pm \triangleq \{(b_2, \lambda_n^\pm(b_2)) : b_2 \in [0, b_n^\star]\}, \quad \mathscr{C}_n = \mathscr{C}_n^- \cup \mathscr{C}_n^+,$$

and the graph of the first eigenvalue λ_1^+ is given by

$$\mathscr{C}_1^+ \triangleq \{(b_2, 1 + b_2^2 - b_1^2) : b_2 \in [0, b_1]\}.$$

As we have already mentioned, it is not clear whether the bifurcation occurs with λ_1^- because the range of the linearized operator has an infinite codimension. The main result reads as follows.

Proposition 20. *Let $b_1 \in]0, 1[$ and $n \ge 2$. Then the following hold true*:

(i) *For any $0 < b_2 < b_1$, we have $\lambda_1^- < \lambda_n^\pm$.*

(ii) *If $n < b_1^{-2}$, then*

$$\mathscr{C}_n \cap \mathscr{C}_1^+ = \varnothing.$$

(iii) *If $n \ge b_1^{-2}$, then $\mathscr{C}_n \cap \mathscr{C}_1^+$ is a single point, that is, there exists $x_n \in [0, b_n^\star]$ such that*

$$\mathscr{C}_n \cap \mathscr{C}_1^+ = \{(x_n, \lambda_1^+(x_n))\}.$$

(iv) *If $b_2 \notin \{x_m : m \ge b_1^{-2}\}$, then for all $n \ge 2$, $\lambda_1^+ \ne \lambda_n^\pm$.*

(v) *The sequence $\{x_m\}_{m \ge b_1^{-2}}$ is increasing and converges to b_1.*

Proof. (i) This follows easily from the monotonicity of the eigenvalue $n \mapsto \lambda_n^-$ and the fact that $\lambda_n^- \le \lambda_n^+$. Indeed, for all $n \ge 2$,

$$\lambda_1^- = (b_2/b_1)^2 = \lim_{n \to +\infty} \lambda_n^- < \lambda_n^- \le \lambda_n^+.$$

(ii) In view of (v) from Proposition 19, the mapping $b_2 \in [0, b_n^\star] \mapsto \lambda_n^+(b_2) - \lambda_1^+(b_2)$ is strictly decreasing, and therefore, for $b_2 \in]0, b_n^\star]$,

$$\lambda_n^+(b_2) - \lambda_1^+(b_2) < \lambda_n^+(0) - \lambda_1^+(0) = b_1^2 - \frac{1}{n}.$$

Therefore, for $n < b_1^{-2}$, the last term in the right-hand side is negative and consequently

$$\lambda_n^-(b_2) \le \lambda_n^+(b_2) < \lambda_1^+(b_2) \quad \text{for all } b_2 \in]0, b_n^\star].$$

(iii) When $n \ge b_1^{-2}$, then $\lambda_n^+(0) - \lambda_1^+(0) \ge 0$, and since $b_2 \in [0, b_n^\star] \mapsto \lambda_n^+(b_2) - \lambda_1^+(b_2)$ is strictly decreasing, the equation $\lambda_n^+(b_2) - \lambda_1^+(b_2) = 0$ has at most one solution in $[0, b_n^\star]$. We shall distinguish three cases. The first one is when $\lambda_n^+(b_n^\star) - \lambda_1^+(b_n^\star) < 0$, in which case the foregoing equation admits a unique solution denoted by x_n. This implies that $\mathscr{C}_n^+ \cap \mathscr{C}_1^+$ is a single point whose abscissa is x_n, and the next step is to check that $\mathscr{C}_n^- \cap \mathscr{C}_1^+$ is empty. Thus,

$$\lambda_n^+(b_n^\star) - \lambda_1^+(b_n^\star) \le \lambda_n^+(x_n) - \lambda_1^+(x_n) = 0.$$

Combining the last inequality with the fact that $\lambda_n^+(b_n^\star) = \lambda_n^-(b_n^\star)$ and the monotonicity of the mapping $b_2 \in [0, b_n^\star] \mapsto \lambda_n^-(b_2) - \lambda_1^+(b_2)$, which follows from (iv) of Proposition 19, we conclude that for all $b_2 \in]0, b_n^\star]$

$$\begin{aligned} \lambda_n^-(b_2) - \lambda_1^+(b_2) &\le \lambda_n^-(b_n^\star) - \lambda_1^+(b_n^\star) \\ &\le \lambda_n^+(b_n^\star) - \lambda_1^+(b_n^\star) \\ &< 0. \end{aligned}$$

Therefore, $\mathscr{C}_n^- \cap \mathscr{C}_1^+ = \varnothing$ and the set $\mathscr{C}_n \cap \mathscr{C}_1^+$ reduces to a single point. The second case is when $\lambda_n^+(b_n^\star) - \lambda_1^+(b_n^\star) > 0$; then $\mathscr{C}_n^+ \cap C_1^+$ is empty, and we shall prove that $\mathscr{C}_n^- \cap \mathscr{C}_1^+$ is a single point. Observe first that

$$\lambda_n^-(b_n^\star) - \lambda_1^+(b_n^\star) = \lambda_n^+(b_n^\star) - \lambda_1^+(b_n^\star) > 0.$$

Moreover,

$$\lambda_n^-(0) - \lambda_1^+(0) = \frac{1 - b_1^{2n}}{n} - (1 - b_1^2) < 0 \quad \text{for all } n \ge 2.$$

Since $b_2 \mapsto \lambda_n^-(b_2) - \lambda_1^+(b_2)$ is strictly increasing, by the intermediate value theorem, there exists only one solution $x_n \in]0, b_n^\star[$ of the equation $\lambda_n^-(b_2) - \lambda_1^+(b_2) = 0$. The third and last case to analyze is when $\lambda_n^+(b_n^\star) - \lambda_1^+(b_n^\star) = 0$. This means that all the curves $\mathscr{C}_n^+$, $\mathscr{C}_n^-$ and $\mathscr{C}_1^+$ meet each other at the single point of abscissa $b_n^\star$.

(iv) It follows immediately from (ii) and (iii).

(v) Let $n \ge b_1^{-1}$, and define the set enclosed by $\mathscr{C}_n$ and located at the first quadrant of the plane:

$$\widehat{\mathscr{C}}_n \triangleq \{(x, y) \in \mathbb{R}^2 : x \in [0, b_n^\star],\ \lambda_n^-(x) \le y \le \lambda_n^+(x)\}.$$

From the monotonicity of the eigenvalues $n \mapsto \lambda_n^{\pm}$ seen in Lemma 18, we note that,

$$\text{for all } (x, y) \in \widehat{\mathscr{C}}_n, \quad \lambda_{n+1}^-(x) < \lambda_n^-(x) \le y \le \lambda_n^+(x) < \lambda_{n+1}^+(x).$$

Hence,

$$\widehat{\mathscr{C}}_n \Subset \widehat{\mathscr{C}}_{n+1}, \qquad \mathscr{C}_{n+1} \cap \widehat{\mathscr{C}}_n = \varnothing. \tag{61}$$

Now, from (iii) and the monotonicity of the mappings $b_2 \mapsto \lambda_n^{\pm}(b_2) - \lambda_1^+(b_2)$ stated in Proposition 19, we deduce that,

$$\text{for all } x \in [0, x_n[, \quad \lambda_n^-(x) < \lambda_1^+(x) < \lambda_n^+(x).$$

Then we have the inclusion

$$\mathscr{C}_{1,n}^+ \triangleq \{(x, \lambda_1^+(x)) : x \in [0, x_n]\} \subset \widehat{\mathscr{C}}_n.$$

It follows from (61) that $\mathscr{C}_{n+1} \cap \mathscr{C}_{1,n}^+ = \varnothing$ and consequently the abscissa of the single point intersection $\mathscr{C}_{n+1} \cap \mathscr{C}_1^+$ must satisfy $x_{n+1} > x_n$. This proves that $\{x_n\}_{n \ge b_1^{-2}}$ is strictly increasing, and thereby this sequence converges to some value $x_\star \le b_1$. Assume that $x_\star < b_1$, and define the subsequences

$$\{x_n^{\pm}\}_{n \ge b_1^{-2}} \triangleq \{x_n : \lambda_n^{\pm}(x_n) = \lambda_1^+(x_n)\}.$$

Clearly one of the two sequences is infinite. Assume first that $\{x_n^+\}$ is infinite and up to an extraction this sequence converges also to $x_\star$, and for simplicity, we still denote this sequence by $\{x_n\}_{n \ge b_1^{-2}}$. Then from the definition of λ_n^+, we can easily check that

$$\begin{aligned}\lim_{n \to +\infty} \lambda_n^+(x_n) &= \frac{1 + (x_\star/b_1)^2}{2} + \frac{1 - (x_\star/b_1)^2}{2} \\ &= 1.\end{aligned}$$

On the other hand,

$$\lim_{n \to +\infty} \lambda_1^+(x_n) = 1 + x_\star^2 - b_1^2.$$

This is possible only if $x_\star = b_1$, which is a contradiction, and thus, $x_\star = b_1$. Now in the case where only the sequence $\{x_n^-\}$ is infinite, then we follow the same reasoning as before. We suppose that $x_\star < b_1$, and one can verify that

$$\begin{aligned}\lim_{n \to +\infty} \lambda_n^-(x_n) &= (x_\star/b_1)^2, \\ \lim_{n \to +\infty} \lambda_1^+(x_n) &= 1 + x_\star^2 - b_1^2.\end{aligned}$$

By equating these numbers, we obtain

$$(1 - b_1^2)(x_\star^2 - b_1^2) = 0,$$

which is impossible since $b_1 < 1$ and consequently $x_\star = b_1$. Hence, the proof of (v) is finished. $\square$

4.5. *Bifurcation for $m \geq 1$.* Now we shall see how to implement the preceding results to prove Theorems 6 and 9 by using the Crandall–Rabinowitz theorem. The proofs will be broken into several steps. First, we introduce the spaces of bifurcation which capture the m-fold symmetry, and they are of Hölderian type. Second, we rewrite Proposition 13 dealing with the regularity of the nonlinear functional defining the V-states in the new setting. We end this section with the proofs of the properties of the linearized operator around the annulus required by the Crandall–Rabinowitz theorem.

4.5.1. *Function spaces.* We shall make use of the same spaces as [de la Hoz et al. 2016b]. For $m \geq 1$, we introduce the spaces X_m and Y_m as follows:

$$X_m = C_m^{1+\alpha}(\mathbb{T}) \times C_m^{1+\alpha}(\mathbb{T}),$$

where $C_m^{1+\alpha}(\mathbb{T})$ is the space of the 2π-periodic functions $f \in C^{1+\alpha}(\mathbb{T})$ whose Fourier series is given by

$$f(w) = \sum_{n=1}^{\infty} a_n \overline{w}^{nm-1}, \quad w \in \mathbb{T},\ a_n \in \mathbb{R}.$$

This space is equipped with the usual strong topology of $C^{1+\alpha}(\mathbb{T})$. We can easily see that X_m is identified as

$$X_m = \left\{ f \in (C^{1+\alpha}(\mathbb{T}))^2 : f(w) = \sum_{n=1}^{\infty} A_n \overline{w}^{nm-1},\ A_n \in \mathbb{R}^2 \right\}. \tag{62}$$

We define the ball of radius $r \in (0,1)$ by

$$B_r^m = \{ f \in (C_m^{1+\alpha}(\mathbb{T}))^2 : \|f\|_{C^{1+\alpha}(\mathbb{T})} < r \}.$$

Take $(f_1, f_2) \in B_r^m$; then the expansions of the associated conformal mappings ϕ_1 and ϕ_2 in the exterior unit disc $\{w \in \mathbb{C} : |w| > 1\}$ are given by

$$\phi_1(w) = b_1 w + f_1(w) = w\left(b_1 + \sum_{n=1}^{\infty} \frac{a_{1,n}}{w^{nm}} \right),$$

$$\phi_2(w) = b_2 w + f_2(w) = w\left(b_2 + \sum_{n=1}^{\infty} \frac{a_{2,n}}{w^{nm}} \right).$$

This captures the m-fold symmetry of the associated boundaries $\phi_1(\mathbb{T})$ and $\phi_2(\mathbb{T})$ via the relation

$$\phi_j(e^{2i\pi/m} w) = e^{2i\pi/m} \phi_j(w), \quad j = 1, 2,\ w \in \mathbb{T}. \tag{63}$$

Set

$$Y_m = \left\{ g \in (C^{\alpha}(\mathbb{T}))^2 : g = \sum_{n \geq 1} C_n e_{nm},\ C_n \in \mathbb{R}^2 \right\}. \tag{64}$$

With the help of Proposition 13, we deduce that the functional $G = (G_1, G_2)$ is well defined and smooth from $\mathbb{R} \times B_r^m$ to Y_m with r small enough. The only thing that one should care about, which has already been discussed in the simply connected case, is the persistence of the symmetry which comes from the rotational invariance of the functional G. As the proofs are very close to the simply connected case without any substantial difficulties, we prefer to skip them and only state the desired results.

Proposition 21. *Let $b \in]0,1[$ and $0 < r < \min(b, 1-b)$; then the following hold true*:

(i) $G : \mathbb{R} \times B_r^m \to Y_m$ *is* C^1 (*it is in fact* C^∞).

(ii) *The partial derivative* $\partial_\lambda DG : \mathbb{R} \times B_r^m \to \mathscr{L}(X_m, Y_m)$ *exists and is continuous* (*it is in fact* C^∞).

Now using (49) and (50), we deduce that the restriction of $DG(\lambda, 0)$ to the space X_m leads to a well defined continuous operator $DG(\lambda, 0) : X_m \to Y_m$. It takes the form

$$DG(\lambda,0)(h_1,h_2) = \sum_{n\geq 1} M_{nm}(\lambda) \begin{pmatrix} a_{1,n} \\ a_{2,n} \end{pmatrix} e_{nm}, \tag{65}$$

with $(h_1, h_2) \in X_m$ having the expansion

$$h_j(w) = \sum_{n\geq 1} a_{j,n} \overline{w}^{nm-1}$$

and the matrix M_n given for $n \geq 1$ by

$$M_n(\lambda) \triangleq \begin{pmatrix} b_1[n\lambda - 1 + b_1^{2n} - n(b_2/b_1)^2] & b_2[(b_2/b_1)^n - (b_1b_2)^n] \\ -b_1[(b_2/b_1)^n - (b_1b_2)^n] & b_2[n\lambda - n + 1 - b_2^{2n}] \end{pmatrix}. \tag{66}$$

4.5.2. *Proof of Theorem 6.* The main goal of this paragraph is to prove Theorem 6. This will be an immediate consequence of the Crandall–Rabinowitz theorem as soon as we check its conditions, which require a careful study. Concerning the regularity assumptions, they were discussed in Proposition 21. As to the properties required for the linearized operator, they are the object of following proposition.

Proposition 22. *Let $0 < b_2 < b_1 < 1$, and set $b \triangleq b_2/b_1$. Let $m \geq 2$ satisfy*

$$m \geq \frac{2 + 2b^m - (b_1^m + b_2^m)^2}{1 - b^2}.$$

Then the following results hold true:

(i) *The kernel of $DG(\lambda_m^\pm, 0)$ is one-dimensional and generated by the vector*

$$v_m(w) = \begin{pmatrix} b_2[m\lambda_m^\pm - m + 1 - b_2^{2m}] \\ b_1[b^m - (b_1b_2)^m] \end{pmatrix} \overline{w}^{m-1}.$$

(ii) *The range of $DG(\lambda_m^\pm, 0)$ is closed and of codimension* 1.

(iii) *The transversality assumption holds*: *the condition*

$$\partial_\lambda DG(\lambda_m^\pm, 0) v_m \notin R(DG(\lambda_m^\pm, 0))$$

is satisfied if and only if

$$m > \frac{2 + 2b^m - (b_1^m + b_2^m)^2}{1 - b^2}.$$

Proof. (i) According to (55), the positivity of the discriminant Δ_n that guarantees the existence of real eigenvalues is equivalent for $m \geq 2$ to

$$m \geq \frac{2 + 2b^m - (b_1^m + b_2^m)^2}{1 - b^2}.$$

To prove that the kernel of $DG(\lambda_m^\pm, 0)$ is one-dimensional, it suffices to check that for $n \geq 2$ the matrix $M_{nm}(\lambda_m^\pm)$ defined in (66) is invertible. This follows from Lemma 18, which asserts that $\lambda_{nm}^\pm \neq \lambda_m^\pm$ for $n \geq 2$ and therefore

$$\det M_{nm}(\lambda_m^\pm) \neq 0.$$

To get a generator for the kernel, it suffices to take a vector orthogonal to the second row of $M_m(\lambda_m^\pm)$.

(ii) We are going to show that for any $m \geq 2$ the range $R(DG(\lambda_m^\pm, 0))$ coincides with the subspace

$$\mathscr{Z}_m \triangleq \left\{ g \in Y_m : g(w) = \sum_{n \geq 1} C_n e_{nm},\ C_1 \in R(M_m),\ C_n \in \mathbb{R}^2 \text{ for all } n \geq 2 \right\}. \tag{67}$$

Assume for now this result; then it is easy to check that $R(DG(\lambda_m^\pm, 0))$ is closed in Y_m and is of codimension 1. Now to get the description of the range, we first observe that from (65) and (66) the range is included in the space $\mathscr{Z}_m$. Therefore, what is left is to check is the inclusion $\mathscr{Z}_m \subset R(DG(\lambda_m^\pm, 0))$. Take $g = (g_1, g_2) \in \mathscr{Z}_m$ with the form

$$g_j(w) = \sum_{n \geq 1} c_{j,n} e_{nm},$$

and let us prove that the equation

$$DG(\lambda_m^\pm, 0)h = g$$

admits a solution $h = (h_1, h_2)$ in the space X_m. Note that h_j has the structure

$$h_j(w) = \sum_{n \geq 1} a_{j,n} \overline{w}^{nm-1}.$$

According to (65), the preceding equation is equivalent to

$$M_{mn} \begin{pmatrix} a_{1,n} \\ a_{2,n} \end{pmatrix} = \begin{pmatrix} c_{1,n} \\ c_{2,n} \end{pmatrix} \quad \text{for all } n \geq 1.$$

For $n = 1$, this equation is satisfied because from the definition of $\mathscr{Z}_m$ we assume that the vector $C_1 \triangleq \binom{c_{1,n}}{c_{2,n}}$ belongs to the range of the matrix M_m. With regard to $n \geq 2$, we use the fact that M_{nm} is invertible, and therefore, the sequences $(a_{j,n})_{n \geq 2}$ are uniquely determined by

$$\begin{pmatrix} a_{1,n} \\ a_{2,n} \end{pmatrix} = M_{nm}^{-1} \begin{pmatrix} c_{1,n} \\ c_{2,n} \end{pmatrix}, \quad n \geq 2. \tag{68}$$

By computing the matrix $M_{mn}^{-1}(\lambda_m^{\pm})$, we deduce that for all $n \geq 2$

$$\begin{aligned} a_{1,n} &= \frac{b_2[nm(\lambda_m^{\pm}-1)+1-b_2^{2nm}]}{\det(M_{nm}(\lambda_m^{\pm}))}c_{1,n} - \frac{b_2[(b_2/b_1)^{nm}-(b_1b_2)^{nm}]}{\det(M_{nm}(\lambda_m^{\pm}))}c_{2,n}, \\ a_{2,n} &= \frac{b_1[(b_2/b_1)^{nm}-(b_1b_2)^{nm}]}{\det(M_{nm}(\lambda_m^{\pm}))}c_{1,n} + \frac{b_1[nm(\lambda_m^{\pm}-(b_2/b_1)^2)-1+b_1^{2nm}]}{\det(M_{nm}(\lambda_m^{\pm}))}c_{2,n}. \end{aligned} \tag{69}$$

Hence, the proof of $(h_1, h_2) \in X_m$ amounts to showing that

$$w \mapsto \begin{pmatrix} h_1(w) - a_{1,1}\overline{w}^{m-1} \\ h_2(w) - a_{2,1}\overline{w}^{m-1} \end{pmatrix} \in C^{1+\alpha}(\mathbb{T}) \times C^{1+\alpha}(\mathbb{T}).$$

We shall develop the computations only for the first component, and the second one can be done in a similar way. Notice that $\det(M_{nm}(\lambda_m^{\pm}))$ does not vanish for $n \geq 2$ and behaves for large n like

$$\det(M_{nm}(\lambda_m^{\pm})) = b_1b_2m^2(\lambda_m^{\pm}-1)[\lambda_m^{\pm}-(b_2/b_1)^2]n^2 + b_1b_2m(1-(b_2/b_1)^2)n - 1 + o(1).$$

Since $\lambda_m^{\pm} \notin \{1, (b_2/b_1)^2\}$, by Taylor expansion,

$$a_{1,n} = \frac{1}{b_1 m(\lambda_m^{\pm}-(b_2/b_1)^2)}\frac{c_{1,n}}{n} + \gamma_{1,n}c_{1,n} + \gamma_{2,n}c_{2,n}$$

with

$$|\gamma_{j,n}| \leq \frac{C}{n^2}.$$

Set $\tilde{h}_1(w) = h_1(w) - a_{1,1}\overline{w}^{m-1}$, and define the functions

$$K_j(w) = \sum_{n\geq 2} n\gamma_{j,n}\overline{w}^{nm}, \qquad \tilde{g}_j = \sum_{n\geq 2}\frac{c_{j,n}}{n}e_{nm}.$$

Then one can check that

$$\overline{w}\tilde{h}_1(w) = \frac{1}{mb_1(\lambda_m^{\pm}-(b_2/b_1)^2)}\sum_{n\geq 2}\frac{c_{1,n}}{n}\overline{w}^{nm} + \{K_1 \star (\Pi\tilde{g}_1)\}(w) + \{K_2 \star (\Pi\tilde{g}_2)\}(w). \tag{70}$$

The convolution is understood to be the usual one: for two continuous functions $f, g : \mathbb{T} \to \mathbb{C}$, we define,

$$\text{for all } w \in \mathbb{T}, \quad f \star g(w) = \fint_{\mathbb{T}} f(\tau)g(\tau\overline{w})\,\frac{d\tau}{\tau}.$$

The notation Π is used for the Szegő projection defined by

$$\Pi\Big(\sum_{n\in\mathbb{Z}} c_n w^n\Big) = \sum_{n\in-\mathbb{N}} c_n w^n,$$

which acts continuously on $C^{1+\alpha}(\mathbb{T})$. One can easily see that the first term in the right-hand side of (70) belongs to $C^{1+\alpha}(\mathbb{T})$. With regard to the last two terms, note that $K_j \in L^2(\mathbb{T}) \subset L^1(\mathbb{T})$ and $\tilde{g}_j \in C^{1+\alpha}(\mathbb{T})$; then using the classical convolution law $L^1(\mathbb{T}) \star C^{1+\alpha}(\mathbb{T}) \to C^{1+\alpha}(\mathbb{T})$ combined with the continuity of Π, we deduce that those terms belong to $C^{1+\alpha}(\mathbb{T})$ and the function $w \mapsto \overline{w}\tilde{h}_1(w)$ belongs to this space too. This finishes the proof of the range of $DG(\lambda_m^{\pm}, 0)$.

(iii) Recall from part (i) that the kernel of $DG(\lambda_m^{\pm}, 0)$ is one-dimensional and generated by the vector v_m defined by

$$w \in \mathbb{T} \mapsto v_m(w) = \begin{pmatrix} b_2[m\lambda_m^{\pm} - m + 1 - b_2^{2m}] \\ b_1[(b_2/b_1)^m - (b_1b_2)^m] \end{pmatrix} \overline{w}^{m-1}.$$

We shall prove that

$$\partial_\lambda DG(\lambda_m^{\pm}, 0)v_m \notin R(DG(\lambda_m^{\pm}, 0))$$

if and only if $\lambda_m^+ \neq \lambda_m^-$, which is equivalent to

$$m > \frac{2 + 2b^m - (b_1^m + b_2^m)^2}{1 - b^2}.$$

Let $(h_1, h_2) \in X_m$ with the expansion

$$h_j(w) = \sum_{n\geq 1} a_{j,n}\overline{w}^{nm-1}.$$

Then differentiating (65) with respect to λ,

$$\partial_\lambda DG(\lambda, 0)(h_1, h_2) = m \sum_{n\geq 1} n \begin{pmatrix} b_1a_{1,n} \\ b_2a_{2,n} \end{pmatrix} e_{nm}. \tag{71}$$

Hence,

$$\begin{aligned} \partial_\lambda DG(\lambda_m^{\pm}, 0)v_m &= mb_1b_2 \begin{pmatrix} m\lambda_m^{\pm} - m + 1 - b_2^{2m} \\ (b_2/b_1)^m - (b_1b_2)^m \end{pmatrix} e_m \\ &\triangleq mb_1b_2\mathbb{W}_m e_m. \end{aligned}$$

This pair of functions is in the range of $DG(\lambda_m^{\pm}, 0)$ if and only if the vector $\mathbb{W}_m$ is a scalar multiple of the second column of the matrix $M_m(\lambda_m^{\pm})$ defined by (66). This happens if and only if

$$(m\lambda_m^{\pm} - m + 1 - b_2^{2m})^2 - \left(\left(\frac{b_2}{b_1}\right)^m - (b_1b_2)^m\right)^2 = 0. \tag{72}$$

Combining this equation with $\det M_m = 0$, we find

$$(m\lambda_m^{\pm} - m + 1 - b_2^{2m})^2 + (m\lambda_m^{\pm} - m + 1 - b_2^{2m})\left(m\lambda_m^{\pm} - 1 + b_1^{2m} - m\left(\frac{b_2}{b_1}\right)^2\right) = 0,$$

which is equivalent to

$$(m\lambda - m + 1 - b_2^{2m})\left(2m\lambda - m\left(1 + \left(\frac{b_2}{b_1}\right)^2\right) - b_2^{2m} + b_1^{2m}\right) = 0.$$

Thus, we find that

$$m\lambda_m^{\pm} - m + 1 - b_2^{2m} = 0 \quad \text{or} \quad 2m\lambda_m^{\pm} - m\left(1 + \left(\frac{b_2}{b_1}\right)^2\right) - b_2^{2m} + b_1^{2m} = 0.$$

The first possibility is excluded by (72), and the second one corresponds to a multiple eigenvalue condition: $\lambda_m^+ = \lambda_m^-$, that is, $\Delta_m = 0$. This completes the proof of Proposition 22. □

4.5.3. *Proof of Theorem 9.* Our next task is to study the bifurcation of 1-fold rotating patches. Recall from Section 4.4.3 that for $m=1$ there are two different eigenvalues given by

$$\lambda_1^- = (b_2/b_1)^2, \qquad \lambda_1^+ = 1 + b_2^2 - b_1^2.$$

In that paragraph, we observed significant differences in their behaviors, and we shall see next how this fact does affect the bifurcation problem. It appears that the bifurcation with λ_1^- is very complicate due to the range of the linearized operator which is of infinite codimension. Nevertheless, with λ_1^+, the situation is actually more tractable and the bifurcation occurs frequently. Before stating the basic results of this section, we need to define some notation. Let $b_1 \in]0, 1[$ be a fixed real number, and define the set

$$\mathscr{E}_{b_1} \triangleq \{b_2 \in]0, b_1[: \text{there exists } m \geq 2 \text{ such that } P_m(\lambda_1^+) = 0\}.$$

The polynomial P_m was defined in (51), which is up to a factor the characteristic polynomial of the matrix $M_m(\lambda)$. The set $\mathscr{E}_{b_1}$ corresponds to the abscissa of the points of intersection between the collection of the curves $\{\mathscr{C}_m : m \geq 2\}$ and $\mathscr{C}_1^+$, which were defined in (53). Recall from Proposition 20(ii–iii) that for each $m \geq 2$ there is at most one value x_m of b_2 such that $P_m(\lambda_1^+) = 0$. Moreover, the sequence $(x_m)_{m \geq b_1^{-2}}$ is strictly increasing and converges to b_1. Now we will prove the following result.

Proposition 23. *The following assertions hold true.*

(i) *The range of $DG(\lambda_1^-, 0)$ has an infinite codimension.*

(ii) *If $b_2 \in \mathscr{E}_{b_1}$, then the kernel of $DG(\lambda_1^+, 0)$ is two-dimensional and generated by the vectors $v_1 = \binom{1}{1}$ and v_m of Proposition 22, with $m \geq 2$ being the only integer such that $P_m(\lambda_1^+) = 0$. In addition, the range of $DG(\lambda_1^+, 0)$ is closed and has codimension 2.*

(iii) *If $b_2 \notin \mathscr{E}_{b_1}$, then the kernel of $DG(\lambda_1^+, 0)$ is one-dimensional and is generated by the vector v_1 seen before. Furthermore, the range of $DG(\lambda_1^+, 0)$ has codimension 1 and the transversality assumption is satisfied:*

$$\partial_\lambda DG(\lambda_1^+, 0) v_1 \notin R(DG(\lambda_1^+, 0)).$$

Proof. (i) According to (66), we obtain

$$M_n(\lambda_1^-) \triangleq \begin{pmatrix} b_1[-1 + b_1^{2n}] & b_2[(b_2/b_1)^n - (b_1 b_2)^n] \\ -b_1[(b_2/b_1)^n - (b_1 b_2)^n] & b_2[n((b_2/b_1)^n - 1) + 1 - b_2^{2n}] \end{pmatrix}.$$

In this case, we get that the determinant of $M_n(\lambda_1^-)$ behaves for large n like $b_1 b_2 n$. Consequently, we deduce from (69) the existence of $\alpha \neq 0$ such that

$$a_{1,n} = \alpha c_{1,n} + o(1),$$

which means that the preimage of an element of Y_m by $DG(\lambda_1^-, 0)$ is not in general better than $C^\alpha(\mathbb{T})$. This implies that the range of the linearized operator is of infinite codimension. It follows that one important condition of the Crandall–Rabinowitz theorem is violated, and therefore, the bifurcation in this special case still unsolved.

(ii) Let $b_2 \in \mathscr{E}_{b_1}$. Then by definition, there exists $m \geq 2$ such that $P_m(\lambda_1^+) = 0$. This means that λ_1^+ coincides with one of the two numbers $\lambda_m^\pm$. Therefore, the kernel of $DG(\lambda_1^+, 0)$ is given by the two-dimensional vector space

$$\operatorname{Ker} DG(\lambda_1^+, 0) = \operatorname{Ker} M_1(\lambda_1^+) \oplus \operatorname{Ker} M_m(\lambda_1^+)\overline{w}^{m-1}.$$

Easy computations give the expression

$$M_1(\lambda_1^+) = b_2(1-b_1^2)\begin{pmatrix} -b_2/b_1 & b_2/b_1 \\ -1 & 1 \end{pmatrix}.$$

Obviously the kernel of $M_1(\lambda_1^+)$ is spanned by the vector $v_1 = \binom{1}{1}$. However, we know that $\operatorname{Ker} M_m(\lambda_1^+)$ is spanned by the vector v_m already seen in Proposition 22. To prove that the range is of codimension 2, we follow the same arguments of Proposition 22 bearing in mind that the determinant of $M_n(\lambda_1^+)$ behaves for large n like cn^2 with $c \neq 0$. We skip the details which are left to the reader.

(iii) Let $b_2 \notin \mathscr{E}_{b_1}$; then $P_m(\lambda_1^+)$ does not vanish for any $m \geq 2$. This means that the matrix $M_m(\lambda_1^+)$ is invertible, and therefore, the kernel of $DG(\lambda_1^+, 0)$ is one-dimensional and given by

$$\operatorname{Ker} DG(\lambda_1^+, 0) = \operatorname{Ker} M_1(\lambda_1^+) = \langle v_1 \rangle.$$

Similarly to Proposition 22, we get that the range is of codimension 1. In addition, the transversality condition is satisfied since the eigenvalue λ_1^+ is simple ($\lambda_1^+ \neq \lambda_1^-$) as has been discussed in the proof of Proposition 22(iii). The proof of Proposition 23 is now finished, and the result of Theorem 9 follows. □

5. Numerical experiments

In order to obtain the V-states, we follow a similar procedure to that in [de la Hoz et al. 2016a; 2016b]; therefore, we shall omit some details, which can be consulted in those references.

5.1. *Simply connected V-states.*

5.1.1. *Numerical derivation.* Given a simply connected domain D with boundary $z(\theta)$, where $\theta \in [0, 2\pi[$ is the Lagrangian parameter and z is counterclockwise parametrized, the condition of D being a V-state rotating with angular velocity Ω is given by (15), i.e.,

$$\operatorname{Re}\left\{\left(2\Omega\overline{z(\theta)} + \frac{1}{2\pi i}\int_0^{2\pi} \frac{\overline{z(\theta)-z(\phi)}}{z(\theta)-z(\phi)} z_\phi(\phi)\,d\phi - \frac{1}{2\pi i}\int_0^{2\pi} \frac{|z(\phi)|^2}{1-\overline{z(\theta)}z(\phi)} z_\phi(\phi)\,d\phi\right) z_\theta(\theta)\right\} = 0. \quad (73)$$

As in [de la Hoz et al. 2016a; 2016b], we use a pseudospectral method to find m-fold V-states from (73). We discretize $\theta \in [0, 2\pi[$ in N equally spaced nodes $\theta_i = 2\pi i/N$, $i = 0, 1, \ldots, N-1$. Observe that the integrand in the first integral in (73) satisfies

$$\lim_{\phi\to\theta} \left.\frac{\overline{z(\theta)-z(\phi)}}{z(\theta)-z(\phi)}\right|_{\theta=\phi} = \frac{\overline{z_\theta(\theta)}}{z_\theta(\theta)}. \quad (74)$$

Therefore, bearing in mind (74), we can evaluate numerically with spectral accuracy the integrals in (73) at a node $\theta = \theta_i$ by means of the trapezoidal rule, provided that N is large enough:

$$\begin{aligned}
&\frac{1}{2\pi}\int_0^{2\pi}\frac{\overline{z(\theta_i)-z(\phi_j)}}{z(\theta_i)-z(\phi_j)}z_\phi(\phi_j)\,d\phi \approx \frac{1}{N}\Bigg(\overline{z_\theta(\theta_i)}+\sum_{\substack{j=0\\ j\neq i}}^{N-1}\frac{\overline{z(\theta_i)-z(\phi_j)}}{z(\theta_i)-z(\phi_j)}z_\phi(\phi_j)\Bigg),\\
&\frac{1}{2\pi}\int_0^{2\pi}\frac{|z(\phi)|^2}{1-z(\theta_i)z(\phi)}z_\phi(\phi)\,d\phi \approx \frac{1}{N}\sum_{j=0}^{N-1}\frac{|z(\phi_j)|^2}{1-z(\theta_i)z(\phi_j)}z_\phi(\phi_j).
\end{aligned}\tag{75}$$

In order to obtain m-fold V-states, we approximate the boundary z as

$$z(\theta) = e^{i\theta}\Bigg[b + \sum_{k=1}^{M} a_k\cos(mk\theta)\Bigg],\tag{76}$$

where the mean radius is b, and we are imposing that $z(-\theta) = \bar{z}(\theta)$; i.e., we are looking for V-states symmetric with respect to the x-axis. For sampling purposes, N has to be chosen such that $N \geq 2mM+1$; additionally, it is convenient to take N a multiple of m, in order to be able to reduce the N-element discrete Fourier transforms to N/m-element discrete Fourier transforms. If we write $N = m2^r$, then $M = \lfloor (m2^r - 1)/(2m)\rfloor = 2^{r-1}-1$.

We introduce (76) into (73) and approximate the error in (73) by an M-term sine expansion:

$$\begin{aligned}
\operatorname{Re}\Bigg\{\Bigg(2\Omega\overline{z(\theta)} + \frac{1}{2\pi i}\int_0^{2\pi}\frac{\overline{z(\theta)-z(\phi)}}{z(\theta)-z(\phi)}z_\phi(\phi)\,d\phi - \frac{1}{2\pi i}\int_0^{2\pi}\frac{|z(\phi)|^2}{1-z(\theta)z(\phi)}z_\phi(\phi)\,d\phi\Bigg)z_\theta(\theta)\Bigg\}&\\
\approx \sum_{k=1}^{M} b_k\sin(mk\theta).&
\end{aligned}\tag{77}$$

This last expression can be represented in a very compact way as

$$\mathcal{F}_{b,\Omega}(a_1,\ldots,a_M) = (b_1,\ldots,b_M)\tag{78}$$

for a certain $\mathcal{F}_{b,\Omega} : \mathbb{R}^M \to \mathbb{R}^M$. Remark that, for any Ω and any $b \in]0,1[$, we trivially have $\mathcal{F}_{b,\Omega}(\mathbf{0}) = \mathbf{0}$, i.e., the circumference of radius b is a solution of the problem. Therefore, obtaining a simply connected V-state is reduced to numerically finding a nontrivial root $(a_1,\ldots,a_M)$ of (78). To do so, we discretize the $(M\times M)$-dimensional Jacobian matrix $\mathcal{J}$ of $\mathcal{F}_{b,\Omega}$ using first-order approximations. Fixing $|h| \ll 1$ (we have chosen $h = 10^{-10}$), we have that

$$\frac{\partial\mathcal{F}_{b,\Omega}(a_1,\ldots,a_M)}{\partial a_1} \approx \frac{\mathcal{F}_{b,\Omega}(a_1+h,\ldots,a_M) - \mathcal{F}_{b,\Omega}(a_1,\ldots,a_M)}{h}.\tag{79}$$

Hence, the first M coefficients of the sine expansion of (79) form the first row of $\mathcal{J}$, and so on. Therefore, if the n-th iteration is denoted by $(a_1,\ldots,a_M)^{(n)}$, then the $(n+1)$-th iteration is given by

$$(a_1,\ldots,a_M)^{(n+1)} = (a_1,\ldots,a_M)^{(n)} - \mathcal{F}_{b,\Omega}((a_1,\ldots,a_M)^{(n)})\cdot[\mathcal{J}^{(n)}]^{-1},$$

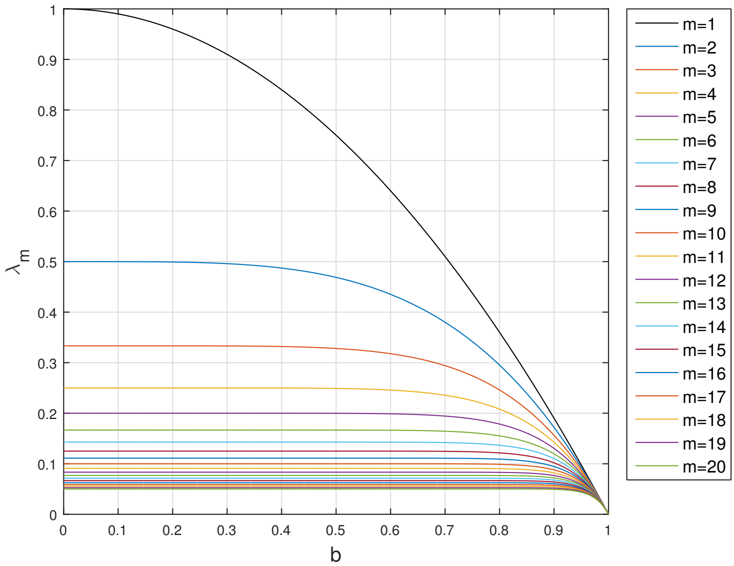


Figure 2. λ_m as a function of b, for $m = 1, \ldots, 20$.

where $[\mathcal{J}^{(n)}]^{-1}$ denotes the inverse of the Jacobian matrix at $(a_1, \ldots, a_M)^{(n)}$. This iteration converges in a small number of steps to a nontrivial root for a large variety of initial data $(a_1, \ldots, a_M)^{(0)}$. In particular, it is usually enough to perturb the unit circumference by assigning a small value to $a_1^{(0)}$ and leave the other coefficients equal to zero. Our stopping criterion is

$$\max\left|\sum_{k=1}^{M} b_k \sin(mk\theta)\right| < \text{tol},$$

where $\text{tol} = 10^{-13}$. For the sake of coherence, we eventually change the sign of all the coefficients $\{a_k\}$, in order for, without loss of generality, $a_1 > 0$.

5.1.2. *Numerical discussion.* Given m and b, Proposition 14 defines the value λ_m at which we bifurcate from the circumference of radius b. Let us recall that $\lambda_m = 1 - 2\Omega_m$. Although working with λ is more convenient from an analytical point of view, we use $\Omega = (1-\lambda)/2$ in the graphical representations of the V-states that follow because Ω is a more natural parameter from a physical point of view. Therefore, we bifurcate at $\Omega_m = (m - 1 + b^{2m})/(2m)$.

In Figure 2, we have plotted λ_m as a function of b, for $m = 1, \ldots, 20$. Figure 2 suggests that there are two different situations: b *close* to 1 and b *not so close* to 1. Note that, in the latter case, the curves can be approximated by $\lambda_m \approx 1/m$, i.e., $\Omega_m \approx (m-1)/(2m)$, which is in agreement with [Deem and Zabusky 1978].

In order to illustrate how the shape of the simply connected V-states depends on b, we consider the cases $1 \le m \le 4$; observe that everything said for $m = 3$ and $m = 4$ is valid for all $m \ge 3$. In general, fixing m and b, we bifurcate from the circumference with radius b at Ω_m. During the bifurcation process, there may be saddle-node bifurcation points [Kielhöfer 2012] appearing; in that case, we use the techniques

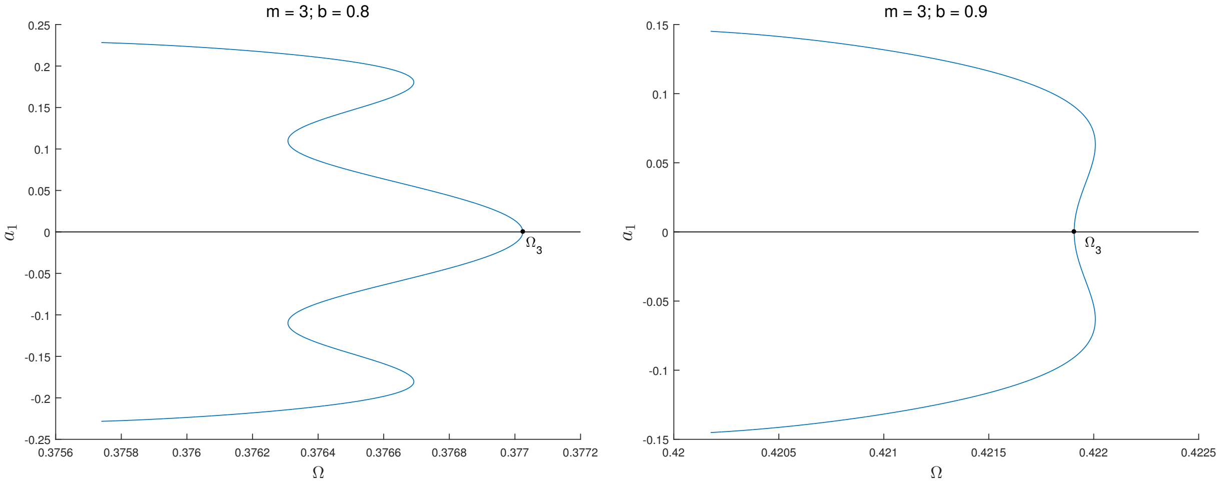


Figure 3. Bifurcation diagrams corresponding to $m = 3$ and $b = 0.8$ (left) and to $m = 3$ and $b = 0.9$ (right) with $N = 384$.

described in [de la Hoz et al. 2016a]. For instance, in Figure 3, we have plotted the bifurcation diagrams of the coefficient a_1 in (76) against Ω, for $m = 3$ and $b = 0.8$ (left) and for $m = 3$ and $b = 0.9$ (right). Note that, in the bifurcation diagrams, when starting to bifurcate at Ω_m, we sometimes take $\Omega < \Omega_m$ (left) and other times $\Omega > \Omega_m$ (right) although the latter case may appear only when b is *large enough*. Note also that we may have several saddle-node bifurcation points in the same bifurcation diagram, and hence more than two V-states corresponding to the same Ω, and in the same bifurcation branch. For instance, the left-hand side of Figure 3 tells us that there are three V-states corresponding to $m = 3$, $b = 0.8$ and $\Omega = 0.3765$, which we have plotted in Figure 4.

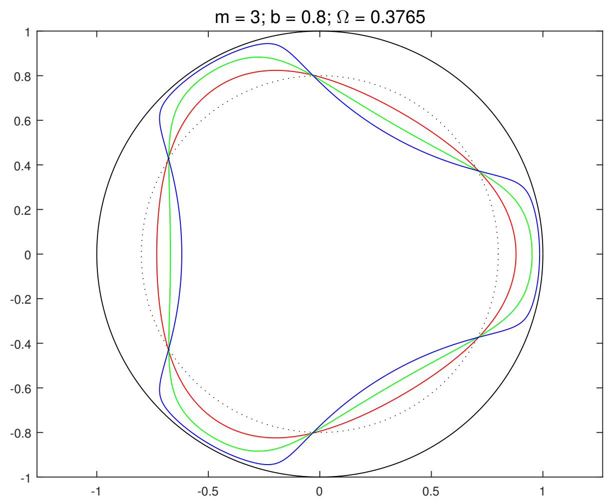


Figure 4. V-states from the same bifurcation branch (left side of Figure 3) corresponding to $m = 3$, $b = 0.8$ and $\Omega = 0.3765$ with $N = 768$.

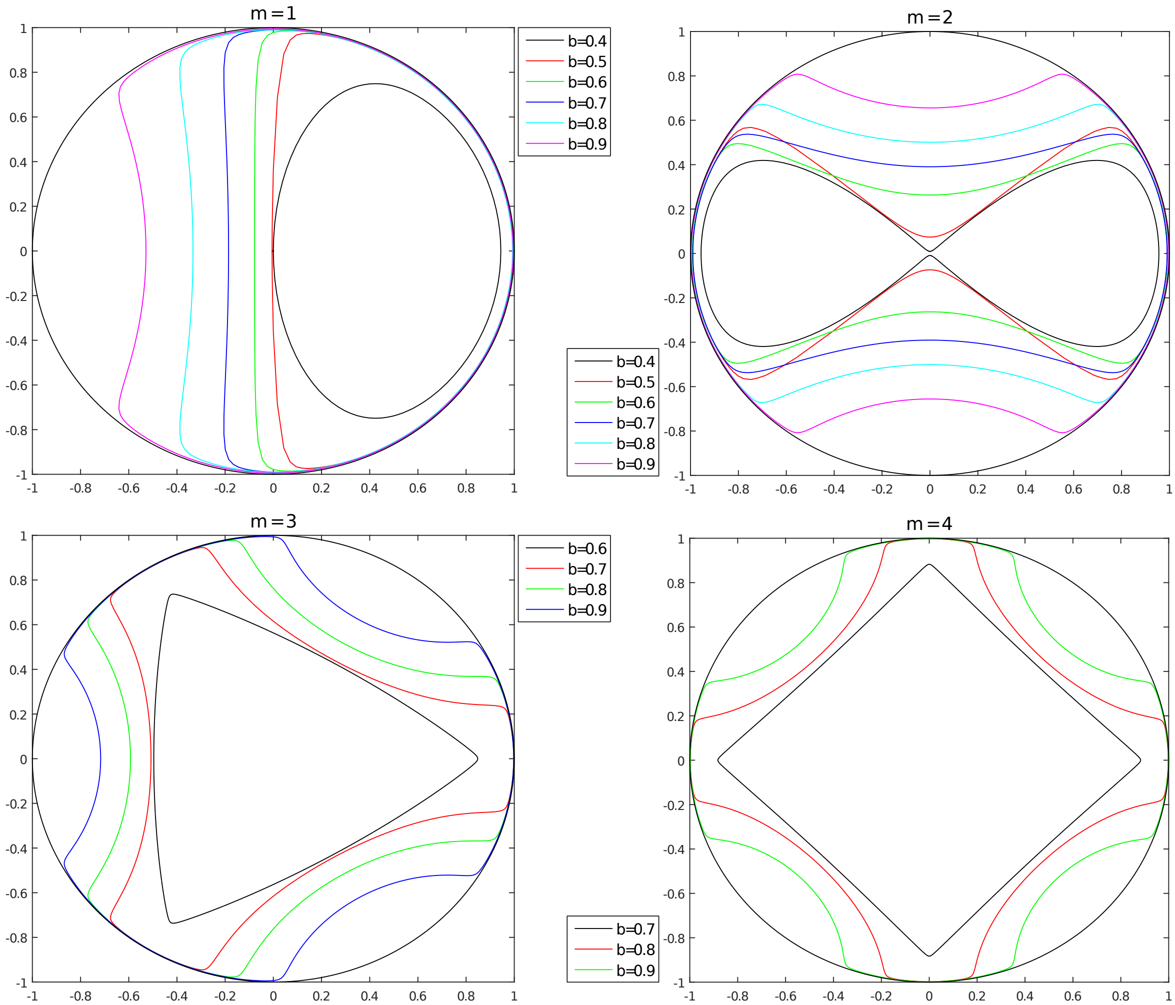


Figure 5. Approximations to the limiting V-states corresponding to $1 \leq m \leq 4$, for different b with $N = 256 \times m$. The values of Ω corresponding to the plots are given in Table 1.

We have approximated the limiting V-states occurring for $1 \leq m \leq 4$, which are depicted in Figure 5. Figure 5 confirms the observation on the size of b made from Figure 2. Loosely speaking, when b is *far enough* from 1, the rigid boundary does not have any remarkable effect on the shape of the V-states. Take for instance the cases $m = 1$ with $b = 0.4$, $m = 2$ with $b = 0.4$, $m = 3$ with $b = 0.6$ and $m = 4$ with $b = 0.7$: the approximations to the respective limiting V-states are clearly *far away* from the unit circumference whereas, in all the other cases, the distance to the unit circumference is smaller than 10^{-2}. In fact, Figure 5 suggests that, from a certain b on, we can obtain V-states arbitrarily close to the unit circumference and that the limiting V-state is precisely the one whose distance to the unit circumference is zero in the limit. Moreover, as b grows towards 1, the limiting V-states tend to cover an increasingly larger part of the unit circumference.

$b \downarrow \quad m \rightarrow$	1	2	3	4
0.9	0.3749	0.4057	0.4199	0.4283
0.8	0.3251	0.3589	0.3755	0.3859
0.7	0.2900	0.3163	0.3321	0.3650
0.6	0.2640	0.2731	0.3144	0.3572
0.5	0.2459	0.2363		
0.4	0.1964	0.2018		

Table 1. Values of Ω for the V-states plotted in Figure 5.

Continuing with Figure 5, the cases $m = 1$ and $m = 2$ are pretty different from the other cases. Indeed, when $m \geq 3$ and b is small enough, the limiting V-states very closely resemble those in [Deem and Zabusky 1978] and corner-shaped singularities seem to develop. It is remarkable that the rigid boundary only affects the shape of the V-states for b pretty close to 1; furthermore, the larger m is, the larger b has to be, in order for the influence of the rigid boundary to become noticeable. On the other hand, when $m = 2$ and b is small enough, the limiting V-states are lemniscate-shaped; whether some self-intersection actually occurs deserves further study. Finally, when $m = 1$ and b is small enough, the limiting V-states seem to resemble an asymmetrical oval.

5.2. *Doubly connected V-states.*

5.2.1. *Numerical derivation.* Given a doubly connected domain D with outer boundary $z_1(\theta)$ and inner boundary $z_2(\theta)$, where $\theta \in [0, 2\pi[$ is the Lagrangian parameter and z_1 and z_2 are parametrized, D is a V-state if and only if its boundaries satisfy

$$\begin{aligned}\operatorname{Re}\bigg\{\bigg(2\Omega\overline{z_1(\theta)} + \frac{1}{2\pi i}\int_0^{2\pi}\frac{\overline{z_1(\theta)-z_1(\phi)}}{z_1(\theta)-z_1(\phi)}z_{1,\phi}(\phi)\,d\phi - \frac{1}{2\pi i}\int_0^{2\pi}\frac{\overline{z_1(\theta)-z_2(\phi)}}{z_1(\theta)-z_2(\phi)}z_{2,\phi}(\phi)\,d\phi \\ - \frac{1}{2\pi i}\int_0^{2\pi}\frac{|z_1(\phi)|^2}{1-z_1(\theta)z_1(\phi)}z_{1,\phi}(\phi)\,d\phi \\ + \frac{1}{2\pi i}\int_0^{2\pi}\frac{|z_2(\phi)|^2}{1-z_1(\theta)z_2(\phi)}z_{2,\phi}(\phi)\,d\phi\bigg)z_{1,\theta}(\theta)\bigg\} = 0, \quad (80)\end{aligned}$$

$$\begin{aligned}\operatorname{Re}\bigg\{\bigg(2\Omega\overline{z_2(\theta)} + \frac{1}{2\pi i}\int_0^{2\pi}\frac{\overline{z_2(\theta)-z_1(\phi)}}{z_2(\theta)-z_1(\phi)}z_{1,\phi}(\phi)\,d\phi - \frac{1}{2\pi i}\int_0^{2\pi}\frac{\overline{z_2(\theta)-z_2(\phi)}}{z_2(\theta)-z_2(\phi)}z_{2,\phi}(\phi)\,d\phi \\ - \frac{1}{2\pi i}\int_0^{2\pi}\frac{|z_1(\phi)|^2}{1-z_2(\theta)z_1(\phi)}z_{1,\phi}(\phi)\,d\phi \\ + \frac{1}{2\pi i}\int_0^{2\pi}\frac{|z_2(\phi)|^2}{1-z_2(\theta)z_2(\phi)}z_{2,\phi}(\phi)\,d\phi\bigg)z_{2,\theta}(\theta)\bigg\} = 0. \quad (81)\end{aligned}$$

As in the simply connected case, we use a pseudospectral method to find V-states. We discretize $\theta \in [0, 2\pi[$ in N equally spaced nodes $\theta_i = 2\pi i/N$, $i = 0, 1, \ldots, N-1$, where N has to be large enough. Then since z_1 and z_2 never intersect, all the integrals in (80) and (81) can be evaluated numerically with spectral accuracy at a node $\theta = \theta_i$ by means of the trapezoidal rule, exactly as in (75).

In order to obtain doubly connected m-fold V-states, we approximate z_1 and z_2 as in (76):

$$z_1(\theta)=e^{i\theta}\left[b_1+\sum_{k=1}^{M}a_{1,k}\cos(mk\theta)\right],\qquad z_2(\theta)=e^{i\theta}\left[b_2+\sum_{k=1}^{M}a_{2,k}\cos(mk\theta)\right],\tag{82}$$

where the mean outer and inner radii are b_1 and b_2, respectively, and we are imposing that $z_1(-\theta)=\bar{z}_1(\theta)$ and $z_2(-\theta)=\bar{z}_2(\theta)$, i.e., looking for V-states symmetric with respect to the x-axis. Again, if we choose N of the form $N=m2^r$, then $M=\lfloor(m2^r-1)/(2m)\rfloor=2^{r-1}-1$.

We introduce (82) into (80) and (81), and as in (77), we approximate the errors in (80) and (81) by their M-term sine expansions, which are respectively $\sum_{k=1}^{M}b_{1,k}\sin(mk\theta)$ and $\sum_{k=1}^{M}b_{2,k}\sin(mk\theta)$. Then as in (78), the resulting systems of equations can be represented in a very compact way as

$$\mathcal{F}_{b_1,b_2,\Omega}(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})=(b_{1,1},\dots,b_{1,M},b_{2,1},\dots,b_{2,M})\tag{83}$$

for a certain $\mathcal{F}_{b_1,b_2,\Omega}:\mathbb{R}^{2M}\to\mathbb{R}^{2M}$. Remark that, for any Ω and any $0<b_2<b_1<1$, we have $\mathcal{F}_{b_1b_2,\Omega}(\mathbf{0})=\mathbf{0}$ trivially; i.e., any circular annulus is a solution of the problem. Therefore, obtaining a doubly connected V-state is reduced to numerically finding $\{a_{1,k}\}$ and $\{a_{2,k}\}$ such that $(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})$ is a nontrivial root of (83). To do so, we discretize the $(2M\times 2M)$-dimensional Jacobian matrix $\mathcal{J}$ of $\mathcal{F}_{b_1,b_2,\Omega}$ as in (79), taking $h=10^{-9}$:

$$\begin{aligned}&\frac{\partial\mathcal{F}_{b_1,b_2,\Omega}(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})}{\partial a_{1,1}}\\&\quad\approx\frac{\mathcal{F}_{b_1,b_2,\Omega}(a_{1,1}+h,a_{1,2},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})-\mathcal{F}_{b_1,b_2,\Omega}(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})}{h}.\end{aligned}\tag{84}$$

Then the sine expansion of (84) gives us the first row of $\mathcal{J}$, and so on. Hence, if the n-th iteration is denoted by $(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})^{(n)}$, then the $(n+1)$-th iteration is given by

$$\begin{aligned}&(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})^{(n+1)}\\&\quad=(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})^{(n)}-\mathcal{F}_{b_1,b_2,\Omega}((a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})^{(n)})\cdot[\mathcal{J}^{(n)}]^{-1},\end{aligned}$$

where $[\mathcal{J}^{(n)}]^{-1}$ denotes the inverse of the Jacobian matrix at $(a_{1,1},\dots,a_{1,M},a_{2,1},\dots,a_{2,M})^{(n)}$. To make this iteration converge, it is usually enough to perturb the annulus by assigning a small value to $a_{1,1}^{(0)}$ or $a_{2,1}^{(0)}$ and leave the other coefficients equal to zero. Our stopping criterion is

$$\left(\max\left|\sum_{k=1}^{M}b_{1,k}\sin(mk\theta)\right|<\mathrm{tol}\right)\wedge\left(\max\left|\sum_{k=1}^{M}b_{2,k}\sin(mk\theta)\right|<\mathrm{tol}\right),$$

where $\mathrm{tol}=10^{-13}$. As in [de la Hoz et al. 2016a; 2016b], $a_{1,1}\cdot a_{2,1}<0$, so for the sake of coherence, we eventually change the sign of all the coefficients $\{a_{1,k}\}$ and $\{a_{2,k}\}$, in order for, without loss of generality, $a_{1,1}>0$ and $a_{2,1}<0$.

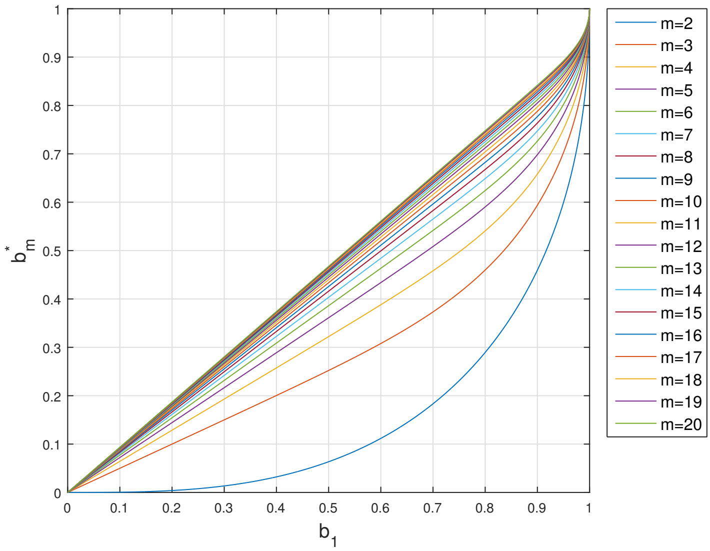


Figure 6. $b_m^\star$ as a function of b_1, for $m = 2, \dots, 20$.

5.2.2. *Numerical discussion.* Proposition 19 states that, given $b_1 \in]0, 1[$ and $m \geq 2$, there is a certain $b_m^\star$ such that $b_2 \in [0, b_m^\star]$. Let us recall that $b_m^\star$ is the only solution of

$$m = \frac{2 + 2(x/b_1)^m - (b_1^m + x^m)^2}{1 - (x/b_1)^2}.$$

In Figure 6, we have plotted $b_m^\star$ as a function of b_1, for $m = 2, \dots, 20$.

If we make $b_2 = b_m^\star$, then the discriminant Δ_m defined in Theorem 6 is equal to zero, and in that case, $\Omega_m^+ = \Omega_m^-$ or, equivalently, $\lambda_m^+ = \lambda_m^-$. Note that the relation between $\Omega^\pm$ and $\lambda_m^\pm$ is given by

$$\Omega_m^\pm = \tfrac{1}{2}(1 - \lambda_m^\mp).$$

In Figure 7, we plotted $\lambda_m^\pm$ as a function of $b_2 \in [0, b_m^\star]$, for $m = 2, \dots, 20$ and $b_1 \in \{0.25, 0.5, 0.75, 0.99\}$. We have also plotted in black the special case $m = 1$, where $b_2 \in [0, b_1]$, $\lambda_1^+ = 1 + b_2^2 - b_1^2$ and $\lambda_1^- = (b_2/b_1)^2$. Observe that, whereas the curves λ_m^+ and λ_m^- are disjoint for $m \geq 2$, λ_1^+ may intersect λ_m^+ or λ_m^-. It is particularly interesting to see what happens when b_1 is close to 1; indeed, when $b_1 = 0.99$, the curves λ_m^- become practically indistinguishable.

Although Figure 7 gives a fairly good idea of the structure of $\lambda_m^\pm$, it may be clarifying to show globally how the curves in Figure 7 behave as b_1 changes, for a fixed m. In Figure 8, we have plotted $\lambda_m^\pm$ as a function of $b_2 \in [0, b_m^\star]$, for $m = 2, 3, 4$ and for all $b_1 \in]0, 1[$, in such a way that, for a given b_1, the intersection between $z = b_1$ and the resulting surfaces yields curves equivalent to those in Figure 8. In general, the surfaces corresponding to $m \geq 3$ are very similar. On the other hand, Figure 8 shows that, when $m = 2$ and b_1 is *not too large*, the size of the curves $(b_2, \lambda_2^\pm)$ is *very small*; indeed, in Figure 7, $(b_2, \lambda_2^\pm)$ is hardly visible when $b_1 = 0.25$. A similar observation can be made with respect to the case $m = 2$ in Figure 6, which is markedly different from the others.

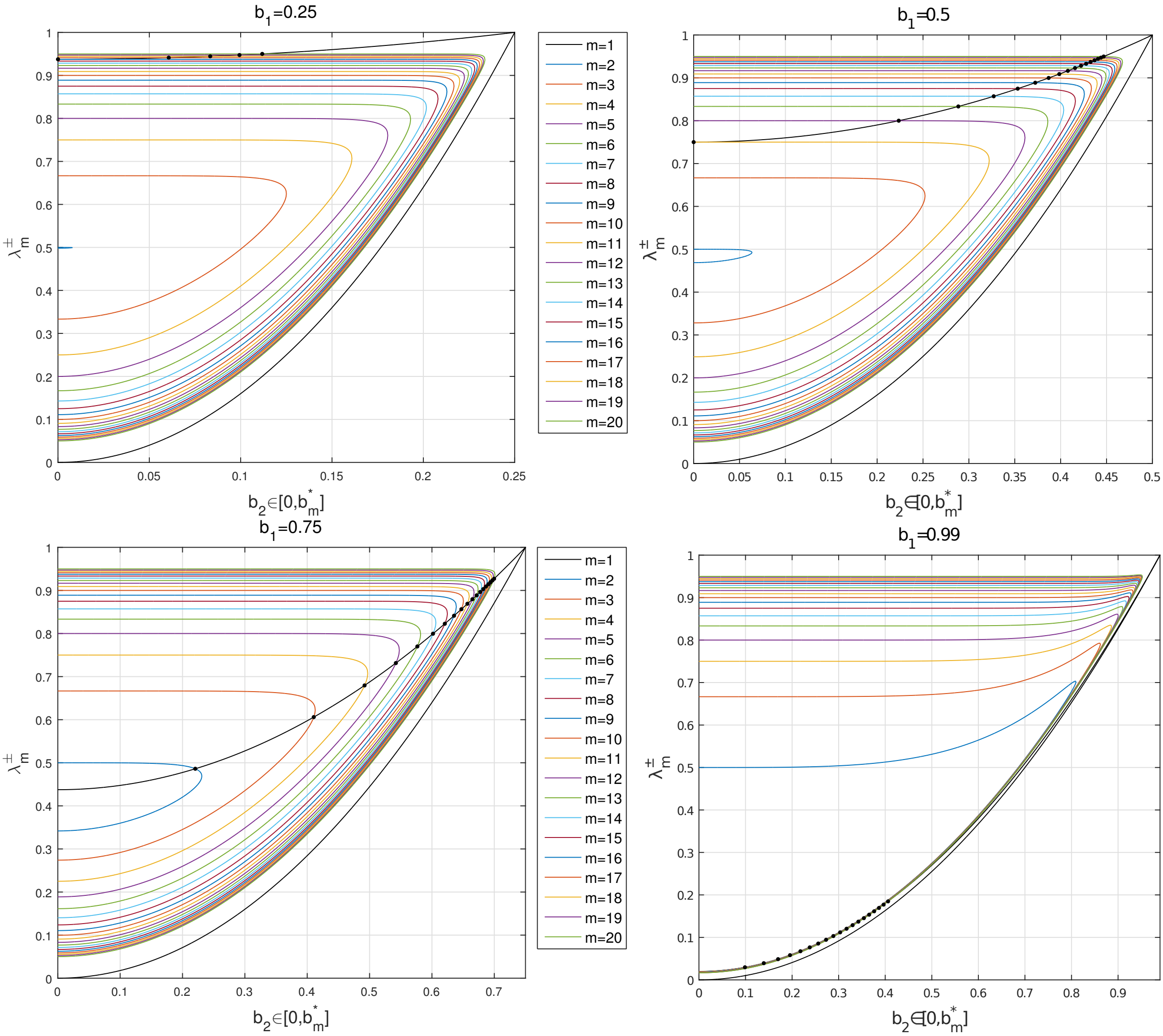


Figure 7. $\lambda_m^{\pm}$ as a function of $b_2 \in [0, b_m^\star]$, for $m = 2, \ldots, 20$, together with the case $m = 1$ (black), for $b_1 \in \{0.25, 0.5, 0.75, 0.99\}$. We have marked with a small black dots the intersections happening between the case $m = 1$ and the other cases.

As in the simply connected case, we use $\Omega = (1 - \lambda)/2$ as our bifurcation parameter. In order to treat the saddle-node bifurcation points [Kielhöfer 2012] that may appear during the bifurcation process, we again use the techniques described in [de la Hoz et al. 2016a].

Before illustrating the shape of the doubly connected V-states, let us mention that the situation is much more involved than in the simply connected case, where there were roughly two situations for all m: *b close* to 1 and *b not so close* to 1. Indeed, we have to play now with both the proximity of b_1 to 1 and that of b_2 to $b_m^\star$. Furthermore, we can start the bifurcation from the annulus of radii b_1 and b_2 at two

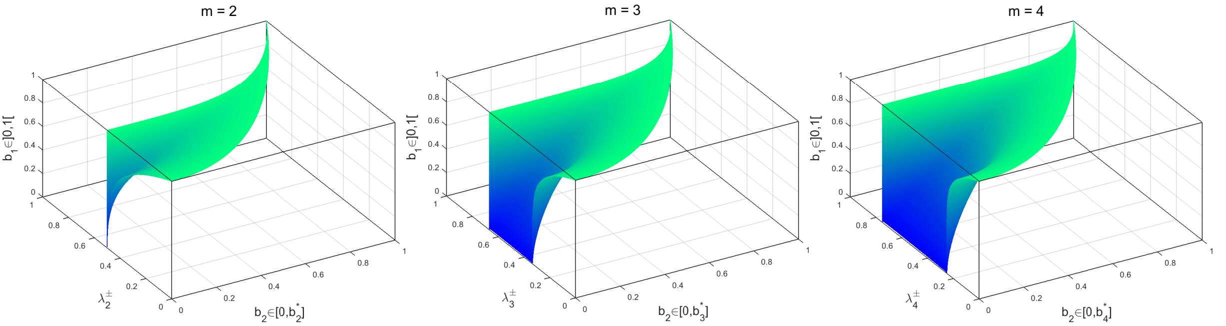


Figure 8. $\lambda_m^{\pm}$ as a function of $b_2 \in [0, b_m^\star]$, for $m = 2, 3, 4$ and for all $b_1 \in]0, 1[$.

different values of Ω, i.e., Ω_m^+ and Ω_m^-. Finally, the case $m = 1$ needs to be studied individually. All in all, we have detected the following scenarios.

When $m \geq 3$, there are roughly three cases when starting to bifurcate at Ω_m^+ and two cases when starting to bifurcate at Ω_m^-. More precisely, if we start to bifurcate at Ω_m^+, we have to distinguish between the following:

- b_2 is *very close* to $b_m^\star$. In that case, it seems possible to obtain V-states for all $\Omega \in]\Omega_m^-, \Omega_m^+[$, very much like in [de la Hoz et al. 2016b], irrespective of the size of b_1. For example, in Figure 9, we have calculated the V-states corresponding to $m = 4$, $b_1 = 0.8$ and $b_2 = 0.53$. Observe that $b_4^\star = 0.5407\ldots$, i.e., we have chosen b_2 *close enough* to $b_4^\star$. On the right-hand side, we have plotted the bifurcation diagram of the coefficients $a_{1,1}$ and $a_{2,1}$ in (82) against Ω, which shows that there is indeed a continuous bifurcation branch that joins Ω_m^- and Ω_m^+, where $\Omega_4^- = 0.1335\ldots$ and $\Omega_4^+ = 0.1671\ldots$. On the left-hand side, we have plotted V-states for four different values of $\Omega \in]\Omega_m^-, \Omega_m^+[$.

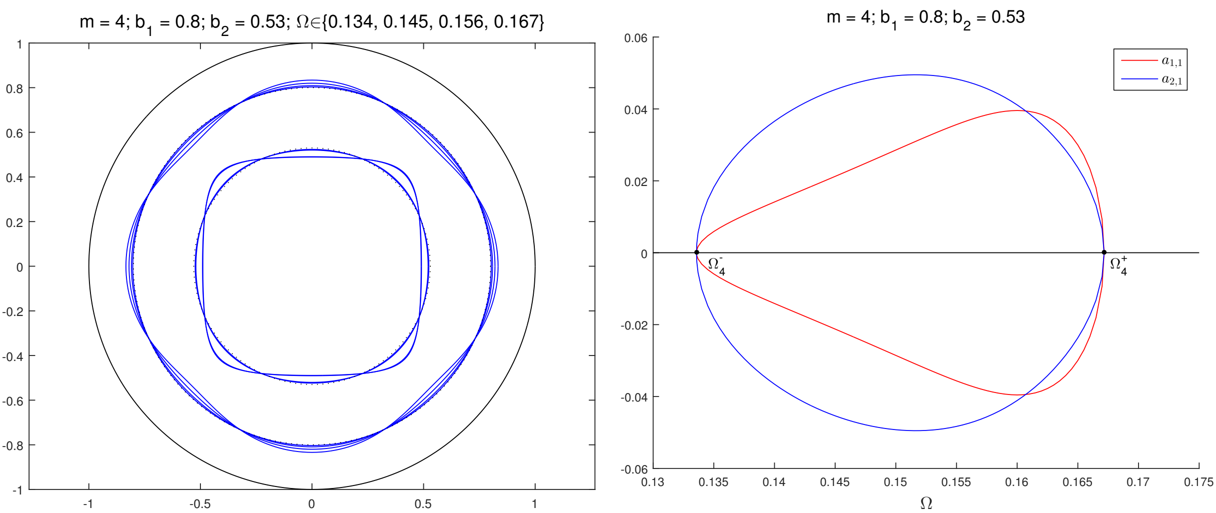


Figure 9. Left: V-states corresponding to $m = 4$, $b_1 = 0.8$, $b_2 = 0.53$ and several values of Ω. Right: bifurcation diagram. Here $N = 256$.

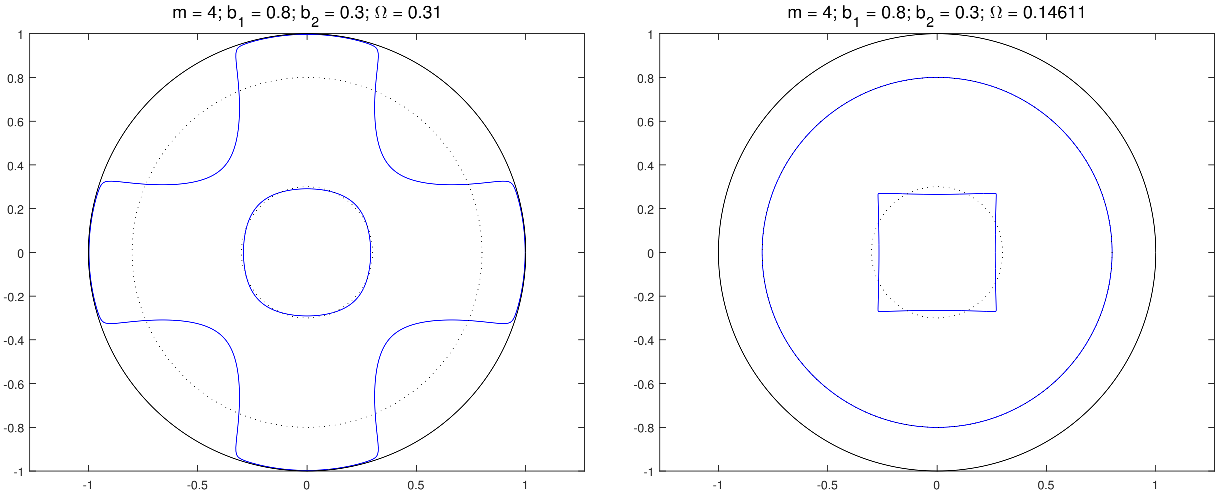


Figure 10. Approximation to the limiting V-states corresponding to $m=4$, $b_1=0.8$ and $b_2=0.3$. Left: we have started to bifurcate at $\Omega_4^+=0.3256\ldots$, taking $\Omega<\Omega_4^+$. Right: we have started to bifurcate at $\Omega_4^-=0.1250\ldots$, taking $\Omega>\Omega_4^-$. Here $N=1024$.

- b_1 is *close* to 1, and b_2 is *small enough*. There are limiting V-states, for which the distance between the outer boundary z_1 and the unit circumference tends to zero, but the inner boundary z_2 does not deviate greatly from the circumference of radius b_2. On the left-hand side of Figure 10, we have approximated the limiting V-state corresponding to $m=4$, $b_1=0.8$ and $b_2=0.3$. The shape of z_1 is not very far from the case $m=4$ and $b=0.8$ of Figure 5.

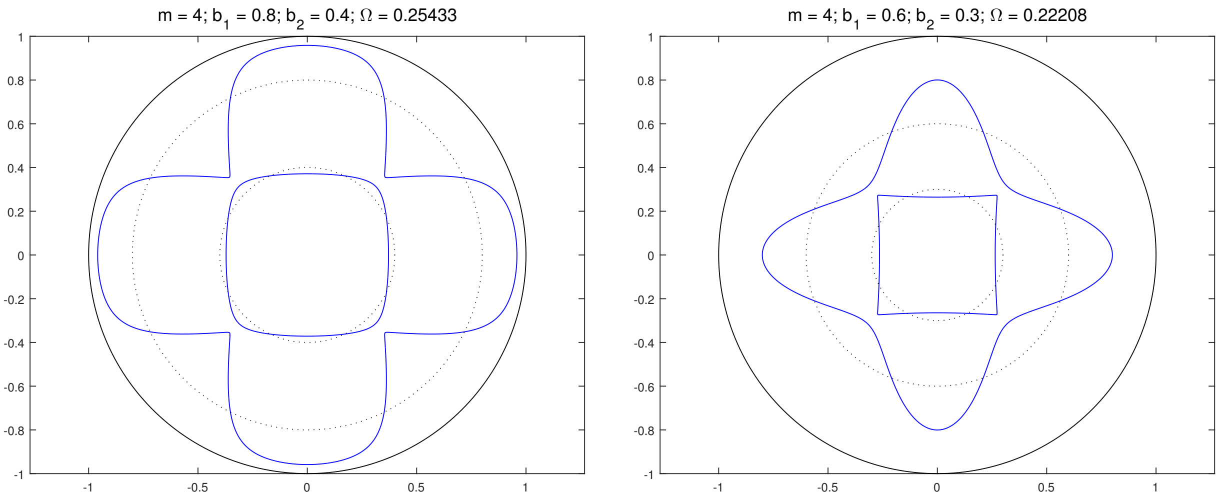


Figure 11. Left: approximation to the limiting V-state corresponding to $m=4$, $b_1=0.8$ and $b_2=0.4$, starting to bifurcate at $\Omega_4^+=0.2706$, taking $\Omega<\Omega_4^+$. Right: approximation to the limiting V-state corresponding to $m=4$, $b_1=0.6$ and $b_2=0.3$, starting to bifurcate at $\Omega_4^+=0.2516$, taking $\Omega<\Omega_4^+$. Here $N=1024$.

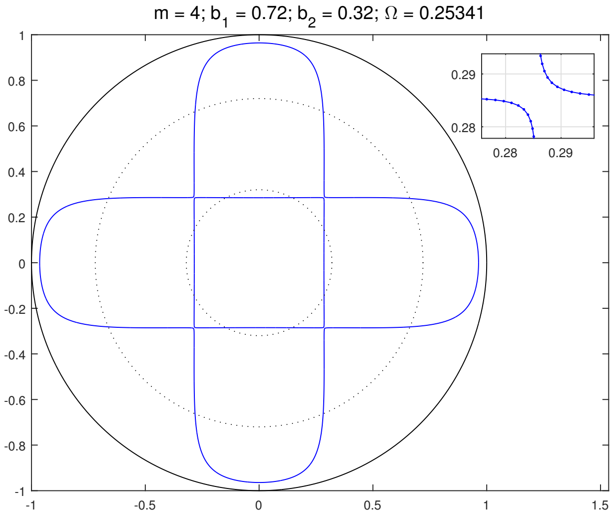


Figure 12. Approximation to the limiting V-state corresponding to $m = 4$, $b_1 = 0.72$ and $b_2 = 0.32$, starting to bifurcate at $\Omega_4^+ = 0.2851$, taking $\Omega < \Omega_4^+$. Here $N = 2048$. The zoom shows that the boundaries are very close from each other, but there is no intersection.

- b_1 and b_2 do not fit in the previous two cases. In that case, there are also limiting V-states, characterized by the appearance of corner-shaped singularities in z_1 or z_2. In Figure 11, we have approximated the limiting V-states corresponding to $m = 4$, $b_1 = 0.8$ and $b_2 = 0.4$ (left) and to $m = 4$, $b_1 = 0.6$ and $b_2 = 0.3$ (right). Observe that the influence of the rigid boundary seems less perceptible in the second example, which accordingly does not differ too much from those in [de la Hoz et al. 2016b].

  Although the distance between z_1 and the unit circumference is always strictly positive, the distance between z_1 and z_2 is sometimes very small, and we cannot exclude in advance the existence of limiting V-states where z_1 and z_2 actually touch each other. For instance, after playing with the values of b_1 and b_2, we have found that the choice of $b_1 = 0.72$ and $b_2 = 0.32$ enables us to find a V-state such that the distance between z_1 and z_2 is of about 7×10^{-3}. This V-state is plotted in Figure 12, together with a zoom of one apparent intersection of the boundaries that shows that there is really no intersection and that the nodal resolution is adequate.

On the other hand, if we start to bifurcate at Ω_m^-, we have to distinguish between the following:

- b_2 is *very close* to $b_m^\star$. This case has been explained above. In fact, it is irrelevant whether we start to bifurcate at Ω_m^- or at Ω_m^+.
- b_2 is not *close enough* to $b_m^\star$. In that case, there are limiting V-states, characterized by the appearance of corner-shaped singularities in z_2 whereas the outer boundary z_1 does not deviate greatly from the circumference of radius b_1. On the right-hand side of Figure 10, we have approximated the limiting V-state corresponding to $m = 4$, $b_1 = 0.8$ and $b_2 = 0.3$. We have not bothered to plot the V-states

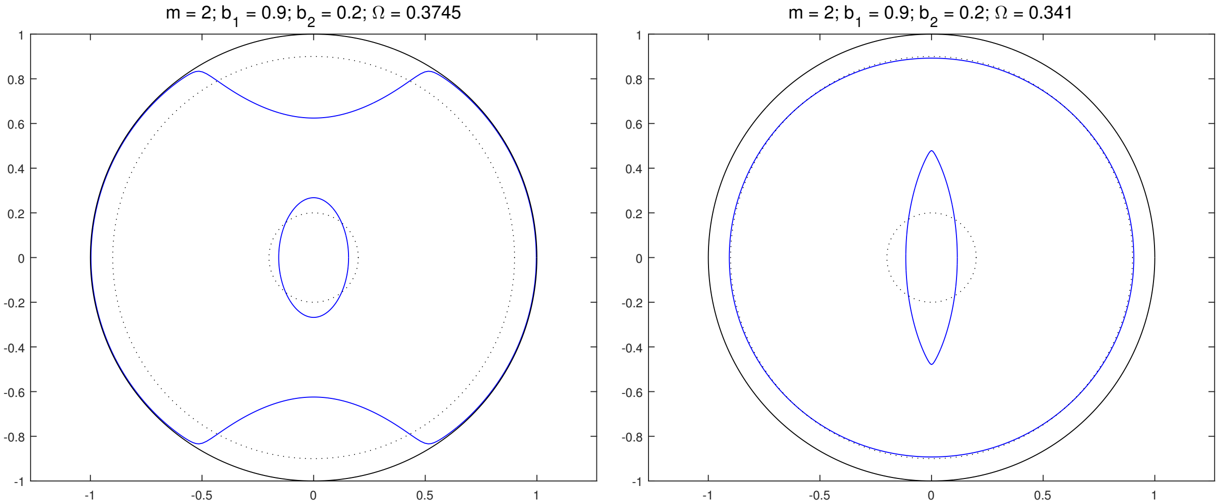


Figure 13. Left: approximation to the limiting V-states corresponding to $m = 2$, $b_1 = 0.9$ and $b_2 = 0.2$, starting to bifurcate at $\Omega_2^+ = 0.3892\ldots$, taking $\Omega < \Omega_2^+$. Right: we have started to bifurcate at $\Omega_2^- = 0.2497\ldots$, taking $\Omega > \Omega_2^-$. Here $N = 512$.

corresponding to those in Figures 11 and 12 but starting to bifurcate at Ω_m^- because they are virtually identical, up to a scaling of z_2. This case closely matches that in [de la Hoz et al. 2016b], and the inner boundary resembles the simply connected V-states in [Deem and Zabusky 1978].

Summarizing, if we compare the doubly connected V-states just described with those in [de la Hoz et al. 2016b], we conclude that the truly unique case here is when b_1 is *close* to 1 and b_2 is *small enough*.

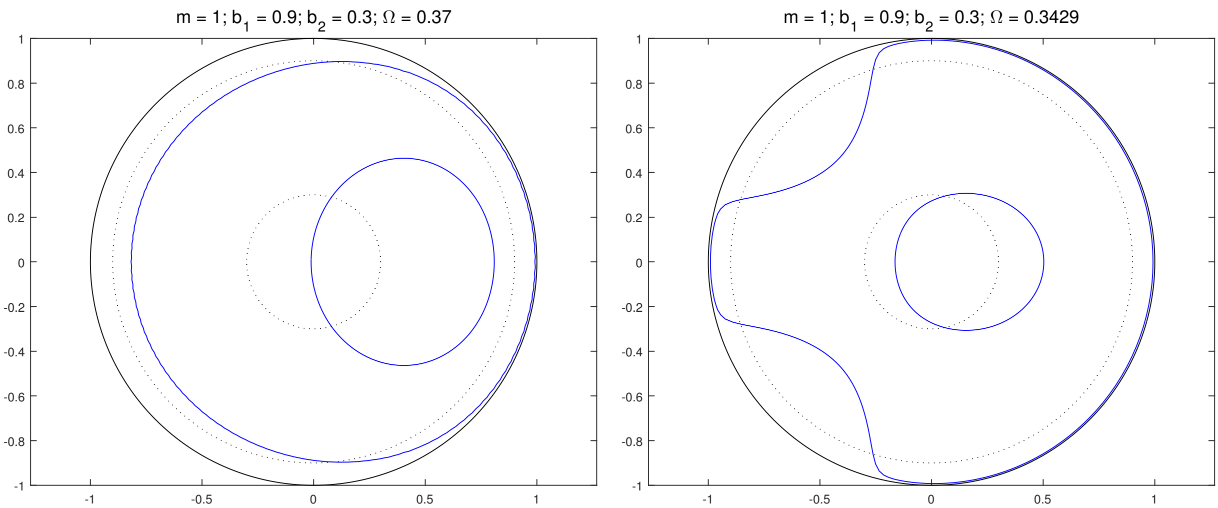


Figure 14. Approximation to the limiting V-states corresponding to $m = 1$, $b_1 = 0.9$ and $b_2 = 0.3$. Left: we have started to bifurcate at $\Omega_1^+ = \frac{4}{9}$, taking $\Omega < \Omega_1^+$. Right: we have started to bifurcate at $\Omega_1^- = 0.36$, taking $\Omega > \Omega_1^-$. Here $N = 256$.

Regarding the case $m=2$, everything said above is applicable. For example, in Figure 13, we have taken $b_1=0.9$ and $b_1=0.2$, i.e., a value of b_1 *close* to 1 and a value of b_2 *small enough*. On the left-hand side, we show an approximation to the limiting V-state appearing when starting to bifurcate at Ω_2^+; note the clear parallelism with the case $m=2$ and $b=0.9$ of Figure 5 and with the left-hand side of Figure 10. On the right-hand side, we show an approximation to the limiting V-state appearing when starting to bifurcate at Ω_2^-; as in the right-hand side of Figure 10, corner-shaped singularities seem to develop in z_2 whereas z_1 has barely deviated from a circumference.

The case $m=1$ also deserves a comment. In Figure 14, we have approximated the limiting V-states corresponding to $m=1$, taking again a value of b_1 *close* to 1 and a value of b_2 *small enough*, more precisely, $b_1=0.9$ and $b_2=0.3$. On the left-hand side, we have started to bifurcate at Ω_1^+, and on the right-hand side, we have started to bifurcate at Ω_1^-. It is remarkable that, in both cases, the distance of z_1 to the unit circumference is smaller than 10^{-2}. Moreover, even if the V-state on the left-hand side is roughly in agreement with Figure 5 and with the left-hand sides of Figures 10 and 13, the V-state on the right-hand side exhibits a completely different, unexpected behavior.

Acknowledgements

Francisco de la Hoz was supported by the Basque Government, through the project IT641-13, and by the Ministerio de Economía y Competitividad, through the project MTM2014-53145-P. Taoufik Hmidi was partially supported by the Agence Nationale de la Recherche project Dyficolti ANR-13-BS01-0003-01. Joan Mateu was partially supported by the grants of Generalitat de Catalunya 2014SGR7, Ministerio de Economía y Competitividad MTM2013-4469 and ECPP7 Marie Curie project Metric Analysis for Emergent Technologies.

References

[Bertozzi and Constantin 1993] A. L. Bertozzi and P. Constantin, "Global regularity for vortex patches", *Comm. Math. Phys.* **152**:1 (1993), 19–28. MR 1207667 Zbl 0771.76014

[Burbea 1980] J. Burbea, "Vortex motions and conformal mappings", pp. 276–298 in *Nonlinear evolution equations and dynamical systems* (Lecce, Italy, 1979), edited by M. Boiti et al., Lecture Notes in Phys. **120**, Springer, Berlin, 1980. MR 581901

[Burbea 1982] J. Burbea, "Motions of vortex patches", *Lett. Math. Phys.* **6**:1 (1982), 1–16. MR 646163 Zbl 0484.76031

[Burbea and Landau 1982] J. Burbea and M. Landau, "The Kelvin waves in vortex dynamics and their stability", *J. Comput. Phys.* **45**:1 (1982), 127–156. MR 650429 Zbl 0488.76023

[Castro et al. 2016a] A. Castro, D. Córdoba, and J. Gómez-Serrano, "Existence and regularity of rotating global solutions for the generalized surface quasi-geostrophic equations", *Duke Math. J.* **165**:5 (2016), 935–984. MR 3482335 Zbl 1339.35234

[Castro et al. 2016b] A. Castro, D. Córdoba, and J. Gómez-Serrano, "Uniformly rotating analytic global patch solutions for active scalars", *Ann. PDE* **2**:1 (2016), Art. 1, 34. MR 3462104

[Cerretelli and Williamson 2003] C. Cerretelli and C. H. K. Williamson, "A new family of uniform vortices related to vortex configurations before merging", *J. Fluid Mech.* **493** (2003), 219–229. MR 2017951 Zbl 1063.76008

[Chemin 1998] J.-Y. Chemin, *Perfect incompressible fluids*, Oxford Lecture Series in Mathematics and its Applications **14**, Clarendon, New York, 1998. MR 1688875 Zbl 0927.76002

[Crandall and Rabinowitz 1971] M. G. Crandall and P. H. Rabinowitz, "Bifurcation from simple eigenvalues", *J. Functional Analysis* **8** (1971), 321–340. MR 0288640 Zbl 0219.46015

[Deem and Zabusky 1978] G. S. Deem and N. J. Zabusky, "Vortex waves: stationary "V states", interactions, recurrence, and breaking", *Phys. Rev. Lett.* **40**:13 (1978), 859–862.

[Depauw 1999] N. Depauw, "Poche de tourbillon pour Euler 2D dans un ouvert à bord", *J. Math. Pures Appl.* (9) **78**:3 (1999), 313–351. MR 1687165 Zbl 0927.76014

[Dritschel 1986] D. G. Dritschel, "The nonlinear evolution of rotating configurations of uniform vorticity", *J. Fluid Mech.* **172** (1986), 157–182. Zbl 0616.76069

[Farwig and Hishida 2011] R. Farwig and T. Hishida, "Asymptotic profile of steady Stokes flow around a rotating obstacle", *Manuscripta Math.* **136**:3–4 (2011), 315–338. MR 2844813 Zbl 1229.35172

[Fraenkel 2000] L. E. Fraenkel, *An introduction to maximum principles and symmetry in elliptic problems*, Cambridge Tracts in Mathematics **128**, Cambridge University, 2000. MR 1751289 Zbl 0947.35002

[Guo et al. 2004] Y. Guo, C. Hallstrom, and D. Spirn, "Dynamics near an unstable Kirchhoff ellipse", *Comm. Math. Phys.* **245**:2 (2004), 297–354. MR 2039699 Zbl 1061.76012

[Hassainia and Hmidi 2015] Z. Hassainia and T. Hmidi, "On the V-states for the generalized quasi-geostrophic equations", *Comm. Math. Phys.* **337**:1 (2015), 321–377. MR 3324164 Zbl 1319.35188

[Hmidi 2015] T. Hmidi, "On the trivial solutions for the rotating patch model", *J. Evol. Equ.* **15**:4 (2015), 801–816. MR 3427065 Zbl 1338.35365

[Hmidi and Mateu 2016] T. Hmidi and J. Mateu, "Bifurcation of rotating patches from Kirchhoff vortices", *Discrete Contin. Dyn. Syst.* **36**:10 (2016), 5401–5422. MR 3543554

[Hmidi et al. 2013] T. Hmidi, J. Mateu, and J. Verdera, "Boundary regularity of rotating vortex patches", *Arch. Ration. Mech. Anal.* **209**:1 (2013), 171–208. MR 3054601 Zbl 1286.35201

[Hmidi et al. 2015] T. Hmidi, J. Mateu, and J. Verdera, "On rotating doubly connected vortices", *J. Differential Equations* **258**:4 (2015), 1395–1429. MR 3294351 Zbl 06396994

[de la Hoz et al. 2016a] F. de la Hoz, Z. Hassainia, and T. Hmidi, "Doubly connected V-states for the generalized surface quasi-geostrophic equations", *Arch. Ration. Mech. Anal.* **220**:3 (2016), 1209–1281. MR 3466846 Zbl 1334.35263

[de la Hoz et al. 2016b] F. de la Hoz, T. Hmidi, J. Mateu, and J. Verdera, "Doubly connected V-states for the planar Euler equations", *SIAM J. Math. Anal.* **48**:3 (2016), 1892–1928. MR 3507551 Zbl 1342.35239

[Kamm 1987] J. R. Kamm, *Shape and stability of two-dimensional uniform vorticity regions*, Ph.D. thesis, California Institute of Technology, 1987, Available at http://search.proquest.com/docview/303551772. MR 2635730

[Kielhöfer 2012] H. Kielhöfer, *Bifurcation theory: an introduction with applications to partial differential equations*, 2nd ed., Applied Mathematical Sciences **156**, Springer, New York, 2012. MR 2859263 Zbl 1230.35002

[Kirchhoff 1876] G. Kirchhoff, *Vorlesungen über mathematische Physik: Mechanik*, Teubner, Leipzig, 1876. JFM 08.0542.01

[Lamb 1945] H. Lamb, *Hydrodynamics*, 6th ed., Dover, Mineola, NY, 1945.

[Love 1893] A. E. H. Love, "On the stability of certain vortex motions", *Proc. London Math. Soc.* **25**:1 (1893), 18–43.

[Luzzatto-Fegiz and Williamson 2010] P. Luzzatto-Fegiz and C. H. K. Williamson, "Stability of elliptical vortices from "Imperfect–Velocity–Impulse" diagrams", *Theor. Comp. Fluid Dyn.* **24**:1 (2010), 181–188. Zbl 1191.76056

[Majda and Bertozzi 2002] A. J. Majda and A. L. Bertozzi, *Vorticity and incompressible flow*, Cambridge Texts in Applied Mathematics **27**, Cambridge University, 2002. MR 1867882 Zbl 0983.76001

[Mitchell and Rossi 2008] T. B. Mitchell and L. F. Rossi, "The evolution of Kirchhoff elliptic vortices", *Phys. Fluids* **20** (2008), 054103. Zbl 1182.76523

[Overman 1986] E. A. Overman, II, "Steady-state solutions of the Euler equations in two dimensions, II: Local analysis of limiting V-states", *SIAM J. Appl. Math.* **46**:5 (1986), 765–800. MR 858995 Zbl 0608.76018

[Płotka and Dritschel 2012] H. Płotka and D. G. Dritschel, "Quasi-geostrophic shallow-water vortex-patch equilibria and their stability", *Geophys. Astrophys. Fluid Dyn.* **106**:6 (2012), 574–595. MR 2982526

[Polvani and Flierl 1986] L. M. Polvani and G. R. Flierl, "Generalized Kirchhoff vortices", *Phys. Fluids* **29**:8 (1986), 2376–2379. Zbl 0601.76015

[Pommerenke 1992] C. Pommerenke, *Boundary behaviour of conformal maps*, Grundlehren der mathematischen Wissenschaften **299**, Springer, Berlin, 1992. MR 1217706 Zbl 0762.30001

[Saffman 1992] P. G. Saffman, *Vortex dynamics*, Cambridge University, New York, 1992. MR 1217252 Zbl 0777.76004

[Tang 1987] Y. Tang, "Nonlinear stability of vortex patches", *Trans. Amer. Math. Soc.* **304**:2 (1987), 617–638. MR 911087 Zbl 0636.76019

[Wan 1986] Y. H. Wan, "The stability of rotating vortex patches", *Comm. Math. Phys.* **107**:1 (1986), 1–20. MR 861881 Zbl 0624.76055

[Warschawski 1935] S. E. Warschawski, "On the higher derivatives at the boundary in conformal mapping", *Trans. Amer. Math. Soc.* **38**:2 (1935), 310–340. MR 1501813 Zbl 0014.26707

[Wu et al. 1984] H. M. Wu, E. A. Overman, II, and N. J. Zabusky, "Steady-state solutions of the Euler equations in two dimensions: rotating and translating V-states with limiting cases, I: Numerical algorithms and results", *J. Comput. Phys.* **53**:1 (1984), 42–71. MR 734586 Zbl 0524.76029

[Yudovich 1963] V. I. Yudovich, "Nonstationary flow of an ideal incompressible liquid", *USSR Comp. Math. Math. Phys.* **3**:6 (1963), 1407–1456. Zbl 0129.19402

Received 20 Oct 2015. Revised 27 Apr 2016. Accepted 28 May 2016.

FRANCISCO DE LA HOZ: francisco.delahoz@ehu.eus
Department of Applied Mathematics and Statistics and Operations Research, Faculty of Science and Technology, University of the Basque Country UPV/EHU, Barrio Sarriena S/N, 48940 Leioa, Spain

ZINEB HASSAINIA: zineb.hassainia@cims.nyu.edu
Courant Institute for Mathematical Sciences, New York University, New York, NY 10012-1185, United States

TAOUFIK HMIDI: thmidi@univ-rennes1.fr
Institut de Recherche Mathématique de Rennes, Université de Rennes 1, Campus de Beaulieu, 35 042 Rennes CEDEX, France

JOAN MATEU: mateu@mat.uab.cat
Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain

ANALYSIS AND PDE
Vol. 9, No. 7, 2016
dx.doi.org/10.2140/apde.2016.9.1671

ISOLATED SINGULARITIES OF POSITIVE SOLUTIONS OF ELLIPTIC EQUATIONS WITH WEIGHTED GRADIENT TERM

PHUOC-TAI NGUYEN

Let $\Omega \subset \mathbb{R}^N$ $(N > 2)$ be a C^2 bounded domain containing the origin 0. We study the behavior near 0 of positive solutions of equation (E) $-\Delta u + |x|^\alpha u^p + |x|^\beta |\nabla u|^q = 0$ in $\Omega \setminus \{0\}$, where $\alpha > -2$, $\beta > -1$, $p > 1$, and $q > 1$. When $1 < p < (N+\alpha)/(N-2)$ and $1 < q < (N+\beta)/(N-1)$, we provide a full classification of positive solutions of (E) vanishing on $\partial\Omega$. On the contrary, when $p \geq (N+\alpha)/(N-2)$ or $(N+\beta)/(N-1) \leq q \leq 2+\beta$, we show that any isolated singularity at 0 is removable.

1. Introduction

Let $\Omega \subset \mathbb{R}^N$ $(N > 2)$ be a C^2 bounded domain containing the origin 0. In this paper, we study isolated singularities at 0 of nonnegative solutions of the quasilinear equation

$$-\Delta u + |x|^\alpha u^p + |x|^\beta |\nabla u|^q = 0 \tag{1-1}$$

in $\Omega \setminus \{0\}$ where $\alpha > -2$, $\beta > -1$, $p > 1$, and $q > 1$. By a nonnegative solution of (1-1) we mean a nonnegative function $u \in C^2(\Omega \setminus \{0\})$ which satisfies (1-1) in the classical sense.

Equation (1-1) consists of two mechanisms: the semilinear equation

$$-\Delta u + |x|^\alpha u^p = 0 \tag{1-2}$$

in $\Omega \setminus \{0\}$ and the quasilinear equation

$$-\Delta u + |x|^\beta |\nabla u|^q = 0 \tag{1-3}$$

in $\Omega \setminus \{0\}$. For the sake of simplicity, in the sequel, we use the notation

$$(F \circ u)(x) = |x|^\alpha u(x)^p + |x|^\beta |\nabla u(x)|^q. \tag{1-4}$$

MSC2010: 35A20, 35J60.
Keywords: gradient terms, weak singularities, strong singularities, removability.

In the literature, many results concerning isolated singularities for (1-2) with $\alpha = 0$ have been published, among which we refer to [Brézis and Véron 1980/81; Vázquez and Véron 1985; Véron 1981; 1996; Baras and Pierre 1984, Marcus 2013] and references therein. Marcus and Véron [2014] provided a full description of isolated singularities of positive solutions of (1-2) (with $\alpha > -2$) when $1 < p < p_{c,\alpha}$ with

$$p_{c,\alpha} := \frac{N+\alpha}{N-2}. \tag{1-5}$$

More precisely, in this range, if v is a positive solution of (1-2) vanishing on $\partial\Omega$, then:

- either $v = v_k^\Omega$ ($k > 0$), the solution of

$$-\Delta v + |x|^\alpha v^p = k\delta_0 \qquad \text{in } \Omega, \text{ with } v = 0 \text{ on } \partial\Omega \tag{1-6}$$

  (here δ_0 is the Dirac measure concentrated at the origin) and $v(x) = kc_N(1+o(1))|x|^{2-N}$ as $|x| \to 0$ where $c_N = 1/(N(N-2)\omega_N)$ with ω_N being the volume of the unit ball in $\mathbb{R}^N$;
- or $v = v_\infty^\Omega := \lim_{k\to\infty} v_k^\Omega$ and $v(x) = \vartheta(1+o(1))|x|^{-\frac{2+\alpha}{p-1}}$ as $|x| \to 0$ with

$$\vartheta := \left[\left(\frac{2+\alpha}{p-1}\right)\left(\frac{2p+\alpha}{p-1} - N\right)\right]^{\frac{1}{p-1}}. \tag{1-7}$$

When $p \ge p_{c,\alpha}$, they showed that there is no positive solution of (1-2) vanishing on $\partial\Omega$.

Classification of interior isolated singularities in the general framework (where the nonlinearity does not depend on gradient term) was established in [Friedman and Véron 1986], in [Cîrstea and Du 2010] (for the p-laplacian), and in [Cîrstea 2014] (for elliptic equations with inverse square potentials). A deep existence and uniqueness result for a more general class of semilinear equations was given in [Marcus 2013].

Much less work concerning the behavior near the origin of positive solutions of equations with the nonlinearity depending mostly on the gradient term has been investigated. See Serrin [1965] and, more recently, Bidaut-Véron, García-Huidobro, and Véron [Bidaut-Véron et al. 2014].

Recently, boundary trace problem for semilinear equation with gradient terms were studied by P. T. Nguyen and L. Véron [2012] and by M. Marcus and Nguyen [2015].

When the nonlinearity is of the form (1-4), i.e., it depends on both u and ∇u, as well as weights, one encounters the following difficulties:

(i) The first one stems from the competition of two terms $|x|^\alpha u^p$ and $|x|^\beta |\nabla u|^q$. When $\frac{2+\alpha}{p-1} \neq \frac{2+\beta-q}{q-1}$, (1-1) admits no *similarity transformation* (see Section 2). Moreover, in this framework, the Keller–Osserman estimate is no longer a sharp upper bound for solutions of (1-1).

(ii) The second one comes from the lack of monotonicity property of the nonlinearity. Furthermore, it is noteworthy that in general the sum of two solution of (1-1) is not a supersolution.

(iii) The presence of the weights $|x|^\alpha$ and $|x|^\beta$, which may vanish or be singular at 0 according to the value of α and β, make the asymptotic behavior near 0 of solutions of (1-1) more intricate.

Fix $d_1 \in (0,1)$ such that $B_{3d_1}(0) \Subset \Omega$ and put $d_2 = \operatorname{diam}(\Omega)$. Set

$$\tau = \min\left\{\frac{2+\alpha}{p-1}, \frac{2+\beta-q}{q-1}\right\} \quad \text{with } q < 2+\beta. \tag{1-8}$$

We first give sharp estimates on solutions of (1-1) and their gradient. These estimates are obtained due to a combination of Bernstein's method, Keller–Osserman estimates, and a transformation argument.

Proposition 1.1. *Let $\alpha > -2$, $\beta > -1$, $p > 1$, and $1 < q < 2+\beta$. There exists a positive constant $c_i = c_i(\alpha, \beta, N, p, q, d_1, d_2)$ $(i = 1, 2)$ such that if u is a positive solution of (1-1) in $\Omega \setminus \{0\}$ vanishing on $\partial\Omega$, then*

$$u(x) \le c_1 |x|^{-\tau} \quad \textit{for all } x \in \Omega \setminus \{0\}, \tag{1-9}$$

and

$$|\nabla u(x)| \le c_2 |x|^{-\tau-1} \quad \textit{for all } x \in \overline{B_{d_1}(0)} \setminus \{0\}. \tag{1-10}$$

Estimates (1-9) and (1-10) give an upper bound of $F \circ u$ but do not ensure that $F \circ u \in L^1(\Omega)$. While investigating the integrability of $F \circ u$ we are led to the following definition.

Definition 1.2. A nonnegative solution u of (1-1) is called a *weakly singular solution* if $F \circ u \in L^1(B_\varepsilon)$ for some $\varepsilon > 0$. Otherwise, u is called a *strongly singular solution*.

We next introduce the definition of solutions to

$$\begin{cases} -\Delta u + F \circ u = k\delta_0 & \text{in } \Omega, \\ \qquad\qquad u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1-11}$$

Definition 1.3. Let $k \ge 0$. A nonnegative function u is a solution of (1-11) if $u \in L^1(\Omega)$, $F \circ u \in L^1(\Omega)$, and

$$\int_\Omega \big(-u\Delta\zeta + (F \circ u)\zeta\big)\, dx = k\zeta(0) \quad \text{for all } \zeta \in C_0^2(\overline{\Omega}). \tag{1-12}$$

Remark. Clearly, if u is a solution of (1-11) then u is a weakly singular solution of (1-1).

Let Γ_N $(N > 2)$ be the Newtonian kernel in $\mathbb{R}^N$ defined by

$$\Gamma_N(x) := c_N |x|^{2-N} = \frac{1}{N(N-2)\omega_N} |x|^{2-N}, \quad x \neq 0 \tag{1-13}$$

with ω_N the volume of the unit ball in $\mathbb{R}^N$. Denote by G^Ω the Green kernel of $(-\Delta)$ in Ω and by $\mathbb{G}^\Omega$ the corresponding operator.

The study of (1-1) is strongly linked to that of (1-3). As we will see in the sequel there exists an exponent

$$q_{c,\beta} = \frac{N+\beta}{N-1} \tag{1-14}$$

such that if $1 < q < q_{c,\beta}$, the problem (1-3) admits weakly and strongly singular solutions; while if $q_{c,\beta} < q < 2+\beta$, then such solutions don't exist. When both equations (1-2) and (1-3) are combined in (1-1), the existence result for (1-1) is valid in the range $(p,q) \in (1, p_{c,\alpha}) \times (1, q_{c,\beta})$. This is reflected in the following theorems.

Theorem A. *Assume $\alpha > -2$, $\beta > -1$, $1 < p < p_{c,\alpha}$, and $1 < q < q_{c,\beta}$. For any $k > 0$, there exists a unique solution $u_k^\Omega \in C^2(\Omega \setminus \{0\}) \cap C(\bar{\Omega} \setminus \{0\})$ of (1-11). Moreover,*

$$u_k^\Omega(x) = kG^\Omega(x,0) - \mathbb{G}^\Omega[F \circ u_k^\Omega](x) \quad \text{for all } x \in \Omega \setminus \{0\}, \tag{1-15}$$

$$u_k^\Omega(x) = k(1+o(1))\,\Gamma_N(x) \quad \text{as } x \to 0, \tag{1-16}$$

$$\lim_{|x|\to 0} \left(|x|^{N-1} \nabla u_k^\Omega(x) + \frac{k}{N\omega_N} \frac{x}{|x|} \right) = 0. \tag{1-17}$$

Due to (1-16) and the comparison principle [Gilbarg and Trudinger 2001, Theorem 9.2], the sequence $\{u_k^\Omega\}$ is increasing. Denote $u_\infty^\Omega := \lim_{k\to\infty} u_k^\Omega$. The asymptotic behaviors of u_∞^Ω and its gradient are given in the following theorem.

Theorem B. *Assume $\alpha > -2$, $\beta > -1$, $1 < p < p_{c,\alpha}$, and $1 < q < q_{c,\beta}$. Then u_∞^Ω is a strongly singular solution of (1-1) vanishing on $\partial\Omega$. Moreover,*

$$\lim_{|x|\to 0} |x|^\tau u_\infty^\Omega(x) = \Theta, \tag{1-18}$$

$$\lim_{|x|\to 0} \left(|x|^{\tau+1} \nabla u_\infty^\Omega(x) + \Theta\tau \frac{x}{|x|} \right) = 0, \tag{1-19}$$

where τ is defined in (1-8) and Θ is a positive constant depending on N, α, β, p, q.

Remark. The value of Θ varies according to the relationship between the parameters α, β, p, and q. For simplicity, set

$$D := \frac{2+\alpha}{p-1} \times \frac{q-1}{2+\beta-q} \quad \text{with } q < 2+\beta. \tag{1-20}$$

In Theorem B, Θ is the unique solution of

$$\lambda t^{p-1} + j\tau^q t^{q-1} - \tau(\tau+2-N) = 0, \tag{1-21}$$

where j and λ are given by

$$\begin{cases} j = 0 \text{ and } \lambda = 1 & \text{if } D < 1 \text{ (hence } \Theta = \vartheta \text{ defined in (1-7));} \\ j = 1 \text{ and } \lambda = 0 & \text{if } D > 1 \text{ (hence } \Theta = \theta_0 \text{ defined in (4-3));} \\ j = \lambda = 1 & \text{if } D = 1 \text{ (hence } \Theta = \theta_1 \text{, the solution of } g_1(t) = 0, \\ & \qquad\qquad \text{where } g_\lambda \text{ is defined defined in (4-2)).} \end{cases} \tag{1-22}$$

Theorem B shows the competition between two terms $|x|^\alpha u^p$ and $|x|^\beta |\nabla u|^q$: if $D < 1$ then $|x|^\alpha u^p$ plays a dominant role, otherwise $|x|^\beta |\nabla u|^q$ plays a dominant role.

As a consequence of Theorems A and B, we obtain a description of nonnegative singular solutions of (1-1) vanishing on $\partial\Omega$.

Theorem C. *Assume $\alpha > -2$, $\beta > -1$, $1 < p < p_{c,\alpha}$, and $1 < q < q_{c,\beta}$. Let $u \in C^2(\Omega \setminus \{0\}) \cap C(\bar{\Omega} \setminus \{0\})$ be a nonnegative solution of (1-1) in $\Omega \setminus \{0\}$ vanishing on $\partial\Omega$. Then either $u \equiv 0$, or $u \equiv u_k^\Omega$ for some $k > 0$, or $u \equiv u_\infty^\Omega$.*

On the contrary, the next theorem states that when $p \ge p_{c,\alpha}$ or $q_{c,\beta} \le q < 2+\beta$ there exists no positive singular solution.

Theorem D. *Assume $\alpha > -2$, $\beta > -1$, $p > 1$, and $1 < q \le 2+\beta$. If $p \ge p_{c,\alpha}$ or $q \ge q_{c,\beta}$ then any nonnegative solution $u \in C^2(\Omega \setminus \{0\}) \cap C(\bar{\Omega} \setminus \{0\})$ of (1-1) in $\Omega \setminus \{0\}$ vanishing on $\partial\Omega$ must be zero.*

The paper is organized as follows. In Section 2, we prove Proposition 1.1 by treating successively the equations (1-3) and (1-1). Section 3 is devoted to the proof of Theorem A. Construction of weakly singular solutions u_k^Ω is based on an approximation method and delicate estimates on approximating solutions and on their gradient. In Section 4, the existence of a strongly singular solution u_∞^Ω (Theorem B) is obtained due to the monotonicity of the sequence $\{u_k^\Omega\}$ and a priori estimates established in Section 2. In Section 5, by combining Harnack's inequality, a scaling argument, and the asymptotic behavior of weakly singular solutions and a strongly singular solution, we obtain a complete description of isolated singularities (Theorem C). Finally, Theorem D is proved thanks to a nonexistence result for suitable equations on the unit sphere S^{N-1}.

Notation and terminology. Denote by $B_r(x_0)$ the ball of center $x_0 \in \mathbb{R}^N$ and radius r. Henceforth, we simply write B_r for $B_r(0)$. Unless otherwise stated, Ω is a C^2 bounded domain containing the origin 0. Fix $d_1 \in (0,1)$ such that $B_{3d_1} \Subset \Omega$ and put $d_2 = \operatorname{diam}(\Omega)$.

Define, for $\ell > 0$ and $x \in \Omega_\ell := \ell^{-1}\Omega$,

$$R_\ell[u](x) = \ell^{N-2} u(\ell x), \quad S_\ell[u](x) = \ell^{\frac{2+\alpha}{p-1}} u(\ell x), \quad T_\ell[u](x) = \ell^{\frac{2+\beta-q}{q-1}} u(\ell x). \tag{1-23}$$

If u is a solution of (1-2) (resp., (1-3)) in $\Omega\setminus\{0\}$ then $S_\ell[u]$ (resp., $T_\ell[u]$) is a solution of (1-2) (resp., (1-3)) in $\Omega_\ell \setminus \{0\}$. If $\Omega = \Omega_\ell$ and $u = S_\ell[u]$ (resp., $u = T_\ell[u]$) for every $\ell > 0$, we say that S_ℓ (resp., T_ℓ) is a *similarity transformation* and u is a *self-similar solution* of (1-2) (resp., (1-3)).

2. A priori estimates

2.1. *A priori estimates on solutions of* (1-3). Let us start this section by recalling the comparison principle [Gilbarg and Trudinger 2001, Theorem 10.1].

Proposition 2.1. *Let $\mathcal{O}$ be a bounded domain in $\mathbb{R}^N$. Assume $H : \mathcal{O} \times \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R}_+$ is nondecreasing with respect to u for any $(x,\xi) \in \mathcal{O} \times \mathbb{R}^N$, continuously differentiable with respect to ξ, and $H(x,0,0) = 0$. Let $u_1, u_2 \in C^2(\mathcal{O}) \cap C(\bar{\mathcal{O}})$ be two nonnegative functions satisfying*

$$-\Delta u_1 + H(x,u_1,\nabla u_1) \le -\Delta u_2 + H(x,u_2,\nabla u_2) \quad \text{in } \mathcal{O}$$

and $u_1 \le u_2$ on $\partial\mathcal{O}$. Then $u_1 \le u_2$ in $\mathcal{O}$.

We shall establish a priori estimates on solutions of (1-3) and on their gradients. By using Bernstein's method (see [Lasry and Lions 1989; Lions 1985]), we derive estimates on the gradients of solutions of (1-3).

Lemma 2.2. *Assume $\beta > -1$ and $q > 1$. There exists $c_3 = c_3(N,q,\beta)$ such that if $u \in C^2(\Omega\setminus\{0\})$ is a solution of (1-3) in $\Omega \setminus \{0\}$ then*

$$|\nabla u(x)| \le c_3 |x|^{-\frac{1+\beta}{q-1}} \quad \text{for all } x \in \bar{B}_{d_1} \setminus \{0\}. \tag{2-1}$$

Proof. Pick an arbitrary point $x_0 \in \overline{B_{d_1}} \setminus \{0\}$ and denote $\rho_0 = |x_0|$. Take $\eta \in C^\infty(\mathbb{R}^N)$ such that $0 \le \eta \le 1$, $\operatorname{supp} \eta \subset B_{1/2}$ and $\eta \equiv 1$ in $B_{1/3}$. Put $\phi(x) = \eta(\rho_0^{-1}(x - x_0))$; then $|D^2\phi| \le c_3' \rho_0^{-2}$ and $|\nabla \phi| \le c_3' \rho_0^{-1} \phi^{\frac{1}{2}}$ with $c_3' = c_3'(N)$. Set $w = \phi^{2m} |\nabla u|^2$ with $m = \frac{1}{2(q-1)}$ and define the operator

$$\mathcal{L}[w] := -\Delta w + q|x|^\beta |\nabla u|^{q-2} \nabla u \cdot \nabla w.$$

Due to (1-3) we get

$$\mathcal{L}[w] = -2m(2m-1)\phi^{2(m-1)} |\nabla \phi|^2 |\nabla u|^2 - 2m\phi^{2m-1} \Delta\phi |\nabla u|^2 - 8m\phi^{2m-1} \sum_{i,j} \partial_i \phi \, \partial_j u \, \partial_{ij} u$$
$$- 2\phi^{2m} |D^2 u|^2 - 2\beta |x|^{\beta-2} \phi^{2m} |\nabla u|^q x \nabla u + 2mq|x|^\beta \phi^{2m-1} |\nabla u|^q \nabla\phi \nabla u.$$

By virtue of the inequality $N|D^2u|^2 \ge (\Delta u)^2$ and the inequality $2ab \le a^2 + b^2$ for any $a, b \in \mathbb{R}$, we obtain, in $B_{\rho_0/2}(x_0)$,

$$\mathcal{L}[w] \le c_4 \Big(\rho_0^{-2} \phi^{2m-1} |\nabla u|^2 + \rho_0^{\beta-1} \phi^{2m} |\nabla u|^{q+1} + \rho_0^{\beta-1} \phi^{2m-\frac{1}{2}} |\nabla u|^{q+1} \Big) - \frac{\phi^{2m} |x|^{2\beta} |\nabla u|^{2q}}{N} \tag{2-2}$$

where $c_4 = c_4(\beta, q, N)$. Denote by x^* a maximizer of w then $\mathcal{L}[w](x^*) \ge 0$. In light of the relation $|\nabla u| = \phi^{-m} w^{\frac{1}{2}}$, the fact that $\frac{1}{2}\rho_0 \le |x| \le \frac{3}{2}\rho_0$ with $x \in B_{\rho_0/2}(x_0)$ and (2-2), we deduce

$$w(x^*)^{q-1} \le c_5 \big(\rho_0^{-2(\beta+1)} + \rho_0^{-(\beta+1)} w(x^*)^{\frac{q-1}{2}} \big),$$

where $c_5 = c_5(\beta, q, N)$. Consequently,

$$\max_{x \in B_{\rho_0/2}(x_0)} (\phi^{2m} |\nabla u|^2) \le w(x^*) \le c_5' \rho_0^{-\frac{2(1+\beta)}{q-1}}.$$

Therefore, $|\nabla u(x_0)| \le c_6 |x_0|^{-\frac{1+\beta}{q-1}}$, where c_6 depends on N, q, and β. □

Remark. From Lemma 2.2, one can verify that if $u \in C^2(\Omega \setminus \{0\})$ is a positive solution of (1-3) then, for every $x \in B_{d_1} \setminus \{0\}$,

$$u(x) \le \max\{u(x) : x \in \partial B_{d_1}\} + c_3 \frac{q-1}{2+\beta-q} \Big(|x|^{-\frac{2+\beta-q}{q-1}} - d_1^{-\frac{2+\beta-q}{q-1}} \Big)$$

if $q \ne 2 + \beta$, and

$$u(x) \le \max\{u(x) : x \in \partial B_{d_1}\} + c_3 (\ln d_1 - \ln|x|) \tag{2-3}$$

if $q = 2 + \beta$. Consequently, when $q > 2 + \beta$, we can conclude that u remains bounded. Therefore, in the sequel, we consider the case $q \le 2 + \beta$.

We next derive an upper bound for subsolutions of (1-3) with $\beta \ge 0$.

Lemma 2.3. *Assume $K > 0$, $\beta \ge 0$, and $1 < q < 2 + \beta$. If $u \in C^2(\Omega \setminus \{0\}) \cap C(\overline{\Omega} \setminus \{0\})$ is a positive function such that*

$$-\Delta u + K|x|^\beta |\nabla u|^q \le 0 \tag{2-4}$$

in $\Omega \setminus \{0\}$ *and vanishing on* $\partial\Omega$, *then*

$$u(x) \le c_7 |x|^{-\frac{2+\beta-q}{q-1}} \tag{2-5}$$

for every $x \in \Omega \setminus \{0\}$, *where* $c_7 = K^{-\frac{1}{q-1}}(1+\beta)^{\frac{1}{q-1}}(q-1)^{\frac{q-2}{q-1}}(2+\beta-q)^{-1}$.

Proof. Let $\epsilon > 0$ be small, and put $\Phi_\epsilon(x) = c_7(|x|-\epsilon)^{-\frac{2+\beta-q}{q-1}} + \epsilon$ with $x \in B_\epsilon^c$. By a simple computation, we get $-\Delta\Phi_\epsilon + K|x|^\beta|\nabla\Phi_\epsilon|^q \ge 0$ in $\Omega \setminus \overline{B}_\epsilon$. Since Φ_ϵ dominates u on $\partial\Omega \cup \partial B_\epsilon$, it follows from Proposition 2.1 that $\Phi_\epsilon \ge u$ in $\Omega \setminus B_\epsilon$. Letting $\epsilon \to 0$ leads to (2-5). □

Combining Lemmas 2.2 and 2.3 we get:

Lemma 2.4. *Let* $\beta > -1$ *and* $1 < q < 2+\beta$. *There exists a constant* $c_8 = c_8(N, q, \beta, d_1, d_2)$ *such that if* $u \in C^2(\Omega \setminus \{0\}) \cap C(\overline{\Omega} \setminus \{0\})$ *is a solution of* (1-3) *vanishing on* $\partial\Omega$ *then*

$$u(x) \le c_8 |x|^{-\frac{2+\beta-q}{q-1}} \quad \textit{for all } x \in \Omega \setminus \{0\}. \tag{2-6}$$

Proof. If $\beta \ge 0$ then (2-6) follows from (2-5). Next we consider $\beta \in (-1, 0)$. Fix $x \in B_{d_1} \setminus \{0\}$ and pick $z \in \partial B_{d_1}$ such that $|z - x| = d_1 - |x|$. By Lemmas 2.2 and 2.3,

$$u(x) \le c_7 d_1^{-\frac{2+\beta-q}{q-1}} + c_3 \frac{q-1}{2+\beta-q} |x|^{-\frac{2+\beta-q}{q-1}} \le c_9 |x|^{-\frac{2+\beta-q}{q-1}} \quad \text{for all } x \in B_{d_1} \setminus \{0\}, \tag{2-7}$$

where $c_9 = c_9(N, q, \beta, d_1, d_2)$. Next put $c_9' > \max\{c_9, c_7\}$ so that the function $x \mapsto c_9'|x|^{-\frac{2+\beta-q}{q-1}}$ is a supersolution of (1-3) in $\Omega \setminus B_{d_1/2}$ which dominates u on $\partial\Omega \cup \partial B_{d_1/2}$. By Proposition 2.1, $u(x) \le c_9'|x|^{-\frac{2+\beta-q}{q-1}}$ for every $x \in \Omega \setminus B_{d_1/2}$. This, together with (2-7), implies (2-6). □

By a similar argument, we obtain the following result.

Lemma 2.5. *Let* $\beta > -1$ *and* $1 < q < 2+\beta$. *There exist* $c_i = c_i(N, q, \beta)$ *with* $i = 10, 11$ *such that if* $u \in C^2(\mathbb{R}^N \setminus \{0\})$ *is a solution of* (1-3) *in* $\mathbb{R}^N \setminus \{0\}$ *satisfying* $\lim_{|x|\to\infty} u(x) = 0$ *then*

$$u(x) \le c_{10} |x|^{-\frac{2+\beta-q}{q-1}} \quad \textit{and} \quad |\nabla u(x)| \le c_{11} |x|^{-\frac{1+\beta}{q-1}} \quad \textit{for all } x \in \mathbb{R}^N \setminus \{0\}. \tag{2-8}$$

2.2. ***A priori estimates on solutions of* (1-1).** We recall that τ is defined in (1-8). Due to the Keller–Osserman estimate and the above result, we obtain a sharp upper bound for solutions of (1-1).

Lemma 2.6. *Let* $\alpha > -2$, $\beta > -1$, $p > 1$, *and* $1 < q < 2+\beta$. *There exists* $c_{12} = c_{12}(\alpha, \beta, N, p, q, d_1, d_2)$ *such that if* u *is a positive solution of* (1-1) *in* $\Omega \setminus \{0\}$ *vanishing on* $\partial\Omega$ *then*

$$u(x) \le c_{12} |x|^{-\tau} \quad \textit{for all } x \in \Omega \setminus \{0\}. \tag{2-9}$$

Proof. Since u is a positive subsolution of (1-2), due to Keller–Osserman estimate, there exists a constant $c_{13} = c_{13}(N, p, \alpha)$ such that

$$u(x) \le c_{13} |x|^{-\frac{2+\alpha}{p-1}} \quad \text{for all } x \in \Omega \setminus \{0\}.$$

We consider two cases: $D \le 1$ and $D > 1$ where D is defined in (1-20).

Case 1: $D \le 1$. In this case, $\tau = \frac{2+\alpha}{p-1}$ and hence we obtain (2-9).

Case 2: $D > 1$. Notice that in this case $\tau = \frac{2+\beta-q}{q-1}$. For $\epsilon \in (0, d_1)$, let w_ϵ be the solution of

$$-\Delta w + |x|^\beta |\nabla w|^q = 0 \quad \text{in } \Omega \setminus \bar{B}_\epsilon, \quad \text{such that } w = \begin{cases} u & \text{on } \partial B_\epsilon, \\ 0 & \text{on } \partial\Omega. \end{cases} \tag{2-10}$$

By Proposition 2.1, $u \le w_\epsilon$ in $\Omega \setminus B_\epsilon$. Therefore, $u \le w_{\epsilon'} \le w_\epsilon$ in $\Omega \setminus B_{\epsilon'}$ for $0 < \epsilon < \epsilon'$. It can be checked that the function $x \mapsto c_{14}|x|^{-\frac{2+\alpha}{p-1}}$ (with $c_{14} > c_{13}$ large, depending on N, p, q, α, β, and d_2) is a supersolution of (1-3) which dominates w_ϵ on $\partial\Omega \cup \partial B_\epsilon$. By the comparison principle, $w_\epsilon(x) \le c_{14}|x|^{-\frac{2+\alpha}{p-1}}$ for $x \in \Omega \setminus B_\epsilon$. Consequently, the sequence $\{w_\epsilon\}$ is locally uniformly bounded in $\Omega \setminus \{0\}$. In light of local regularity results for elliptic equations [DiBenedetto 1983], for every compact subset $\mathcal{O} \Subset \Omega \setminus \{0\}$, there exist constants $M > 0$ and $\mu \in (0, 1)$ depending on N, p, q, α, β, d_2, and $\operatorname{dist}(0, \mathcal{O})$ such that $\|w_\epsilon\|_{C^{1,\mu}(\mathcal{O})} \le M$. Therefore, $\{w_\epsilon\}$ converges to a function $\tilde{w}$ in $C^1_{\text{loc}}(\Omega \setminus \{0\})$ which is a solution of (1-3) in $\Omega \setminus \{0\}$, vanishing on $\partial\Omega$, and satisfying $\tilde{w} \ge u$ in $\Omega \setminus \{0\}$. By virtue of Lemma 2.4, $\tilde{w} \le c_8|x|^{-\frac{2+\beta-q}{q-1}}$ for every $x \in \Omega \setminus \{0\}$. Consequently, $u \le c_8|x|^{-\frac{2+\beta-q}{q-1}}$ for every $x \in \Omega \setminus \{0\}$. This completes the proof. □

We next establish a sharp estimate from above for the gradient of solutions of (1-1).

Proposition 2.7. *Let $\alpha > -2$, $\beta > -1$, $p > 1$, and $1 < q < 2+\beta$. There exists $c_{15} = c_{15}(\alpha, \beta, N, p, q, d_1, d_2)$ such that if u is a nonnegative solution of (1-1) in $\Omega \setminus \{0\}$ vanishing on $\partial\Omega$ then*

$$|\nabla u(x)| \le c_{15}|x|^{-(\tau+1)} \quad \textit{for all } x \in B_{d_1} \setminus \{0\}. \tag{2-11}$$

Proof. Let x_0, ρ_0, η, ϕ, w, m, $\mathcal{L}[w]$, and x^* as in the proof of Lemma 2.2. Then we get

$$\begin{aligned} \mathcal{L}[w] = &-2m(2m-1)\phi^{2(m-1)}|\nabla\phi|^2|\nabla u|^2 - 2m\phi^{2m-1}\Delta\phi|\nabla u|^2 - 8m\phi^{2m-1}\sum_{i,j}\partial_i\phi\,\partial_j u\,\partial_{ij}u \\ &- 2\phi^{2m}|D^2u|^2 - 2\alpha|x|^{\alpha-2}\phi^{2m}u^p\, x\nabla u - 2p|x|^\alpha\phi^{2m}u^{p-1}|\nabla u|^2 \\ &- 2\beta|x|^{\beta-2}\phi^{2m}|\nabla u|^q\, x\nabla u + 2mq|x|^\beta\phi^{2m-1}|\nabla u|^q\nabla\phi\nabla u. \end{aligned}$$

Case 1: $D \ge 1$. In this case, we have

$$\frac{(\beta+1)(1-2q)}{q-1} \le \alpha - 2\beta - 1 - \tau p, \tag{2-12}$$

where τ is defined in (1-8). By Lemma 2.6 and Young's inequality, proceeding as in the proof of Lemma 2.2, we obtain in $B_{\rho_0/2}(x_0)$

$$w(x^*)^{q-\frac{1}{2}} \le c_{16}\big(\rho_0^{-2(\beta+1)}w(x^*)^{\frac{1}{2}} + \rho_0^{\alpha-2\beta-1-\tau p} + \rho_0^{-(\beta+1)}w(x^*)^{\frac{q}{2}}\big), \tag{2-13}$$

where $c_{16} = c_{16}(\alpha, \beta, p, q, N, d_1, d_2)$. By Young's inequality, we get

$$\rho_0^{-2(\beta+1)}w(x^*)^{\frac{1}{2}} \le \frac{1}{q}\rho_0^{-(\beta+1)}w(x^*)^{\frac{q}{2}} + \frac{q-1}{q}\rho_0^{\frac{(\beta+1)(1-2q)}{q-1}}. \tag{2-14}$$

From (2-12), (2-13), and (2-14), we deduce

$$w(x^*)^{q-\frac{1}{2}} \le c_{17}\big(\rho_0^{-(\beta+1)}w(x^*)^{\frac{q}{2}} + \rho_0^{\frac{(\beta+1)(1-2q)}{q-1}}\big), \tag{2-15}$$

which implies

$$\rho_0^{\beta+1} w(x^*)^{\frac{q-1}{2}} \le c_{17}\big(\rho_0^{-\frac{(\beta+1)q}{q-1}} w(x^*)^{-\frac{q}{2}} + 1\big), \tag{2-16}$$

where $c_{17} = c_{17}(\alpha, \beta, p, q, N, d_1, d_2)$. Consequently, $w(x^*) \le c_{18}\rho_0^{-\frac{2(1+\beta)}{q-1}}$, and therefore

$$|\nabla u(x)| \le c_{19}|x|^{-\frac{1+\beta}{q-1}} \quad \text{for all } x \in B_{d_1} \setminus \{0\}, \tag{2-17}$$

where $c_i = c_i(\alpha, \beta, N, p, q, d_1, d_2)$ with $i = 18, 19$. Notice that $\dfrac{1+\beta}{q-1} = \tau + 1$; hence we obtain (2-11).

Case 2: $D < 1$. Take $x_0 \in B_{d_1} \setminus \{0\}$. Put $\ell = |x_0| \in (0, d_1)$ then $S_\ell[u]$ is a solution of

$$-\Delta v + |x|^\alpha v^p + \ell^{\frac{p(2+\beta-q)-\alpha(q-1)-q-\beta}{p-1}} |x|^\beta |\nabla v|^q = 0 \quad \text{in } \Omega_\ell \setminus \{0\}. \tag{2-18}$$

By the regularity result in [DiBenedetto 1983], there exists $c_{20} = c_{20}(\alpha, \beta, p, q)$ such that

$$\sup\big\{|\nabla S_\ell[u](x)| : x \in B_{3/2} \setminus B_{3/4}\big\} \le c_{20}.$$

Consequently,

$$\ell^{\frac{1+p+\alpha}{p-1}} |\nabla u(\ell x)| \le c_{21} \quad \text{for all } x \in B_{3/2} \setminus B_{3/4}.$$

By choosing $x = \ell^{-1} x_0$, we derive $|\nabla u(x_0)| \le c_{22}|x_0|^{-\frac{1+p+\alpha}{p-1}}$. This completes the proof since

$$\frac{1+p+\alpha}{p-1} = \tau + 1. \qquad \square$$

Proof of Proposition 1.1. Estimates (1-9) and (1-10) follow directly from Lemmas 2.2, 2.4, and 2.6, as well as Proposition 2.7. $\square$

3. Weakly singular solutions

We start with the existence of weakly singular solutions of (1-1). The construction is based on approximation method.

Proof of Theorem A. We prove the theorem in five steps.

Step 1: Construction of solutions. Let $k > 0$. For every $\epsilon > 0$, let $u_{k,\epsilon}^\Omega$ be the unique solution of

$$\begin{cases} -\Delta u + |x|^\alpha u^p + |x|^\beta |\nabla u|^q = 0 & \text{in } \Omega \setminus \overline{B}_\epsilon, \\ \hfill u = 0 & \text{on } \partial\Omega, \\ \hfill u = k\Gamma_N(\epsilon) & \text{on } \partial B_\epsilon. \end{cases} \tag{3-1}$$

The existence of $u_{k,\epsilon}^\Omega$ can be obtained by using an argument similar to the proof of [Gilbarg and Trudinger 2001, Theorem 11.4] and the uniqueness follows from the comparison principle Proposition 2.1. Moreover, by the comparison principle, $0 \le u_{k,\epsilon}^\Omega \le k\Gamma_N$ in $\overline{\Omega} \setminus B_\epsilon$ and $u_{k,\epsilon}^\Omega \le u_{k,\epsilon'}^\Omega$ in $\overline{\Omega} \setminus B_{\epsilon'}$ for every $0 < \epsilon < \epsilon'$. Therefore, $u_k^\Omega := \lim_{\epsilon \to 0} u_{k,\epsilon}^\Omega$ satisfies

$$u_k^\Omega(x) \le k\Gamma_N(x) \quad \text{for all } x \in \Omega \setminus \{0\}. \tag{3-2}$$

By regularity results for elliptic equations, u_k^Ω is a solution of (1-1) in $\Omega \setminus \{0\}$ vanishing on $\partial\Omega$.

Fix an arbitrary point $x_0 \in \overline{B_{d_1}} \setminus \overline{B_\epsilon}$ and put $\ell = |x_0| \in (\epsilon, d_1]$. Note that $R_\ell[u^\Omega_{k,\epsilon}]$ solves

$$\begin{cases} -\Delta v + \ell^{N+\alpha-p(N-2)}|x|^\alpha v^p + \ell^{N+\beta-q(N-1)}|x|^\beta|\nabla v|^q = 0 & \text{in } \Omega_\ell \setminus \overline{B_{\epsilon/\ell}}, \\ v = 0 & \text{on } \partial\Omega_\ell, \\ v = k\Gamma_N\big(\frac{\epsilon}{\ell}\big) & \text{on } \partial B_{\epsilon/\ell}. \end{cases} \tag{3-3}$$

Since $1 < p < p_{c,\alpha}$ and $1 < q < q_{c,\beta}$, it follows that

$$\ell^{N+\alpha-p(N-2)}|x|^\alpha < \max\{1, 3^\alpha\} \quad \text{and} \quad \ell^{N+\beta-q(N-1)}|x|^\beta < \max\{1, 3^\beta\} \quad \text{for all } x \in B_3 \setminus B_1.$$

By the maximum principle, $R_\ell[u^\Omega_{k,\epsilon}] \le k\Gamma_N$ in $\Omega_\ell \setminus \overline{B_{\epsilon/\ell}}$, which implies $R_\ell[u^\Omega_{k,\epsilon}] \le k\Gamma_N(1)$ in $B_3 \setminus B_1$. Due to local regularity for elliptic equations (see, e.g., [DiBenedetto 1983]), there exist constants $c_{23} = c_{23}(N, \alpha, \beta, p, q, k)$ and $\mu = \mu(N, \alpha, \beta, p, q, k) \in (0, 1)$ such that

$$\|R_\ell[u^\Omega_{k,\epsilon}]\|_{C^{1,\mu}(B_{5/2}\setminus\overline{B_{3/2}})} \le c_{23}.$$

Again by the regularity results (see [Lieberman 1988, Theorem 1] and [DiBenedetto 1983]), there exists $c_{24} = c_{24}(\alpha, \beta, N, p, q, k)$ such that

$$\ell^{N-1}\sup\big\{|\nabla u^\Omega_{k,\epsilon}(\ell x)| : |x| = 1\big\} \le c_{24}.$$

By choosing $x = \ell^{-1}x_0$, we deduce $|\nabla u^\Omega_{k,\epsilon}(x_0)| \le c_{24}|x_0|^{1-N}$. Thus

$$|\nabla u^\Omega_{k,\epsilon}(x)| \le c_{25}|x|^{1-N} \quad \text{for all } x \in \Omega \setminus B_\epsilon \tag{3-4}$$

with $c_{25} = c_{25}(\alpha, \beta, N, p, q, k, d_1, d_2)$.

Step 2: Proof of (1-16). The solution $u^\Omega_{k,\epsilon}$ can be written in the form

$$u^\Omega_{k,\epsilon}(x) = k\Gamma_N(\epsilon) - \mathbb{G}^{\Omega\setminus\overline{B_\epsilon}}[F \circ u^\Omega_{k,\epsilon}](x),$$

where $\mathbb{G}^{\Omega\setminus\overline{B_\epsilon}}$ is the Green operator in $\Omega \setminus \overline{B_\epsilon}$ [Marcus and Véron 2014, Theorem 1.2.2]. Hence, by (3-4),

$$k\Gamma_N(x) \ge u^\Omega_{k,\epsilon}(x) \ge k\Gamma_N(x) - c_{26}\mathbb{G}^\Omega[|\cdot|^{\alpha+p(2-N)} + |\cdot|^{\beta+q(1-N)}](x) \quad \text{for all } x \in \Omega \setminus \overline{B_\epsilon}.$$

By letting $\epsilon \to 0$, we get

$$k\Gamma_N(x) \ge u^\Omega_k(x) \ge k\Gamma_N(x) - c\mathbb{G}^\Omega[|\cdot|^{\alpha+p(2-N)} + |\cdot|^{\beta+q(1-N)}](x) \quad \text{for all } x \in \Omega \setminus \{0\}. \tag{3-5}$$

It is classical (see [op. cit.]) that

$$G^\Omega(x, y) \sim \min\big\{|x-y|^{2-N}, \rho(x)\rho(y)|x-y|^{-N}\big\}$$

for every $x, y \in \Omega, x \neq y$, where $\rho(x) = \operatorname{dist}(x, \partial\Omega)$. Therefore there exists $c_{27} = c_{27}(N, \Omega)$ such that, for x near 0,

$$\frac{\mathbb{G}^\Omega[|\cdot|^{\alpha+p(2-N)} + |\cdot|^{\beta+q(1-N)}](x)}{\Gamma_N(x)} \le c_{27}|x|^{N-2}\int_\Omega |x-y|^{2-N}\big(|y|^{\alpha-p(N-2)} + |y|^{\beta-q(N-1)}\big)\,dy. \tag{3-6}$$

Choose α' and β' such that $p(N-2)-N < \alpha' < \min\{\alpha, p(N-2)-2\}$ and $q(N-1)-N < \beta' < \min\{\beta, q(N-1)-2\}$. Then by [Lieb and Loss 1997, Corollary 5.10],

$$\begin{aligned}\int_\Omega |x-y|^{2-N}|y|^{\alpha-p(N-2)}\,dy &\le c_{28}\,d_2^{\alpha-\alpha'}|x|^{2+\alpha'-p(N-2)},\\ \int_\Omega |x-y|^{2-N}|y|^{\beta-q(N-1)}\,dy &\le c_{28}\,d_2^{\beta-\beta'}|x|^{2+\beta'-q(N-1)}.\end{aligned} \tag{3-7}$$

Combining (3-6) and (3-7) yields

$$\lim_{|x|\to 0}\frac{\mathbb{G}^\Omega[|\cdot|^{\alpha+p(2-N)}+|\cdot|^{\beta+q(1-N)}](x)}{\Gamma_N(x)}=0. \tag{3-8}$$

From (3-5) and (3-8), we obtain (1-16).

Step 3: Proof of (1-17). For $\ell\in(0,1)$, put $v_\ell = R_\ell[u_k^\Omega]$ then v_ℓ is the solution of

$$\begin{cases}-\Delta v+\ell^{N+\alpha-p(N-2)}|x|^\alpha v^p+\ell^{N+\beta-q(N-1)}|x|^\beta|\nabla v|^q=0, & \text{in } \Omega_\ell\setminus\{0\}\\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad v=0 & \text{on } \partial\Omega_\ell.\end{cases} \tag{3-9}$$

Since $0<u_k^\Omega<k\Gamma_N$ in $\Omega\setminus\{0\}$, it follows that $0<v_\ell<k\Gamma_N$ in $\Omega_\ell\setminus\{0\}$.

Since $1<p<p_{c,\alpha}$ and $1<q<q_{c,\beta}$, by local regularity for elliptic equations [DiBenedetto 1983], the Arzelà–Ascoli theorem, and a standard diagonalization argument, there exists a subsequence $\{v_{\ell_n}\}$ converging to a positive harmonic function in $C^1_{\mathrm{loc}}(\mathbb{R}^N\setminus\{0\})$ as $\ell_n\to 0$. On the other hand, from (1-16), we deduce that $\{v_\ell\}$ converges to $k\Gamma_N$ uniformly in $B_2\setminus B_{1/2}$ as $\ell\to 0$. Therefore, the whole sequence $\{v_\ell\}$ converges to $k\Gamma_N$ in $C^1_{\mathrm{loc}}(\mathbb{R}^N\setminus\{0\})$ as $\ell\to 0$. In particular, $\nabla v_\ell\to k\nabla\Gamma_N$ in $B_2\setminus B_{1/2}$, which implies (1-17).

Step 4: u_k^Ω is a weak solution of (1-11). By a similar argument as in Step 1, we derive

$$|\nabla u_k^\Omega(x)|\le c_{29}k|x|^{1-N}\quad\text{for all } x\in\Omega\setminus\{0\} \tag{3-10}$$

where $c_{29}=c_{29}(\alpha,\beta,N,p,q,d_1,d_2)$. This, together with (3-2), implies $u_k^\Omega\in L^1(\Omega)$ and $F\circ u_k^\Omega\in L^1(\Omega)$.

For every $\epsilon>0$, by Green's formula, one gets

$$\int_{\Omega\setminus\overline{B}_\epsilon}\bigl(-u_k^\Omega\Delta\zeta+(F\circ u_k^\Omega)\zeta\bigr)\,dx=-\int_{\partial B_\epsilon}\frac{\partial u_k^\Omega}{\partial\boldsymbol{n}}\zeta\,dS+\int_{\partial\overline{B}_\epsilon}u_k^\Omega\frac{\partial\zeta}{\partial\boldsymbol{n}}\,dS, \tag{3-11}$$

where $\boldsymbol{n}$ is the outward normal unit vector on ∂B_ϵ. Due to (1-17), the right-hand side of (3-11) converges to $k\zeta(0)$. Therefore, thanks to dominated convergence theorem, by letting $\epsilon\to 0$, we obtain (1-12). Finally, by [Marcus and Véron 2014, Theorem 1.2.2], we get (1-15).

Step 5: Uniqueness. Assume u' is a positive solutions of (1-1) satisfying (1-16); then

$$\lim_{|x|\to 0}\frac{u_k^\Omega(x)}{u'(x)}=1.$$

Hence, for every $\delta>0$, there exists $r(\delta)>0$ such that $(1+\delta)u_k^\Omega+\delta\ge u'$ on $\partial B_{r(\delta)}$. The function $(1+\delta)u_k^\Omega+\delta$ is a supersolution of (1-1) which dominates u' on $\partial\Omega\cup\partial B_{r(\delta)}$; therefore, by the comparison principle, $(1+\delta)u_k^\Omega+\delta \ge u'$ in $\Omega\setminus B_{r(\delta)}$. Letting $\delta\to 0$ yields $u_k^\Omega \ge u'$ in $\Omega\setminus\{0\}$. By permuting u_k^Ω and u', we derive $u'=u_k^\Omega$. ☐

If Ω is replaced by $\mathbb{R}^N$, we have the following variant of Theorem A.

Proposition 3.1. *Assume $\alpha>-2$, $\beta>-1$, $1<p<p_{c,\alpha}$, and $1<q<q_{c,\beta}$. Then for any $k>0$, there exists a unique solution $u_k^{\mathbb{R}^N}\in C^2(\mathbb{R}^N\setminus\{0\})$ of (1-1) in $\mathbb{R}^N\setminus\{0\}$ satisfying*

$$\lim_{|x|\to\infty} u_k^{\mathbb{R}^N}(x)=0 \quad \textit{and} \quad u_k^{\mathbb{R}^N}(x)=k(1+o(1))\Gamma_N(x) \ \textit{as}\ |x|\to 0. \tag{3-12}$$

Moreover, $u_k^{\mathbb{R}^N}\in L^1_{\mathrm{loc}}(\mathbb{R}^N)$, $F\circ u_k^{\mathbb{R}^N}\in L^1_{\mathrm{loc}}(\mathbb{R}^N)$, and there holds

$$\int_{\mathbb{R}^N}\big(-u_k^{\mathbb{R}^N}\Delta\zeta+(F\circ u_k^{\mathbb{R}^N})\zeta\big)\,dx=k\zeta(0) \quad \textit{for all}\ \zeta\in C_c^2(\mathbb{R}^N). \tag{3-13}$$

Proof. For each $R>0$, let $u_k^{B_R}$ be the unique solution of (1-1) in $B_R\setminus\{0\}$, vanishing on ∂B_R and satisfying

$$\lim_{|x|\to 0}\frac{u_k^{B_R}(x)}{\Gamma_N(x)}=k. \tag{3-14}$$

By the comparison principle, $u_k^{B_R}\le u_k^{B_{R'}}\le k\Gamma_N$ in $B_R\setminus\{0\}$ for every $R<R'$. In light of local regularity [DiBenedetto 1983] and a standard argument,

$$u_k^{\mathbb{R}^N}:=\lim_{R\to\infty}u_k^{B_R}\in C^2(\mathbb{R}^N\setminus\{0\})$$

is a solution of (1-1) in $\mathbb{R}^N\setminus\{0\}$. By combining (3-14) and the fact that $u_k^{B_R}\le u_k^{\mathbb{R}^N}\le k\Gamma_N$ in $B_R\setminus\{0\}$ for every $R>0$, we derive (3-12). Uniqueness follows from the comparison principle. By proceeding as in the proof of Theorem A, one can verify (3-13). ☐

By a similar, and more simpler, argument as in the proof of Theorem A, one can easily obtain the existence of weakly singular solutions of (1-3).

Proposition 3.2. *Assume $\beta>-1$ and $1<q<q_{c,\beta}$ with $q_{c,\beta}$ defined in (1-14). For any $k>0$, there exists a unique solution $w_k^\Omega\in C^2(\Omega\setminus\{0\})\cap C(\overline{\Omega}\setminus\{0\})$ of*

$$-\Delta w+|x|^\beta|\nabla w|^q=k\delta_0 \quad \textit{in}\ \ \Omega,\ \textit{with}\ w=0\ \textit{on}\ \partial\Omega. \tag{3-15}$$

Moreover,

$$w_k^\Omega=kG^\Omega(\cdot,0)-\mathbb{G}^\Omega[|\cdot|^\beta|\nabla w_k^\Omega|^q]; \tag{3-16}$$

$$w_k^\Omega(x)=k(1+o(1))\Gamma_N(x) \quad \textit{as}\ |x|\to 0; \tag{3-17}$$

$$\lim_{|x|\to 0}\left(|x|^{N-1}\nabla w_k^\Omega(x)+\frac{k}{N\omega_N}\frac{x}{|x|}\right)=0. \tag{3-18}$$

Remark. In addition, by proceeding as in the proof of Proposition 3.1, we obtain the existence of the weak singular solutions $w_k^{\mathbb{R}^N}$ of (1-3) in $\mathbb{R}^N\setminus\{0\}$.

4. Strongly singular solutions

Denote by S^{N-1} the unit sphere in $\mathbb{R}^N$ and let $(r,\sigma) \in (0,\infty)\times S^{N-1}$ be the spherical coordinates in $\mathbb{R}^N \setminus \{0\}$. Let ∇' and Δ' denote respectively the covariant gradient and the Laplace–Beltrami operator on S^{N-1}. In order to characterize strongly singular solutions of (1-1), we study the following quasilinear equation on S^{N-1}:

$$-\Delta'\omega + \lambda\omega^{\frac{\alpha(q-1)+q+\beta}{2+\beta-q}} + \left(\left(\frac{2+\beta-q}{q-1}\right)^2\omega^2 + |\nabla'\omega|^2\right)^{\frac{q}{2}} - \Lambda\omega = 0, \tag{4-1}$$

where

$$\lambda \ge 0, \quad \text{and} \quad \Lambda = \Lambda(N,q,\beta) := \frac{2+\beta-q}{q-1}\left(\frac{q+\beta}{q-1} - N\right).$$

We introduce an auxiliary function

$$g_\lambda(t) = \lambda t^{\frac{(2+\alpha)(q-1)}{2+\beta-q}} + \left(\frac{2+\beta-q}{q-1}\right)^q t^{q-1} - \Lambda, \quad t \in (0,\infty), \quad \lambda \ge 0. \tag{4-2}$$

It is easy to see that if $1 < q < q_{c,\beta}$ then $\Lambda > 0$; therefore, the algebraic equation $g_\lambda(t) = 0$ admits a unique positive solution θ_λ. Obviously, θ_λ is a positive solution of (4-1), and θ_0 is explicitly given by

$$\theta_0 = \frac{q-1}{2+\beta-q}\left(\frac{q+\beta}{q-1} - N\right)^{\frac{1}{q-1}}. \tag{4-3}$$

Proposition 4.1. *Let $\alpha > -2$, $\beta > -1$, $1 < q < 2+\beta$, and $\lambda \ge 0$. Denote by $\mathcal{E}_\lambda$ the set of C^2 positive solutions of (4-1) in S^{N-1}.*

(i) *If $q \ge q_{c,\beta}$, then $\mathcal{E}_\lambda = \varnothing$.*

(ii) *If $1 < q < q_{c,\beta}$, then $\mathcal{E}_\lambda = \{\theta_\lambda\}$.*

Proof. (i) Suppose by contradiction that ω is a positive solution of (4-1) and $\omega(\sigma_{\max}) = \max_{S^{N-1}} \omega > 0$ with $\sigma_{\max} \in S^{N-1}$. From (4-1), we get

$$\lambda\omega(\sigma_{\max})^{\frac{\alpha(q-1)+q+\beta}{2+\beta-q}} + \left(\frac{2+\beta-q}{q-1}\right)^q \omega(\sigma_{\max})^q - \Lambda\omega(\sigma_{\max}) \le 0. \tag{4-4}$$

Since $q \ge q_{c,\beta}$, we know $\Lambda \le 0$. Therefore, the left hand side is positive, which is a contradiction.

(ii) If ω is a positive solution of (4-1), let $\sigma_{\max}, \sigma_{\min} \in S^{N-1}$ such that

$$\omega(\sigma_{\max}) = \max_{S^{N-1}} \omega \ge \min_{S^{N-1}} \omega = \omega(\sigma_{\min}) > 0.$$

Clearly, $\sigma_{\max}$ satisfies (4-4) and $\sigma_{\min}$ satisfies

$$\lambda\omega(\sigma_{\min})^{\frac{\alpha(q-1)+q+\beta}{2+\beta-q}} + \left(\frac{2+\beta-q}{q-1}\right)^q \omega(\sigma_{\min})^q - \Lambda\omega(\sigma_{\min}) \ge 0. \tag{4-5}$$

Consequently, $g_\lambda(\omega(\sigma_{\max})) \le 0 \le g_\lambda(\omega(\sigma_{\min}))$. Since g_λ is strictly increasing in $(0,\infty)$, it follows that $\omega(\sigma_{\max}) \le \theta_\lambda \le \omega(\sigma_{\min})$. Thus, $\omega \equiv \theta_\lambda$. This completes the proof. □

The next lemma states existence result for both equations (1-3) and (1-1).

Lemma 4.2. *Let Ω be either a smooth bounded domain containing the origin 0 or $\mathbb{R}^N$.*

(i) *Assume $\beta > -1$ and $1 < q < q_{c,\beta}$. Then $w_\infty^\Omega := \lim_{k\to\infty} w_k^\Omega$ is a nonnegative solution of (1-3) in $\Omega \setminus \{0\}$ satisfying either $w_\infty^\Omega = 0$ on $\partial\Omega$ if Ω is bounded or $\lim_{|x|\to\infty} w_\infty^\Omega(x) = 0$ if $\Omega = \mathbb{R}^N$.*

(ii) *Assume $\alpha > -2,\ \beta > -1,\ 1 < p < p_{c,\alpha}$, and $1 < q < q_{c,\beta}$. Then $u_\infty^\Omega := \lim_{k\to\infty} u_k^\Omega$ is a nonnegative solution of (1-1) in $\Omega \setminus \{0\}$ satisfying either $u_\infty^\Omega = 0$ on $\partial\Omega$ if Ω is bounded or $\lim_{|x|\to\infty} u_\infty^\Omega(x) = 0$ if $\Omega = \mathbb{R}^N$.*

Proof. We only demonstrate (ii) since the proof of (i) is similar and simpler. It follows from Theorem A and Proposition 3.1 that $\{u_k^\Omega\}$ is increasing and bounded from above by the function $\overline{U}(x) = c_{30}|x|^{-\frac{2+\alpha}{p-1}}$ where c_{30} is a large positive constant depending on N, p, and α. Therefore, $u_\infty^\Omega := \lim_{k\to\infty} u_k^\Omega$ is a solution of (1-1) in $\Omega \setminus \{0\}$ and $u_\infty^\Omega \le \overline{U}$ in $\Omega \setminus \{0\}$. □

The asymptotic behavior of w_∞^Ω near the origin 0 is analyzed in the following result.

Proposition 4.3. *Assume $\beta > -1,\ 1 < q < q_{c,\beta}$, and Ω is either a smooth bounded domain containing the origin 0 or $\mathbb{R}^N$. Let w_∞^Ω be the solution in Lemma 4.2(i). Then w_∞^Ω is a strongly singular solution of (1-3). Moreover, with θ_0 as in (4-3),*

$$\lim_{|x|\to 0} |x|^{\frac{2+\beta-q}{q-1}} w_\infty^\Omega(x) = \theta_0 \tag{4-6}$$

$$\lim_{|x|\to 0}\left(|x|^{\frac{1+\beta}{q-1}} \nabla w_\infty^\Omega(x) + \left(\frac{q+\beta}{q-1} - N\right)^{\frac{1}{q-1}} \frac{x}{|x|}\right) = 0. \tag{4-7}$$

Proof. The proof is based upon the similarity argument.

Case 1: $\Omega = \mathbb{R}^N$. For $k > 0$, recall that w_k^Ω is the weakly singular solution of (1-3) in $\mathbb{R}^N$. For every $\ell > 0$, $T_\ell[w_k^{\mathbb{R}^N}]$ is a solution of (1-3) in $\mathbb{R}^N \setminus \{0\}$ which satisfies

$$\lim_{|x|\to 0} \frac{T_\ell[w_k^{\mathbb{R}^N}](x)}{\Gamma_N(x)} = \ell^{\frac{2+\beta-q}{q-1}+2-N} k.$$

Due to the uniqueness,

$$T_\ell[w_k^{\mathbb{R}^N}] = w^{\mathbb{R}^N}_{\ell^{(2+\beta-q)/(q-1)+2-N} k}.$$

By letting $k \to \infty$, we deduce that $T_\ell[w_\infty^{\mathbb{R}^N}] = w_\infty^{\mathbb{R}^N}$, i.e., $w_\infty^{\mathbb{R}^N}$ is self-similar. Consequently, $w_\infty^{\mathbb{R}^N}$ can be written in the form

$$w_\infty^{\mathbb{R}^N}(x) = |x|^{-\frac{2+\beta-q}{q-1}} \omega(x/|x|) \quad \text{for all } x \ne 0, \tag{4-8}$$

where ω is a positive solution of (4-1) with $\lambda = 0$. Since $1 < q < q_{c,\beta}$, by Proposition 4.1, $\omega \equiv \theta_0$. Therefore,

$$w_\infty^{\mathbb{R}^N}(x) = \theta_0 |x|^{-\frac{2+\beta-q}{q-1}} =: W_0(x) \quad \text{for all } x \ne 0,$$

the unique self-similar solution of (1-3) in $\mathbb{R}^N \setminus \{0\}$.

Case 2: Ω is a bounded smooth domain. Since $T_\ell[w_k^\Omega]=w^{\Omega_\ell}_{\ell^{(2+\beta-q)/(q-1)+2-N}k}$ by uniqueness, it follows that

$$T_\ell[w_\infty^\Omega]=w_\infty^{\Omega_\ell}. \tag{4-9}$$

Since $w_\infty^\Omega(x)\le c_8|x|^{-\frac{2+\beta-q}{q-1}}$ in $\Omega\setminus\{0\}$, $w_\infty^{\Omega_\ell}$ satisfies the same estimate in $\Omega_\ell\setminus\{0\}$ for every $\ell\in(0,1)$. By local regularity for elliptic equations and Arzelà–Ascoli theorem, there exists a subsequence $\{w_\infty^{\Omega_{\ell_n}}\}$ converging in $C^1_{\mathrm{loc}}(\mathbb{R}^N\setminus\{0\})$ to a function w_0 which is a solution of (1-3) in $\mathbb{R}^N\setminus\{0\}$.

If Ω is star-shaped with respect to the origin 0 then we get $w_k^{\Omega_\ell}\le w_k^{\Omega_{\ell'}}$ for every $k>0$ and $0<\ell'<\ell<1$, which implies that $w_\infty^{\Omega_\ell}\le w_\infty^{\Omega_{\ell'}}$ for every $0<\ell'<\ell<1$. Therefore, the whole sequence $\{w_\infty^{\Omega_\ell}\}$ converges to w_0 in $C^1_{\mathrm{loc}}(\mathbb{R}^N\setminus\{0\})$ as $\ell\to 0$. By (4-9), for any $\ell,\ell'>0$,

$$T_\ell[T_{\ell'}[w_\infty^\Omega]]=T_\ell[w_\infty^{\Omega_{\ell'}}]=w_\infty^{\Omega_{\ell'\ell}}.$$

By letting $\ell'\to 0$, we obtain $T_\ell[w_0]=w_0$ for every $\ell>0$, namely w_0 is a self-similar solution of (1-3) in $\mathbb{R}^N\setminus\{0\}$. Therefore, $w_0=w_\infty^{\mathbb{R}^N}=W_0$ and consequently,

$$\lim_{\ell\to 0}\ell^{\frac{2+\beta-q}{q-1}}w_\infty^\Omega(\ell x)=\theta_0|x|^{-\frac{2+\beta-q}{q-1}}.$$

By putting $y=\ell x$ with $|x|=1$, we get (4-6).

In general, if Ω is not necessarily star-shaped with respect to the origin 0, since $\overline{B_{3d_1}}\subset\Omega\subset B_{d_2}$, it follows that $w_\infty^{B_{3d_1}}\le w_\infty^\Omega\le w_\infty^{B_{d_2}}$. As (4-6) holds for $w_\infty^{B_{3d_1}}$ (i.e., Ω is replaced by B_{3d_1}) and $w_\infty^{B_{d_2}}$, we derive (4-6). Consequently, for every $x\ne 0$,

$$w_0(x)=\lim_{n\to\infty}w_\infty^{\Omega_{\ell_n}}(x)=\lim_{n\to\infty}\ell_n^{\frac{2+\beta-q}{q-1}}w_\infty^\Omega(\ell_n x)=\theta_0|x|^{-\frac{2+\beta-q}{q-1}}=W_0(x).$$

Hence the whole sequence $\{w_\infty^{\Omega_\ell}\}_\ell$ converges to W_0 in $C^1_{\mathrm{loc}}(\mathbb{R}^N\setminus\{0\})$ as $\ell\to 0$. By using a similar argument as in Step 3 of the proof of Theorem A, we obtain (4-7). This implies $|x|^\beta|\nabla w_\infty^\Omega|^q\notin L^1(B_\epsilon)$ for every $\epsilon>0$. Thus w_∞^Ω is a strongly singular solution of (1-3). □

Note that (1-1) does not admit any similarity transformation, except when $D=1$. However, due to the asymptotic behavior of v_∞^Ω (the strongly singular solution of (1-2)) and of w_∞^Ω near 0, we can establish the asymptotic behavior of u_∞^Ω. Put

$$\Theta=\begin{cases}\vartheta & \text{if } D<1,\\ \theta_1 & \text{if } D=1,\\ \theta_0 & \text{if } D>1,\end{cases} \tag{4-10}$$

where ϑ is as in (1-7) and θ_λ $(\lambda=0,1)$ is given in (4-2).

Now we are ready to deal with strongly singular solution of (1-1).

Proposition 4.4. *Assume $\alpha>-2$, $\beta>-1$, $1<p<p_{c,\alpha}$, and $1<q<q_{c,\beta}$. Let Ω be either a smooth bounded domain containing the origin 0 or $\mathbb{R}^N$ and u_∞^Ω be the solution of (1-1) defined in Lemma 4.2. Then u_∞^Ω is a strongly singular solution of (1-1). Moreover (1-18) and (1-19) hold.*

Proof. We consider three cases.

Case 1: $D = 1$. In this case, S_ℓ is a similarity transformation for (1-1). Therefore, (1-18) and (1-19) can be obtained by proceeding as in the proof of Proposition 4.3 and consequently u_∞^Ω is a strongly singular solution of (1-1). Notice that if $\Omega = \mathbb{R}^N$ then $\Omega_\ell = \mathbb{R}^N$ and $u = 0$ on $\partial\Omega_\ell$ is understood as $u(x) \to 0$ as $|x| \to \infty$.

Case 2: $D > 1$. For every $\ell \in (0, 1)$, put $W_\ell = T_\ell[u_\infty^\Omega]$. Then W_ℓ is a solution of

$$-\Delta u + \ell^{\frac{\alpha(q-1)+q+\beta-p(2+\beta-q)}{q-1}}|x|^\alpha u^p + |x|^\beta |\nabla u|^q = 0 \quad \text{in } \Omega_\ell \setminus \{0\}, \text{ with } u = 0 \text{ on } \partial\Omega_\ell. \tag{4-11}$$

By the regularity result [DiBenedetto 1983], for every $R > 1$ there exist $M = M(\alpha, \beta, p, q, N, R, d_1, d_2)$ and $\mu = \mu(\alpha, \beta, p, q, N, d_1, d_2) \in (0, 1)$ such that

$$\|W_\ell\|_{C^{1,\mu}(B_R \setminus B_{R^{-1}})} < M.$$

Consequently, there exists a subsequence $\{W_{\ell_n}\}$ which converges to a function $\widetilde{W}$ in $C^1_{\mathrm{loc}}(\mathbb{R}^N \setminus \{0\})$ as $\ell_n \to 0$. The function $\widetilde{W}$ is a solution of (1-3) in $\mathbb{R}^N \setminus \{0\}$ satisfying $\lim_{|x|\to\infty} \widetilde{W}(x) = 0$. By Proposition 2.1, $w_\infty^{\mathbb{R}^N} \ge \widetilde{W} \ge u_k^{\mathbb{R}^N}$ for every $k > 0$. Therefore, thanks to (3-12), we get

$$\liminf_{x\to 0} \frac{\widetilde{W}(x)}{w_k^{\mathbb{R}^N}(x)} = \liminf_{x\to 0} \frac{\widetilde{W}(x)}{k\Gamma_N(x)} = \liminf_{x\to 0} \frac{\widetilde{W}(x)}{u_k^{\mathbb{R}^N}(x)} \ge 1.$$

By using a similar argument as in the proof Proposition 3.1, together with the comparison principle, we deduce that $\widetilde{W} \ge w_k^{\mathbb{R}^N}$ in $\mathbb{R}^N \setminus \{0\}$ for every $k > 0$. It follows that $\widetilde{W} \ge w_\infty^{\mathbb{R}^N}$ in $\mathbb{R}^N \setminus \{0\}$ and hence $\tilde{W} = w_\infty^{\mathbb{R}^N}$ in $\mathbb{R}^N \setminus \{0\}$. Thus the whole sequence $\{W_\ell\}$ converges to $w_\infty^{\mathbb{R}^N}$ in $C^1_{\mathrm{loc}}(\mathbb{R}^N \setminus \{0\})$ as $\ell \to 0$. This leads to (1-18) and (1-19). Consequently u_∞^Ω is a strongly singular solution.

Case 3: $D < 1$. For every $\ell \in (0, 1)$, put $V_\ell = S_\ell[u_\infty^\Omega]$. Similarly, we can show that the sequence $\{V_\ell\}$ converges to $v_\infty^{\mathbb{R}^N}$ (the strongly singular solution of (1-2)) in $C^1_{\mathrm{loc}}(\mathbb{R}^N \setminus \{0\})$ as $\ell \to 0$. This yields the desired result. □

Proof of Theorem B. The theorem follows from Lemma 4.2 and Proposition 4.4. □

5. Classification and removability of isolated singularities

5.1. ***Classification of isolated singularities.*** The following lemma plays an important role in proving the classification result.

Lemma 5.1. *Assume Ω is a bounded domain containing the origin 0, $\alpha > -2$, $\beta > -1$, $1 < p < p_{c,\alpha}$, and $1 < q < q_{c,\beta}$. Let $u \in C^2(\Omega \setminus \{0\}) \cap C(\overline{\Omega} \setminus \{0\})$ be a nonnegative solution of (1-1) in $\Omega \setminus \{0\}$ vanishing on $\partial\Omega$. Then there exists $c_{31} = c_{31}(N, \alpha, \beta, p, q, d_1, d_2)$ such that for any $\delta \in \left(0, \frac{1}{4}d_1\right)$, there holds*

$$\sup\{u(x) : x \in \partial B_\delta\} \le c_{31} \inf\{u(x) : x \in \partial B_\delta\}. \tag{5-1}$$

Proof. Fix $\delta \in \left(0, \frac{1}{4}d_1\right)$ and take $x_0 \in \partial B_\delta \setminus \{0\}$. Put $r_0 = |x_0|$, $y_0 = r_0^{-1}x_0 \in \partial B_1$,

$$\varphi_{r_0} = \begin{cases} S_{r_0}[u] & \text{if } D \le 1, \\ T_{r_0}[u] & \text{if } D > 1. \end{cases}$$

It is easy to see that φ_{r_0} is a nonnegative solution of one of the following equations

$$\begin{cases} -\Delta\varphi + |x|^{\alpha}\varphi^{p} + r_0^{\frac{p(2+\beta-q)-\alpha(q-1)-q-\beta}{p-1}}|x|^{\beta}|\nabla\varphi|^{q} = 0 & \text{if } D < 1, \\ -\Delta\varphi + r_0^{\frac{\alpha(q-1)+q+\beta-p(2+\beta-q)}{q-1}}|x|^{\alpha}\varphi^{p} + |x|^{\beta}|\nabla\varphi|^{q} = 0 & \text{if } D > 1, \\ -\Delta\varphi + |x|\alpha\varphi^{p} + |x|^{\beta}|\nabla\varphi|^{q} = 0 & \text{if } D = 1. \end{cases}$$

in $\Omega_{r_0} = r_0^{-1}\Omega$. By Lemma 2.6, for every $y \in B_{1/4}(y_0)$,

$$\varphi_{r_0}(y) = r_0^{\tau}u(r_0 y) \le c_{12}|y|^{-\tau} < c_{12}2^{\tau}.$$

By Harnack's inequality (see, e.g., [Trudinger 1980; 1967]) there exists $c_{32} = c_{32}(\alpha, \beta, p, q, N, d_1, d_2)$ such that

$$\sup\{\varphi_{r_0}(y) : y \in B_{1/8}(y_0)\} \le c_{32}\inf\{\varphi_{r_0}(y) : y \in B_{1/8}(y_0)\}.$$

As ∂B_δ can be covered by a finite number (depending only on N) of balls of center on ∂B_δ and of radius $\frac{1}{4}\delta$, we obtain (5-1). □

Proof of Theorem C. The proof is based on Lemma 5.1, scaling argument and asymptotic behavior of weakly singular solutions and strongly singular solutions. Put

$$L := \limsup_{|x|\to 0} \frac{u(x)}{\Gamma_N(x)} \ge 0. \tag{5-2}$$

Case 1: $L = 0$. Then for every $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that $\delta \to 0$ as $\epsilon \to 0$ and $u \le \epsilon\Gamma_N$ on ∂B_δ. Thanks to Proposition 2.1, $u \le \epsilon\Gamma_N$ in $\Omega \setminus B_\delta$. Letting $\epsilon \to 0$ yields $u \equiv 0$.

Case 2: $L = \infty$. By (5-1),

$$\liminf_{|x|\to 0} \frac{u(x)}{\Gamma_N(x)} = \infty,$$

which along with (1-16) implies

$$\liminf_{|x|\to 0} \frac{u(x)}{u_k^{\Omega}(x)} = \infty \quad \text{for all } k > 0.$$

By the comparison principle, $u \ge u_k^{\Omega}$ in $\Omega \setminus \{0\}$ for every $k > 0$. Hence $u \ge u_\infty^{\Omega}$ in $\Omega \setminus \{0\}$. Consequently, by Theorem B, we derive

$$\liminf_{|x|\to 0}|x|^{\tau}u(x) \ge \lim_{|x|\to 0}|x|^{\tau}u_\infty^{\Omega}(x) = \Theta. \tag{5-3}$$

We next prove that[1]

$$\limsup_{|x|\to 0}|x|^{\tau}u(x) \le \Theta. \tag{5-4}$$

For any $\epsilon > 0$, it can be checked that there exists $\Theta_\epsilon > 0$ with $\Theta_\epsilon \to \Theta$ as $\epsilon \to 0$ such that $\Theta_\epsilon|x|^{-\tau-\epsilon}$ is a supersolution of (1-1) in $B_{d_1} \setminus \{0\}$ when $D = 1$ (respectively, of (1-2) in $B_{d_1} \setminus \{0\}$ when $D < 1$ and of

[1] The proof of (5-4) was proposed by a referee.

(1-3) in $B_{d_1} \setminus \{0\}$ when $D > 1$). Then by (2-9) and the comparison principle, we find that

$$u(x) \le \Theta_\epsilon |x|^{-\tau-\epsilon} + \max_{\partial B_{d_1}} u$$

in $B_{d_1} \setminus \{0\}$ for every $\epsilon > 0$. Letting $\epsilon \to 0$ for fixed $x \in B_{d_1} \setminus \{0\}$, then $|x| \to 0$, we obtain (5-4).

Case 3: $0 < L < \infty$. In light of (5-1), there is a positive number k such that

$$\liminf_{|x|\to 0} \frac{u(x)}{\Gamma_N(x)} = k > c_{34}^{-1} L, \tag{5-5}$$

here $c_{34} = c_{34}(N, \alpha, \beta, p, q, d_1, d_2) > 1$, which implies

$$\liminf_{|x|\to 0} \frac{u(x)}{u_k^\Omega(x)} = 1. \tag{5-6}$$

By Proposition 2.1, $u \ge u_k^\Omega$ in $\Omega \setminus \{0\}$. From (5-6), there exists a sequence $\{x_n\}$ converging to 0 such that

$$\lim_{n\to\infty} \frac{u(x_n)}{u_k^\Omega(x_n)} = 1.$$

Put $r_n = |x_n|$, $v_{k,n} = R_{r_n}[u_k^\Omega]$ and $v_n = R_{r_n}[u]$ in $\Omega_{r_n} = r_n^{-1}\Omega$. Then both $v_{k,n}$ and v_n are solutions of

$$-\Delta v + r_n^{N+\alpha-p(N-2)} |x|^\alpha v^p + r_n^{N+\beta-q(N-1)} |x|^\beta |\nabla v|^q = 0 \quad \text{in } \Omega_{r_n} \setminus \{0\}.$$

By the Arzelà–Ascoli theorem, regularity theory of elliptic equations and a standard diagonalization argument, up to subsequences, $\{v_{k,n}\}$ and $\{v_n\}$ converge respectively in $C^1_{\mathrm{loc}}(\mathbb{R}^N \setminus \{0\})$ to nonnegative harmonic functions V_k^* and V^* in $\mathbb{R}^N \setminus \{0\}$. Since $u \ge u_k^\Omega$, it follows that $V^* \ge V_k^*$. Put

$$\kappa_n = \sup\{\frac{u(x)}{u_k^\Omega(x)} : x \in \partial B_{r_n}\} \in [1, c_{34}]$$

and $y_n = r_n^{-1} x_n \in \partial B_1$. Therefore, up to subsequences, $\kappa_n \to \kappa \in [1, c_{34}]$ and $y_n \to y^* \in \partial B_1$. Consequently, $V^*(y^*) = V_k^*(y^*)$. By the strong maximum principle, we deduce that $V^* = V_k^*$ in $\mathbb{R}^N \setminus \{0\}$, which implies $\kappa = 1$. Thus, for every $\epsilon > 0$, there exists $n_\epsilon > 0$ such that

$$n \ge n_\epsilon \implies u_k^\Omega \le u \le (1+\epsilon) u_k^\Omega \quad \text{in } \partial B_{r_n}.$$

The comparison principle implies $u \le (1+\epsilon) u_k^\Omega$ in $\Omega \setminus B_{r_n}$. Letting $\epsilon \to 0$ yields $u \le u_k^\Omega$ in $\Omega \setminus \{0\}$. Thus $u \equiv u_k^\Omega$. □

5.2. *Removability.* We shall treat successively two cases: $q_{c,\beta} \le q < 2 + \beta$ and $q = 2 + \beta$.

Proof of Theorem D with $q_{c,\beta} \le q < 2 + \beta$. The proof is divided into three cases and strongly based upon Proposition 4.1 and self-similarity arguments.

Case 1: If $D = 1$ then $p \ge p_{c,\alpha}$ and $q \ge q_{c,\beta}$. For $0 < \delta < \frac{1}{2}d_1$ and $R > d_2 = \operatorname{diam}(\Omega)$, let $u_{\delta,R}$ be the solution of

$$\begin{cases} -\Delta u + F \circ u = 0 & \text{in } B_R \setminus \overline{B_\delta}, \\ \qquad u = c_{33}\delta^{-\tau} & \text{on } \partial B_\delta, \\ \qquad u = 0 & \text{on } \partial B_R, \end{cases} \tag{5-7}$$

where $c_{33} = \max\{c_8, c_{12}, \Theta\}$. By the comparison principle, $u \le u_{\delta,R} \le u_{\delta',R'}$ in $\Omega \setminus B_{\delta'}$ for every $0 < \delta \le \delta'$ and $0 < R \le R'$. Put $\tilde{u} := \lim_{R\to\infty} \lim_{\delta\to 0} u_{\delta,R}$; then $\tilde{u}$ is a solution of (1-1) in $\mathbb{R}^N \setminus \{0\}$ and $u \le \tilde{u}$ in $\Omega \setminus \{0\}$. By uniqueness, $T_\ell[u_{\delta,R}] = u_{\delta/\ell,R/\ell}$ for every $\ell > 0$. Letting $\delta \to 0$ and $R \to \infty$ successively implies $T_\ell[\tilde{u}] = \tilde{u}$ for every $\ell > 0$. Hence $\tilde{u}$ is a self-similar solution of (1-1) in $\mathbb{R}^N \setminus \{0\}$ and can be represented in the form

$$\tilde{u}(x) = |x|^{-\frac{2+\beta-q}{q-1}}\omega(x/|x|) \quad \text{for all } x \in \mathbb{R}^N \setminus \{0\},$$

where ω is a solution of (4-1). Since $q_{c,\beta} \le q < 2+\beta$, from Proposition 4.1 we deduce that $\omega \equiv 0$. It follows that $\tilde{u} \equiv 0$ and thus $u \equiv 0$.

Case 2: If $D > 1$ then we must have $q \ge q_{c,\beta}$. For any $0 < \delta < R$, let $w_{\delta,R}$ be the solution of

$$\begin{cases} -\Delta w + |x|^\beta |\nabla w|^q = 0 & \text{in } B_R \setminus \overline{B_\delta}, \\ \qquad w = c_{33}\delta^{-\frac{2+\beta-q}{q-1}} & \text{on } \partial B_\delta, \\ \qquad w = & \text{on } \partial B_R. \end{cases} \tag{5-8}$$

By the comparison principle, $u \le w_{\delta,R} \le w_{\delta',R'}$ in $\Omega \setminus B_{\delta'}$ for every $0 < \delta \le \delta'$ and $0 < R \le R'$. Put $\tilde{w} := \lim_{R\to\infty} \lim_{\delta\to 0} w_{\delta,R}$ then $\tilde{w}$ is a solution of (1-3) in $\mathbb{R}^N \setminus \{0\}$ and $u \le \tilde{w}$ in $\Omega \setminus \{0\}$. By uniqueness, $T_\ell[w_{\delta,R}] = w_{\delta/\ell,R/\ell}$ for every $\ell > 0$. Letting $\delta \to 0$ and $R \to \infty$ successively implies $T_\ell[\tilde{w}] = \tilde{w}$ for every $\ell > 0$. Hence $\tilde{w}$ is a self-similar solution of (1-3) in $\mathbb{R}^N \setminus \{0\}$ and can be represented in the form

$$\tilde{w}(x) = |x|^{-\frac{2+\beta-q}{q-1}}\omega(x/|x|) \quad \text{for all } x \in \mathbb{R}^N \setminus \{0\},$$

where ω is a solution of (4-1) with $\lambda = 0$. Since $q_{c,\beta} \le q < 2+\beta$, from Proposition 4.1 we deduce that $\omega \equiv 0$. It follows that $\tilde{w} \equiv 0$ and thus $u \equiv 0$.

Case 3: If $D < 1$ then we must have $p \ge p_{c,\alpha}$. One can use an argument similar to the proof in Case 2 to obtain $u \equiv 0$. □

Remark. Theorem D with $q < 2+\beta$ can be obtained by a different way which is suggested by the referee. The proof, that we present below, is more direct, independent of Proposition 4.1 and does not require any self-similarity arguments.

Assume that either $p \ge p_{c,\alpha}$ or $q \ge q_{c,\beta}$. We distinguish two cases:

Case 1: If $D \ge 1$ then we must have $q \ge q_{c,\beta}$.

Case 2: If $D < 1$ then we must have $p \ge p_{c,\alpha}$.

If $q > q_{c,\beta}$ in Case 1 or $p > p_{c,\alpha}$ in Case 2, then by (1-13) and (2-9), we deduce that

$$\lim_{|x|\to 0} \frac{u(x)}{\Gamma_N(x)} = 0.$$

Since $u = 0$ on $\partial\Omega$, the comparison principle gives that $u \equiv 0$ in $\Omega \setminus \{0\}$.

If $q = q_{c,\beta}$ in Case 1 or $p = p_{c,\alpha}$ in Case 2 then by (1-13) and (2-9), we deduce that

$$\lim_{|x|\to 0} \frac{u(x)}{\Gamma_N(x)} < \infty.$$

For every $\epsilon > 0$ small, it can be easily checked that there exists $C_\epsilon > 0$ with $C_\epsilon \to 0$ as $\epsilon \to 0$ such that $S_\epsilon(x) := C_\epsilon |x|^{2-N-\epsilon}$ is a supersolution of (1-3) in $B_1 \setminus \{0\}$ when $q = q_{c,\beta}$ in Case 1 (respectively, a supersolution of (1-2) in $B_1 \setminus \{0\}$ when $p = p_{c,\alpha}$ in Case 2). Since

$$\lim_{|x|\to 0} \frac{u(x)}{S_\epsilon(x)} = 0,$$

by the comparison principle, $u(x) \le S_\epsilon(x) + \max_{\partial B_{d_1}} u$ in $B_{d_1} \setminus \{0\}$. Letting $\epsilon \to 0$, we get $u \le \max_{\partial B_{d_1}} u$. Since $u = 0$ on $\partial\Omega \setminus \{0\}$, we find that $u \equiv 0$ in $\Omega \setminus \{0\}$.

In order to prove Theorem D in the case $q = 2 + \beta$ we need the following lemma.

Lemma 5.2. *Let $\beta > -1$. If $w \in C^2(\Omega \setminus \{0\}) \cap C(\bar{\Omega} \setminus \{0\})$ is a nonnegative solution of*

$$-\Delta w + |x|^\beta |\nabla w|^{2+\beta} = 0 \quad \text{in } \Omega \setminus \{0\}, \tag{5-9}$$

which vanishes on $\partial\Omega$ then $w \equiv 0$.

Proof. By (2-3), there exists a positive constant $c_{35} = c_{35}(N, q, \beta, d_1, d_2, \|w\|_{L^\infty(\partial B_{d_1})})$ such that $w(x) \le c_{35} - c_3 \ln|x|$ in $B_{d_1} \setminus \{0\}$. The constant c_{35} can be chosen such that $\Phi(x) := c_{35} - c_3 \ln|x|$ is a positive superharmonic function in $\Omega \setminus \{0\}$.

For $\epsilon \in (0, d_1)$, let h_ϵ be the harmonic function in $\Omega \setminus B_\epsilon$ such that $h_\epsilon = w$ on ∂B_ϵ and $h_\epsilon = 0$ on $\partial\Omega$. By the comparison principle, $w \le h_\epsilon$ in $\Omega \setminus B_\epsilon$ for every $\epsilon \in (0, d_1)$. Consequently, $h_\epsilon \le h_{\epsilon'}$ for $0 < \epsilon' < \epsilon$. On the other hand, since Φ is a positive superharmonic function in $\Omega \setminus B_\epsilon$ which dominates h_ϵ on $\partial\Omega \cup \partial B_\epsilon$, by the comparison principle, $h_\epsilon \le \Phi$ in $\Omega \setminus B_\epsilon$. Therefore, $\{h_\epsilon\}$ converges, as $\epsilon \to 0$, to a harmonic function $\hat{h}$ in $\Omega \setminus \{0\}$ which vanishes on $\partial\Omega$ and satisfies $w \le \hat{h} \le \Phi$ in $\Omega \setminus \{0\}$. Since $N > 2$, we deduce that $\hat{h}(x) = o(\Gamma_N(x))$ as $|x| \to 0$. Therefore $\hat{h} \equiv 0$. Thus $w \equiv 0$. □

Proof of Theorem D with $q = 2 + \beta$.

For $\epsilon \in (0, d_1)$, let w_ϵ be the solution of (2-10) with $q = 2 + \beta$. The sequence $\{w_\epsilon\}$ converges, as $\epsilon \to 0$, to a solution $\hat{w}$ of (5-9) in $\Omega \setminus \{0\}$ which vanishes on $\partial\Omega$. Since $u \le w_\epsilon$ for every $\epsilon \in (0, d_1)$, it follows that $u \le \hat{w}$. By Lemma 5.2, $\hat{w} \equiv 0$ and thus $u \equiv 0$. □

Acknowledgements

This research was supported by Fondecyt Grant 3160207. The author would like to thank the anonymous referees for a careful reading of the manuscript and helpful comments.

References

[Baras and Pierre 1984] P. Baras and M. Pierre, "Singularités éliminables pour des équations semi-linéaires", *Ann. Inst. Fourier (Grenoble)* **34**:1 (1984), 185–206. MR 743627

[Bidaut-Véron et al. 2014] M.-F. Bidaut-Véron, M. Garcia-Huidobro, and L. Véron, "Local and global properties of solutions of quasilinear Hamilton–Jacobi equations", *J. Funct. Anal.* **267**:9 (2014), 3294–3331. MR 3261111

[Brézis and Véron 1980/81] H. Brézis and L. Véron, "Removable singularities for some nonlinear elliptic equations", *Arch. Rational Mech. Anal.* **75**:1 (1980/81), 1–6. MR 592099

[Cîrstea 2014] F. C. Cîrstea, *A complete classification of the isolated singularities for nonlinear elliptic equations with inverse square potentials*, Mem. Amer. Math. Soc. **227**, 2014. MR 3135311

[Cîrstea and Du 2010] F. C. Cîrstea and Y. Du, "Isolated singularities for weighted quasilinear elliptic equations", *J. Funct. Anal.* **259**:1 (2010), 174–202. MR 2610383

[DiBenedetto 1983] E. DiBenedetto, "$C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations", *Nonlinear Anal.* **7**:8 (1983), 827–850. MR 709038

[Friedman and Véron 1986] A. Friedman and L. Véron, "Singular solutions of some quasilinear elliptic equations", *Arch. Rational Mech. Anal.* **96**:4 (1986), 359–387. MR 855755

[Gilbarg and Trudinger 2001] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Springer, Berlin, 2001. MR 1814364

[Lasry and Lions 1989] J.-M. Lasry and P.-L. Lions, "Nonlinear elliptic equations with singular boundary conditions and stochastic control with state constraints, I: The model problem", *Math. Ann.* **283**:4 (1989), 583–630. MR 990591

[Lieb and Loss 1997] E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics **14**, American Mathematical Societyi, Providence, RI, 1997. MR 1415616

[Lieberman 1988] G. M. Lieberman, "Boundary regularity for solutions of degenerate elliptic equations", *Nonlinear Anal.* **12**:11 (1988), 1203–1219. MR 969499

[Lions 1985] P.-L. Lions, "Quelques remarques sur les problèmes elliptiques quasilinéaires du second ordre", *J. Analyse Math.* **45** (1985), 234–254. MR 833413

[Marcus 2013] M. Marcus, "Remarks on nonlinear equations with measures", *Commun. Pure Appl. Anal.* **12**:4 (2013), 1745–1753. MR 2997539

[Marcus and Nguyen 2015] M. Marcus and P.-T. Nguyen, "Elliptic equations with nonlinear absorption depending on the solution and its gradient", *Proc. Lond. Math. Soc.* (*3*) **111**:1 (2015), 205–239. MR 3404781

[Marcus and Véron 2014] M. Marcus and L. Véron, *Nonlinear second order elliptic equations involving measures*, De Gruyter Series in Nonlinear Analysis and Applications **21**, De Gruyter, Berlin, 2014. MR 3156649

[Nguyen Phuoc and Véron 2012] T. Nguyen Phuoc and L. Véron, "Boundary singularities of solutions to elliptic viscous Hamilton–Jacobi equations", *J. Funct. Anal.* **263**:6 (2012), 1487–1538. MR 2948222

[Serrin 1965] J. Serrin, "Isolated singularities of solutions of quasilinear equations", *Acta Math.* **113** (1965), 219–240. MR 0176219

[Trudinger 1967] N. S. Trudinger, "On Harnack type inequalities and their application to quasilinear elliptic equations", *Comm. Pure Appl. Math.* **20** (1967), 721–747. MR 0226198

[Trudinger 1980] N. S. Trudinger, "Local estimates for subsolutions and supersolutions of general second order elliptic quasilinear equations", *Invent. Math.* **61**:1 (1980), 67–79. MR 587334

[Vázquez and Véron 1985] J. L. Vázquez and L. Véron, "Isolated singularities of some semilinear elliptic equations", *J. Differential Equations* **60**:3 (1985), 301–321. MR 811769

[Véron 1981] L. Véron, "Singular solutions of some nonlinear elliptic equations", *Nonlinear Anal.* **5**:3 (1981), 225–242. MR 607806

[Véron 1996] L. Véron, *Singularities of solutions of second order quasilinear equations*, Pitman Research Notes in Mathematics Series **353**, Longman, Harlow, 1996. MR 1424468

Received 20 Oct 2015. Revised 21 Apr 2016. Accepted 6 Jun 2016.

PHUOC-TAI NGUYEN: nguyenphuoctai.hcmup@gmail.com
Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860, 6904441 Santiago, Chile

ANALYSIS AND PDE
Vol. 9, No. 7, 2016
dx.doi.org/10.2140/apde.2016.9.1693

A SECOND ORDER ESTIMATE FOR GENERAL COMPLEX HESSIAN EQUATIONS

DUONG H. PHONG, SEBASTIEN PICARD AND XIANGWEN ZHANG

We consider the general complex Hessian equations with right-hand sides depending on gradients, which are motivated by the Fu–Yau equations arising from the study of Strominger systems. The second order estimate for the solution is crucial to solving the equation by the method of continuity. We obtain such an estimate for the χ-plurisubharmonic solutions.

1. Introduction

Let (X, ω) be a compact Kähler manifold of dimension $n \geq 2$. Let $u \in C^\infty(X)$ and consider a $(1, 1)$-form $\chi(z, u)$ possibly depending on u and satisfying the positivity condition $\chi \geq \varepsilon\omega$ for some $\varepsilon > 0$. We define

$$g = \chi(z, u) + i\partial\bar{\partial}u, \tag{1-1}$$

and u is called χ-plurisubharmonic if $g > 0$ as a $(1, 1)$ form. In this paper, we are concerned with the complex Hessian equation

$$(\chi(z, u) + i\partial\bar{\partial}u)^k \wedge \omega^{n-k} = \psi(z, Du, u)\,\omega^n \tag{1-2}$$

for $1 \leq k \leq n$, where $\psi(z, v, u) \in C^\infty(T^{1,0}(X)^* \times \mathbb{R})$ is a given strictly positive function.

The complex Hessian equation can be viewed as an intermediate equation between the Laplace equation and the complex Monge–Ampère equation. It encompasses the most natural invariants of the complex Hessian matrix of a real valued function, namely the elementary symmetric polynomials of its eigenvalues. When $k = 1$, the equation (1-2) is quasilinear, and the estimates follow from the classical theory of quasilinear PDEs. The real counterparts of (1-2) for $1 < k \leq n$, with ψ not depending on the gradient of u, have been studied extensively in the literature (see the survey paper [Wang 2009] and more recent related work [Guan 2014]), as these equations appear naturally and play very important roles in both classical and conformal geometry. When the right-hand side ψ depends on the gradient of the solution, even the real case has been a long-standing problem due to substantial difficulties in obtaining a priori C^2 estimates. This problem has recently been solved by Guan, Ren and Wang [Guan et al. 2015] for convex solutions of real Hessian equations.

In the complex case, equation (1-2) with $\psi = \psi(z, u)$ has been extensively studied in recent years, due to its appearance in many geometric problems, including the J-flow [Song and Weinkove 2008] and

Work supported in part by the National Science Foundation under Grant DMS-1266033, DMS-1605968 and DMS-1308136.
MSC2010: 35J15, 35J60, 58J05, 35J96.
Keywords: complex Hessian equation, second order estimate.

quaternionic geometry [Alesker and Verbitsky 2010]. The related Dirichlet problem for (1-2) on domains in $\mathbb{C}^n$ has been studied by Li [2004] and Błocki [2005]. The corresponding problem on compact Kähler or Hermitian manifolds has also been studied extensively; see, for example, [Dinew and Kołodziej 2014; Hou 2009; Kołodziej and Nguyen 2016; Lu and Nguyen 2015; Zhang 2010]. In particular, as a crucial step in the continuity method, C^2 estimates for complex Hessian type equations have been studied in various settings; see [Hou et al. 2010; Sun 2014; Székelyhidi 2015; Székelyhidi et al. 2015; Zhang 2015].

However, (1-2) with $\psi = \psi(z, Du, u)$ has been much less studied. An important case corresponding to $k = n = 2$, so that it is actually a Monge–Ampère equation in two dimensions, is central to the solution by Fu and Yau [2008; 2007] of a Strominger system on a toric fibration over a K3 surface. A natural generalization of this case to general dimension n was suggested by Fu and Yau [2008] and can be expressed as

$$((e^u + f e^{-u})\omega + n i \partial\bar\partial u)^2 \wedge \omega^{n-2} = \psi(z, Du, u)\, \omega^n, \tag{1-3}$$

where $\psi(z, v, u)$ is a function on $T^{1,0}(X)^* \times \mathbb{R}$ with a particular structure, and (X, ω) is a compact Kähler manifold. A priori estimates for this equation were obtained by the authors in [Phong et al. 2015].

In this paper, motivated by our previous work [Phong et al. 2015], we study a priori C^2 estimates for the equation (1-2) with general $\chi(z, u)$ and general right-hand side $\psi(z, Du, u)$. Building on the techniques developed in [Guan et al. 2015] (see also [Li et al. 2016] for real Hessian equations), we can prove the following theorem.

Theorem 1. *Let (X, ω) be a compact Kähler manifold of complex dimension n. Suppose $u \in C^4(X)$ is a solution of (1-2) with $g = \chi + i\partial\bar\partial u > 0$ and $\chi(z, u) \geq \varepsilon\omega$. Let $0 < \psi(z, v, u) \in C^\infty(T^{1,0}(X)^* \times \mathbb{R})$. Then we have the uniform second order derivative estimate*

$$|D\bar{D}u|_\omega \leq C, \tag{1-4}$$

where C is a positive constant depending only on $\varepsilon, n, k, \sup_X |u|, \sup_X |Du|$, and the C^2 norm of χ as a function of (u, z), the infimum of ψ, and the C^2 norm of ψ as a function of (z, Du, u), all restricted to the ranges in Du and u defined by the uniform upper bounds on $|u|$ and $|Du|$.

We remark that the above estimate is stated for χ-plurisubharmonic solutions, that is, $g = \chi + i\partial\bar\partial u > 0$. Actually, we only need to assume that g is in the Γ_{k+1} cone (see (3-11) below for the definition of the Garding cone Γ_k and also the discussion in Remark 2 at the end of the paper). However, a better condition would be $g \in \Gamma_k$, which is the natural cone for ellipticity. In fact, this is still an open problem even for real Hessian equations when $2 < k < n$. If $k = 2$, [Guan et al. 2015] removed the convexity assumption by investigating the structure of the operator. A simpler argument was given recently by Spruck and Xiao [2015]. However, the arguments are not applicable to the complex case due to the difference between the terms $|DDu|^2$ and $|D\bar{D}u|^2$ in the complex setting. When $k = 2$ in the complex setting, C^2 estimates for (1-3) were obtained in [Phong et al. 2015] without the plurisubharmonicity assumption, but the techniques rely on the specific right-hand side $\psi(z, Du, u)$ studied there.

We also note that if $k = n$, the condition $g = \chi + i\partial\bar\partial u > 0$ is the natural assumption for the ellipticity of equation (1-2). Thus, our result implies the a priori C^2 estimate for complex Monge–Ampère equations with right-hand side depending on gradients:

$$(\chi(z,u)+i\partial\bar{\partial}u)^n=\psi(z,Du,u)\,\omega^n.$$

This generalizes the C^2 estimate for the equation studied by Fu and Yau [2008; 2007] mentioned above, which corresponds to $n=2$ and a specific form $\chi(z,u)$ as well as a specific right-hand side $\psi(z,Du,u)$. For dimension $n\geq 2$ and $k=n$, the estimate was obtained by Guan and Ma, in unpublished notes, using a different method where the structure of the Monge–Ampère operator plays an important role.

Compared to the estimates when $\psi=\psi(z,u)$, the dependence on the gradient of u in (1-2) creates substantial new difficulties. The main obstacle is the appearance of terms such as $|DDu|^2$ and $|D\bar{D}u|^2$ when one differentiates the equation twice. We adapt the techniques used in [Guan et al. 2015] and [Li et al. 2016] for real Hessian equations to overcome these difficulties. Furthermore, we also need to handle properly some subtle issues when dealing with the third-order terms due to complex conjugacy.

2. Preliminaries

Let σ_k be the k-th elementary symmetric function; that is, for $1\leq k\leq n$ and $\lambda=(\lambda_1,\dots,\lambda_n)\in\mathbb{R}^n$,

$$\sigma_k(\lambda)\;=\sum_{1<i_1<\cdots<i_k<n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_n}.$$

Let $\lambda(a_{\bar{j}i})$ denote the eigenvalues of a Hermitian symmetric matrix $(a_{\bar{j}i})$ with respect to the background Kähler metric ω. We define $\sigma_k(a_{\bar{j}i})=\sigma_k(\lambda(a_{\bar{j}i}))$. This definition can be naturally extended to complex manifolds. Denoting by $A^{1,1}(X)$ the space of smooth real $(1,1)$-forms on a compact Kähler manifold (X,ω), we define, for any $g\in A^{1,1}(X)$,

$$\sigma_k(g)=\binom{n}{k}\frac{g^k\wedge\omega^{n-k}}{\omega^n}.$$

Using the above notation, we can rewrite (1-2) as follows:

$$\sigma_k(g)=\sigma_k(\chi_{\bar{j}i}+u_{\bar{j}i})=\psi(z,Du,u). \tag{2-1}$$

We use the notation

$$\sigma_k^{p\bar{q}}=\frac{\partial\sigma_k(g)}{\partial g_{\bar{q}p}},\qquad \sigma_k^{p\bar{q},r\bar{s}}=\frac{\partial^2\sigma_k(g)}{\partial g_{\bar{q}p}\bar{\partial}g_{\bar{s}r}}.$$

The symbol D indicates the covariant derivative with respect to the given metric ω. All norms and inner products are with respect to ω unless denoted otherwise. We denote by $\lambda_1,\lambda_2,\dots,\lambda_n$ the eigenvalues of $g_{\bar{j}i}=\chi_{\bar{j}i}+u_{\bar{j}i}$ with respect to ω, and use the ordering $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n>0$. Our calculations are carried out at a point z on the manifold X, and we use coordinates such that at this point $\omega=i\sum\delta_{\ell k}\,dz^k\wedge d\bar{z}^\ell$ and $g_{\bar{j}i}$ is diagonal. We also use the notation

$$\mathcal{F}=\sum_p\sigma_k^{p\bar{p}}.$$

Differentiating (2-1) yields

$$\sigma_k^{p\bar{q}}D_{\bar{j}}g_{\bar{q}p}=D_{\bar{j}}\psi. \tag{2-2}$$

Differentiating the equation a second time gives

$$\begin{aligned}\sigma_k^{p\bar q} D_i D_{\bar j} g_{\bar q p} + \sigma_k^{p\bar q, r\bar s} D_i g_{\bar q p} D_{\bar j} g_{\bar s r} &= D_i D_{\bar j} \psi \\ &\geq -C(1+|DDu|^2+|D\bar D u|^2) + \sum_\ell \psi_{v_\ell} u_{\ell \bar j i} + \sum_\ell \psi_{\bar v_\ell} u_{\bar\ell \bar j i}. \end{aligned} \tag{2-3}$$

We denote by C a uniform constant which depends only on (X,ω), n, k, $\|\chi\|_{C^2}$, $\inf\psi$, $\|u\|_{C^1}$ and $\|\psi\|_{C^2}$. We now compute the operator $\sigma_k^{p\bar q} D_p D_{\bar q}$ acting on $g_{\bar j i} = \chi_{\bar j i} + u_{\bar j i}$. Recalling that $\chi_{\bar j i}$ depends on u, we estimate

$$\begin{aligned}\sigma_k^{p\bar q} D_p D_{\bar q} g_{\bar j i} &= \sigma_k^{p\bar q} D_p D_{\bar q} D_i D_{\bar j} u + \sigma_k^{p\bar q} D_p D_{\bar q} \chi_{\bar j i} \\ &\geq \sigma_k^{p\bar q} D_p D_{\bar q} D_i D_{\bar j} u - C(1+\lambda_1)\mathcal{F}. \end{aligned} \tag{2-4}$$

Commuting derivatives

$$\begin{aligned} D_p D_{\bar q} D_i D_{\bar j} u &= D_i D_{\bar j} D_p D_{\bar q} u - R_{\bar q i \bar j}{}^{\bar a} u_{\bar a p} + R_{\bar q p \bar j}{}^{\bar a} u_{\bar a i} \\ &= D_i D_{\bar j} g_{\bar q p} - D_i D_{\bar j} \chi_{\bar q p} - R_{\bar q i \bar j}{}^{\bar a} u_{\bar a p} + R_{\bar q p \bar j}{}^{\bar a} u_{\bar a i}. \end{aligned} \tag{2-5}$$

Therefore, by (2-3),

$$\begin{aligned}\sigma_k^{p\bar q} D_p D_{\bar q} g_{\bar j i} \geq -\sigma_k^{p\bar q, r\bar s} D_j g_{\bar q p} D_{\bar j} g_{\bar s r} + \sum \psi_{v_\ell} g_{\bar j i \ell} + \sum \psi_{\bar v_\ell} g_{\bar j i \bar\ell} \\ -C(1+|DDu|^2+|D\bar D u|^2 + (1+\lambda_1)\mathcal{F}). \end{aligned} \tag{2-6}$$

We next compute the operator $\sigma_k^{p\bar q} D_p D_{\bar q}$ acting on $|Du|^2$, introducing the notation

$$|DDu|^2_{\sigma\omega} = \sigma_k^{p\bar q} \omega^{m\bar\ell} D_p D_m u D_{\bar q} D_{\bar \ell} u, \quad |D\bar D u|^2_{\sigma\omega} = \sigma_k^{p\bar q} \omega^{m\bar\ell} D_p D_{\bar\ell} u D_m D_{\bar q} u. \tag{2-7}$$

Then

$$\begin{aligned}\sigma_k^{p\bar q} |Du|^2_{\bar q p} &= \sigma_k^{p\bar q}(D_p D_{\bar q} D_m u D^m u + D_m u D_p D_{\bar q} D^m u) + |DDu|^2_{\sigma\omega} + |D\bar D u|^2_{\sigma\omega} \\ &= \sigma_k^{p\bar q}\big(D_m(g_{\bar q p} - \chi_{\bar q p}) D^m u + D_m u D^m (g_{\bar q p} - \chi_{\bar q p})\big) + \sigma_k^{p\bar q} R_{\bar q p}{}^{m\bar\ell} u_{\bar\ell} u_m \\ &\qquad + |DDu|^2_{\sigma\omega} + |D\bar D u|^2_{\sigma\omega}. \end{aligned} \tag{2-8}$$

Using the differentiated equation, we obtain

$$\begin{aligned}\sigma_k^{p\bar q}|Du|^2_{\bar q p} &\geq 2\,\mathrm{Re}\langle Du, D\psi\rangle - C(1+\mathcal{F}) + |DDu|^2_{\sigma\omega} + |D\bar D u|^2_{\sigma\omega} \\ &\geq 2\,\mathrm{Re}\bigg(\sum_{p,m}(D_p D_m u D_{\bar p} u + D_p u D_{\bar p} D_m u)\psi_{v_m}\bigg) - C(1+\mathcal{F}) + |DDu|^2_{\sigma\omega} + |D\bar D u|^2_{\sigma\omega}. \end{aligned}$$

To simplify the expression, we introduce the notation

$$\langle D|Du|^2, D_{\bar v}\psi\rangle = \sum_m (D_m D_p u D^p u\, \psi_{v_m} + D_p u D_m D^p u\, \psi_{v_m}). \tag{2-9}$$

We obtain

$$\sigma_k^{p\bar q}|Du|^2_{\bar q p} \geq 2\,\mathrm{Re}\langle D|Du|^2, D_{\bar v}\psi\rangle - C(1+\mathcal{F}) + |DDu|^2_{\sigma\omega} + |D\bar D u|^2_{\sigma\omega}. \tag{2-10}$$

We also compute

$$-\sigma_k^{p\bar q} u_{\bar q p} = \sigma_k^{p\bar q}(\chi_{\bar q p} - g_{\bar q p}) \geq \varepsilon\mathcal{F} - k\psi. \tag{2-11}$$

3. The C^2 estimate

In this section, we give the proof of the estimate stated in the theorem. When $k=1$, (1-2) becomes

$$\Delta_\omega u + \mathrm{Tr}_\omega \chi(z,u) = n\psi(z, Du, u), \tag{3-1}$$

where Δ_ω and Tr_ω are the Laplacian and trace with respect to the background metric ω. It follows that $\Delta_\omega u$ is bounded, and the desired estimate follows in turn from the positivity of the metric g. Henceforth, we assume that $k \geq 2$. Motivated by the idea from [Guan et al. 2015] for real Hessian equations, we apply the maximum principle to the test function

$$G = \log P_m + mN|Du|^2 - mMu, \tag{3-2}$$

where $P_m = \sum_j \lambda_j^m$. Here, m, M and N are large positive constants to be determined later. We may assume that the maximum of G is achieved at some point $z \in X$. After rotating the coordinates, we may assume that the matrix $g_{\bar{j}i} = \chi_{\bar{j}i} + u_{\bar{j}i}$ is diagonal.

Recall that if $F(A) = f(\lambda_1, \dots, \lambda_n)$ is a symmetric function of the eigenvalues of a Hermitian matrix $A = (a_{\bar{j}i})$, then at a diagonal matrix A with distinct eigenvalues, we have

$$F^{i\bar{j}} = \delta_{ij} f_i, \tag{3-3}$$

$$F^{i\bar{j},r\bar{s}} w_{i\bar{j}k} w_{r\bar{s}\bar{k}} = \sum f_{ij} w_{i\bar{i}k} w_{j\bar{j}\bar{k}} + \sum_{p \neq q} \frac{f_p - f_q}{\lambda_p - \lambda_q} |w_{p\bar{q}k}|^2, \tag{3-4}$$

where

$$F^{i\bar{j}} = \frac{\partial F}{\partial a_{\bar{j}i}}, \qquad F^{i\bar{j},r\bar{s}} = \frac{\partial^2 F}{\partial a_{\bar{j}i} \partial a_{\bar{s}r}},$$

and $w_{i\bar{j}k}$ is an arbitrary tensor; see [Ball 1984]. Using these identities to differentiate G, we first obtain the critical equation

$$\frac{DP_m}{P_m} + mND|Du|^2 - mMDu = 0. \tag{3-5}$$

Differentiating G a second time and contracting with $\sigma_k^{p\bar{q}}$ yields

$$\begin{aligned} 0 \geq \frac{m}{P_m}\biggl(\sum_j \lambda_j^{m-1} \sigma_k^{p\bar{p}} D_p D_{\bar{p}} g_{\bar{j}j}\biggr) - \frac{|DP_m|_\sigma^2}{P_m^2} + mN\sigma_k^{p\bar{p}} |Du|^2_{\bar{p}p} - mM\sigma_k^{p\bar{p}} u_{\bar{p}p} \\ + \frac{m}{P_m}\biggl((m-1)\sum_j \lambda_j^{m-2} \sigma_k^{p\bar{p}} |D_p g_{\bar{j}j}|^2 + \sigma_k^{p\bar{p}} \sum_{i\neq j} \frac{\lambda_i^{m-1} - \lambda_j^{m-1}}{\lambda_i - \lambda_j} |D_p g_{\bar{j}i}|^2\biggr). \end{aligned} \tag{3-6}$$

Here, we use the notation $|\eta|_\sigma^2 = \sigma_k^{p\bar{q}} \eta_p \eta_{\bar{q}}$. Substituting (2-6), (2-10) and (2-11), we obtain

$$
\begin{aligned}
0 \geq \frac{1}{P_m}\Big(-C\sum_j \lambda_j^{m-1}(1+|DDu|^2+|D\bar{D}u|^2+(1+\lambda_1)\mathcal{F})\Big)& \\
+\frac{1}{P_m}\Big(\sum_j \lambda_j^{m-1}(-\sigma_k^{p\bar{q},r\bar{s}} D_j g_{\bar{q}p} D_{\bar{j}} g_{\bar{s}r} + \sum_\ell \psi_{v_\ell} g_{\bar{j}j\ell} + \sum_\ell \psi_{\bar{v}_\ell} g_{\bar{j}j\bar{\ell}})\Big)& \\
+\frac{1}{P_m}\Big((m-1)\sum_j \lambda_j^{m-2}\sigma_k^{p\bar{p}}|D_p g_{\bar{j}j}|^2 + \sigma_k^{p\bar{p}}\sum_{i\neq j}\frac{\lambda_i^{m-1}-\lambda_j^{m-1}}{\lambda_i-\lambda_j}|D_p g_{\bar{j}i}|^2\Big)& \\
-\frac{|DP_m|^2_\sigma}{mP_m^2} + N(|DDu|^2_{\sigma\omega}+|D\bar{D}u|^2_{\sigma\omega})& \\
+N\langle D|Du|^2, D_{\bar{v}}\psi\rangle + N\langle D_{\bar{v}}\psi, D|Du|^2\rangle + (M\varepsilon - CN)\mathcal{F} - kM\psi.& \qquad (3\text{-}7)
\end{aligned}
$$

From the critical equation (3-5), we have

$$
\frac{1}{P_m}\sum_{j,\ell}\lambda_j^{m-1} g_{\bar{j}j\ell}\psi_{v_\ell} = \frac{1}{m}\Big\langle \frac{DP_m}{P_m}, D_{\bar{v}}\psi\Big\rangle = -N\langle D|Du|^2, D_{\bar{v}}\psi\rangle + M\langle Du, D_{\bar{v}}\psi\rangle.
$$

It follows that

$$
\begin{aligned}
\frac{1}{P_m}\sum_{j,\ell}(\psi_{v_\ell} g_{\bar{j}j\ell} + \psi_{\bar{v}_\ell} g_{\bar{j}j\bar{\ell}}) + N\langle D|Du|^2, D_{\bar{v}}\psi\rangle + N\langle D_{\bar{v}}\psi, D|Du|^2\rangle& \\
= M(\langle Du, D_{\bar{v}}\psi\rangle + \langle D_{\bar{v}}\psi, Du\rangle) \geq -CM.&
\end{aligned}
$$

Using (3-4), one can obtain the well-known identity

$$
-\sigma_k^{p\bar{q},r\bar{s}} D_j g_{\bar{q}p} D_{\bar{j}} g_{\bar{s}r} = -\sigma_k^{p\bar{p},q\bar{q}} D_j g_{\bar{p}p} D_{\bar{j}} g_{\bar{q}q} + \sigma_k^{p\bar{p},q\bar{q}}|D_j g_{\bar{p}q}|^2, \qquad (3\text{-}8)
$$

where

$$
\sigma_k^{p\bar{p},q\bar{q}} = \frac{\partial}{\partial\lambda_p}\frac{\partial}{\partial\lambda_q}\sigma_k(\lambda).
$$

We assume that $\lambda_1 \gg 1$, otherwise the C^2 estimate is complete. The main inequality (3-7) becomes

$$
\begin{aligned}
0 \geq \frac{-C}{\lambda_1}(1+|DDu|^2+|D\bar{D}u|^2) + \frac{1}{P_m}\Big(\sum_j \lambda_j^{m-1}\big(-\sigma_k^{p\bar{p},q\bar{q}} D_j g_{\bar{p}p} D_{\bar{j}} g_{\bar{q}q} + \sigma_k^{p\bar{p},q\bar{q}}|D_j g_{\bar{p}q}|^2\big)\Big)& \\
+\frac{1}{P_m}\Big((m-1)\sum_j \lambda_j^{m-2}\sigma_k^{p\bar{p}}|D_p g_{\bar{j}j}|^2 + \sigma_k^{p\bar{p}}\sum_{i\neq j}\frac{\lambda_i^{m-1}-\lambda_j^{m-1}}{\lambda_i-\lambda_j}|D_p g_{\bar{j}i}|^2\Big)& \\
-\frac{|DP_m|^2_\sigma}{mP_m^2} + N(|DDu|^2_{\sigma\omega}+|D\bar{D}u|^2_{\sigma\omega}) + (M\varepsilon - CN - C)\mathcal{F} - CM.& \qquad (3\text{-}9)
\end{aligned}
$$

The main objective is to show that the third-order terms on the right-hand side of (3-9) are nonnegative. To deal with this issue, we need a lemma from [Guan et al. 2015]; see also [Guan et al. 2012; Li et al. 2016].

Lemma 1 [Guan et al. 2015]. *Suppose* $1 \le \ell < k \le n$, *and let* $\alpha = 1/(k-\ell)$. *Let* $W = (w_{\bar{q}p})$ *be a Hermitian tensor in the* Γ_k *cone. Then for any* $\theta > 0$,

$$-\sigma_k^{p\bar{p},q\bar{q}}(W) w_{\bar{p}pi} w_{\bar{q}q\bar{i}} + \Big(1-\alpha+\frac{\alpha}{\theta}\Big)\frac{|D_i\sigma_k(W)|^2}{\sigma_k(W)} \ge \sigma_k(W)(\alpha+1-\alpha\theta)\left|\frac{D_i\sigma_\ell(W)}{\sigma_\ell(W)}\right|^2 - \frac{\sigma_k}{\sigma_\ell}(W)\sigma_\ell^{p\bar{p},q\bar{q}}(W) w_{\bar{p}pi} w_{\bar{q}q\bar{i}}. \tag{3-10}$$

Here the Γ_k cone is defined as

$$\Gamma_k = \{\lambda \in \mathbb{R}^n \mid \sigma_m(\lambda) > 0,\ m = 1, \dots, k\}. \tag{3-11}$$

We say a Hermitian matrix $W \in \Gamma_k$ if $\lambda(W) \in \Gamma_k$.

It follows from the above lemma that, by taking $\ell = 1$, we have

$$-\sigma_k^{p\bar{p},q\bar{q}} D_i g_{\bar{p}p} D_{\bar{i}} g_{\bar{q}q} + K|D_i\sigma_k|^2 \ge 0, \tag{3-12}$$

for $K > (1-\alpha+\alpha/\theta)(\inf\psi)^{-1}$ if $2 \le k \le n$.

We denote

$$A_i = \frac{\lambda_i^{m-1}}{P_m}\big(K|D_i\sigma_k|^2 - \sigma_k^{p\bar{p},q\bar{q}} D_i g_{\bar{p}p} D_{\bar{i}} g_{\bar{q}q}\big),$$

$$B_i = \frac{1}{P_m}\Big(\sum_p \sigma_k^{p\bar{p},i\bar{i}} \lambda_p^{m-1} |D_i g_{\bar{p}p}|^2\Big), \qquad C_i = \frac{(m-1)\sigma_k^{i\bar{i}}}{P_m}\Big(\sum_p \lambda_p^{m-2}|D_i g_{\bar{p}p}|^2\Big),$$

$$D_i = \frac{1}{P_m}\Big(\sum_{p\neq i} \sigma_k^{p\bar{p}} \frac{\lambda_p^{m-1}-\lambda_i^{m-1}}{\lambda_p-\lambda_i}|D_i g_{\bar{p}p}|^2\Big), \qquad E_i = \frac{m\sigma_k^{i\bar{i}}}{P_m^2}\Big|\sum_p \lambda_p^{m-1} D_i g_{\bar{p}p}\Big|^2.$$

Define $T_{j\bar{p}q} = D_j\chi_{\bar{p}q} - D_q\chi_{\bar{p}j}$. For any $0 < \tau < 1$, we can estimate

$$\begin{aligned}\frac{1}{P_m}\Big(\sum_p \lambda_p^{m-1}\sigma_k^{j\bar{j},i\bar{i}}|D_p g_{\bar{j}i}|^2\Big) &\ge \frac{1}{P_m}\Big(\sum_p \lambda_p^{m-1}\sigma_k^{p\bar{p},i\bar{i}}|D_i g_{\bar{p}p} + T_{p\bar{p}i}|^2\Big)\\ &\ge \frac{1}{P_m}\Big(\sum_p \lambda_p^{m-1}\sigma_k^{p\bar{p},i\bar{i}}\big((1-\tau)|D_i g_{\bar{p}p}|^2 - C_\tau|T_{p\bar{p}i}|^2\big)\Big)\\ &= (1-\tau)\sum_i B_i - \frac{C_\tau}{P_m}\sum_p \lambda_p^{m-2}\big(\lambda_p\sigma_k^{p\bar{p},i\bar{i}}\big)|T_{p\bar{p}i}|^2.\end{aligned}$$

Now we use $\sigma_l(\lambda|i)$ and $\sigma_l(\lambda|ij)$ to denote the l-th elementary functions of

$$(\lambda|i) = (\lambda_1, \dots, \widehat{\lambda_i}, \dots, \lambda_n) \in \mathbb{R}^{n-1} \quad \text{and} \quad (\lambda|ij) = (\lambda_1, \dots, \widehat{\lambda_i}, \dots, \widehat{\lambda_j}, \dots, \lambda_n) \in \mathbb{R}^{n-2},$$

respectively. The following simple identities are used frequently:

$$\sigma_k^{i\bar{i}} = \sigma_{k-1}(\lambda|i) \quad \text{and} \quad \sigma_k^{p\bar{p},i\bar{i}} = \sigma_{k-2}(\lambda|pi).$$

Using the identity $\sigma_l(\lambda)=\sigma_l(\lambda|p)+\lambda_p\sigma_{l-1}(\lambda|p)$ for any $1\le p\le n$, we obtain

$$\begin{aligned}\frac{1}{P_m}\biggl(\sum_p \lambda_p^{m-1}\sigma_k^{j\bar{j},i\bar{i}}|D_p g_{\bar{j}i}|^2\biggr) &\ge (1-\tau)\sum_i B_i-\frac{C_\tau}{P_m}\sum_p\lambda_p^{m-2}(\sigma_k^{i\bar{i}}-\sigma_{k-1}(\lambda|pi))|T_{p\bar{p}i}|^2\\ &\ge (1-\tau)\sum_i B_i-\frac{C_\tau}{\lambda_1^2}\mathcal{F}\ge(1-\tau)\sum_i B_i-\mathcal{F}. \qquad (3\text{-}13)\end{aligned}$$

We use the notation C_τ for a constant depending on τ. To get the last inequality above, we assume that $\lambda_1^2\ge C_\tau$; otherwise, we already have the desired estimate $\lambda_1\le C$. Similarly, we may estimate

$$\begin{aligned}\frac{1}{P_m}\sigma_k^{j\bar{j}}\sum_{i\ne p}\frac{\lambda_i^{m-1}-\lambda_p^{m-1}}{\lambda_i-\lambda_p}|D_j g_{\bar{p}i}|^2 &\ge \frac{1}{P_m}\sigma_k^{p\bar{p}}\sum_{p\ne i}\frac{\lambda_i^{m-1}-\lambda_p^{m-1}}{\lambda_i-\lambda_p}|D_i g_{\bar{p}p}+T_{p\bar{p}i}|^2\\ &\ge \frac{1}{P_m}\sigma_k^{p\bar{p}}\sum_{p\ne i}\frac{\lambda_i^{m-1}-\lambda_p^{m-1}}{\lambda_i-\lambda_p}\bigl((1-\tau)|D_i g_{\bar{p}p}|^2-C_\tau|T_{p\bar{p}i}|^2\bigr)\\ &\ge \sum_i(1-\tau)D_i-\frac{C_\tau}{\lambda_1^2}\mathcal{F}\ge\sum_i(1-\tau)D_i-\mathcal{F}. \qquad (3\text{-}14)\end{aligned}$$

With the introduced notation in place, the main inequality becomes

$$\begin{aligned}0\ge{}&\frac{-C(K)}{\lambda_1}\bigl(1+|DDu|^2+|D\bar{D}u|^2\bigr)-\tau\frac{|DP_m|_\sigma^2}{mP_m^2}\\ &+\sum_i\bigl(A_i+(1-\tau)B_i+C_i+(1-\tau)D_i-(1-\tau)E_i\bigr)\\ &+N\bigl(|DDu|_{\sigma\omega}^2+|D\bar{D}u|_{\sigma\omega}^2\bigr)+(M\varepsilon-CN-C)\mathcal{F}-CM. \qquad (3\text{-}15)\end{aligned}$$

Using the critical equation (3-5), we have

$$\begin{aligned}\tau\frac{|DP_m|_\sigma^2}{mP_m^2}&=\tau m\bigl|ND|Du|^2-MDu\bigr|_\sigma^2\le 2\tau m\bigl(N^2\bigl|D|Du|^2\bigr|_\sigma^2+M^2|Du|_\sigma^2\bigr)\\ &\le C\tau mN^2\bigl(|DDu|_{\sigma\omega}^2+|D\bar{D}u|_{\sigma\omega}^2\bigr)+C\tau mM^2\mathcal{F}. \qquad (3\text{-}16)\end{aligned}$$

We thus have

$$\begin{aligned}0\ge{}&\frac{-C(K)}{\lambda_1}\bigl(1+|DDu|^2+|D\bar{D}u|^2\bigr)+(N-C\tau mN^2)\bigl(|DDu|_{\sigma\omega}^2+|D\bar{D}u|_{\sigma\omega}^2\bigr)\\ &+\sum_i\bigl(A_i+(1-\tau)B_i+C_i+(1-\tau)D_i-(1-\tau)E_i\bigr)\\ &+(M\varepsilon-C\tau mM^2-CN-C)\mathcal{F}-CM. \qquad (3\text{-}17)\end{aligned}$$

3.1. ***Estimating the third-order terms.*** In this subsection, we will adapt the argument in [Li et al. 2016] to estimate the third-order terms.

Lemma 2. *For sufficiently large m, the following estimates hold*:

$$P_m^2(B_1+C_1+D_1-E_1) \geq P_m\lambda_1^{m-2}\sum_{p\neq 1}\sigma_k^{p\bar{p}}|D_1 g_{\bar{p}p}|^2 - \lambda_1^m\sigma_k^{1\bar{1}}\lambda_1^{m-2}|D_1 g_{\bar{1}1}|^2, \tag{3-18}$$

and for any fixed $i \neq 1$,

$$P_m^2(B_i+C_i+D_i-E_i) \geq 0. \tag{3-19}$$

Proof. Fix $i \in \{1, 2, \ldots, n\}$. First, we compute

$$\begin{aligned} P_m(B_i+D_i) &= \sum_{p\neq i}\sigma_k^{p\bar{p},i\bar{i}}\lambda_p^{m-1}|D_i g_{\bar{p}p}|^2 + \sum_{p\neq i}\sigma_k^{p\bar{p}}\frac{\lambda_p^{m-1}-\lambda_i^{m-1}}{\lambda_p-\lambda_i}|D_i g_{\bar{p}p}|^2 \\ &= \sum_{p\neq i}\lambda_p^{m-2}\big((\lambda_p\sigma_k^{p\bar{p},i\bar{i}}+\sigma_k^{p\bar{p}})|D_i g_{\bar{p}p}|^2\big) + \left(\sum_{p\neq i}\sigma_k^{p\bar{p}}\sum_{q=0}^{m-3}\lambda_p{}^q\lambda_i^{m-2-q}|D_i g_{\bar{p}p}|^2\right). \end{aligned}$$

Note that

$$\lambda_p\sigma_k^{p\bar{p},i\bar{i}}+\sigma_k^{p\bar{p}} \geq \sigma_k^{i\bar{i}}.$$

To see this, we write

$$\begin{aligned} \lambda_p\sigma_k^{p\bar{p},i\bar{i}}+\sigma_k^{p\bar{p}} &= \lambda_p\sigma_{k-2}(\lambda|pi)+\sigma_{k-1}(\lambda|p) \\ &= \sigma_{k-1}(\lambda|i)-\sigma_{k-1}(\lambda|ip)+\sigma_{k-1}(\lambda|p) \\ &= \sigma_{k-1}(\lambda|i)+\lambda_i\sigma_{k-2}(\lambda|ip) \geq \sigma_{k-1}(\lambda|i) = \sigma_k^{i\bar{i}}, \end{aligned}$$

where we used the standard identity $\sigma_l(\lambda) = \sigma_l(\lambda|p)+\lambda_p\sigma_{l-1}(\lambda|p)$ twice, to get the second and third equalities. Therefore

$$P_m(B_i+D_i) \geq \sigma_k^{i\bar{i}}\left(\sum_{p\neq i}\lambda_p^{m-2}|D_i g_{\bar{p}p}|^2\right) + \left(\sum_{p\neq i}\sigma_k^{p\bar{p}}\sum_{q=0}^{m-3}\lambda_p{}^q\lambda_i^{m-2-q}|D_i g_{\bar{p}p}|^2\right). \tag{3-20}$$

It follows that

$$\begin{aligned} P_m(B_i+C_i+D_i) \geq m\sigma_k^{i\bar{i}}\sum_{p\neq i}\lambda_p^{m-2}|D_i g_{\bar{p}p}|^2 + (m-1)\sigma_k^{i\bar{i}}\lambda_i^{m-2}|D_i g_{\bar{i}i}|^2 \\ + \sum_{p\neq i}\sigma_k^{p\bar{p}}\sum_{q=0}^{m-3}\lambda_p{}^q\lambda_i^{m-2-q}|D_i g_{\bar{p}p}|^2. \end{aligned} \tag{3-21}$$

Expanding out the definition of E_i,

$$P_m^2 E_i = m\sigma_k^{i\bar{i}}\sum_{p\neq i}\lambda_p^{2m-2}|D_i g_{\bar{p}p}|^2 + m\sigma_k^{i\bar{i}}\lambda_i^{2m-2}|D_i g_{\bar{i}i}|^2 + m\sigma_k^{i\bar{i}}\sum_p\sum_{q\neq p}\lambda_p^{m-1}\lambda_q^{m-1}D_i g_{\bar{p}p}D_{\bar{i}}g_{\bar{q}q}. \tag{3-22}$$

Therefore,

$$
\begin{aligned}
P_m^2&(B_i + C_i + D_i - E_i)\\
&\geq \left(m\sigma_k^{i\bar{i}} \sum_{p\neq i} (P_m - \lambda_p^m)\lambda_p^{m-2} |D_i g_{\bar{p}p}|^2 - m\sigma_k^{i\bar{i}} \sum_{p\neq i}\sum_{q\neq p,i} \lambda_p^{m-1}\lambda_q^{m-1} D_i g_{\bar{p}p} D_{\bar{i}} g_{\bar{q}q} \right)\\
&\quad + P_m \sum_{p\neq i} \sigma_k^{p\bar{p}} \sum_{q=0}^{m-3} {\lambda_p}^q \lambda_i^{m-2-q} |D_i g_{\bar{p}p}|^2 - 2m\sigma_k^{i\bar{i}} \operatorname{Re} \sum_{q\neq i} \lambda_i^{m-1}\lambda_q^{m-1} D_i g_{\bar{i}i} D_{\bar{i}} g_{\bar{q}q}\\
&\quad + ((m-1)P_m - m\lambda_i^m)\sigma_k^{i\bar{i}} \lambda_i^{m-2} |D_i g_{\bar{i}i}|^2. \qquad (3\text{-}23)
\end{aligned}
$$

We now estimate the expression on the second line above. First,

$$
m\sigma_k^{i\bar{i}} \sum_{p\neq i} (P_m - \lambda_p^m)\lambda_p^{m-2} |D_i g_{\bar{p}p}|^2 = m\sigma_k^{i\bar{i}} \sum_{p\neq i}\sum_{q\neq p,i} \lambda_q^m \lambda_p^{m-2} |D_i g_{\bar{p}p}|^2 + m\sigma_k^{i\bar{i}} \sum_{p\neq i} \lambda_i^m \lambda_p^{m-2} |D_i g_{\bar{p}p}|^2.
$$

Next, we can estimate

$$
\begin{aligned}
-m\sigma_k^{i\bar{i}} \sum_{p\neq i}\sum_{q\neq p,i} \lambda_p^{m-1}\lambda_q^{m-1} D_i g_{\bar{p}p} D_{\bar{i}} g_{\bar{q}q} &\geq -m\sigma_k^{i\bar{i}} \sum_{p\neq i}\sum_{q\neq p,i} \frac{1}{2}\big(\lambda_p^{m-2}\lambda_q^m |D_i g_{\bar{p}p}|^2 + \lambda_p^m \lambda_q^{m-2} |D_i g_{\bar{q}q}|^2\big)\\
&= -m\sigma_k^{i\bar{i}} \sum_{p\neq i}\sum_{q\neq p,i} \lambda_p^{m-2}\lambda_q^m |D_i g_{\bar{p}p}|^2. \qquad (3\text{-}24)
\end{aligned}
$$

We arrive at

$$
\begin{aligned}
P_m^2&(B_i + C_i + D_i - E_i)\\
&\geq m\sigma_k^{i\bar{i}} \sum_{p\neq i} \lambda_i^m \lambda_p^{m-2} |D_i g_{\bar{p}p}|^2 + P_m \sum_{p\neq i} \sigma_k^{p\bar{p}} \sum_{q=0}^{m-3} {\lambda_p}^q \lambda_i^{m-2-q} |D_i g_{\bar{p}p}|^2\\
&\quad - 2m\sigma_k^{i\bar{i}} \operatorname{Re}\left(\lambda_i^{m-1} D_i g_{\bar{i}i} \sum_{q\neq i} \lambda_q^{m-1} D_{\bar{i}} g_{\bar{q}q} \right) + ((m-1)P_m - m\lambda_i^m)\sigma_k^{i\bar{i}} \lambda_i^{m-2} |D_i g_{\bar{i}i}|^2. \qquad (3\text{-}25)
\end{aligned}
$$

The next step is to extract good terms from the second summation on the first line. We fix $p \neq i$.

Case 1: $\lambda_i \geq \lambda_p$. Then $\sigma_k^{p\bar{p}} \geq \sigma_k^{i\bar{i}}$. Hence

$$
P_m \sigma_k^{p\bar{p}} \sum_{q=1}^{m-3} {\lambda_p}^q \lambda_i^{m-2-q} \geq \lambda_i^m \sigma_k^{i\bar{i}} \sum_{q=1}^{m-3} {\lambda_p}^q \lambda_p^{m-2-q} = (m-3)\sigma_k^{i\bar{i}} \lambda_i^m \lambda_p^{m-2}. \qquad (3\text{-}26)
$$

Case 2: $\lambda_i \leq \lambda_p$. Then $\lambda_p \sigma_k^{p\bar{p}} = \lambda_i \sigma_k^{i\bar{i}} + (\sigma_k(\lambda|i) - \sigma_k(\lambda|p)) \geq \lambda_i \sigma_k^{i\bar{i}}$, and we obtain

$$
P_m \sigma_k^{p\bar{p}} \sum_{q=1}^{m-3} {\lambda_p}^q \lambda_i^{m-2-q} \geq \lambda_p^m \sigma_k^{i\bar{i}} \sum_{q=1}^{m-3} {\lambda_p}^{q-1} \lambda_i^{m-1-q} \geq (m-3)\sigma_k^{i\bar{i}} \lambda_i^m \lambda_p^{m-2}. \qquad (3\text{-}27)
$$

Combining both cases, we have

$$\begin{aligned}P_m\sigma_k^{p\bar p}\sum_{q=0}^{m-3}\lambda_p{}^q\lambda_i^{m-2-q}|D_ig_{\bar pp}|^2&=P_m\sigma_k^{p\bar p}\sum_{q=1}^{m-3}\lambda_p{}^q\lambda_i^{m-2-q}|D_ig_{\bar pp}|^2+P_m\sigma_k^{p\bar p}\lambda_i^{m-2}|D_ig_{\bar pp}|^2\\&\ge(m-3)\sigma_k^{i\bar i}\lambda_i^m\lambda_p^{m-2}|D_ig_{\bar pp}|^2+P_m\sigma_k^{p\bar p}\lambda_i^{m-2}|D_ig_{\bar pp}|^2.\end{aligned}$$

Substituting this estimate into inequality (3-25), we obtain

$$\begin{aligned}&P_m^2(B_i+C_i+D_i-E_i)\\&\quad\ge(2m-3)\sigma_k^{i\bar i}\sum_{p\ne i}\lambda_i^m\lambda_p^{m-2}|D_ig_{\bar pp}|^2-2m\sigma_k^{i\bar i}\operatorname{Re}\Big(\lambda_i^{m-1}D_ig_{\bar ii}\sum_{p\ne i}\lambda_p^{m-1}D_{\bar i}g_{\bar pp}\Big)\\&\qquad+P_m\lambda_i^{m-2}\sum_{p\ne i}\sigma_k^{p\bar p}|D_ig_{\bar pp}|^2+((m-1)P_m-m\lambda_i^m)\sigma_k^{i\bar i}\lambda_i^{m-2}|D_ig_{\bar ii}|^2.\end{aligned}\tag{3-28}$$

Choose $m\gg 1$ such that

$$m^2\le(2m-3)(m-2).\tag{3-29}$$

We can therefore estimate

$$\begin{aligned}&2m\sigma_k^{i\bar i}\operatorname{Re}\Big(\lambda_i^{m-1}D_ig_{\bar ii}\sum_{p\ne i}\lambda_p^{m-1}D_{\bar i}g_{\bar pp}\Big)\\&\quad\le2\sigma_k^{i\bar i}\sum_{p\ne i}\big((2m-3)^{1/2}\lambda_i^{m/2}\lambda_p^{(m-2)/2}|D_ig_{\bar pp}|\big)\big((m-2)^{1/2}\lambda_i^{(m-2)/2}\lambda_p^{m/2}|D_{\bar i}g_{\bar ii}|\big)\\&\quad\le(2m-3)\sigma_k^{i\bar i}\sum_{p\ne i}\lambda_i^m\lambda_p^{m-2}|D_ig_{\bar pp}|^2+(m-2)\sigma_k^{i\bar i}\sum_{p\ne i}\lambda_i^{m-2}\lambda_p^m|D_{\bar i}g_{\bar ii}|^2.\end{aligned}\tag{3-30}$$

We finally arrive at

$$\begin{aligned}P_m^2(B_i+C_i+D_i-E_i)&\ge P_m\lambda_i^{m-2}\sum_{p\ne i}\sigma_k^{p\bar p}|D_ig_{\bar pp}|^2+((m-1)P_m-m\lambda_i^m)\sigma_k^{i\bar i}\lambda_i^{m-2}|D_ig_{\bar ii}|^2\\&\quad-(m-2)\sigma_k^{i\bar i}\sum_{p\ne i}\lambda_i^{m-2}\lambda_p^m|D_{\bar i}g_{\bar ii}|^2.\end{aligned}\tag{3-31}$$

If we let $i=1$, we obtain inequality (3-18). For any fixed $i\ne1$, this inequality yields

$$\begin{aligned}P_m^2(B_i+C_i+D_i-E_i)&\ge P_m\lambda_i^{m-2}\sum_{p\ne i}\sigma_k^{p\bar p}|D_ig_{\bar pp}|^2+((m-1)\lambda_1^m-\lambda_i^m)\sigma_k^{i\bar i}\lambda_i^{m-2}|D_ig_{\bar ii}|^2\\&\quad+(m-1)\sum_{p\ne1,i}\lambda_p^m\sigma_k^{i\bar i}\lambda_i^{m-2}|D_ig_{\bar ii}|^2-(m-2)\sigma_k^{i\bar i}\sum_{p\ne i}\lambda_i^{m-2}\lambda_p^m|D_{\bar i}g_{\bar ii}|^2\\&\ge P_m\lambda_i^{m-2}\sum_{p\ne i}\sigma_k^{p\bar p}|D_ig_{\bar pp}|^2\ge0.\end{aligned}$$

This completes the proof of Lemma 2. □

We observed in (3-12) that $A_i \geq 0$. Lemma 2 implies that for any $i \neq 1$,

$$A_i + B_i + C_i + D_i - E_i \geq 0.$$

Thus we have shown that for $i \neq 1$, the third-order terms in the main inequality (3-17) are indeed nonnegative. The only remaining case is when $i = 1$. By adapting once again the techniques from [Guan et al. 2015], we obtain the following lemma.

Lemma 3. *Let $1 < k \leq n$. Suppose there exists $0 < \delta \leq 1$ such that $\lambda_\mu \geq \delta\lambda_1$ for some $\mu \in \{1, 2, \ldots, k-1\}$. There exists a small $\delta' > 0$ such that if $\lambda_{\mu+1} \leq \delta'\lambda_1$, then*

$$A_1 + B_1 + C_1 + D_1 - E_1 \geq 0.$$

Proof. By Lemma 2, we have

$$P_m^2(A_1 + B_1 + C_1 + D_1 - E_1) \geq P_m^2 A_1 + P_m \lambda_1^{m-2} \sum_{p \neq 1} \sigma_k^{p\bar{p}} |D_1 g_{\bar{p}p}|^2 - \lambda_1^m \sigma_k^{1\bar{1}} \lambda_1^{m-2} |D_1 g_{\bar{1}1}|^2. \tag{3-32}$$

The key insight in [Guan et al. 2015], used also in [Li et al. 2016], is to extract a good term involving $|D_1 g_{\bar{1}1}|^2$ from A_1. By the inequality in Lemma 1 (with $\theta = 1/2$), we have for $\mu < k$

$$\begin{aligned}
P_m^2 A_1 &\geq \frac{P_m \lambda_1^{m-1} \sigma_k}{\sigma_\mu^2} \left(\left(1 + \frac{\alpha}{2}\right) \left| \sum_p \sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p} \right|^2 - \sigma_\mu \sigma_\mu^{p\bar{p},q\bar{q}} D_1 g_{\bar{p}p} D_{\bar{1}} g_{\bar{q}q} \right) \\
&= \frac{P_m \lambda_1^{m-1} \sigma_k}{\sigma_\mu^2} \left(\sum_p \left(1 + \frac{\alpha}{2}\right) |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2 + \sum_{p \neq q} \frac{\alpha}{2} \sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p} \sigma_\mu^{q\bar{q}} D_{\bar{1}} g_{\bar{q}q} \right. \\
&\qquad \left. + \sum_{p \neq q} (\sigma_\mu^{p\bar{p}} \sigma_\mu^{q\bar{q}} - \sigma_\mu \sigma_\mu^{p\bar{p},q\bar{q}}) D_1 g_{\bar{p}p} D_{\bar{1}} g_{\bar{q}q} \right) \\
&\geq \frac{P_m \lambda_1^{m-1} \sigma_k}{\sigma_\mu^2} \left(\sum_p |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2 - \sum_{p \neq q} |F^{pq} D_1 g_{\bar{p}p} D_{\bar{1}} g_{\bar{q}q}| \right),
\end{aligned} \tag{3-33}$$

where we defined $F^{pq} = \sigma_\mu^{p\bar{p}} \sigma_\mu^{q\bar{q}} - \sigma_\mu \sigma_\mu^{p\bar{p},q\bar{q}}$. Notice that if $\mu = 1$, then $F^{pq} = 1$. If $\mu \geq 2$, then the Newton–Maclaurin inequality implies

$$F^{pq} = \sigma_{\mu-1}^2(\lambda|pq) - \sigma_\mu(\lambda|pq)\sigma_{\mu-2}(\lambda|pq) \geq 0. \tag{3-34}$$

We split the sum involving F^{pq} in the following way:

$$\sum_{p \neq q} |F^{pq} D_1 g_{\bar{p}p} D_{\bar{1}} g_{\bar{q}q}| = \sum_{\substack{p \neq q \\ p,q \leq \mu}} F^{pq} |D_1 g_{\bar{p}p}| |D_{\bar{1}} g_{\bar{q}q}| + \sum_{(p,q) \in J} F^{pq} |D_1 g_{\bar{p}p}| |D_{\bar{1}} g_{\bar{q}q}|, \tag{3-35}$$

where J is the set of indices where at least one of $p \neq q$ is strictly greater than μ. The summation of terms in J can be estimated by

$$
\begin{aligned}
-\sum_{(p,q)\in J} F^{pq}|D_1 g_{\bar{p}p}||D_{\bar{1}} g_{\bar{q}q}| &\geq -\sum_{(p,q)\in J} \sigma_\mu^{p\bar{p}}\sigma_\mu^{q\bar{q}}|D_1 g_{\bar{p}p}||D_{\bar{1}} g_{\bar{q}q}| \\
&\geq -\epsilon \sum_{p\leq\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2 - C\sum_{p>\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2. \qquad (3\text{-}36)
\end{aligned}
$$

If $\mu = 1$, the first term on the right-hand side of (3-35) vanishes and this estimate applies to all terms on the right hand side of (3-35).

If $\mu \geq 2$, we have for $p, q \leq \mu$,

$$
\sigma_{\mu-1}(\lambda|pq) \leq C\frac{\lambda_1\cdots\lambda_{\mu+1}}{\lambda_p\lambda_q} \leq C\frac{\sigma_\mu^{p\bar{p}}\lambda_{\mu+1}}{\lambda_q}. \qquad (3\text{-}37)
$$

Using (3-34) and (3-37), for δ' small enough we can control

$$
\begin{aligned}
-\sum_{\substack{p\neq q\\ p,q\leq\mu}} F^{pq}|D_1 g_{\bar{p}p}||D_{\bar{1}} g_{\bar{q}q}| &\geq -\sum_{\substack{p\neq q\\ p,q\leq\mu}} \sigma_{\mu-1}^2(\lambda|pq)|D_1 g_{\bar{p}p}||D_{\bar{1}} g_{\bar{q}q}| \\
&\geq -C\lambda_{\mu+1}^2 \sum_{\substack{p\neq q\\ p,q\leq\mu}} \frac{\sigma_\mu^{p\bar{p}}}{\lambda_p}|D_1 g_{\bar{p}p}|\frac{\sigma_\mu^{q\bar{q}}}{\lambda_q}|D_{\bar{1}} g_{\bar{q}q}| \\
&\geq -C\sum_{p\leq\mu} \frac{\lambda_{\mu+1}^2}{\lambda_p^2}|\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2 \\
&\geq -C\sum_{p\leq\mu} \frac{\delta'^2}{\delta^2}|\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2 \geq -\epsilon\sum_{p\leq\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2. \qquad (3\text{-}38)
\end{aligned}
$$

Combining all cases, we have

$$
-\sum_{p\neq q} |F^{pq} D_1 g_{\bar{p}p} D_{\bar{1}} g_{\bar{q}q}| \geq -2\epsilon\sum_{p\leq\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2 - C\sum_{p>\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2. \qquad (3\text{-}39)
$$

Using this inequality in (3-33) yields

$$
\begin{aligned}
P_m^2 A_1 &\geq \frac{P_m\lambda_1^{m-1}\sigma_k}{\sigma_\mu^2}\Big((1-2\epsilon)\sum_{p\leq\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2 - C\sum_{p>\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2\Big) \\
&\geq (1-2\epsilon)\frac{P_m\lambda_1^{m-1}\sigma_k}{\sigma_\mu^2}|\sigma_\mu^{1\bar{1}} D_1 g_{\bar{1}1}|^2 - C\frac{P_m\lambda_1^{m-1}\sigma_k}{\sigma_\mu^2}\sum_{p>\mu} |\sigma_\mu^{p\bar{p}} D_1 g_{\bar{p}p}|^2. \qquad (3\text{-}40)
\end{aligned}
$$

We estimate

$$\begin{aligned}(1-2\epsilon)\frac{P_m\lambda_1^{m-1}\sigma_k}{\sigma_\mu^2}|\sigma_\mu^{1\bar{1}}D_1g_{\bar{1}1}|^2 &= (1-2\epsilon)\frac{P_m\lambda_1^{m-2}\sigma_k}{\lambda_1}\left(\frac{\lambda_1\sigma_\mu^{1\bar{1}}}{\sigma_\mu}\right)^2|D_1g_{\bar{1}1}|^2\\ &\geq (1-2\epsilon)P_m\lambda_1^{m-2}\frac{\sigma_k}{\lambda_1}\left(1-C\frac{\lambda_{\mu+1}}{\lambda_1}\right)^2|D_1g_{\bar{1}1}|^2\\ &\geq (1-2\epsilon)(1-C\delta')^2P_m\lambda_1^{m-2}\sigma_k^{1\bar{1}}|D_1g_{\bar{1}1}|^2\\ &\geq (1-2\epsilon)(1-C\delta')^2(1+\delta^m)\lambda_1^{2m-2}\sigma_k^{1\bar{1}}|D_1g_{\bar{1}1}|^2.\end{aligned} \tag{3-41}$$

For δ' and ϵ small enough, we obtain

$$P_m^2A_1 \geq \lambda_1^m\sigma_k^{1\bar{1}}\lambda_1^{m-2}|D_1g_{\bar{1}1}|^2 - C\frac{P_m\lambda_1^{m-1}\sigma_k}{\sigma_\mu^2}\sum_{p>\mu}|\sigma_\mu^{p\bar{p}}D_1g_{\bar{p}p}|^2. \tag{3-42}$$

We see that the $|D_1g_{\bar{1}1}|^2$ term cancels from the inequality (3-32) and we are left with

$$P_m^2(A_1+B_1+C_1+D_1-E_1) \geq P_m\lambda_1^{m-2}\sum_{p>\mu}\left(\sigma_k^{p\bar{p}} - C\frac{\lambda_1\sigma_k(\sigma_\mu^{p\bar{p}})^2}{\sigma_\mu^2}\right)|D_1g_{\bar{p}p}|^2. \tag{3-43}$$

For δ' small enough, the above expression is nonnegative. Indeed, for any $p>\mu$, we have

$$(\lambda_1\sigma_\mu^{p\bar{p}})^2 \leq \frac{1}{\delta^2}(\lambda_\mu\sigma_\mu^{p\bar{p}})^2 \leq C\frac{(\sigma_\mu)^2}{\delta^2}. \tag{3-44}$$

Therefore,

$$C\frac{\lambda_1\sigma_k(\sigma_\mu^{p\bar{p}})^2}{\sigma_\mu^2} \leq \frac{C}{\delta^2}\frac{\sigma_k}{\lambda_1}. \tag{3-45}$$

On the other hand, we notice that if $p>k$, then

$$\sigma_k^{p\bar{p}} \geq \lambda_1\cdots\lambda_{k-1} \geq c_n\frac{\sigma_k}{\lambda_k} \geq \frac{c_n}{\delta'}\frac{\sigma_k}{\lambda_1}.$$

If $\mu<p\leq k$, then

$$\sigma_k^{p\bar{p}} \geq \frac{\lambda_1\cdots\lambda_k}{\lambda_p} \geq c_n\frac{\sigma_k}{\lambda_p} \geq \frac{c_n}{\delta'}\frac{\sigma_k}{\lambda_1}.$$

It follows that for δ' small enough we have

$$\sigma_k^{p\bar{p}} \geq C\frac{\lambda_1\sigma_k(\sigma_\mu^{p\bar{p}})^2}{\sigma_\mu^2}. \tag{3-46}$$

This completes the proof of Lemma 3. □

3.2. ***Completing the proof.*** With Lemma 2 and Lemma 3 at our disposal, we claim that we may assume in inequality (3-17) that

$$A_i + B_i + C_i + D_i - E_i \geq 0, \quad \forall i = 1, \ldots, n. \tag{3-47}$$

Indeed, first set $\delta_1 = 1$. If $\lambda_2 \leq \delta_2 \lambda_1$ for $\delta_2 > 0$ small enough, then by Lemma 3 we see that (3-47) holds. Otherwise, $\lambda_2 \geq \delta_2 \lambda_1$. If $\lambda_3 \leq \delta_3 \lambda_1$ for $\delta_3 > 0$ small enough, then by Lemma 3 we see that (3-47) holds. Otherwise, $\lambda_3 \geq \delta_3 \lambda_1$. Proceeding iteratively, we may arrive at $\lambda_k \geq \delta_k \lambda_1$. But in this case, the C^2 estimate follows directly from the equation as

$$C \geq \sigma_k \geq \lambda_1 \cdots \lambda_k \geq (\delta_k)^{k-1} \lambda_1. \tag{3-48}$$

Therefore we may assume (3-47), and inequality (3-17) becomes

$$\begin{aligned} 0 \geq \frac{-C(K)}{\lambda_1}\big(1 + |DDu|^2 + |D\bar{D}u|^2\big) + (N - C\tau m N^2)(|DDu|^2_{\sigma\omega} + |D\bar{D}u|^2_{\sigma\omega}) \\ + (M\varepsilon - C\tau m M^2 - CN - C)\mathcal{F} - CM. \end{aligned} \tag{3-49}$$

Since $\sigma_k^{i\bar{i}} \geq \sigma_k^{1\bar{1}} \geq \dfrac{k}{n}\dfrac{\sigma_k}{\lambda_1} \geq \dfrac{1}{C\lambda_1}$ for fixed i, we can estimate

$$|DDu|^2_{\sigma\omega} + |D\bar{D}u|^2_{\sigma\omega} \geq \frac{1}{C\lambda_1}(|DDu|^2 + |D\bar{D}u|^2) \geq \frac{1}{C\lambda_1}|DDu|^2 + \frac{\lambda_1}{C}. \tag{3-50}$$

This leads to

$$\begin{aligned} 0 \geq \Big(\frac{N}{C} - C\tau m N^2 - C(K)\Big)\lambda_1 + \frac{1}{\lambda_1}\Big(\frac{N}{C} - C\tau m N^2 - C(K)\Big)(1 + |DDu|^2) \\ + (M\varepsilon - C\tau m M^2 - CN - C)\mathcal{F} - CM. \end{aligned}$$

By choosing τ small, for example $\tau = 1/(NM)$, we have

$$\begin{aligned} 0 \geq \Big(\frac{N}{C} - \frac{Cm}{M}N - C(K)\Big)\lambda_1 + \frac{1}{\lambda_1}\Big(\frac{N}{C} - \frac{Cm}{M}N - C(K)\Big)(1 + |DDu|^2) \\ + \Big(M\varepsilon - \frac{Cm}{N}M - CN - C\Big)\mathcal{F} - CM. \end{aligned}$$

Taking N and M large enough, we can ensure that the coefficients of the first three terms are positive. For example, if we let $M = N^2$ for N large, then

$$\begin{aligned} \frac{N}{C} - \frac{Cm}{M}N - C(K) &= \frac{N}{C} - \frac{Cm}{N} - C(K) > 0, \\ M\varepsilon - \frac{Cm}{N}M - CN - C &= N^2\varepsilon - CmN - CN - C > 0. \end{aligned}$$

Thus, an upper bound of λ_1 follows. □

Remark 2. In the above estimate, we assume that $\lambda = (\lambda_1, \ldots, \lambda_n) \in \Gamma_n$. Indeed, our estimate still works with $\lambda \in \Gamma_{k+1}$. It was observed in [Li et al. 2016, Lemma 7] that if $\lambda \in \Gamma_{k+1}$, then $\lambda_1 \geq \cdots \geq \lambda_n > -K_0$ for some positive constant K_0. Thus, we can replace λ by $\tilde{\lambda} = \lambda + K_0 I$ in our test function G in (3-2).

Acknowledgements

The authors would like to thank Pengfei Guan for stimulating conversations and for sharing his unpublished notes. The authors are also very grateful to the referees for an exceptionally careful reading and for many helpful suggestions.

References

[Alesker and Verbitsky 2010] S. Alesker and M. Verbitsky, "Quaternionic Monge–Ampère equation and Calabi problem for HKT-manifolds", *Israel J. Math.* **176** (2010), 109–138. MR 2653188 Zbl 1193.53118

[Ball 1984] J. M. Ball, "Differentiability properties of symmetric and isotropic functions", *Duke Math. J.* **51**:3 (1984), 699–728. MR 757959 Zbl 0566.73001

[Błocki 2005] Z. Błocki, "Weak solutions to the complex Hessian equation", *Ann. Inst. Fourier* (*Grenoble*) **55**:5 (2005), 1735–1756. MR 2172278 Zbl 1081.32023

[Dinew and Kołodziej 2014] S. Dinew and S. Kołodziej, "A priori estimates for complex Hessian equations", *Anal. PDE* **7**:1 (2014), 227–244. MR 3219505 Zbl 1297.32020

[Fu and Yau 2007] J.-X. Fu and S.-T. Yau, "A Monge–Ampère-type equation motivated by string theory", *Comm. Anal. Geom.* **15**:1 (2007), 29–75. MR 2301248 Zbl 1122.35145

[Fu and Yau 2008] J.-X. Fu and S.-T. Yau, "The theory of superstring with flux on non-Kähler manifolds and the complex Monge–Ampère equation", *J. Differential Geom.* **78**:3 (2008), 369–428. MR 2396248 Zbl 1141.53036

[Guan 2014] B. Guan, "Second-order estimates and regularity for fully nonlinear elliptic equations on Riemannian manifolds", *Duke Math. J.* **163**:8 (2014), 1491–1524. MR 3284698 Zbl 1296.58012

[Guan et al. 2012] P. Guan, J. Li, and Y. Li, "Hypersurfaces of prescribed curvature measure", *Duke Math. J.* **161**:10 (2012), 1927–1942. MR 2954620 Zbl 1254.53073

[Guan et al. 2015] P. Guan, C. Ren, and Z. Wang, "Global C^2-estimates for convex solutions of curvature equations", *Comm. Pure Appl. Math.* **68**:8 (2015), 1287–1325. MR 3366747 Zbl 1327.53009

[Hou 2009] Z. Hou, "Complex Hessian equation on Kähler manifold", *Int. Math. Res. Not.* **2009**:16 (2009), 3098–3111. MR 2533797 Zbl 1177.32013

[Hou et al. 2010] Z. Hou, X.-N. Ma, and D. Wu, "A second order estimate for complex Hessian equations on a compact Kähler manifold", *Math. Res. Lett.* **17**:3 (2010), 547–561. MR 2653687 Zbl 1225.32026

[Kołodziej and Nguyen 2016] S. Kołodziej and N. C. Nguyen, "Weak solutions of complex Hessian equations on compact Hermitian manifolds", *Compos. Math.* (online publication September 2016).

[Li 2004] S.-Y. Li, "On the Dirichlet problems for symmetric function equations of the eigenvalues of the complex Hessian", *Asian J. Math.* **8**:1 (2004), 87–106. MR 2128299 Zbl 1068.32024

[Li et al. 2016] M. Li, C. Ren, and Z. Wang, "An interior estimate for convex solutions and a rigidity theorem", *J. Funct. Anal.* **270**:7 (2016), 2691–2714. MR 3464054 Zbl 1341.35055

[Lu and Nguyen 2015] C. H. Lu and V.-D. Nguyen, "Degenerate complex Hessian equations on compact Kähler manifolds", *Indiana Univ. Math. J.* **64**:6 (2015), 1721–1745. MR 3436233 Zbl 1341.32033

[Phong et al. 2015] D. H. Phong, S. Picard, and X. Zhang, "On estimates for the Fu–Yau generalization of a Strominger system", preprint, 2015. arXiv 1507.08193

[Song and Weinkove 2008] J. Song and B. Weinkove, "On the convergence and singularities of the J-flow with applications to the Mabuchi energy", *Comm. Pure Appl. Math.* **61**:2 (2008), 210–229. MR 2368374 Zbl 1135.53047

[Spruck and Xiao 2015] J. Spruck and L. Xiao, "A note on starshaped compact hypersurfaces with a prescribed scalar curvature in space forms", preprint, 2015. arXiv 1505.01578

[Sun 2014] W. Sun, "On uniform estimate of complex elliptic equations on closed Hermitian manifolds", preprint, 2014. arXiv 1412.5001

[Székelyhidi 2015] G. Székelyhidi, "Fully non-linear elliptic equations on compact Hermitian manifolds", preprint, 2015. arXiv 1501.02762v3

[Székelyhidi et al. 2015] G. Székelyhidi, V. Tosatti, and B. Weinkove, "Gauduchon metrics with prescribed volume form", preprint, 2015. arXiv 1503.04491

[Wang 2009] X.-J. Wang, "The k-Hessian equation", pp. 177–252 in *Geometric analysis and PDEs*, edited by A. Ambrosetti et al., Lecture Notes in Mathematics **1977**, Springer, 2009. MR 2500526 Zbl 1196.35093

[Zhang 2010] X. Zhang, "A priori estimates for complex Monge–Ampère equation on Hermitian manifolds", *Int. Math. Res. Not.* **2010**:19 (2010), 3814–3836. MR 2725515 Zbl 1209.32025

[Zhang 2015] D. Zhang, "Hessian equations on closed Hermitian manifolds", preprint, 2015. arXiv 1501.03553

Received 4 Jan 2016. Revised 29 Apr 2016. Accepted 28 May 2016.

DUONG H. PHONG: phong@math.columbia.edu
Department of Mathematics, Columbia University, New York, NY 10027, United States

SEBASTIEN PICARD: picard@math.columbia.edu
Department of Mathematics, Columbia University, New York, NY 10027, United States

XIANGWEN ZHANG: xiangwen@math.uci.edu
Department of Mathematics, University of California, Irvine, CA 92697, United States

ANALYSIS AND PDE
Vol. 9, No. 7, 2016
dx.doi.org/10.2140/apde.2016.9.1711

PARABOLIC WEIGHTED NORM INEQUALITIES AND PARTIAL DIFFERENTIAL EQUATIONS

JUHA KINNUNEN AND OLLI SAARI

We introduce a class of weights related to the regularity theory of nonlinear parabolic partial differential equations. In particular, we investigate connections of the parabolic Muckenhoupt weights to the parabolic BMO. The parabolic Muckenhoupt weights need not be doubling and they may grow arbitrarily fast in the time variable. Our main result characterizes them through weak- and strong-type weighted norm inequalities for forward-in-time maximal operators. In addition, we prove a Jones-type factorization result for the parabolic Muckenhoupt weights and a Coifman–Rochberg-type characterization of the parabolic BMO through maximal functions. Connections and applications to the doubly nonlinear parabolic PDE are also discussed.

1. Introduction

Muckenhoupt's seminal result characterizes weighted norm inequalities for the Hardy–Littlewood maximal operator through the so-called A_p condition

$$\sup_Q \fint_Q w \left(\fint_Q w^{1-p'} \right)^{p-1} < \infty, \quad 1 < p < \infty.$$

Here the supremum is taken over all cubes $Q \subset \mathbb{R}^n$, and $w \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is a nonnegative weight. These weights exhibit many properties that are powerful in applications, such as reverse Hölder inequalities, a factorization property, and characterizability through BMO, where BMO refers to the set of functions of bounded mean oscillation. Moreover, the Muckenhoupt weights play a significant role in the theory of Calderón–Zygmund singular integral operators; see [García-Cuerva and Rubio de Francia 1985].

Another important aspect of the Muckenhoupt weights and BMO is that they also arise in the regularity theory of nonlinear PDEs. More precisely, the logarithm of a nonnegative solution to any PDE of the type

$$\operatorname{div}(|\nabla u|^{p-2} \nabla u) = 0, \quad 1 < p < \infty,$$

belongs to BMO and the solution itself is a Muckenhoupt weight. This was the crucial observation in [Moser 1961], where he proved the celebrated Harnack inequality for nonnegative solutions of such equations.

The research is supported by the Academy of Finland and the Väisälä Foundation.
MSC2010: 42B25, 42B37, 35K55.
Keywords: parabolic BMO, weighted norm inequalities, parabolic PDE, doubly nonlinear equations, one-sided weight.

Even though the theory of the Muckenhoupt weights is well established by now, many questions related to higher-dimensional versions of the one-sided Muckenhoupt condition

$$\sup_{x\in\mathbb{R},h>0} \frac{1}{h}\int_{x-h}^{x} w\left(\frac{1}{h}\int_{x}^{x+h} w^{1-p'}\right)^{p-1} < \infty$$

remain open. This condition was introduced by Sawyer [1986] in connection with ergodic theory. Since then these weights and the one-sided maximal functions have been a subject of intense research; see [Aimar and Crescimbeni 1998; Aimar et al. 1997; Cruz-Uribe et al. 1995; Martín-Reyes 1993; Martín-Reyes et al. 1990; 1993; Martín-Reyes and de la Torre 1993; 1994; Sawyer 1986]. In comparison with the classical A_p weights, the one-sided A_p^+ weights can be quite general. For example, they may grow exponentially, since any increasing function belongs to A_p^+. It is remarkable that this class of weights still allows for weighted norm inequalities for some special classes of singular integral operators (see [Aimar et al. 1997]), but the methods are limited to the dimension one.

The first extensions to the higher dimensions of the one-sided weights are by Ombrosi [2005]. The subsequent research in [Berkovits 2011; Forzani et al. 2011; Lerner and Ombrosi 2010] contains many significant advances, but even in the plane many of the most important questions, such as getting the full characterization of the strong-type weighted norm inequalities for the corresponding maximal functions, have not received satisfactory answers yet.

In this paper, we propose a new approach which enables us to solve many of the previously unreachable problems. In contrast with the earlier attempts, our point of view is related to Moser's work [1964; 1967] on the parabolic Harnack inequality. More precisely, in the regularity theory for the doubly nonlinear parabolic PDEs of the type

$$\frac{\partial(|u|^{p-2}u)}{\partial t} - \operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0, \qquad 1 < p < \infty \tag{1-1}$$

(see [Gianazza and Vespri 2006; Kinnunen and Kuusi 2007; Kuusi et al. 2012; Trudinger 1968; Vespri 1992]), there is a condition (Definition 3.2) that plays a role identical to that of the classical Muckenhoupt condition in the corresponding elliptic theory. Starting from the parabolic Muckenhoupt condition

$$\sup_{R} ⨍_{R^-} w\left(⨍_{R^+} w^{1-q'}\right)^{q-1} < \infty, \qquad 1 < q < \infty, \tag{1-2}$$

where $R^\pm$ are space-time rectangles with a time lag, we create a theory of parabolic weights. Here we use q to distinguish from p in the doubly nonlinear equation. Indeed, they are not related to each other.

The time variable scales as the modulus of the space variable raised to the power p in the geometry natural for (1-1). Consequently, the Euclidean balls and cubes have to be replaced by parabolic rectangles respecting this scaling in all estimates. In order to generalize the one-sided theory of weighted norm inequalities, it would be sufficient to work with the case $p = 2$. However, in view of the connections to nonlinear PDEs (see [Saari 2016; Kinnunen and Saari 2016]), we have decided to develop a general theory for $1 < p < \infty$. As far as we know, the results in this work are new even for the heat equation with $p = 2$. There are no previous studies about weighted norm inequalities with the same optimal relation to solutions of parabolic partial differential equations.

Observe that the theory of parabolic weights contains the classical A_p theory as a special case. However, the difference between elliptic and parabolic weights is not only a question of switching from cubes to parabolic rectangles. There is an extra challenge in the regularity theory of (1-1) because of the time lag appearing in the estimates. A similar phenomenon also occurs in the harmonic analysis with one-sided weights, and it has been the main obstacle in the previous approaches [Berkovits 2011; Forzani et al. 2011; Lerner and Ombrosi 2010; Ombrosi 2005]. Except for the one-dimensional case, an extra time lag appears in the arguments. Roughly speaking, a parabolic Muckenhoupt condition without a time lag implies boundedness of maximal operators with a time lag. In our approach, both the maximal operator and the Muckenhoupt condition have a time lag. This allows us to prove the necessity and sufficiency of the parabolic Muckenhoupt condition for both weak- and strong-type weighted norm inequalities of the corresponding maximal function. Our main technical tools are covering arguments related to the work of Ombrosi [2005] and Forzani et al. [2011]; parabolic chaining arguments from [Saari 2016], and a Calderón–Zygmund argument based on a slicing technique.

Starting from the parabolic Muckenhoupt condition (1-2), we build a complete parabolic theory of one-sided weighted norm inequalities and BMO in the multidimensional case. Our main results are a reverse Hölder inequality (Theorem 5.2), strong-type characterizations for weighted norm inequalities for a parabolic forward-in-time maximal function (Theorem 5.4), a Jones-type factorization result for parabolic Muckenhoupt weights (Theorem 6.3) and a Coifman–Rochberg-type characterization of parabolic BMO through maximal functions (Theorem 7.5). In Section 8, we explain in detail the connection between parabolic Muckenhoupt weights and the doubly nonlinear equation. We refer to [Aimar 1988; Fabes and Garofalo 1985; Kinnunen and Kuusi 2007; Moser 1964; 1967; Saari 2016; Trudinger 1968] for more on parabolic BMO and its applications to PDEs.

2. Notation

Throughout the paper, the n first coordinates of $\mathbb{R}^{n+1}$ will be called *spatial* and the last one *temporal*. The temporal translations will be important in what follows. Given a set $E \subset \mathbb{R}^{n+1}$ and $t \in \mathbb{R}$, we define

$$E + t := \{e + (0, \dots, 0, t) : e \in E\}.$$

The exponent p, with $1 < p < \infty$, related to the doubly nonlinear equation (1-1) will be fixed throughout the paper.

Constants C without subscript will be generic and the dependencies will be clear from the context. We also write $K \lesssim 1$ for $K \le C$ with C as above. The dependencies can occasionally be indicated by subscripts or parentheses, such as $K = K(n, p) \lesssim_{n,p} 1$.

A weight will always mean a real-valued positive locally integrable function on $\mathbb{R}^{n+1}$. Any such function w defines a measure absolutely continuous with respect to Lebesgue measure, and for any measurable $E \subset \mathbb{R}^{n+1}$, we define

$$w(E) := \int_E w.$$

We often omit mentioning that a set is assumed to be measurable. They are always assumed to be. For a locally integrable function f, the integral average is denoted as

$$\frac{1}{|E|}\int_E f = \fint_E f = f_E.$$

The positive part of a function f is $(f)^+ = (f)_+ = 1_{\{f>0\}} f$ and the negative part $(f)^- = (f)_- = -1_{\{f<0\}} f$.

3. Parabolic Muckenhoupt weights

Before the definition of the parabolic Muckenhoupt weights, we introduce the parabolic space-time rectangles in the natural geometry for the doubly nonlinear equation.

Definition 3.1. Let $Q(x,l) \subset \mathbb{R}^n$ be a cube with center x and side length l and sides parallel to the coordinate axes. Let $p > 1$ and $\gamma \in [0,1)$. We define

$$R(x,t,l) = Q(x,l) \times (t - l^p, t + l^p)$$

and

$$R^+(\gamma) = Q(x,l) \times (t + \gamma l^p, t + l^p).$$

The set $R(x,t,l)$ is called a (x,t)-centered parabolic rectangle with side l. We define $R^-(\gamma)$ as the reflection of $R^+(\gamma)$ with respect to $\mathbb{R}^n \times \{t\}$. The shorthand $R^\pm$ will be used for $R^\pm(0)$.

Now we are ready for the definition of the parabolic Muckenhoupt classes. Observe that there is a time lag in the definition for $\gamma > 0$.

Definition 3.2. Let $q > 1$ and $\gamma \in [0,1)$. A weight $w > 0$ belongs to the parabolic Muckenhoupt class $A_q^+(\gamma)$, if

$$\sup_R \biggl(\fint_{R^-(\gamma)} w\biggr)\biggl(\fint_{R^+(\gamma)} w^{1-q'}\biggr)^{q-1} =: [w]_{A_q^+(\gamma)} < \infty. \tag{3-1}$$

If the condition above is satisfied with the direction of the time axis reversed, we say $w \in A_q^-(\gamma)$. If γ is clear from the context or unimportant, it will be omitted in the notation.

The case $A_2^+(\gamma)$ occurs in the regularity theory of parabolic equations; see [Moser 1964; Trudinger 1968]. Before investigating the properties of parabolic weights, we briefly discuss how they differ from the ones already present in the literature. The weights of [Forzani et al. 2011; Lerner and Ombrosi 2010] were defined on the plane, and the sets $R^\pm(\gamma)$ in Definition 3.2 were replaced by two squares that share exactly one corner point. The definition used in [Berkovits 2011] is precisely the same as our Definition 3.2 with $p = 1$ and $\gamma = 0$.

An elementary but useful property of the parabolic Muckenhoupt weights is that they can effectively be approximated by bounded weights.

Proposition 3.3. *Assume $u, v \in A_q^+(\gamma)$. Then $f = \min\{u, v\} \in A_q^+(\gamma)$ and*

$$[f]_{A_q^+} \lesssim [u]_{A_q^+} + [v]_{A_q^+}.$$

The corresponding result holds for $\max\{u, v\}$ *as well.*

Proof. A direct computation gives

$$
\begin{aligned}
&\left(\fint_{R^-(\gamma)} f\right)\left(\fint_{R^+(\gamma)} f^{1-q'}\right)^{q-1} \\
&\quad\lesssim \left(\fint_{R^-(\gamma)} f\right)\left(\frac{1}{|R^+(\gamma)|}\int_{R^+(\gamma)\cap\{u>v\}} f^{1-q'}\right)^{q-1} + \left(\fint_{R^-(\gamma)} f\right)\left(\frac{1}{|R^+(\gamma)|}\int_{R^+(\gamma)\cap\{u\le v\}} f^{1-q'}\right)^{q-1} \\
&\quad\le \left(\fint_{R^-(\gamma)} v\right)\left(\frac{1}{|R^+(\gamma)|}\int_{R^+(\gamma)\cap\{u>v\}} v^{1-q'}\right)^{q-1} + \left(\fint_{R^-(\gamma)} u\right)\left(\frac{1}{|R^+(\gamma)|}\int_{R^+(\gamma)\cap\{u\le v\}} u^{1-q'}\right)^{q-1} \\
&\quad\le [u]_{A_q^+} + [v]_{A_q^+}.
\end{aligned}
$$

The result for $\max\{u, v\}$ is proved in a similar manner. $\square$

Properties of parabolic Muckenhoupt weights. The special role of the time variable makes the parabolic Muckenhoupt weights quite different from the classical ones. For example, the doubling property does not hold, but it can be replaced by a weaker forward-in-time comparison condition. The next proposition is a collection of useful facts about the parabolic Muckenhoupt condition, the most important of which is the property that the value of $\gamma \in [0, 1)$ does not play as big a role as one might guess. This is crucial in our arguments. The same phenomenon occurs later in connection with the parabolic BMO.

Proposition 3.4. *Let $\gamma \in [0, 1)$. Then the following properties hold true*:

(i) *If $1 < q < r < \infty$, then $A_q^+(\gamma) \subset A_r^+(\gamma)$.*

(ii) *Let $\sigma = w^{1-q'}$. Then σ is in $A_{q'}^-(\gamma)$ if and only if $w \in A_q^+(\gamma)$.*

(iii) *Let $w \in A_q^+(\gamma)$, $\sigma = w^{1-q'}$ and $t > 0$. Then*

$$\fint_{R^-(\gamma)} w \le C_t \fint_{t+R^-(\gamma)} w \quad \textit{and} \quad \fint_{R^+(\gamma)} \sigma \le C_t \fint_{-t+R^+(\gamma)} \sigma.$$

(iv) *If $w \in A_q^+(\gamma)$, then we may replace $R^-(\gamma)$ by $R^-(\gamma) - a$ and $R^+(\gamma)$ by $R^+(\gamma) + b$ for any $a, b \ge 0$ in the definition of the parabolic Muckenhoupt class. The new condition is satisfied with a different constant $[w]_{A_q^+}$.*

(v) *If $1 > \gamma' > \gamma$, then $A_q^+(\gamma) \subset A_q^+(\gamma')$.*

(vi) *Let $w \in A_q^+(\gamma)$. Then*

$$w(R^-(\gamma)) \le C\left(\frac{|R^-(\gamma)|}{|S|}\right)^q w(S)$$

for every $S \subset R^+(\gamma)$.

(vii) *If $w \in A_q^+(\gamma)$ with some $\gamma \in [0, 1)$, then $w \in A_q^+(\gamma')$ for all $\gamma' \in (0, 1)$.*

Proof. First we observe that (i) follows from Hölder's inequality and (ii) is obvious. For the case $t + R^-(\gamma) = R^+(\gamma)$ the claim (iii) follows from Jensen's inequality. For a general t, the result follows from subdividing the rectangles $R^{\pm}(\gamma)$ into smaller and possibly overlapping subrectangles and applying the result to them. The property (iv) follows directly from (iii), as does (v) from (iv).

For (vi), take $S \subset R^+(\gamma)$ and let $f = 1_S$. Apply the $A_q^+(\gamma)$ condition to see that

$$\begin{aligned}\left(\frac{|S|}{|R^+(\gamma)|}\right)^q w(R^-(\gamma)) &= (f_{R^+(\gamma)})^q w(R^-(\gamma)) \\ &\le \left(\fint_{R^+(\gamma)} f^q w\right)\left(\fint_{R^+(\gamma)} w^{1-q'}\right)^{q/q'} w(R^-(\gamma)) \\ &\le C w(S).\end{aligned}$$

For the last property (vii), take $R = Q(x,l) \times (t-l^p, t+l^p)$. Let $\gamma \in (0,1)$ and suppose $w \in A_q^+(\gamma)$. We will prove that the condition $A_q^+(2^{-1}\gamma)$ is satisfied. We subdivide Q into 2^{nk} dyadic subcubes $\{Q_i\}_{i=1}^{2^{nk}}$. This gives dimensions for the lower halves of parabolic rectangles $R_i^-(\gamma)$. For a given Q_i, we stack a minimal amount of the rectangles $R_i^-(\gamma)$ so that they almost pairwise disjointly cover $Q_i \times (t-l^p, t-2^{-1}\gamma l^p)$. The number of $R_i^-(\gamma)$ needed to cover $Q \times (t-l^p, t-2^{-1}\gamma l^p)$ is bounded by

$$2^{nk} \cdot \frac{(1-2^{-1}\gamma)l^p}{2^{-nkp}(1-\gamma)l^p} = 2^{nk(p+1)}\frac{2-\gamma}{2(1-\gamma)}.$$

Corresponding to each Q_i, there is a sequence of at most $2^k - 1$ vectors $d_j = 2^{-k-1} l e_j$ with $e_j \in \{0,1\}^n$ such that

$$Q_i + \sum_j d_j = 2^{-k} Q.$$

Next we show how every rectangle $R_i(\gamma)$ can be transported to the same spatially central position $2^{-k}Q$ without losing too much information about their measures. By (vi) we have

$$w(R_i^-(\gamma)) \le C\left(\frac{|R_i^-(\gamma)|}{|S|}\right)^q w(S)$$

for any $S \subset R_i^+(\gamma)$. We choose S such that its projection onto space variables is $(Q_i + d_1) \cap Q_i$, and its projection onto time variables has full length $(1-\gamma)(2^{-k}l)^p$. Then

$$w(R_i^-(\gamma)) \le C_0 w(S) \le C_0 w(R_i^{1-}(\gamma)),$$

where $R_i^{1-}(\gamma) \supset S$ is $Q_i + d_1$ spatially and coincides with S as a temporal projection. The constant C_0 depends on n and q.

Next we repeat the argument to obtain a similar estimate for $R_i^{1-}(\gamma)$ in the place of $R_i^-(\gamma)$. We obtain a new rectangle on the right-hand side, on which we repeat the argument again. With k iterations, we reach the inequality

$$w(R_i^-(\gamma)) \le C_0^{2^k-1} w(R_i^{*-}(\gamma)),$$

where $R_i^{*-}(\gamma)$ is the parabolic box whose projection onto the coordinates corresponding to the space variables is $2^{-k}Q$. The infimum of time coordinates of points in $R_i^{*-}(\gamma)$ equals

$$\inf\{t : (x,t) \in R_i^-\} + (2^k - 1)(1+\gamma)(2^{-k}l)^p.$$

As $p>1$, the second term in this sum can be made arbitrarily small. In particular, for a large enough k, we have

$$(2^k-1)(1+\gamma)(2^{-k}l)^p \le 2\cdot 2^{-k(p-1)}l^p \le \tfrac{1}{100}\gamma l^p.$$

In this fashion, we may choose a suitable finite k and divide the sets $R^{\pm}(2^{-1}\gamma)$ into $N \lesssim_{n,\gamma} 2^{nkp}$ parts $R_i^{\pm}(\gamma)$. They satisfy

$$w(R_i^-(\gamma)) \le C_0^{2^k-1} w(R_i^{*-}(\gamma))$$

and

$$\sigma(R_i^+(\gamma)) \le C_0^{2^k-1} \sigma(R_i^{*+}(\gamma)),$$

where all starred rectangles have their projections onto space variables centered at $2^{-k}Q$; they have equal side length $2^{-kp}l^p$, and

$$\tfrac{1}{2}\gamma l^p \le d(R_i^{*-}(\gamma), R_j^{*+}(\gamma)) < 2l^p$$

for all i, j. All this can be done by a choice of k which is uniform for all rectangles.

It follows that

$$\begin{aligned}\left(\fint_{R^-(2^{-1}\gamma)} w\right)\left(\fint_{R^+(2^{-1}\gamma)} w^{1-q'}\right)^{q-1} &\lesssim \sum_{i,j=1}^N \left(\fint_{R_i^-(\gamma)} w\right)\left(\fint_{R_j^+(\gamma)} w^{1-q'}\right)^{q-1}\\ &\lesssim \sum_{i,j=1}^N \left(\fint_{R_i^{*-}(\gamma)} w\right)\left(\fint_{R_j^{*+}(\gamma)} w^{1-q'}\right)^{q-1}\\ &\lesssim \sum_{i,j=1}^N C = C(n,p,k,\gamma,q,[w]_{A_q^+}(\gamma)),\end{aligned}$$

where in the last inequality we used (iv). Since the estimate is uniform in R, the claim follows. □

4. Parabolic maximal operators

In this section, we will study parabolic forward-in-time maximal operators, which are closely related to the one-sided maximal operators studied in [Berkovits 2011; Forzani et al. 2011; Lerner and Ombrosi 2010]. The class of weights in [Forzani et al. 2011], originally introduced by Ombrosi [2005], characterizes the weak-type inequality for the corresponding maximal operator, but the question about the strong-type inequality remains open. On the other hand, Lerner and Ombrosi [2010] managed to show that the same class of weights supports strong-type boundedness for another class of operators with a time lag. For the boundedness of these operators, however, the condition on weights is not necessary. Later the techniques developed by Berkovits [2011] showed that a weight condition without a time lag implies boundedness of maximal operators with a time lag. That approach applied to all dimensions. In our case both the maximal operator and the Muckenhoupt condition have a time lag. This approach, together with scaling of parabolic rectangles, allows us to prove both the necessity and sufficiency of the parabolic Muckenhoupt condition for weak- and strong-type weighted norm inequalities for the maximal function to be defined next.

Definition 4.1. Let $\gamma \in [0,1)$. For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$ define the parabolic maximal function

$$M^{\gamma+} f(x,t) = \sup_{R(x,t)} \fint_{R^+(\gamma)} |f|,$$

where the supremum is taken over all parabolic rectangles centered at (x,t). If $\gamma = 0$, it will be omitted in the notation. The operator $M^{\gamma-}$ is defined analogously.

The necessity of the A_q^+ condition can be proved in a similar manner to its analogue in the classical Muckenhoupt theory, but already here the geometric flexibility of Definition 3.2 simplifies the statement.

Lemma 4.2. *Let w be a weight such that the operator $M^{\gamma+} : L^q(w) \to L^{q,\infty}(w)$ is bounded. Then $w \in A_q^+(\gamma)$.*

Proof. Take $f > 0$ and choose R such that $f_{S^+} > 0$, where $S^+ = R^+$ if $\gamma = 0$. If $\gamma > 0$,

$$S^+ = R^-(\gamma) + (1-\gamma) l^p + 2^p \gamma l^p$$

will do. Redefine $f = \chi_{S^+} f$. Take a positive $\lambda < C_\gamma f_{S^+}$. With a suitably chosen C_γ, we have

$$w(R^-) \leq w\big(\{x \in \mathbb{R}^{n+1} : M^{\gamma+} f > \lambda\}\big) \leq \frac{C}{\lambda^q} \int_{R^+} f^q w.$$

The claim follows letting $\lambda \to C_\gamma f = C_\gamma (w+\epsilon)^{1-q'}$ and $\epsilon \to 0$, and concluding by argumentation similar to Proposition 3.4. □

Covering lemmas. The converse claim requires a couple of special covering lemmas. It is not clear whether the main covering lemma in [Forzani et al. 2011] extends to dimensions higher than two. However, in our geometry the halves of parabolic rectangles are indexed along their spatial centers instead of corner points, which was the case in [Forzani et al. 2011]. This fact will be crucial in the proof of Lemma 4.4, and this enables us to obtain results in the multidimensional case as well.

Lemma 4.3. *Let R_0 be a parabolic rectangle, and let $\mathcal{F}$ be a countable collection of parabolic rectangles with dyadic side lengths such that for each $i \in \mathbb{Z}$ we have*

$$\sum_{\substack{P \in \mathcal{F} \\ l(P) = 2^i}} 1_{P^-} \lesssim 1.$$

Moreover, assume $P^- \not\subseteq R^-$ for all distinct $P, R \in \mathcal{F}$. Then

$$\sum_{P \in \mathcal{G}} |P| \lesssim |R_0|,$$

where $\mathcal{G} = \{P \in \mathcal{F} : P^+ \cap R_0^+ \neq \varnothing, |P| < |R_0|\}$.

Proof. Recall that $R^\pm = R^\pm(0)$. We may write $\mathcal{G} \subset \mathcal{G}_0(R_0) \cup \mathcal{G}_1$, where

$$\mathcal{G}_0(R) = \{P \in \mathcal{F} : P \cap \partial R^+, |P| < |R|\}$$

and

$$\mathcal{G}_1 = \{P \in \mathcal{F} : P \subset R_0^+, |P| < |R_0|\}.$$

That is, the rectangles having their upper halves in R_0^+ are either contained in it or they meet its boundary. An estimate for $\mathcal{G}_0(R)$ with an arbitrary parabolic rectangle R instead of R_0 will be needed, so we start with it. Let P be a parabolic rectangle with the spatial side length $l(P) = 2^{-i}$. If $P \cap \partial R^+ \neq \varnothing$, then $P \subset A_i$, where A_i can be realized as a collection of $2(n+1)$ rectangles corresponding to each face of R such that

$$|A_i| \lesssim 2l(R)^n \cdot 2^{-ip} + 2nl(R)^{p+n-1} \cdot 2^{-i}.$$

Now choosing $k_0 \in \mathbb{Z}$ such that $2^{-k_0} < l(R) < 2^{-k_0+1}$, we get, by the bounded overlap,

$$\sum_{P \in \mathcal{G}_0(R)} |P| = \sum_{i=k_0}^{\infty} \sum_{\substack{P \in \mathcal{G}_0(R) \\ l(P)=2^{-i}}} |P| \lesssim \sum_{i=k_0}^{\infty} |A_i| \lesssim |R|.$$

Once the rectangles meeting the boundary are clear, we proceed to $\mathcal{G}_1$. The side lengths of rectangles in $\mathcal{G}_1$ are bounded from above. Hence there is at least one rectangle with the maximal side length. Let Σ_1 be the collection of $R \in \mathcal{G}_1$ with the maximal side length. We continue recursively. Once Σ_j with $j = 1, \ldots, k$ have been chosen, take the rectangles R with the maximal side length among the rectangles in $\mathcal{G}_1$ satisfying

$$R^- \cap \bigcup_{P \in \bigcup_{j=1}^{k} \Sigma_j} P^- = \varnothing.$$

Let them form the collection Σ_{k+1}. Define the limit collection to be

$$\Sigma = \bigcup_j \Sigma_j.$$

Each $P \in \mathcal{G}_1$ is either in Σ or P^- meets R^- with $R \in \Sigma$ and $l(P) < l(R)$. Otherwise P would have been chosen to be an element of Σ. This implies

$$\sum_{R \in \mathcal{G}_1} |R| \le \sum_{R \in \mathcal{G}_1 \cap \Sigma} \Bigg(|R| + \sum_{\substack{P \in \mathcal{G}_1 : P^- \cap R^- \neq \varnothing \\ |P| < |R|}} |P| \Bigg).$$

In the second sum, both P and R are in $\mathcal{F}$, so $P^- \nsubseteq R^-$ by assumption. Thus $P \cap \partial R^- \neq \varnothing$, and the sum in the parentheses is controlled by a constant multiple of $|R|$ (by applying the estimate we have for $\mathcal{G}_0(\widetilde{R})$, where $\widetilde{R}$ is a parabolic rectangle with upper half R^-). The rectangles in each Σ_j have equal side length so that

$$\sum_{R \in \mathcal{G}_1} |R| \lesssim \sum_{R \in \mathcal{G}_1 \cap \Sigma} |R| = \sum_j \sum_{R \in \mathcal{G}_1 \cap \Sigma_j} |R| \lesssim \sum_j \Bigg| \bigcup_{R \in \Sigma_j} R \Bigg| \le \Bigg| \bigcup_{R \in \mathcal{G}_1} R \Bigg| \le |R_0|. \qquad \square$$

The hypotheses of the next lemma correspond to a covering obtained using the parabolic maximal function, and the conclusion provides us with a covering that has bounded overlap. This fact is analogous to the two-dimensional Lemma 3.1 in [Forzani et al. 2011].

Lemma 4.4. *Let $\lambda > 0$, $f \in L^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$ be nonnegative, and $A \subset \mathbb{R}^{n+1}$ be a set of finitely many points such that for each $x \in A$ there is a parabolic rectangle R_x with dyadic side length satisfying*

$$\fint_{R_x^+(\gamma)} f \eqsim \lambda. \tag{4-1}$$

Then there is $\Gamma \subset A$ such that for each $x \in \Gamma$ there is $F_x \subset R_x^+(\gamma)$ with

(i) $A \subset \bigcup_{x\in\Gamma} \overline{R_x^-}$,

(ii) $(1/|R_x|)\int_{F_x} f \gtrsim \lambda$ *and* $\sum_{x\in\Gamma} 1_{F_x} \lesssim 1$.

Proof. To simplify the notation, we identify the sets R_x^- with their closures. Their side lengths are denoted by l_x. Let $x_1 \in A$ be a point with maximal temporal coordinate. Recursively, choose $x_{k+1} \in A \setminus \bigcup_{j=1}^k R_x^-$. Define $\Delta = \{x_i\}_i$. This is a finite set. Take $x \in \Delta$ with maximal l_x and define $\Gamma_1 = \{x\}$. Let $\Gamma_{k+1} = \Gamma_k \cup \{y\}$, where $R_y^- \not\subseteq R_x^-$ for all $x \in \Gamma_k$ and l_y is maximal among the l_y satisfying the criterion. By finiteness, the process will stop and let Γ be the final collection.

Given $x, y \in \Gamma$ with $l_x = l_y =: r$ and $x \neq y$, their Euclidean distance satisfies

$$|x - y| \geq \min\left\{\tfrac{1}{2}r, r^p\right\}.$$

There is a dimensional constant $\alpha \in (0, 1)$ such that $\alpha R_x \cap \alpha R_y = \varnothing$, and, given $z \in \mathbb{R}^{n+1}$, there is a dimensional constant $\beta > 0$ such that

$$\bigcup_{x\in\Gamma : z\in R_x} R_x \subset R(z, \beta r).$$

Thus

$$(\beta r)^n (2\beta r)^p = |R(z, \beta r)| \geq \sum_{\substack{x\in\Gamma : l_x = r, \\ z\in R_x}} |\alpha R_x| = (\alpha r)^n (2\alpha r)^p \sum_{x\in\Gamma : l_x = r} 1_{R_x}(z),$$

and consequently

$$\sum_{x\in\Gamma : l_x = r} 1_{R_x} \lesssim 1. \tag{4-2}$$

Define

$$\mathcal{G}_x = \left\{ y \in \Gamma : R_x^+(\gamma) \cap R_y^+(\gamma) \neq \varnothing, |R_y| < |R_x| \right\}.$$

By inequality (4-2), the assumptions of Lemma 4.3 are fulfilled. Hence

$$\sum_{y\in\mathcal{G}_x} |R_y^+(\gamma)| \lesssim |R_x^+(\gamma)|.$$

By (4-1), we have

$$\sum_{y\in\mathcal{G}_x} \int_{R_y^+(\gamma)} f \lesssim \lambda \sum_{y\in\mathcal{G}_x} |R_y^+(\gamma)| \lesssim \lambda |R_x^+(\gamma)| \lesssim \int_{R_x^+(\gamma)} f.$$

Let the constant in this inequality be N.

Define $s := \#\mathcal{G}_x$. When $s \le 2N$, we choose $F_x = R_x^+(\gamma)$. If $s > 2N$, we define

$$E_i^x = \Big\{ z \in R_x^+ : \sum_{y\in\Gamma : l_y < l_x} 1_{R_y^+(\gamma)}(z) \ge i \Big\}.$$

Thus $\sum_i 1_{E_i^x}(z)$ counts the points $y \in \mathcal{G}_x$ whose rectangles contain z. Hence

$$2N \int_{E_{2N}^x} f \le \sum_{i=1}^{s} \int_{E_i^x} f = \int_{R_x^+(\gamma)} f \sum_{i=1}^{s} 1_{E_i^x} \le \int_{R_x^+(\gamma)} f \sum_{y\in\mathcal{G}_x} 1_{R_y^+(\gamma)} = \sum_{y\in\mathcal{G}_x} \int_{R_y^+(\gamma)} f \le N \int_{R_x^+(\gamma)} f.$$

For the set $F_x = R_x^+(\gamma) \setminus E_{2N}^x$, we have

$$\int_{F_x} f = \int_{R_x^+(\gamma)} f - \int_{E_{2N}^x} f \ge \frac{1}{2} \int_{R_x^+(\gamma)} f \gtrsim \lambda |R_x^+(\gamma)|.$$

It remains to prove the bounded overlap of F_x. Take $z \in \bigcap_{i=1}^k F_{x_i}$. Take x_j so that l_{x_j} is maximal among l_{x_i}, $i = 1, \dots, k$. By (4-2) there are at most C_n rectangles with this maximal side length that contain z. Moreover, their subsets F_x meet at most $2N$ upper halves of smaller rectangles so that $k \le 2NC_n$. □

Weak-type inequalities. Now we can proceed to the proof of the weak-type inequality. The proof makes use of a covering argument as in [Forzani et al. 2011] adjusted to the present setting.

Lemma 4.5. *Let $q \ge 1$, $w \in A_q^+(\gamma)$ and $f \in L^q(w)$. There is a constant $C = C(n, \gamma, p, w, q)$ such that*

$$w\big(\{x \in \mathbb{R}^{n+1} : M^{\gamma+} f > \lambda\}\big) \le \frac{C}{\lambda^p} \int |f|^p w$$

for every $\lambda > 0$.

Proof. We first assume $f > 0$ is bounded and compactly supported. Since

$$\begin{aligned} M^{\gamma+} f(x) &= \sup_{h>0} \frac{1}{R(x,h,\gamma)^+} \int_{R(x,h,\gamma)^+} f \\ &\lesssim \sup_{i\in\mathbb{Z}} \frac{1}{R(x,2^i,2^{-2}\gamma)^+} \int_{R(x,2^i,2^{-2}\gamma)^+} f \\ &= \lim_{j\to-\infty} \sup_{i\in\mathbb{Z}, i>j} \frac{1}{R(x,2^i,\gamma')^+} \int_{R(x,2^i,\gamma')^+} f, \end{aligned}$$

it suffices to consider rectangles with dyadic side lengths bounded from below provided that we use smaller γ, and the claim will follow from monotone convergence. The actual value of γ is not important because of Proposition 3.4. We may assume w is bounded from above and from below (see Proposition 3.3).

Moreover, it suffices to estimate $w(E)$, where

$$E = \{x \in \mathbb{R}^{n+1} : \lambda < M^{\gamma+} f \le 2\lambda\}.$$

Once this has been done, we may sum up the estimates to get

$$w\big(\mathbb{R}^{n+1} \cap \{M^{\gamma+} f > \lambda\}\big) = \sum_{i=0}^{\infty} w\big(\mathbb{R}^{n+1} \cap \{2^i\lambda < M^{\gamma+} f \le 2^{i+1}\lambda\}\big)$$
$$\le \sum_{i=0}^{\infty} \frac{1}{2^i} \frac{C}{\lambda^p} \int |f|^p w \le \frac{C}{\lambda^p} \int |f|^p w.$$

Let $K \subset E$ be an arbitrary compact subset. Denote the lower bound for the side lengths of the parabolic rectangles in the basis of the maximal operator by $\xi < 1$. For each $x \in K$, there is dyadic $l_x > \xi$ such that

$$\fint_{R^+(x,l_x,\gamma)} f \approx \lambda.$$

Define $R_x := R(x, l_x)$. Since $f \in L^1$, we have

$$|R_x^+(\gamma)| < \frac{1}{\lambda} \int f = C(\lambda, \|f\|_{L^1}) < \infty.$$

Thus $\sup_{x \in K} l_x < \infty$. Let $a = \min w$. There is $\epsilon > 0$, uniform in x, such that

$$w((1+\epsilon)R_x^- \setminus R_x^-) \le a\xi^{n+p} \le w(R_x^-)$$

and $w((1+\epsilon)R_x^-) \le 2w(R_x^-)$ hold for all $x \in K$. By compactness, there is a finite collection of balls $B(x, \xi^p\epsilon/2)$ to cover K. Denote the set of centers by A, and apply Lemma 4.4 to extract the subcollection Γ. Each $y \in K$ is in $B(x, \xi^p\epsilon/2)$ with $x \in A$. Each $x \in A$ is in R_z^- with $z \in \Gamma$, so each $y \in K$ is in $B(x, \xi^p\epsilon/2) \subset (1+\epsilon)R_z^-$. Thus

$$w(K) \le \sum_{z\in\Gamma} w((1+\epsilon)R_z^-) \le 2\sum_{z\in\Gamma} w(R_z^-)$$
$$\le \frac{C}{\lambda^q} \sum_{z\in\Gamma} w(R_z^-)\bigg(\frac{1}{|R_z^+(\gamma)|}\int_{F_z} f\bigg)^q$$
$$\le \frac{C}{\lambda^q} \sum_{z\in\Gamma} \frac{w(R_z^-)}{|R_z^-|}\bigg(\fint_{R_z^+(\gamma)} w^{1-q'}\bigg)^{q-1} \int_{F_z} f^q w$$
$$\le \frac{C}{\lambda^q} \int f^q w.$$

In the last inequality we used the A_q^+ condition together with a modified configuration justified in Proposition 3.4, and the bounded overlap of the sets F_z. □

Now we are in a position to summarize the first results about the parabolic Muckenhoupt weights. We begin with the weak-type characterization for the operator studied in [Berkovits 2011]. Along with this result, the definition in [Berkovits 2011] leads to all same results in $\mathbb{R}^{n+1}$ as the other definition from [Forzani et al. 2011] does in $\mathbb{R}^2$. The next theorem holds even in the case $p = 1$, which is otherwise excluded in this paper.

Theorem 4.6. *Let w be a weight and $q > 1$. Then $w \in A_q^+(\gamma)$ with $\gamma = 0$ if and only if M^+ is of w-weighted weak type (q, q).*

Proof. Combine Lemma 4.2 and Lemma 4.5. □

The next theorem is the first main result of this paper. Observe that all the parabolic operators $M^{\gamma+}$ with $\gamma \in (0, 1)$ have the same class of good weights. This interesting phenomenon seems to be related to the fact that $p > 1$.

Theorem 4.7. *Let w be a weight and $q > 1$. Then the following conditions are equivalent*:

(i) *$w \in A_q^+$ for some $\gamma \in (0, 1)$.*

(ii) *$w \in A_q^+$ for all $\gamma \in (0, 1)$.*

(iii) *There is $\gamma \in (0, 1)$ such that the operator $M^{\gamma+}$ is of weighted weak type (q, q) with the weight w.*

(iv) *The operator $M^{\gamma+}$ is of weighted weak type (q, q) with the weight w for all $\gamma \in (0, 1)$.*

Proof. This follows from Lemma 4.2, Lemma 4.5 and Proposition 3.4(vii). □

5. Reverse Hölder inequalities

Parabolic reverse Hölder inequalities have already been studied in [Berkovits 2011], and they were used to prove sufficiency of the nonlagged Muckenhoupt condition for the lagged strong-type inequality. The proof included the classical argument with self-improving properties and interpolation. Our reverse Hölder inequality will lead to an even stronger self-improving property, and this will give us a characterization of the strong-type inequality. We will encounter several challenges. For example, our ambient space does not have the usual dyadic structure. In the classical Muckenhoupt theory this would not be a problem, but here the forwarding in time gives new complications. We will first prove an estimate for the level sets, and then we will use it to conclude the reverse Hölder inequality.

Lemma 5.1. *Let $w \in A_q^+(\gamma)$, $\widetilde{R}_0 = Q_0 \times \bigl(\tau, \tau + \frac{3}{2} l_0^p\bigr)$ and $\widehat{R}_0 = Q_0 \times (\tau, \tau + l_0^p)$. Then there exist $C = C([w]_{A_q^+(\gamma)}, n, p)$ and $\beta \in (0, 1)$ such that for every $\lambda \geq w_{R_0^-}$, we have*

$$w(\widehat{R}_0 \cap \{w > \lambda\}) \leq C\lambda \bigl|\widetilde{R}_0 \cap \{w > \beta\lambda\}\bigr|.$$

Proof. We introduce some notation first. For a parabolic rectangle $R = Q \times (t_0, t_0 + 2l(Q)^p)$, we define

$$\widehat{R} = Q \times (t_0, t_0 + l(Q)^p) \tag{5-1}$$

and

$$\check{R} = Q \times \bigl(t_0 + (1+\gamma)l(Q)^p, \tfrac{3}{2} l(Q)^p\bigr). \tag{5-2}$$

Here $\gamma \in \bigl(0, \frac{1}{2}\bigr)$, and by Proposition 3.4, we may replace the sets $R^{\pm}(\gamma)$ everywhere by the sets $\widehat{R}$ and $\check{R}$. Note that $\widehat{R} = R^-$. The hats are used to emphasize that $\widehat{R}$ and $\check{R}$ are admissible in the A_q^+ condition, whereas R^- is used as the set should be interpreted as a part of a parabolic rectangle. For $\beta \in (0, 1)$, the condition $A_q^+(\gamma)$ gives

$$\big|\check{R}\cap\{w\le \beta w_{\widehat{R}}\}\big| \le \beta^{p'-1}\int_{\check{R}}\frac{w^{1-p'}}{w_{\widehat{R}}^{1-p'}} \le (\beta C)^{p'-1}|\check{R}|.$$

Taking $\alpha\in(0,1)$, we may choose β such that

$$\big|\check{R}\cap\{w> \beta w_{\widehat{R}}\}\big| > \alpha|\check{R}|. \tag{5-3}$$

Let

$$\mathcal{B}=\big\{Q\times\big(t-\tfrac{1}{2}l(Q)^p,t+\tfrac{1}{2}l(Q)^p\big): Q\subset Q_0 \text{ dyadic},\ t\in(0,l^p)\big\}.$$

Here dyadic means dyadic with respect to Q_0, and hence the collection $\mathcal{B}$ consists of the lower parts $\widehat{R}$ of spatially dyadic short parabolic rectangles interpreted as metric balls with respect to

$$d((x,t),(x',t'))=\max\big\{|x-x'|_\infty, C_p|t-t'|^{1/p}\big\}.$$

Notice that the $(n+1)$-dimensional Lebesgue measure is doubling with respect to d.

We define a noncentered maximal function with respect to $\mathcal{B}$ as

$$M_{\mathcal{B}}f(x)=\sup_{\{x\}\subset B\in\mathcal{B}} \fint_B f,$$

where the supremum is taken over all sets in $\mathcal{B}$ that contain x. By the Lebesgue differentiation theorem, we have

$$\widehat{R}_0\cap\{w>\lambda\}\subset\{M_{\mathcal{B}}(1_{\widehat{R}_0}w)>\lambda\}=:E$$

up to a null set. Next we will construct a Calderón–Zygmund-type cover. The idea is to use dyadic structure to deal with spatial coordinates, then separate the scales, and finally conclude, with one-dimensional arguments, with the assumptions of Lemma 4.3.

Define the slice $E_t=E\cap(\mathbb{R}^n\times\{t\})$ for fixed t. Since $\lambda\ge w_{\widehat{R}_0}$, we may find a collection of maximal dyadic cubes $Q_i^t\times\{t\}\subset E_t$ such that for each Q_i there is $B_i^t\in\mathcal{B}$ with

$$B_i^t\cap(Q_0\times\{t\})=Q_i^t \quad\text{and}\quad \fint_{B_i^t} w>\lambda.$$

Clearly $\{B_i^t\}_i$ is pairwise disjoint and covers E_t. Moreover, since Q_i^t is maximal, the dyadic parent $\widehat{Q}_i^t$ of Q_i^t satisfies

$$\fint_{\widehat{Q}_i^t\times I} w\le\lambda$$

for all intervals $I\ni t$ with $|I|=l(\widehat{Q}_i^t)^p$ and especially for the ones with $\widehat{Q}_i^t\times I\supset B_i^t$. Hence

$$\lambda<\fint_{B_i^t} w\lesssim \fint_{\widehat{Q}_i^t\times I} w\le\lambda. \tag{5-4}$$

We gather the collections corresponding to $t\in(\tau,\tau+l_0^p)$ together, and separate the resulting collection into subcollections as

$$\mathcal{Q}=\{B_i^t: i\in\mathbb{Z}, t\in(0,l^p)\}=\bigcup_{j\in\mathbb{Z}}\mathcal{Q}_j,$$

where $\mathcal{Q}_j = \{Q \times I \in \mathcal{Q} : |Q| = 2^{-jn}|Q_0|\}$. Each $\mathcal{Q}_j$ can be partitioned into subcollections corresponding to different spatial dyadic cubes $\mathcal{Q}_j = \bigcup_i \mathcal{Q}_{ji}$. Here

$$\mathcal{Q}_{ji} = \{Q \times I \in \mathcal{Q}_j : Q = Q_i^t, t \in (\tau, \tau + l^p)\}.$$

If needed, we may reindex the Calderón–Zygmund cubes canonically with j and i such that j tells the dyadic generation and i specifies the cube such that $Q_{ji}^t = Q_{ji}^{t'}$. Then

$$\bigcup_{B \in \mathcal{Q}_{ij}} B \cap \bigcup_{B' \in \mathcal{Q}_{i'j}} B' = \varnothing$$

whenever $i \neq i'$. Thus we may identify $\mathcal{Q}_{ji}$ with a collection of intervals and extract a covering subcollection with an overlap bounded by 2. Hence we get a covering subcollection of $\mathcal{Q}_j$ with an overlap bounded by 2, and hence a countable covering subcollection of $\mathcal{Q}$ such that its restriction to any dyadic scale has an overlap bounded by 2. Denote the final collection by $\mathcal{F}$. Its elements are interpreted as lower halves of parabolic rectangles; that is, there are parabolic rectangles P with $P^- \in \mathcal{F}$.

Collect the parabolic halves $P^- \in \mathcal{F}$ with maximal side length in the collection Σ_1. Recursively, if Σ_k is chosen, collect $P^- \in \mathcal{F}$ with equal maximal size such that

$$P^+ \cap \bigcup_{Q^- \in \bigcup_{i=1}^k \Sigma_i} Q^+ = \varnothing$$

in the collection Σ_{k+1}. The collections Σ_k share no elements, and their internal overlap is bounded by 2. Since each $A \in \Sigma_k$ has equal size, the bounded overlap is inherited by the collection

$$\Sigma_k^+ := \{A^+ : A^- \in \Sigma_k\}.$$

Moreover, by construction, if $A^+ \in \Sigma_i^+$ and $B^+ \in \Sigma_j^+$ with $i \neq j$ then $A^+ \cap B^+ = \varnothing$. Hence

$$\mathcal{F}' := \bigcup_i \Sigma_i$$

is a collection such that

$$\sum_{P^- \in \mathcal{F}'} 1_{P^+} \leq 2.$$

According to (5-4) and Lemma 4.3, we get

$$w(E) \leq \sum_{B \in \mathcal{F}} w(B) \lesssim \sum_{B \in \mathcal{F}} \lambda |B| \leq \sum_{P^- \in \mathcal{F}'} \Biggl(\lambda |P^-| + \sum_{\substack{B \in \mathcal{F} \\ B^+ \cap P^+ \neq \varnothing \\ |B| < |P|}} \lambda |B| \Biggr) \lesssim \lambda \sum_{P^- \in \mathcal{F}'} |P^+|.$$

Then

$$w(E) \lesssim_\gamma \lambda \sum_{P^- \in \mathcal{F}'} |\check{P}| \lesssim \sum_{P^- \in \mathcal{F}'} \lambda \bigl| \check{P} \cap \{w > \beta\lambda\} \bigr| \leq \int_{\bigcup_{S^- \in \mathcal{F}'} \check{S} \cap \{w > \beta\lambda\}} \sum_{P^- \in \mathcal{F}'} 1_{P^+} \lesssim \lambda \bigl| \widetilde{R}_0 \cap \{w > \beta\lambda\} \bigr|. \qquad \square$$

The fact that the sets in the estimate given by the above lemma are not equal is reflected in the reverse Hölder inequality as a time lag. This phenomenon is unavoidable, and it was noticed already in the one-dimensional case; see, for instance, [Martín-Reyes 1993].

Theorem 5.2. *Let $w \in A_q^+(\gamma)$ with $\gamma \in (0,1)$. Then there exist $\delta > 0$ and a constant C independent of R such that*

$$\left(\fint_{R^-(0)} w^{\delta+1} \right)^{1/(1+\delta)} \le C \fint_{R^+(0)} w.$$

Furthermore, there exists $\epsilon > 0$ such that $w \in A_{q-\epsilon}^+(\gamma)$.

Proof. We will consider a truncated weight $w := \min\{w, m\}$ in order to make quantities bounded. At the end, the claim for general weights will follow by passing to the limit as $m \to \infty$. Without loss of generality, we may take $R^- = Q \times (0, l^p)$. Define $\widehat{R}$ and $\check{R}$ as in the previous lemma (see (5-1) and (5-2)). In addition, let $\widetilde{R}$ be the convex hull of $\widehat{R} \cup \check{R}$.

Let $E = \{w > w_{R^-}\}$. By Lemma 5.1,

$$\begin{aligned}
\int_{R^- \cap E} w^{\delta+1} &= |R^- \cap E| w_{R^-}^{\delta+1} + \delta \int_{w_{R^-}}^{\infty} \lambda^{\delta-1} w(\{R^- \cap \{w > \lambda\}\})\, \mathrm{d}\lambda \\
&\le |R^- \cap E| w_{R^-}^{\delta+1} + C\delta \int_{w_{R^-}}^{\infty} \lambda^{\delta-1} |\{R \cap \{w > \beta\lambda\}\}|\, \mathrm{d}\lambda \\
&\le |R^- \cap E| w_{R^-}^{\delta+1} + C\delta \int_{\widetilde{R} \cap E} w^{\delta+1},
\end{aligned}$$

which implies

$$\int_{R^- \cap E} w^{\delta+1} \le \frac{1}{1-\delta C} \left(|R^- \cap E| w_{R^-}^{\delta+1} + C\delta \int_{\widetilde{R} \setminus (R^- \cap E)} w^{\delta+1} \right).$$

Consequently

$$\begin{aligned}
\int_{R^-} w^{\delta+1} &\le \frac{2-\delta C}{1-\delta C} |R^-| w_{R^-}^{\delta+1} + \frac{C\delta}{1-\delta C} \int_{\widetilde{R} \setminus R^-} w^{\delta+1} \\
&= C_0 |R^-| w_{R^-}^{\delta+1} + C_1 \delta \int_{\widetilde{R} \setminus R^-} w^{\delta+1}.
\end{aligned} \tag{5-5}$$

Then we choose $l_1^p = 2^{-1} l^p$. We can cover Q by M_{np} subcubes $\{Q_i^1\}_{i=1}^{M_{np}}$ with $l(Q_i^1) = l_1$. Their overlap is bounded by M_{np}, and so is the overlap of the rectangles

$$\{R_i^{1-}\} = Q_i \times \left(l^p, \tfrac{3}{2} l^p\right)$$

that cover $\widetilde{R} \setminus R^-$ and share the dimensions of the original R^-. Hence we are in position to iterate. The rectangles $R_{ij}^{(k+1)-}$ are obtained from R_i^{k-} as R_i^{1-} were obtained from $R^- =: R_i^{0-}$, $i = 1, \ldots, M_{np}$. Thus

$$\begin{aligned}
\int_{R^-} w^{\delta+1} &\le C_0 |R^-| w_{R^-}^{\delta+1} + C_1 \delta \sum_{i=1}^{M_{np}} \int_{R_i^1} w^{\delta+1} \\
&\le \sum_{j=0}^{N} \left(C_0^{j+1} (C_1\delta)^j \sum_{i=1}^{M_{np}} |R_i^{j-}| w_{R_i^{j-}}^{\delta+1} \right) + (C_1 \delta M_{np})^N \int_{\bigcup_{i=1}^{M_{np}} \widetilde{R}_i^N \setminus R_i^{N-}} w^{\delta+1} \\
&= \mathrm{I} + \mathrm{II}.
\end{aligned}$$

For the inner sum in the first term we have

$$\sum_{i=1}^{M_{np}} |R_i^{j-}| w_{R_i^{j-}}^{\delta+1} \le \sum_{i=1}^{M_{np}} 2^{-j\delta n} l^{-\delta(n+p)} \left(\int_{R_i^{j-}} w \right)^{\delta+1} \le 2^{-j\delta n} l^{n+p} M_{np}^{\delta+1} w_R^{\delta+1}.$$

Thus

$$\mathrm{I} \le \left(\fint_R w \right)^{1+\delta} C_0 M_{np}^{\delta+1} l^{n+p} \sum_{j=0}^{N} (C_1 C_0 \delta)^j 2^{-j\delta n},$$

where the series converges as $N \to \infty$ if δ is small enough. On the other hand, if w is bounded, it is clear that $\mathrm{II} \to 0$ as $N \to \infty$. This proves the claim for bounded w, hence for truncations $\min\{w, m\}$, and the general case follows from the monotone convergence theorem as $m \to \infty$. The self-improving property of $A_q^+(\gamma)$ follows from applying the reverse Hölder inequality coming from the $A_{q'}^-(\gamma)$ condition satisfied by $w^{1-q'}$ and using Proposition 3.4. □

Remark 5.3. An easy subdivision argument shows that the reverse Hölder inequality can be obtained for any pair $R, t+R$ where $t > 0$. Namely, we can divide R into arbitrarily small, possibly overlapping, subrectangles. Then we may apply the estimate to them and sum up. This kind of procedure has been carried out explicitly in [Berkovits 2011].

Now we are ready to state the analogue of Muckenhoupt's theorem in its complete form. Once it is established, many results familiar from classical Muckenhoupt theory follow immediately.

Theorem 5.4. *Let* $\gamma_i \in (0, 1)$, $i = 1, 2, 3$. *Then the following conditions are equivalent*:

(i) $w \in A_q^+(\gamma_1)$.

(ii) *The operator* M^{γ_2+} *is of weighted weak type* (q, q) *with the weight* w.

(iii) *The operator* M^{γ_3+} *is of weighted strong type* (q, q) *with the weight* w.

Proof. Equivalence of A_q^+ and weak type follows from Theorem 4.7. Theorem 5.2 gives $A_{q-\epsilon}^+$, so (iii) follows from Marcinkiewicz interpolation and the final implication (iii) ⇒ (ii) is clear. □

6. Factorization and A_1^+ weights

In contrast with the classical case, it is not clear what is the correct definition of the parabolic Muckenhoupt class A_1^+. One option is to derive an A_1^+ condition from the inequality of weak type $(1, 1)$ for $M^{\gamma+}$, and get a condition that coincides with the formal limit of A_q^+ conditions. We propose a slightly different approach and consider the class arising from factorization of the parabolic Muckenhoupt weights and characterization of the parabolic BMO.

Definition 6.1. Let $\gamma \in [0, 1)$. A weight $w > 0$ is in $A_1^+(\gamma)$ if for almost every $z \in \mathbb{R}^{n+1}$, we have

$$M^{\gamma-} w(z) \le [w]_{A_1^+(\gamma)} w(z). \tag{6-1}$$

The class $A_1^-(\gamma)$ is defined by reversing the direction of time.

The following proposition shows that, in some cases, the A_1^+ condition implies the A_1-type condition equivalent to the inequality of weak type $(1,1)$. Moreover, if $\gamma = 0$, then the two conditions are equivalent.

Proposition 6.2. *Let $w \in A_1^+(\gamma)$ with $\gamma < 2^{1-p}$.*

(i) *For every parabolic rectangle R, it holds that*

$$\fint_{R^-(2^{p-1}\gamma)} w \lesssim_{\gamma,[w]_{A_1^+}} \inf_{z\in R^+(2^{p-1}\gamma)} w(z). \tag{6-2}$$

(ii) *For all $q > 1$, we have $w \in A_q^+$.*

Proof. Define $\delta = 2^{p-1}\gamma$. Take a parabolic rectangle R_0. We see that every $z \in R_0^+(\delta)$ is a center of a parabolic rectangle with $R^-(z,\gamma) \supset R_0^-(\delta)$ such that

$$\fint_{R^-(\delta)} w \lesssim \fint_{R^-(z,\gamma)} w \le M^{\gamma-}w(z) \lesssim w(z),$$

where the last inequality used (6-1). This proves (i). The statement (ii) follows from the fact that (6-2) is an increasing limit of $A_q^+(\gamma)$ conditions; see Definition 3.2. □

Now we will state the main result of this section, that is, the factorization theorem for the parabolic Muckenhoupt weights corresponding to the classical results, for example, in [Jones 1980; Coifman et al. 1983].

Theorem 6.3. *Let $\delta \in (0,1)$ and $\gamma \in (0, \delta 2^{1-p})$. A weight $w \in A_q^+(\delta)$ if and only if $w = uv^{1-p}$, where $u \in A_1^+(\gamma)$ and $v \in A_1^-(\gamma)$.*

Proof. Let $u \in A_1^+(\gamma)$, $v \in A_1^-(\gamma)$ and fix a parabolic rectangle R. By Proposition 6.2, for all $x \in R^+(\delta)$, we have

$$u(x)^{-1} \le \sup_{x\in R^+(\delta)} u(x)^{-1} = \Bigl(\inf_{x\in R^+(\delta)} u(x)\Bigr)^{-1} \lesssim \Bigl(\fint_{R^-(\delta)} u\Bigr)^{-1},$$

and, for all $y \in R^-(\delta)$, we have the corresponding inequality for v, that is,

$$v(y)^{-1} \le \sup_{y\in R^-(\delta)} v(y)^{-1} = \Bigl(\inf_{y\in R^-(\delta)} v(y)\Bigr)^{-1} \lesssim \Bigl(\fint_{R^+(\delta)} v\Bigr)^{-1}.$$

Hence

$$\Bigl(\fint_{R^-(\delta)} uv^{1-q}\Bigr)\Bigl(\fint_{R^+(\delta)} u^{1-q'}v\Bigr)^{p-1} \lesssim \Bigl(\fint_{R^-(\delta)} u\Bigr)\Bigl(\fint_{R^+(\delta)} v\Bigr)^{1-q}\Bigl(\fint_{R^+(\delta)} v\Bigr)^{q-1}\Bigl(\fint_{R^-(\delta)} u\Bigr)^{-1} = C,$$

which proves that $uv^{1-q} \in A_q^+(\delta)$. The finite constant C depends only on γ, δ, $[u]_{A_1^+(\gamma)}$ and $[v]_{A_1^-(\gamma)}$.

For the other direction, fix $q \ge 2$ and $w \in A_q^+$. Define an operator T as

$$Tf = \bigl(w^{-1/q}M^{\gamma-}(f^{q-1}w^{1/q})\bigr)^{1/(q-1)} + w^{1/q}M^{\gamma+}(fw^{-1/q}).$$

By boundedness of the operators

$$M^{\gamma+} : L^q(w) \to L^q(w) \quad \text{and} \quad M^{\gamma-} : L^{q'}(w^{1-p'}) \to L^{q'}(w^{1-p'}),$$

we conclude that $T : L^q \to L^q$ is bounded. Let

$$B(w) := \|T\|_{L^q \to L^q} \eqsim_{[w]_{A_q^+}} 1.$$

Take $f_0 \in L^q$ with $\|f_0\|_{L^q} = 1$. Let

$$\phi = \sum_{i=1}^{\infty} (2B(w))^{-i} T^i f_0,$$

where T^i simply means the i-th iterate of T. We define

$$u = w^{1/q} \phi^{q-1} \quad \text{and} \quad v = w^{-1/q} \phi.$$

Clearly $w = uv^{1-q}$. We claim that $u \in A_1^+$ and $v \in A_1^-$. Since $q \geq 2$, the operator T is sublinear, and we obtain

$$\begin{aligned} T(\phi) &\leq 2B(w) \sum_{i=1}^{\infty} (2B(w))^{-(i+1)} T^{i+1}(f_0) \\ &= 2B(w) \left(\phi - \frac{T(f_0)}{2B(w)} \right) \leq 2B(w)\phi. \end{aligned}$$

Noting that $\phi = (w^{-1/q} u)^{1/(q-1)} = w^{1/q} v$ and inserting the above inequality into the definition of T, we obtain

$$M^{\gamma-} u \leq (2B(w))^{q-1} u \quad \text{and} \quad M^{\gamma+} v \leq 2B(w) v.$$

This implies $u \in A_1^+$ and $v \in A_1^-$, so the proof is complete for $q \geq 2$. Once the claim is known for $q \geq 2$, the complementary case $1 < q < 2$ follows from Proposition 3.4(ii). □

Next we will characterize A_1^+ weights as small powers of maximal functions up to a multiplication by bounded functions. The following result looks very much like the classical characterization of Muckenhoupt A_1 weights. However, we emphasize that even if the maximal operator $M^{\gamma+}$ is dominated by the Hardy–Littlewood maximal operator, the assumptions of the following lemma are not restrictive at all when it comes to the measure μ. Indeed, the condition $M^{\gamma-}\mu < \infty$ almost everywhere still includes rather rough measures. For instance, their growth towards the positive time direction can be almost arbitrary, and the same property is carried over to the A_1^+ weights.

Lemma 6.4. (i) *Let μ be a locally finite nonnegative Borel measure on $\mathbb{R}^{n+1}$ such that $M^-\mu < \infty$ almost everywhere. If $\delta \in [0, 1)$, then*

$$w := (M^-\mu)^\delta \in A_1^+(0)$$

with $[w]_{A_1^+(0)}$ independent of μ.

(ii) *Let $w \in A_1^+(\gamma')$. Then there exists a μ as above, $\delta \in [0, 1)$ and K with $K, K^{-1} \in L^\infty$ such that*

$$w = K(M^{\gamma-}\mu)^\delta,$$

where $\gamma' < \gamma$.

Proof. Let $x \in \mathbb{R}^{n+1}$ and fix a parabolic rectangle R_0 centered at x. Define $\widetilde{B} = (2R_0)^-$. Decompose μ as $\mu = \mu_1 + \mu_2$, where $\mu_1 = \mu|_{\widetilde{B}}$ and $\mu_2 = \mu|_{\widetilde{B}^c}$. Kolmogorov's inequality gives

$$\fint_{R_0^-} (M^- \mu_1)^\delta \le C|R_0^-|^{-\delta} \mu_1(\widetilde{B})^\delta \le C\left(\frac{\mu(\widetilde{B})}{|\widetilde{B}|}\right)^\delta \le C M^- \mu(x)^\delta.$$

On the other hand, for any $y \in R_0^-$ and a rectangle $R(y, L) \cap (\widetilde{B})^c \ne \varnothing$, we have $L \gtrsim l(R_0)$. Moreover, $R(y, L) \subset R(x, CL)$ so that

$$M^- \mu_2(y) \lesssim M^- \mu(x)$$

and

$$\fint_{R_0^-} (M^- \mu)^\delta \le \fint_{R_0^-} (M^- \mu_2)^\delta + \fint_{R_0^-} (M^- \mu_1)^\delta \lesssim M^- \mu(x)^\delta.$$

To prove (ii), take $w \in A_1^+(\gamma')$ and a parabolic rectangle R centered at x. By the reverse Hölder property (Theorem 5.2), Remark 5.3, and inequality (6-1), we have

$$\left(\fint_{R^-(\gamma)} w^{1+\epsilon}\right)^{1/(1+\epsilon)} \lesssim w(x).$$

Define $\mu = w^{1+\epsilon}$ and $\delta = 1/(1+\epsilon)$. By the Lebesgue differentiation theorem

$$w(x) \le M^{\gamma -} \mu(x)^\delta \lesssim w(x).$$

Hence

$$K = \frac{w}{(M^{\gamma -} \mu)^\delta}$$

is bounded from above and from below, which proves the claim. □

7. A characterization of the parabolic BMO

In this section we discuss the connection between parabolic Muckenhoupt weights and the parabolic BMO. The parabolic BMO was explicitly defined by Fabes and Garofalo [1985], who gave a simplified proof of the parabolic John–Nirenberg lemma in [Moser 1964]. We consider a slightly modified definition in order to make the parabolic BMO a larger space and a more robust class; see [Saari 2016]. Our definition has essentially the same connections to PDEs as the one in [Fabes and Garofalo 1985]. Moreover, this extends the theory beyond the quadratic growth case and applies to the doubly nonlinear parabolic equations.

Definition 7.1. A function $u \in L^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$ belongs to PBMO^+ if there are constants a_R, that may depend on the parabolic rectangles R, such that

$$\sup_R \left(\fint_{R^+(\gamma)} (u - a_R)^+ + \fint_{R^-(\gamma)} (a_R - u)^+\right) < \infty \tag{7-1}$$

for some $\gamma \in (0, 1)$. If (7-1) holds with the time axis reversed, then $u \in \mathrm{PBMO}^-$.

If (7-1) holds for some $\gamma \in (0, 1)$, then it holds for all of them. Moreover, we can consider prolonged parabolic rectangles $Q \times (t - Tl^p, t + Tl^p)$ with $T > 0$ and still recover the same class of functions. These facts follow from the main result in [Saari 2016], and they can be deduced from results in [Aimar 1988] and in a special case from results in [Fabes and Garofalo 1985].

The fact that $\gamma > 0$ is crucial. For example, the John–Nirenberg inequality (Lemma 7.2) for the parabolic BMO cannot hold without a time lag. Hence a space with $\gamma = 0$ cannot be characterized through the John–Nirenberg inequality. The following lemma can be found in [Saari 2016]. See also [Fabes and Garofalo 1985; Aimar 1988].

Lemma 7.2. *Let $u \in \mathrm{PBMO}^+$ and $\gamma \in (0, 1)$. Then there are $A, B > 0$ depending only on n, γ and u such that*

$$|R^+(\gamma) \cap \{(u - a_R)^+ > \lambda\}| \le Ae^{-B\lambda}|R| \tag{7-2}$$

and

$$|R^-(\gamma) \cap \{(a_R - u)^+ > \lambda\}| \le Ae^{-B\lambda}|R|. \tag{7-3}$$

There are also more elementary properties that can be seen from Definition 7.1. Since we will need them later, they will be stated in the next proposition.

Proposition 7.3. (i) *If $u, v \in \mathrm{PBMO}^+$ and $\alpha, \beta \in (0, \infty)$, then $\alpha u + \beta v \in \mathrm{PBMO}^+$.*

(ii) *$u \in \mathrm{PBMO}^+$ if and only if $-u \in \mathrm{PBMO}^-$.*

Proof. For (i), note that

$$\bigl(u + v - (a_R^u + a_R^v)\bigr)^+ \le (u - a_R^u)^+ + (u - a_R^v)^+,$$

and an analogous estimate holds for the negative part. Hence $\alpha u + \beta v \in \mathrm{PBMO}^+$ with

$$a_R = \frac{a_R^u}{\alpha} + \frac{a_R^v}{\beta}.$$

Since

$$(u - a_R)^+ = \bigl((-u) - (-a_R)\bigr)^- \quad \text{and} \quad (u - a_R)^- = \bigl((-u) - (-a_R)\bigr)^+,$$

the second assertion is clear. □

The goal of this section is to characterize the parabolic BMO in the sense of Coifman and Rochberg [1980]. The Muckenhoupt theory developed so far gives a characterization for the parabolic Muckenhoupt weights, so what remains to do is to prove the equivalence of the parabolic BMO and the A_q^+ condition.

Lemma 7.4. *Let $q \in (1, \infty)$ and $\gamma \in (0, 1)$. Then*

$$\mathrm{PBMO}^+ = \{-\lambda \log w : w \in A_q^+(\gamma), \lambda \in (0, \infty)\}. \tag{7-4}$$

Proof. We abbreviate $R^\pm(\gamma) = R^\pm$ even if $\gamma \ne 0$. For $u \in \mathrm{PBMO}^+$, Lemma 7.2 gives $\epsilon > 0$ such that

$$\fint_{R^-} e^{-\epsilon u} = e^{-a_R\epsilon} \fint_{R^-} e^{\epsilon(a_R - u)} \le e^{-a_R\epsilon} \fint_{R^-} e^{\epsilon(a_R - u)^+} \le C_- e^{-a_R\epsilon}$$

and, for some $q < \infty$,

$$\begin{aligned}\fint_{R^+} e^{\epsilon u/(q-1)} &= e^{a_R \epsilon/(q-1)} \fint_{R^+} e^{(u-a_R)\epsilon/(q-1)} \\ &\le e^{a_R\epsilon/(q-1)} \fint_{R^+} e^{(u-a_R)^+\epsilon/(q-1)} \\ &\le C_+ e^{a_R \epsilon/(q-1)},\end{aligned}$$

so $w := e^{-u\epsilon} \in A_q^+$ and $u = -\epsilon^{-1} \cdot \log w$ as it was claimed.

To prove the other direction, take $w \in A_q^+$ with $q \le 2$. Choose

$$a_R = \log w_{R^-}.$$

Then by Jensen's inequality and the parabolic Muckenhoupt condition, we have

$$\begin{aligned}\exp \fint_{R^+} (a_R - \log w)^+ &\le \fint_{R^+} \exp(a_R - \log w)^+ \\ &\le 1 + \fint_{R^+} \exp\left(a_R - \frac{1}{1-q'} \log w^{1-q'}\right) \\ &\le 1 + \exp(a_R) \left(\fint_{R^+} w^{1-q'}\right)^{q-1} \\ &= 1 + w_{R^-} \left(\fint_{R^+} w^{1-q'}\right)^{q-1} \\ &\le 1 + C_{A_q^+}.\end{aligned}$$

On the other hand, again by Jensen's inequality,

$$\begin{aligned}\exp \fint_{R^-} (\log w - a_R)^+ &\le \fint_{R^-} \exp(\log w - a_R)^+ \\ &\le 1 + \fint_{R^-} \exp(\log w - a_R) \\ &\le 1 + \exp(-a_R) \fint_{R^-} w \\ &\le 1 + w_{R^-}^{-1} w_{R^-} \le 2.\end{aligned}$$

This implies

$$\log(2(1 + C_{A_q^+})) \ge \fint_{R^+} (-\log w - (-a_R))^+ + \fint_{R^-} (-a_R - (-\log w))^+,$$

and $u = -\log w \in \mathrm{PBMO}^+$. Applying the same argument for $A_{q'}^-$ with $q > 2$ shows that $-\log w^{1-q'} \in \mathrm{PBMO}^-$ and consequently Proposition 7.3 implies $-(q'-1)\log w \in \mathrm{PBMO}^+$. □

The following Coifman–Rochberg-type characterization [1980] for the parabolic BMO is the main result of this section. Observe, that it gives us a method to construct functions of parabolic bounded mean oscillation with prescribed singularities.

Theorem 7.5. *If $f \in \mathrm{PBMO}^+$ then there exist $\gamma \in (0,1)$, constants $\alpha, \beta > 0$, a bounded function $b \in L^\infty$ and nonnegative Borel measures μ and ν such that*

$$f = -\alpha \log M^{\gamma -}\mu + \beta \log M^{\gamma +}\nu + b.$$

Conversely, any f of the form above with $\gamma = 0$ and $M^-\mu, M^+\nu < \infty$ belongs to PBMO^+.

Proof. Take first $f \in \mathrm{PBMO}^+$. By Lemma 7.4,

$$f = -C \log w$$

with $C > 0$ and $w \in A_2^+$. By Theorem 6.3, there are $u \in A_1^+$ and $v \in A_1^-$ satisfying the corresponding maximal function estimates (6-1) such that

$$w = uv^{-1}.$$

By Lemma 6.4, there exist functions $K_u, K_v, K_u^{-1}, K_v^{-1} \in L^\infty$ and nonnegative Borel measures μ and ν such that

$$u = K_u(M^{\gamma -}\mu)^\alpha \quad \text{and} \quad v = K_v(M^{\gamma +}\nu)^\beta.$$

Hence f is of the desired form. The other direction follows from Lemma 6.4. □

8. Doubly nonlinear equation

We begin with pointing out that the theory discussed here applies not only to (1-1) but also to the PDEs

$$\frac{\partial(|u|^{p-2}u)}{\partial t} - \operatorname{div} A(x,t,u,Du) = 0, \qquad 1 < p < \infty,$$

where A satisfies the growth conditions

$$A(x,t,u,Du) \cdot Du \geq C_0 |Du|^p$$

and

$$|A(x,t,u,Du)| \leq C_1 |Du|^{p-1}.$$

See [Kinnunen and Kuusi 2007; Saari 2016] for more. For simplicity, we have chosen to focus on the prototype equation (1-1) here.

Supersolutions are weights. We say that

$$v \in L^p_{\mathrm{loc}}\big((-\infty,\infty); W^{1,p}_{\mathrm{loc}}(\mathbb{R}^{n+1})\big)$$

is a supersolution to (1-1) provided

$$\int \left(|\nabla v|^{p-2}\nabla v \cdot \nabla\phi - |v|^{p-2} v \frac{\partial \phi}{\partial t} \right) \geq 0$$

for all nonnegative $\phi \in C_0^\infty(\mathbb{R}^{n+1})$. If the reversed inequality is satisfied, we call u a subsolution. If a function is both sub- and supersolution, it is a weak solution.

The definition above allows us to use the following a priori estimate, which is Lemma 6.1 in [Kinnunen and Kuusi 2007]. Similar results can also be found in [Moser 1964; Trudinger 1968], but we emphasize that the following lemma applies to the full range $1 < p < \infty$ instead of just $p = 2$.

Lemma 8.1 [Kinnunen and Kuusi 2007]. *Suppose $v > 0$ is a supersolution of the doubly nonlinear equation in σR, where $\sigma > 1$ and R is a parabolic rectangle. Then there are constants $C = C(p, \sigma, n)$, $C' = C'(p, \sigma, n)$ and $\beta = \beta(R)$ such that*

$$\left|R^- \cap \{\log v > \lambda + \beta + C'\}\right| \leq \frac{C}{\lambda^{p-1}} |R^-|$$

and

$$\left|R^+ \cap \{\log v < -\lambda + \beta - C'\}\right| \leq \frac{C}{\lambda^{p-1}} |R^+|$$

for all $\lambda > 0$.

Remark 8.2. There is a technical assumption $v > \rho > 0$ in [Kinnunen and Kuusi 2007]. However, this assumption can be removed; see [Ivert et al. 2014]. Indeed, Lemma 2.3 of [Ivert et al. 2014] improves the inequality (3.1) of [Kinnunen and Kuusi 2007] as to make the proof of the above lemma work with general $v > 0$ in the case of (1-1) or more general parabolic quasiminimizers.

Let v be a positive supersolution and set $u = -\log v$. We apply Lemma 8.1 together with Cavalieri's principle to obtain

$$\fint_{R^+} (u - a_R)_+^b + \fint_{R^-} (a_R - u)_+^b < C(p, \sigma, \gamma, n)$$

with $b = \min\{(p-1)/2, 1\}$. A general form of the John–Nirenberg inequality from [Aimar 1988] together with its local-to-global properties from [Saari 2016] can be used to obtain

$$\fint_{R^+(\gamma)} (u - a_R)_+ + \fint_{R^-(\gamma)} (a_R - u)_+ < C(p, \sigma, \gamma, n).$$

Hence $u = -\log v$ belongs to PBMO$^+$ in the sense of Definition 7.1. The computations required in this passage are carried out in detail in Lemma 6.3 of [Saari 2016]. We collect the results into the following proposition, whose content, up to notation, is folklore by now.

Proposition 8.3. *Let $v > 0$ be a supersolution to (1-1) in $\mathbb{R}^{n+1}$. Then*

$$u = -\log v \in \mathrm{PBMO}^+.$$

In addition, $v \in \bigcap_{q>1} A_q^+$.

Remark 8.4. This gives a way to construct nontrivial examples of the parabolic Muckenhoupt weights and parabolic BMO functions.

Since $\log v \in \mathrm{PBMO}^-$, we have that some power of the positive supersolution w satisfies a local $A_2^+(\gamma)$ condition. This follows from Lemma 7.4. However, working a bit more with the PDE, it is possible to prove a weak Harnack estimate which implies the improved weight condition stated in the above proposition. This has been done in [Kinnunen and Kuusi 2007], but the refinement provided in [Ivert et al. 2014] is again needed in order to cover all positive supersolutions.

Applications. The previous proposition asserts that the definitions of parabolic weights and parabolic BMO are correct from the point of view of doubly nonlinear equations. These properties can be used to deduce two interesting results, the second one of which is new. The first one is a global integrability result for supersolutions; see Theorem 6.5 from [Saari 2016]. The second application of the parabolic theory of weights is related to singularities of supersolutions. It follows from Proposition 8.3 and Theorem 7.5. In qualitative terms, the following theorem tells quite explicitly what kind of functions the generic positive supersolutions are.

Theorem 8.5. *Let $v > 0$ be a supersolution to (1-1) in $\mathbb{R}^{n+1}$. Then there are positive Borel measures ν and μ with*

$$M^{\gamma-}\nu < \infty \quad \textit{and} \quad M^{\gamma+}\mu < \infty,$$

numbers $\alpha, \beta > 0$, and a positive function b with $b, b^{-1} \in L^\infty(\mathbb{R}^{n+1})$ so that

$$v = b\frac{(M^{\gamma-}\nu)^\alpha}{(M^{\gamma+}\mu)^\beta}.$$

References

[Aimar 1988] H. Aimar, "Elliptic and parabolic BMO and Harnack's inequality", *Trans. Amer. Math. Soc.* **306**:1 (1988), 265–276. MR Zbl

[Aimar and Crescimbeni 1998] H. Aimar and R. Crescimbeni, "On one-sided BMO and Lipschitz functions", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **27**:3–4 (1998), 437–456. MR Zbl

[Aimar et al. 1997] H. Aimar, L. Forzani, and F. J. Martín-Reyes, "On weighted inequalities for singular integrals", *Proc. Amer. Math. Soc.* **125**:7 (1997), 2057–2064. MR Zbl

[Berkovits 2011] L. Berkovits, "Parabolic Muckenhoupt weights in the Euclidean space", *J. Math. Anal. Appl.* **379**:2 (2011), 524–537. MR Zbl

[Coifman and Rochberg 1980] R. R. Coifman and R. Rochberg, "Another characterization of BMO", *Proc. Amer. Math. Soc.* **79**:2 (1980), 249–254. MR Zbl

[Coifman et al. 1983] R. Coifman, P. W. Jones, and J. L. Rubio de Francia, "Constructive decomposition of BMO functions and factorization of A_p weights", *Proc. Amer. Math. Soc.* **87**:4 (1983), 675–676. MR Zbl

[Cruz-Uribe et al. 1995] D. Cruz-Uribe, C. J. Neugebauer, and V. Olesen, "The one-sided minimal operator and the one-sided reverse Hölder inequality", *Studia Math.* **116**:3 (1995), 255–270. MR Zbl

[Fabes and Garofalo 1985] E. B. Fabes and N. Garofalo, "Parabolic B.M.O. and Harnack's inequality", *Proc. Amer. Math. Soc.* **95**:1 (1985), 63–69. MR Zbl

[Forzani et al. 2011] L. Forzani, F. J. Martín-Reyes, and S. Ombrosi, "Weighted inequalities for the two-dimensional one-sided Hardy–Littlewood maximal function", *Trans. Amer. Math. Soc.* **363**:4 (2011), 1699–1719. MR Zbl

[García-Cuerva and Rubio de Francia 1985] J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies **116**, North-Holland, Amsterdam, 1985. MR Zbl

[Gianazza and Vespri 2006] U. Gianazza and V. Vespri, "A Harnack inequality for solutions of doubly nonlinear parabolic equations", *J. Appl. Funct. Anal.* **1**:3 (2006), 271–284. MR Zbl

[Ivert et al. 2014] P.-A. Ivert, N. Marola, and M. Masson, "Energy estimates for variational minimizers of a parabolic doubly nonlinear equation on metric measure spaces", *Ann. Acad. Sci. Fenn. Math.* **39**:2 (2014), 711–719. MR Zbl

[Jones 1980] P. W. Jones, "Factorization of A_p weights", *Ann. of Math.* (2) **111**:3 (1980), 511–530. MR Zbl

[Kinnunen and Kuusi 2007] J. Kinnunen and T. Kuusi, "Local behaviour of solutions to doubly nonlinear parabolic equations", *Math. Ann.* **337**:3 (2007), 705–728. MR Zbl

[Kinnunen and Saari 2016] J. Kinnunen and O. Saari, "On weights satisfying parabolic Muckenhoupt conditions", *Nonlinear Anal.* **131** (2016), 289–299. MR Zbl

[Kuusi et al. 2012] T. Kuusi, J. Siljander, and J. M. Urbano, "Local Hölder continuity for doubly nonlinear parabolic equations", *Indiana Univ. Math. J.* **61**:1 (2012), 399–430. MR Zbl

[Lerner and Ombrosi 2010] A. K. Lerner and S. Ombrosi, "A boundedness criterion for general maximal operators", *Publ. Mat.* **54**:1 (2010), 53–71. MR Zbl

[Martín-Reyes 1993] F. J. Martín-Reyes, "New proofs of weighted inequalities for the one-sided Hardy-Littlewood maximal functions", *Proc. Amer. Math. Soc.* **117**:3 (1993), 691–698. MR Zbl

[Martín-Reyes and de la Torre 1993] F. J. Martín-Reyes and A. de la Torre, "Two weight norm inequalities for fractional one-sided maximal operators", *Proc. Amer. Math. Soc.* **117**:2 (1993), 483–489. MR Zbl

[Martín-Reyes and de la Torre 1994] F. J. Martín-Reyes and A. de la Torre, "One-sided BMO spaces", *J. London Math. Soc.* (2) **49**:3 (1994), 529–542. MR Zbl

[Martín-Reyes et al. 1990] F. J. Martín-Reyes, P. Ortega Salvador, and A. de la Torre, "Weighted inequalities for one-sided maximal functions", *Trans. Amer. Math. Soc.* **319**:2 (1990), 517–534. MR Zbl

[Martín-Reyes et al. 1993] F. J. Martín-Reyes, L. Pick, and A. de la Torre, "A_∞^+ condition", *Canad. J. Math.* **45**:6 (1993), 1231–1244. MR Zbl

[Moser 1961] J. Moser, "On Harnack's theorem for elliptic differential equations", *Comm. Pure Appl. Math.* **14**:3 (1961), 577–591. MR Zbl

[Moser 1964] J. Moser, "A Harnack inequality for parabolic differential equations", *Comm. Pure Appl. Math.* **17**:1 (1964), 101–134. MR Zbl

[Moser 1967] J. Moser, "Correction to: 'A Harnack inequality for parabolic differential equations' ", *Comm. Pure Appl. Math.* **20**:1 (1967), 231–236. MR Zbl

[Ombrosi 2005] S. Ombrosi, "Weak weighted inequalities for a dyadic one-sided maximal function in $\mathbb{R}^n$", *Proc. Amer. Math. Soc.* **133**:6 (2005), 1769–1775. MR Zbl

[Saari 2016] O. Saari, "Parabolic BMO and global integrability of supersolutions to doubly nonlinear parabolic equations", *Rev. Mat. Iberoam.* **32**:3 (2016), 1001–1018. MR

[Sawyer 1986] E. Sawyer, "Weighted inequalities for the one-sided Hardy–Littlewood maximal functions", *Trans. Amer. Math. Soc.* **297**:1 (1986), 53–61. MR Zbl

[Trudinger 1968] N. S. Trudinger, "Pointwise estimates and quasilinear parabolic equations", *Comm. Pure Appl. Math.* **21**:3 (1968), 205–226. MR Zbl

[Vespri 1992] V. Vespri, "On the local behaviour of solutions of a certain class of doubly nonlinear parabolic equations", *Manuscripta Math.* **75**:1 (1992), 65–80. MR Zbl

Received 15 Feb 2016. Revised 20 Jun 2016. Accepted 28 Aug 2016.

JUHA KINNUNEN: juha.k.kinnunen@aalto.fi
Department of Mathematics, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland

OLLI SAARI: olli.saari@aalto.fi
Department of Mathematics, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland

ANALYSIS AND PDE
Vol. 9, No. 7, 2016
dx.doi.org/10.2140/apde.2016.9.1737

A DOUBLE WELL POTENTIAL SYSTEM

JAEYOUNG BYEON, PIERO MONTECCHIARI AND PAUL H. RABINOWITZ

A semilinear elliptic system of PDEs with a nonlinear term of double well potential type is studied in a cylindrical domain. The existence of solutions heteroclinic to the bottom of the wells as minima of the associated functional is established. Further applications are given, including the existence of multitransition solutions as local minima of the functional.

1. Introduction

In this paper, the system of partial differential equations

$$-\Delta u + V_u(x,u) = 0, \quad x \in \Omega, \tag{PDE}$$

where $\Omega \subset \mathbb{R}^n$ and $u : \Omega \to \mathbb{R}^m$, will be studied. The set Ω is a cylindrical domain in $\mathbb{R}^n$ given by $\Omega = \mathbb{R} \times \mathcal{D}$, where $\mathcal{D}$ is a bounded open set in $\mathbb{R}^{n-1}$ with $\partial\mathcal{D} \in C^1$. On $\partial\Omega$, we require

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial\Omega = \mathbb{R} \times \partial\mathcal{D}, \tag{BC}$$

where ν is the outward-pointing unit normal to $\partial\mathcal{D}$. Later, Ω will be allowed to be a more general cylindrical domain which depends 1-periodically on x_1.

As to the function V, to begin assume:

(V_1) $V \in C^1(\overline{\Omega} \times \mathbb{R}^m, \mathbb{R})$ and $V(x_1+1, x_2, \dots, x_n, u) = V(x,u)$, i.e., V is 1-periodic in x_1.

(V_2) There are points $a^- \neq a^+$ such that $V(x, a^\pm) = 0$ for all $x \in \Omega$ and $V(x,u) > 0$ otherwise.

(V_3) There is a constant $\underline{V} > 0$ such that $\liminf_{|u|\to\infty} V(x,u) \geq \underline{V}$ uniformly in $x \in \Omega$.

(V_4) For $n \geq 2$, there exist constants $c_1, C_1 > 0$ such that

$$|V_u(x,u)| \leq c_1 + C_1 |u|^p,$$

where $1 < p < (n+2)/(n-2)$ for $n \geq 3$ and there is no upper growth restriction on p if $n = 2$.

An example of V satisfying (V_1)–(V_4) is $V(x,u) = |u-a^-|^q |u-a^+|^q$ for $q \in (1, n/(n-2))$ and $a^+ \neq a^- \in \mathbb{R}^n$. By (V_2), V is a double well potential and we are interested in the existence of classical solutions of (PDE) that are heteroclinic in x_1 from a^- to a^+. If $n = 1$ and m is arbitrary, (PDE) reduces to a second-order Hamiltonian system of ordinary differential equations and conditions (V_1)–(V_3) suffice for such an existence result. For arbitrary n and m, conditions (V_1)–(V_4) enable us to show (PDE) possesses

MSC2010: primary 35J47; secondary 35J57, 58E30.
Keywords: elliptic system, double well potential, heteroclinic, minimization.

a weak solution. As is usual, we say that $U \in W^{1,2}_{\text{loc}}(\Omega, \mathbb{R}^m)$ is a weak solution of (PDE) and (BC) when for any $\varphi \in W^{1,2}_{\text{loc}}(\Omega)$ having compact support in $\overline{\Omega}$,

$$\int_\Omega \big(\nabla U \cdot \nabla \varphi + V_u(x, U)\varphi\big)\, dx = 0. \tag{1.1}$$

The weak solution is a classical solution when $n = 1$. However, when $n > 1$, more regularity of V and $\partial\Omega$ is required to get a classical solution.

In Section 2, the functional

$$J(u) = \int_\Omega \big(\tfrac{1}{2}|\nabla u|^2 + V(x, u)\big)\, dx \equiv \int_\Omega L(u)\, dx, \tag{1.2}$$

whose formal Euler–Lagrange equation is (PDE), will be studied. Minimization arguments will be used to show that J has a critical point. In particular when $n = 1$, our first existence result for (PDE) is:

Theorem. *If V satisfies (V_1)–(V_3), then (PDE) possesses a solution heteroclinic from a^- to a^+.*

For $n > 1$, existence of solutions requires more work. In Section 3, a regularity theorem will be stated as a consequence of which we have:

Theorem. *If (V_1)–(V_4) hold, $V \in C^2$, and $\partial\Omega \in C^3$, there is a classical solution U of (PDE) and (BC) such that $\lim_{x_1 \to \pm\infty} U(x_1, \dots, x_n) = a^\pm$ uniformly for $(x_2, \dots, x_n) \in \mathcal{D}$.*

In Section 2, we find the solution by a minimization argument in an appropriate class of functions, Γ, and a detailed proof of the regularity will be given in Section 6.

Four generalizations of our existence results will be given in Section 4. The first, Theorem 4.1, essentially replaces conditions (V_2)–(V_4) by the requirement that V possesses a convex basin containing $a^\pm$ — see hypothesis (V_5) — to get an $L^\infty(\Omega, \mathbb{R}^m)$ bound for the minimizer of Section 2 and this bound leads in turn to the existence of a solution of (PDE) and (BC), which is heteroclinic in x_1 from a^- to a^+. This result gives the existence of the heteroclinic solution of (PDE) and (BC) for the example of $V(x, u) = |u - a^-|^q\, |u - a^+|^q$ mentioned earlier, but now for any $q > 1$.

The second replaces Ω by a more general domain varying periodically in x_1. The third considers a PDE perturbation of the case of $n = 1$. Finally for the fourth, the case of multiwell potentials will be discussed briefly.

In Section 5, it will be shown that variational gluing arguments in the spirit of [Montecchiari and Rabinowitz 2016] together with the basic heteroclinic minimizers of (1.2) as well as their counterparts when the roles of a^- and a^+ are reversed can be used to construct infinitely many multitransition homoclinic and heteroclinic solutions of (PDE). These solutions are local minima of (1.2) that as a function of x_1 transit back and forth between the two global minima, $a^\pm$, of V. Obtaining these solutions requires a mild nondegeneracy condition — see Proposition 5.10(ii) — on the set of heteroclinic minimizers of (1.2). Stated very informally, we will show:

Theorem. *If (V_1)–(V_4) are satisfied and a mild nondegeneracy condition on the heteroclinics in x_1 from $a^\pm$ to $a^\mp$ holds, then for each $k \in \mathbb{N} \cup \{\infty\}$, $k \geq 2$, there exist infinitely many k-transition solutions of (PDE) and (BC).*

As has been noted above, our existence results rely on minimization arguments from the calculus of variations. These arguments are elementary, but often delicately exploit (V_1)–(V_3). The regularity arguments where (V_4) and further smoothness of V and $\partial\mathcal{D}$ play their roles are of necessity rather technical.

To conclude this section, some of the literature on (PDE) and (BC) will be discussed. The earliest work we know of is for the case of $n = 1$, where of course $\mathcal{D} = \varnothing$ and (BC) is vacuous. Thus (PDE) becomes a second-order Hamiltonian system. Using geometrical arguments, the existence of heteroclinic solutions for $V = V(u)$ was studied for a more general class of potentials by Bolotin [1978]. See also the survey article by Kozlov [1985]. Subsequently other work was done, also for the autonomous case where $V \in C^3$ has nondegenerate minima and $m = 2$, by Sternberg [1991]. Rabinowitz [1993] treated $V = V(t,u)$ where $V \in C^2$ is periodic in t. He used minimization arguments from [Rabinowitz 1989], where $V = V(u)$ and is periodic in the components of u. Alikakos and Fusco [2008] also treated the autonomous case for a C^2 potential under a milder condition than the nondegeneracy of the minima.

For $m = 1$ and $n > 1$, where (BC) plays a role, minimization arguments similar to the ones used in [Rabinowitz 1994] were used in [Rabinowitz 2002] and generalized in [Rabinowitz 2004] to obtain heteroclinics in x_1. The case of $m, n > 1$ for (PDE) has been studied extensively in several papers by Alikakos and his collaborators, especially Fusco, mainly in the autonomous setting when V possesses symmetries and one seeks solutions possessing these symmetries [Alikakos 2012; 2013; Alikakos and Fusco 2008; 2009; 2011; 2015; Alikakos and Smyrnelis 2012]. In fact it was their recent paper, [Alikakos and Fusco 2015], together with our work [Montecchiari and Rabinowitz 2016] on systems like (PDE) but with potentials $V(x,u)$ that are periodic in the components of u that led to this paper. Alikakos and Fusco [2015] studied (PDE) and (BC), with Ω periodic in x_1, essentially under the C^2 version of (V_1), and stronger forms of (V_2) and (V_5). See the survey paper [Alikakos 2013] for many more references to and related questions for (PDE). For some other related results on entire solutions of systems of Allen–Cahn-type, see [Alessio 2013; Alessio and Montecchiari 2014; Bronsard and Reitich 1993; Gui and Schatzman 2008; Schatzman 2002].

2. The existence of a minimizer of J

In this section, as a first step towards finding heteroclinic solutions of (PDE), a minimizer will be obtained for the functional J, defined in (1.2). The functional will be studied on the Hilbert space

$$E \equiv \left\{ u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m) \;\middle|\; \|u\|^2 \equiv \int_\Omega |\nabla u|^2\,dx + \int_{T_0} |u|^2\,dx < \infty \right\},$$

where for $i \in \mathbb{Z}$, we set $T_i = (i, i+1) \times \mathcal{D}$. As the class of admissible functions, take

$$\Gamma = \left\{ u \in E \;\middle|\; \|u - a^\pm\|_{L^2(T_i,\mathbb{R}^m)} \to 0,\ i \to \pm\infty \right\}.$$

Define

$$c = \inf_{u \in \Gamma} J(u). \tag{2.1}$$

It is readily seen that $\Gamma \neq \varnothing$ and $0 \le c < \infty$. Then we have:

Theorem 2.2. *Suppose $\Omega = \mathbb{R} \times \mathcal{D}$ with $\mathcal{D} \subset \mathbb{R}^{n-1}$ a bounded domain and $\partial\mathcal{D} \in C^1$. If V satisfies (V_1)–(V_3), then there exists a $U \in \Gamma$ such that $J(U) = c > 0$. Moreover, there is a constant $M > 0$ such that for any minimizer U of (2.1),*

$$\sup_{i\in\mathbb{Z}} \|U\|_{W^{1,2}(T_i,\mathbb{R}^m)} \le M.$$

Before proving Theorem 2.2, the following result is useful.

Proposition 2.3. *Let V satisfy (V_1)–(V_3), $\partial\mathcal{D} \in C^1$, and $v \in E$ with $J(v) < \infty$. Then there are $\varphi^\pm \in \{a^-, a^+\}$ such that $\|v - \varphi^\pm\|_{L^2(T_i,\mathbb{R}^m)} \to 0$ as $i \to \pm\infty$.*

Proof. Their proofs being the same, we will prove the result for φ^+. For $x \in \bar{\Omega}$, set $x = (x_1, \hat{x})$ with $x_1 \in \mathbb{R}$ and $\hat{x} \in \bar{\mathcal{D}}$. For $x \in T_0$ and $k \in \mathbb{Z}$, set $v_k(x) = v(x_1 + k, \hat{x})$ so $v_k \in W^{1,2}(T_0, \mathbb{R}^m)$. Then (V_2) and $J(v) < \infty$ imply

$$\lim_{k\to\infty} \|\nabla v_k\|_{L^2(T_0,\mathbb{R}^m)} = \lim_{k\to\infty} \int_{T_0} V(x, v_k)\, dx = 0. \tag{2.4}$$

Consequently $\{\|\nabla v_k\|_{L^2(T_0,\mathbb{R}^m)}\}$ is bounded independently of $k \in \mathbb{Z}$. By the Poincaré inequality and the fact that $\mathcal{D} \in C^1$, there is a constant b so that

$$\|v_k - [v_k]\|_{L^2(T_0,\mathbb{R}^m)} \le b\|\nabla v_k\|_{L^2(T_0,\mathbb{R}^m)}, \tag{2.5}$$

where $[v_k]$ denotes the mean value of v_k on T_0. We claim that $\{v_k\}$ is bounded in $L^2(T_0, \mathbb{R}^m)$. If not, (2.5) shows $\{[v_k]\}$ is unbounded in $\mathbb{R}$. For a set $S \subset \mathbb{R}^n$, let $|S|$ denote the measure of S. By (2.5) again, the sequence $\{v_k - [v_k]\}$ converges to 0 in measure. Therefore for any $\delta > 0$, the measure of the set in T_0 where $|v_k - [v_k]| \le \delta$ is at least $\frac{1}{2}|T_0|$ for large k. Thus by (V_3), for large $k > 0$,

$$\int_{T_0} V(x, v_k)\, dx \ge \tfrac{1}{4}|T_0|\underline{V}. \tag{2.6}$$

But (2.6) is contrary to (2.4), so $\{v_k\}$ is bounded in $W^{1,2}(T_0, \mathbb{R}^m)$. Hence there is a $v^* \in W^{1,2}(T_0, \mathbb{R}^m)$ such that along a subsequence, v_k converges to v^* weakly in $W^{1,2}(T_0, \mathbb{R}^m)$ and strongly in $L^2(T_0, \mathbb{R}^m)$. By (2.5), $v^* = [v^*]$; i.e., v^* is a constant vector. Again $v_k \to v^*$ in measure along the subsequence as $k \to \infty$, so for any small δ, we have $|v_k - v^*| \le \delta$ on a subset of T_0 of measure $\ge \frac{1}{2}|T_0|$. Therefore

$$\int_{T_0} V(x, v_k)\, dx \ge \tfrac{1}{2}|T_0| \min_{z\in B_\delta(v^*)} V(x, z), \tag{2.7}$$

where $B_\delta(v)$ denotes an open ball of center v and radius δ in $\mathbb{R}^m$. If $v^* \notin \{a^-, a^+\}$, and δ is small enough, the right-hand side of (2.7) is positive. But as $k \to \infty$, the left-hand side of (2.7) goes to 0. Therefore $v^* \in \{a^-, a^+\}$. For notational convenience, suppose $v^* = a^-$.

It remains to show that the entire sequence $\{v_k\}$, rather than a subsequence, converges to a^-, i.e.,

$$\lim_{k\to\infty} v_k = a^-. \tag{2.8}$$

Otherwise, there exist subsequences $\{i_p\}, \{k_q\} \subset \mathbb{N}$, with $i_p \to \infty$ as $p \to \infty$, $k_q \to \infty$ as $q \to \infty$, $i_p < k_p < i_{p+1}$ for all p and such that

$$\lim_{p\to\infty} v_{i_p} = a^-, \quad \lim_{q\to\infty} v_{k_q} = a^+.$$

Set $\varepsilon = \frac{1}{3}|a^+ - a^-|\sqrt{|T_0|}$. Therefore there is a $\underline{p}$ such that for $p \geq \underline{p}$,

$$\|v_{i_p} - a^-\|_{L^2(T_0,\mathbb{R}^m)} < \varepsilon, \quad \|v_{k_p} - a^+\|_{L^2(T_0,\mathbb{R}^m)} < \varepsilon.$$

We claim that for $\underline{p}$ possibly still larger and all $p \geq \underline{p}$, there is an $s_p \in \mathbb{N}$ such that $i_p < s_p < k_p$ and

$$\|v_{s_p} - a^-\|_{L^2(T_0,\mathbb{R}^m)} \geq \varepsilon, \quad \|v_{s_p} - a^+\|_{L^2(T_0,\mathbb{R}^m)} \geq \varepsilon.$$

If not, for every t between i_p and k_p,

$$\|v_t - a^-\|_{L^2(T_0,\mathbb{R}^m)} < \varepsilon \quad \text{or} \quad \|v_t - a^+\|_{L^2(T_0,\mathbb{R}^m)} < \varepsilon.$$

Replace i_p and k_p by the smallest adjacent pair $j, j+1 \in \mathbb{N} \cap [i_p, k_p]$ such that

$$\|v_j - a^-\|_{L^2(T_0,\mathbb{R}^m)} < \varepsilon, \quad \|v_{j+1} - a^+\|_{L^2(T_0,\mathbb{R}^m)} < \varepsilon. \tag{2.9}$$

Next observe that

$$\begin{aligned} |v_{j+1}(x) - v_j(x)| &= \left|\int_0^1 v_{x_1}(x_1 + j + s, \hat{x})\,ds\right| \leq \left|\int_0^2 v_{x_1}(j + s, \hat{x})\,ds\right| \\ &\leq \sqrt{2}\left(\int_0^2 v_{x_1}(j + s, \hat{x})^2\,ds\right)^{1/2}. \end{aligned}$$

Therefore

$$\|v_{j+1} - v_j\|_{L^2(T_0,\mathbb{R}^m)} \leq \sqrt{2}\|v_{x_1}\|_{L^2(T_0 \cup T_1,\mathbb{R}^m)}. \tag{2.10}$$

By (2.4), for $\underline{p}$ still larger, we can assume the right-hand side of (2.10) is $\leq \varepsilon$. On the other hand, by (2.9),

$$\begin{aligned} \|v_{j+1} - v_j\|_{L^2(T_0,\mathbb{R}^m)} &> \|a^+ - a^-\|_{L^2(T_0,\mathbb{R}^m)} - 2\varepsilon \\ &= |a^+ - a^-|\sqrt{|T_0|} - 2\varepsilon. \end{aligned} \tag{2.11}$$

Since $3\varepsilon = |a^+ - a^-|\sqrt{|T_0|}$, (2.11) is not possible and therefore there exists a sequence $\{s_p\}$ as claimed. But then

$$J(v) \geq \sum_{\underline{p}}^{\infty} \int_{T_{s_p}} L(v)\,dx = \infty$$

and we have a contradiction, establishing Proposition 2.3. □

To prove Theorem 2.2, let $\{u_k\}$ be a minimizing sequence for (2.1). Thus there is a constant M_1 such that for all $k \in \mathbb{N}$,

$$J(u_k) \leq M_1. \tag{2.12}$$

Let $\rho \in \left(0, \frac{1}{4}|a^+ - a^-|\sqrt{|T_0|}\right)$. Noting that Γ and J are invariant under a unit phase shift in the x_1-direction, it can be assumed that

$$\|u_k - a^-\|_{L^2(T_i,\mathbb{R}^m)} \le \rho \quad \text{for all } i \le 0 \qquad \text{and} \qquad \|u_k - a^-\|_{L^2(T_1,\mathbb{R}^m)} > \rho \tag{2.13}$$

for all $k \in \mathbb{N}$. Now a few observations about any $u \in \Gamma$ are required. Set

$$\begin{aligned}
\Gamma_1 &\equiv \left\{u \in \Gamma \;\middle|\; \min\{\|u - a^-\|_{L^2(T_0,\mathbb{R}^m)}, \|u - a^+\|_{L^2(T_0,\mathbb{R}^m)}\} \ge \rho\right\},\\
\Gamma_2 &\equiv \left\{u \in \Gamma \;\middle|\; \max\{\|u - a^-\|_{L^2(T_0,\mathbb{R}^m)}, \|u - a^+\|_{L^2(T_1,\mathbb{R}^m)}\} \le \rho\right\},\\
\Gamma_3 &\equiv \left\{u \in \Gamma \;\middle|\; \max\{\|u - a^+\|_{L^2(T_0,\mathbb{R}^m)}, \|u - a^-\|_{L^2(T_1,\mathbb{R}^m)}\} \le \rho\right\}.
\end{aligned}$$

Proposition 2.14. (1) *There is a constant* $\kappa_1 > 0$ *such that*

$$d_1 \equiv \inf_{u \in \Gamma_1} \int_{T_0} L(u)\,dx \ge \kappa_1.$$

(2) *There is a constant* $\kappa > 0$ *such that*

$$d \equiv \inf_{u \in \Gamma_2 \cup \Gamma_3} \int_{T_0 \cup T_1} L(u)\,dx \ge \kappa.$$

Proof. If $\kappa_1 = 0$, there is a sequence $\{v_k\}$ in Γ_1 such that

$$\int_{T_0} L(v_k)\,dx \to 0 \quad \text{as } k \to \infty. \tag{2.15}$$

Arguing as in the proof of Proposition 2.3, we again conclude (2.4)–(2.5) hold and either (i) both $\{v_k\}$ is bounded in $L^2(T_0, \mathbb{R}^m)$ and $\{[v_k]\}$ is bounded in $\mathbb{R}^m$ or (ii) both sequences are unbounded. If (i) occurs, as in the proof of Proposition 2.3, $\{v_k\}$ converges along a subsequence in $L^2(T_0, \mathbb{R}^m)$ to a constant function $v^* = [v^*]$ and for any small δ, for large k, we have $|v_k - v^*| \le \delta$ on a subset of T_0 of measure $\ge \frac{1}{2}|T_0|$. Thus (2.7) again holds. Noting that

$$|v - a^\pm|\,|T_0|^{1/2} = \|v - a^\pm\|_{L^2(T_0,\mathbb{R}^m)} \ge \rho$$

for δ small compared to ρ, (V_2) and (V_3) show the right-hand side of (2.7) is positive independently of v. This contradicts (2.4) and this case is proved.

Next suppose that (ii) occurs. Then the argument centered around (2.6) again applies and this case is impossible. Thus (1) of the proposition is proved.

For the proof of (2), we use a similar argument. Assume to the contrary that $\kappa = 0$. Then there is a sequence $\{v_k\}$ in $\Gamma_2 \cup \Gamma_3$ such that

$$\int_{T_0 \cup T_1} L(v_k)\,dx \to 0 \quad \text{as } k \to \infty. \tag{2.16}$$

Taking a subsequence if necessary, it can be assumed that $\{v_k\} \subset \Gamma_2$ or $\{v_k\} \subset \Gamma_3$. Suppose $\{v_k\} \subset \Gamma_2$. Arguing as in the proof of (1), by (2.16),

$$\lim_{k \to \infty} \|\nabla v_k\|_{L^2(T_0 \cup T_1,\mathbb{R}^m)} = \lim_{k \to \infty} \int_{T_0 \cup T_1} V(x, v_k)\,dx = 0. \tag{2.17}$$

Again by the Poincaré inequality, there is a constant b_1 so that

$$\|v_k - [v_k]_1\|_{L^2(T_0\cup T_1,\mathbb{R}^m)} \le b_1\|\nabla v_k\|_{L^2(T_0\cup T_1,\mathbb{R}^m)} \to 0 \quad \text{as } k\to\infty, \tag{2.18}$$

where $[v_k]_1$ denotes the mean value of v_k on $T_0\cup T_1$. It follows as in case (ii) of (1) that $\{[v_k]_1\}$ is bounded. Taking a subsequence again if necessary, it can be assumed that $\lim_{k\to\infty}[v_k]_1 = a\in\mathbb{R}^m$. Then we see that

$$\begin{aligned} 2\rho &\ge \lim_{k\to\infty}\|v_k - a^-\|_{L^2(T_0,\mathbb{R}^m)} + \lim_{k\to\infty}\|v_k - a^+\|_{L^2(T_1,\mathbb{R}^m)} \\ &= \|a^- - a\|_{L^2(T_0,\mathbb{R}^m)} + \|a^+ - a\|_{L^2(T_1,\mathbb{R}^m)} \\ &= |a^- - a|\sqrt{|T_0|} + |a^+ - a|\sqrt{|T_1|} \\ &\ge |a^- - a^+|\sqrt{|T_0|}, \end{aligned}$$

which contradicts that $\rho < \frac{1}{4}|a^+ - a^-|\sqrt{|T_0|}$. In the remaining case where $\{v_k\}\subset\Gamma_3$, a contradiction follows by the same argument. This proves (2). □

Remark 2.19. Observe that for any $u\in\Gamma$ satisfying (2.12) and any $i\in\mathbb{N}$, either

$$\min\left\{\|u-a^-\|_{L^2(T_i,\mathbb{R}^m)}, \|u-a^+\|_{L^2(T_i,\mathbb{R}^m)}\right\} > \rho \tag{2.20}$$

or

$$\min\left\{\|u-a^-\|_{L^2(T_i,\mathbb{R}^m)}, \|u-a^+\|_{L^2(T_i,\mathbb{R}^m)}\right\} \le \rho. \tag{2.21}$$

Let $l(u)$ be the number of values of i for which (2.20) holds. By (2.12) and Proposition 2.14(1),

$$l(u)\kappa_1 \le M_1. \tag{2.22}$$

Thus (2.22) shows $l(u)$ is bounded from above independently of u; i.e., (2.20) holds for at most M_1/κ_1 values of i. Next let $l^*(u)$ denote the number of values of i for which

$$\max\left\{\|u-a^-\|_{L^2(T_i,\mathbb{R}^m)}, \|u-a^+\|_{L^2(T_{i+1},\mathbb{R}^m)}\right\} \le \rho$$

or

$$\max\left\{\|u-a^+\|_{L^2(T_i,\mathbb{R}^m)}, \|u-a^-\|_{L^2(T_{i+1},\mathbb{R}^m)}\right\} \le \rho.$$

Hence $l^*(u)$ represents the number of transitions of u from being "near" $a^\pm$ on T_i to being "near" $a^\mp$ on T_{i+1}. By Proposition 2.14(2),

$$l^*(u)\kappa \le M_1. \tag{2.23}$$

This means that the number of pairs of consecutive intervals on which u shifts from being near one of a^- or a^+ to the other is uniformly bounded for $u\in\Gamma$ satisfying (2.12).

Bounds for the functions u_k in the minimizing sequence are provided by the next result.

Proposition 2.24. *If V satisfies (V_1)–(V_3), then there is a constant M such that $\|u_k\|_{W^{1,2}(T_i,\mathbb{R}^m)} \le M$ for all $k\in\mathbb{N}$ and $i\in\mathbb{Z}$.*

Proof. We argue as in an analogous situation in [Montecchiari and Rabinowitz 2016]. It can be assumed that u_k satisfies the normalization (2.13). By (2.12),

$$J(u_k) = \sum_{i\in\mathbb{Z}} \int_{T_i} L(u_k)\,dx \le M_1. \tag{2.25}$$

Therefore (2.25) and (2.13) immediately yield the desired bound for some value of M, say M_2, for $i \le 0$. For any $i \in \mathbb{Z}$ for which $\|u_k - a^{\pm}\|_{L^2(T_i,\mathbb{R}^m)} \le \rho$, we get the $\|u_k\|_{W^{1,2}(T_i)}$ bound exactly as was done for $i \le 0$ and obtain the same upper bound, M_2. By Remark 2.19, there are at most l values of i that remain. They lie in

$$A_k = \{i \in \mathbb{N} \mid \|u_k - a^-\|_{L^2(T_i,\mathbb{R}^m)} \ge \rho,\ \|u_k - a^+\|_{L^2(T_i,\mathbb{R}^m)} \ge \rho\}.$$

Note that $A_k \subset \mathbb{N}$. Let $i \notin A_k$ and $i+1 \in A_k$. Let $\hat{x} = (x_2,\dots,x_n)$. For $(s,\hat{x}) \in T_i$ and $(\sigma,\hat{x}) \in T_{i+1}$,

$$u_k(\sigma,\hat{x}) = u_k(s,\hat{x}) + \int_s^{\sigma} \frac{\partial u_k(t,\hat{x})}{\partial t}\,dt,$$

so

$$|u_k(\sigma,\hat{x})|^2 \le 2|u_k(s,\hat{x})|^2 + 4\int_i^{i+2} |\nabla u_k(t,\hat{x})|^2\,dt. \tag{2.26}$$

Integrating (2.26) over $s,\sigma,\hat{x}$ gives

$$\|u_k\|^2_{L^2(T_{i+1},\mathbb{R}^m)} \le 2\|u_k\|^2_{L^2(T_i,\mathbb{R}^m)} + 4\|\nabla u_k\|^2_{L^2(T_i\cup T_{i+1},\mathbb{R}^m)}. \tag{2.27}$$

Therefore by (2.25) and the above remarks,

$$\|u_k\|^2_{W^{1,2}(T_{i+1},\mathbb{R}^m)} \le 2M_2^2 + 8M_1 \equiv M_3. \tag{2.28}$$

Then, if $i+2 \in A_k$, the argument of (2.27)–(2.28) can be repeated. Since the number of elements of A_k is bounded by $l \in \mathbb{N}$, the process stops in at most l steps, giving the desired bound with $M = M(l)$. □

Completion of the proof of Theorem 2.2. It is convenient to introduce some notions. A set $I \subset \mathbb{Z}$ will be called connected if for any $i,j \in I$ with $i \le j$, any integer between i and j is also an element in I. For two connected sets $I_1, I_2 \in \mathbb{Z}$ with $I_1 \cap I_2 = \varnothing$, we write $I_1 < I_2$ if $i_1 < i_2$ for any $i_1 \in I_1$ and $i_2 \in I_2$. For a connected set $I \subset \mathbb{Z}$, the length $|I|$ of I is defined by $|I| = \sup\{|i-j| \mid i,j \in I\}$. Now consider the minimizing sequence $\{u_k\}$ normalized by (2.13). By Remark 2.19, for each k, there are finitely many disjoint connected sets $I_1^k < \cdots < I_{l(k)}^k$ in $\mathbb{Z}$ satisfying

$$\{i \in \mathbb{Z} \mid \|u_k - a^-\|_{L^2(T_i,\mathbb{R}^m)} \le \rho\} = \bigcup_{j=1}^{l(k)} I_j^k.$$

The normalization (2.13) shows that for any integer $i \le 0$, we have $i \in I_1^k$ and $|I_j^k| < \infty$ for $j = 2,\dots,l(k)$. Remark 2.19 also implies that the sequence $\{l(k)\}$ is bounded. Taking a subsequence of $k \in \mathbb{N}$ if necessary, it can be assumed that $l(k)$ is a positive integer l independent of $k \in \mathbb{N}$. Define

$$p_0 \equiv \max\{i \in \{1,\dots,l\} \mid \limsup_{k\to\infty} |I_i^k| = \infty\}.$$

It is well-defined since $|I_1^k| = \infty$. Note that if $p_0 < l$,

$$\limsup_{k\to\infty} \sum_{j=p_0+1}^{l} |I_j^k| < \infty. \tag{2.29}$$

Now define $p(k)$ to be the largest $i \in I_{p_0}^k$. Set $v_k(x) = u_k(x_1 + p(k), x_2, \ldots, x_n)$ for $k \in \mathbb{N}$, so $\{v_k\}$ is a new minimizing sequence. By Proposition 2.24, the set of norms $\{\|v_k\|_{W^{1,2}(T_i,\mathbb{R}^m)} \mid i \in \mathbb{Z},\ k \in \mathbb{N}\}$ is bounded. Since $\partial\Omega \in C^1$, taking a subsequence if necessary, we see that for some $U \in E$ and any $i \in \mathbb{Z}$, v_k converges weakly to U in $W^{1,2}(T_i)$, strongly to U in $L^2(T_i, \mathbb{R}^m)$ and pointwise a.e. to U on T_i as $k \to \infty$. Therefore $V(x, v_k) \to V(x, U)$ pointwise a.e. The weak lower semicontinuity of the $|\nabla u|^2$ term in J on bounded sets implies that for any $p < q \in \mathbb{Z}$,

$$\sum_{i=p}^{q}\left(\int_{T_i} |\nabla U|^2\,dx\right) \le \liminf_{k\to\infty} \sum_{i=p}^{q}\left(\int_{T_i} |\nabla v_k|^2\,dx\right).$$

By Fatou's lemma,

$$\sum_{i=p}^{q}\left(\int_{T_i} V(x,U)\,dx\right) \le \liminf_{k\to\infty} \sum_{i=p}^{q}\left(\int_{T_i} V(x,v_k)\,dx\right).$$

Combining these inequalities yields

$$\sum_{i=p}^{q}\left(\int_{T_i} L(U)\,dx\right) \le \liminf_{k\to\infty} \sum_{i=p}^{q}\left(\int_{T_i} L(v_k)\,dx\right) \le \liminf_{k\to\infty} J(v_k) \le c.$$

Letting $p \to -\infty$ and $q \to \infty$ gives

$$J(U) \le c. \tag{2.30}$$

Since $\lim_{k\to\infty} |I_{p_0}^k| = \infty$, we see that

$$\|U - a^-\|_{L^2(T_i,\mathbb{R}^m)} \le \rho \quad \text{for } i \le 0. \tag{2.31}$$

By (2.30), $\|\nabla U\|_{L^2(T_i,\mathbb{R}^m)} \to 0$ as $i \to \infty$. By the Poincaré inequality, there is a constant b, independent of $i \in \mathbb{Z}$, so that

$$\|U - [U]^i\|_{L^2(T_i,\mathbb{R}^m)} \le b\|\nabla U\|_{L^2(T_i,\mathbb{R}^m)} \to 0 \quad \text{as } i \to \infty, \tag{2.32}$$

where $[U]^i$ is the mean value of U on T_i. Since $\int_\Omega V(x,U)\,dx < \infty$, as in the proof of Proposition 2.14, it follows that $\lim_{i\to\infty}[U]^i = a^-$ or a^+. Thus,

$$\lim_{i\to\infty} \|U - a^-\|_{L^2(T_i,\mathbb{R}^m)} = 0 \quad \text{or} \quad \lim_{i\to\infty} \|U - a^+\|_{L^2(T_i,\mathbb{R}^m)} = 0.$$

If $\lim_{i\to\infty} \|U - a^-\|_{L^2(T_i,\mathbb{R}^m)} = 0$, this contradicts (2.29) since $\lim_{k\to\infty} \|U - v_k\|_{L^2(T_i,\mathbb{R}^m)} = 0$ for each $i \in \mathbb{Z}$. Consequently,

$$\lim_{i\to\infty} \|U - a^+\|_{L^2(T_i,\mathbb{R}^m)} = 0 \tag{2.33}$$

and $U \in \Gamma$. This with (2.30) shows U is a minimizer of J in (2.1). It is clear that $J(U) = c > 0$ and Theorem 2.2 is proved. □

If $n = 1$, then $\mathcal{D} = \varnothing$ and $\Omega = \mathbb{R}$ in the problem (PDE). Thus, in this case, (PDE) reduces to a second-order Hamiltonian system of ordinary differential equations. Moreover, we get a much stronger conclusion than Theorem 2.2:

Theorem 2.34. *Assume $n = 1$. If V satisfies (V_1)–(V_3) with $\mathcal{D} = \varnothing$ and $\Omega = \mathbb{R}$, then any minimizer U of (2.1) is a classical solution of (PDE).*

Proof. Since $n = 1$, the above $W^{1,2}_{\mathrm{loc}}$ bounds imply U is continuous. Its asymptotic behavior then shows $U \in L^\infty(\mathbb{R}, \mathbb{R}^m)$. Consider $\varphi \in W^{1,2}_{\mathrm{loc}}(\mathbb{R})$ having compact support in $\mathbb{R}$ and $t \in \mathbb{R}$. Then for $0 < |t|$ small, $U + t\varphi \in \Gamma$. Consequently, $J(U + t\varphi) \geq J(U)$ or

$$\int_{\operatorname{supp}\varphi} L(U + t\varphi) - L(U)\, dx \geq 0 \tag{2.35}$$

for all such t and φ. Hence

$$\int_\Omega \nabla U \cdot \nabla\varphi + V_u(x, U) \cdot \varphi\, dx = 0 \tag{2.36}$$

for all such φ, so U is a weak solution of (PDE). But for $n = 1$, the weak form of (PDE) implies U is a classical solution of (PDE). □

Remark 2.37. If the minimizer U of Theorem 2.2 lies in $L^\infty(\Omega, \mathbb{R}^m)$, the argument just given in (2.35)–(2.36) shows U is a weak solution of (PDE) even for $n > 1$.

Remark 2.38. This existence result for $n = 1$ under (V_1)–(V_3) seems to be new. It generalizes earlier such results, [Bolotin 1978; Kozlov 1985; Sternberg 1991; Rabinowitz 1989; 1993; 2012; Alikakos and Fusco 2015], which get the existence results under slightly stronger hypotheses on V in terms of smoothness and nondegenerate behavior of V at the equilibrium solutions $a^\pm$.

To conclude this section, as a corollary of Theorem 2.34, an explicit L^∞ bound for any minimizer U will be given. The bound will be useful in Section 4. First some notational preliminaries are needed. Since $U = U(x_1)$, writing t for x_1, by (2.1),

$$J(U) = c = \int_{\mathbb{R}} \left(\tfrac{1}{2}|U_t|^2 + V(t, U)\right) dt, \tag{2.39}$$

so

$$\int_{\mathbb{R}} |U_t|^2\, dt \leq 2c. \tag{2.40}$$

With $\rho \leq \frac{1}{2}|a^+ - a^-|$, let

$$T(\rho) \equiv \left\{t \in \mathbb{R} \;\middle|\; \min\{|U(t) - a^-|, |U(t) - a^+|\} \geq \rho\right\}.$$

By (V_1)–(V_3),

$$\beta(\rho) \equiv \inf\left\{V(t, u) \;\middle|\; t \in \mathbb{R},\ \min\{|u - a^-|, |u - a^+|\} \geq \rho\right\} > 0.$$

Therefore by (2.39),

$$|T(\rho)|\beta(\rho) \leq \int_{\mathbb{R}} V(t, U)\, dt \leq c. \tag{2.41}$$

Corollary 2.42. *If U is a minimizer of (2.1) as in Theorem 2.34, then*

$$\|U\|_{L^\infty(\mathbb{R},\mathbb{R}^m)} \le \rho + \max\{|a^-|,|a^+|\} + \left(\frac{2}{\beta(\rho)}\right)^{\frac{1}{2}} c \equiv K. \tag{2.43}$$

Proof. If $\|U\|_{L^\infty(\mathbb{R},\mathbb{R}^m)} \le \max\{|a^-|,|a^+\}$, the estimate holds. Thus we may assume that $\|U\|_{L^\infty(\mathbb{R},\mathbb{R}^m)} > \max\{|a^-|,|a^+|\}$. Then, the maximum of $|U|$ is achieved at some $z \in \mathbb{R}$. If $z \notin T(\rho)$, it follows that

$$|U(z)| \le \rho + \max\{|a^+|,|a^-|\}.$$

If $z \in T(\rho)$, we take ξ to be the closest boundary point of $T(\rho)$ to z. Then, we see from (2.40)–(2.41) that

$$\begin{aligned}\|U\|_{L^\infty(\mathbb{R},\mathbb{R}^m)} = |U(z)| &\le |U(\xi)| + \left|\int_\xi^z U_t(s)\,ds\right| \\ &\le |U(\xi)| + \left(|z-\xi|\int_\xi^z |U_t(s)|^2\,ds\right)^{\frac{1}{2}} \\ &\le |U(\xi)| + |T(\rho)|^{\frac{1}{2}}(2c)^{\frac{1}{2}} \le |U(\xi)| + \left(\frac{2}{\beta(\rho)}\right)^{\frac{1}{2}} c.\end{aligned}$$

Since $|U(\xi)| \le \rho + \max\{|a^-|,|a^+|\}$, (2.43) now follows. ☐

Remark 2.44. Suppose that V in Theorem 2.34 is modified for $|u| > K$ so that the resulting function, V^*, still satisfies (V_1)–(V_3) (for $n = 1$) and

$$\inf\{V^*(t,u) \mid t \in \mathbb{R},\ \min\{|u-a^-|,|u-a^+|\} \ge \rho\} \ge \beta(\rho).$$

Then the corresponding functional J^* has a minimizer $U^* \in \Gamma$ and since $V^*(t,u) = V(t,u)$ for $|u| \le K$, minimizing sequences $\{u_k\}$ for J^* can be assumed to satisfy $J^*(u_k) \le J(U)$. Consequently

$$J^*(U^*) \le J(U) \tag{2.45}$$

and (2.45) and the derivation of (2.43) show any minimizer U^* of the modified problem is also bounded in L^∞ by K. Thus such a modification produces no new minimizers.

3. The regularity of the weak solution

The regularity of any weak solution U of (PDE) that minimizes J on Γ will be discussed in this section. The special case of $n = 1$ has already been shown in Theorem 2.34. Therefore it will be assumed that $n \ge 2$ in what follows. Using standard terminology, a solution u of (PDE) and (BC) is called a strong solution if $u \in W^{2,2}_{\mathrm{loc}}(\overline{\Omega})$. Our main result is:

Theorem 3.1. *Suppose V satisfies (V_1)–(V_4).*

(1) *If $\partial\Omega = \mathbb{R} \times \partial\mathcal{D} \in C^2$, then any minimizer U of (2.1) is a weak solution of (PDE) and (BC). Moreover, any weak solution $U \in E$ of (PDE) and (BC) is a strong solution of (PDE) and (BC), and $U \in L^\infty(\Omega)$.*

(2) *If* $V_u \in C^1(\overline{\Omega} \times \mathbb{R}^m)$ *and* $\partial\Omega \in C^3$, *then* $U \in C^{2,\alpha}(\overline{\Omega}, \mathbb{R}^m)$ *for any* $\alpha \in (0,1)$ *and* U *is a classical solution of* (PDE) *and* (BC) *with* $\lim_{x_1 \to \pm\infty} U(x_1, \hat{x}) = a^{\pm}$ *uniformly for* $\hat{x} \in \mathcal{D}$.

Regularity results for weak solutions of a single second-order quasilinear elliptic partial differential equation satisfying Dirichlet boundary conditions can be found in the literature; see, e.g., Chapters 8–9 of [Gilbarg and Trudinger 1983]. However, we do not know of a reference for such a result for the system (PDE) with (BC). Therefore for completeness we will provide a proof of Theorem 3.1 but postpone it until Section 6.

4. Some generalizations

In this section, Theorem 2.2 will be generalized in various ways. First we will show that the growth condition, (V_4), can be bypassed when a geometrical condition that leads to an L^∞ bound for minimizers of (2.1) is satisfied. Next the case of a more general domain Ω that is periodic in the x_1-direction will be treated. Then a perturbation result will be given. Lastly, the case when the potential V has multiple minima will be discussed briefly.

To begin the first result, for any set $A \in \mathbb{R}^m$ and $a > 0$, let $A^a \equiv \{y \in \mathbb{R}^m \mid \operatorname{dist}(y, A) < a\}$.

Theorem 4.1. *Suppose that* V *satisfies* (V_1) *and* (V_5), *where*:

(V_5) *There is a convex bounded open set* $O \subset \mathbb{R}^m$ *with* $\partial O \in C^2$ *such that*

(1) *there are two different points* a^- *and* a^+ *in* O *such that* $V(x, a^{\pm}) = 0$ *for all* $x \in \Omega$ *and* $V(x,u) > 0$ *for any* $u \in \overline{O} \setminus \{a^-, a^+\}$, $x \in \Omega$;

(2) *there is a constant* $\delta > 0$ *such that for the outward unit normal vector* $\mu = \mu(u)$ *to* ∂O,

$$V(x,u) \le V(x, u + t\mu(u)) \quad \text{when } x \in \Omega,\ u \in \partial O,\ t \in [0,\delta].$$

Then there is a weak solution, $U \in W^{1,2}_{\mathrm{loc}}(\Omega, \overline{O}) \cap \Gamma$ *of* (PDE) *and* (BC). *If further,* $V_u \in C^1(\overline{\Omega} \times \mathbb{R}^m)$ *and* $\partial\Omega \in C^3$, *the solution* U *is a classical solution of* (PDE) *and* (BC) *with* $U \in C^{2,\alpha}(\overline{\Omega})$ *for any* $\alpha \in (0,1)$ *with* $\lim_{x_1 \to \pm\infty} U(x_1, \hat{x}) = a^{\pm}$ *uniformly for* $\hat{x} \in \mathcal{D}$.

As a first step towards proving Theorem 4.1, a projection map $P : O^\delta \to \overline{O} = O \cup \partial O$ will be defined. Taking a smaller $\delta > 0$ if necessary shows that for each $u \in (\partial O)^\delta$, there exists a unique $s(u) \in \partial O$ with $|u - s(u)| = \min_{w \in \partial O} |u - w|$. This implies $s \in C^1((\partial O)^\delta, \partial O)$. Define a projection map $P : O^\delta \to \overline{O}$ by $P(u) = u$ for $u \in O$ and $P(u) = s(u) \in \partial O$ for $u \in O^\delta \setminus O$. Note that if $u \in O^\delta \setminus O$, then

$$u - s(u) = |u - s(u)|\,\mu(s(u)). \tag{4.2}$$

Making δ smaller if necessary, the implicit function theorem shows $P : O^\delta \setminus O \to \partial O$ is C^1.

Next to prove Theorem 4.1, a property of the function $s(u)$ is needed.

Lemma 4.3. *If* $u \in C^1(\Omega, \mathbb{R}^m)$ *and* $u(x) \in O^\delta$ *for some* $x \in \Omega$, *then, for each* $i = 1, \ldots, n$,

$$\left|\frac{\partial u(x)}{\partial x_i}\right| \ge \left|\frac{\partial (s \circ u)(x)}{\partial x_i}\right|.$$

Proof. It is a well-known result that the function s is a contraction; that is, $|s(z_1)-s(z_2)| \le |z_1-z_2|$ for any $z_1, z_2 \in O^\delta$. Thus, for $y \in \Omega$ close to x,

$$\frac{|u(y)-u(x)|}{|y-x|} \ge \frac{|s(u(y))-s(u(x))|}{|y-x|}.$$

For $y = x + he_i$, $h \in \mathbb{R}$, letting $|h| \to 0$, we get the inequality. $\square$

Proposition 4.4. *For any $u \in C^1(\Omega, O^\delta) \cap W^{1,2}_{\mathrm{loc}}(\Omega, O^\delta)$, it follows that*

$$P \circ u \in W^{1,2}_{\mathrm{loc}}(\Omega, \bar{O}) \quad \textit{and} \quad J(P(u)) \le J(u).$$

Proof. For each $z \in (\partial O)^\delta$, there exists a unique $s(z) \in \partial O$ with $|z - s(z)| = \min_{w \in \partial O} |z-w|$. For each $z \in \partial O$, we have $\mu(z)$ is the outward unit normal vector to ∂O at $z \in \partial O$. For each $z \in (\partial O)^\delta$, we define

$$\lambda(z) = \begin{cases} |z-s(z)| & \text{for } z \in (\partial O)^\delta \setminus \bar{O}, \\ -|z-s(z)| & \text{for } z \in (\partial O)^\delta \cap O, \end{cases}$$

and

$$\lambda_-(z) \equiv \min\{\lambda(z), 0\}.$$

Observe that

$$s \in C^1((\partial O)^\delta, \partial O), \quad \lambda \in C^1((\partial O)^\delta, \mathbb{R}), \quad \mu \in C^1(\partial O, \mathbb{R}^m),$$

and

$$z = s(z) + \lambda(z)\mu(s(z)).$$

For $z = u(x) \in (\partial O)^\delta$, we see that

$$P(u(x)) = s(u(x)) + \lambda_-(u(x))\mu(s(u(x))).$$

Define

$$f_\varepsilon(\lambda) \equiv \begin{cases} 0 & \text{for } \lambda \ge 0, \\ -(\lambda^2+\varepsilon^2)^{\frac{1}{2}} + \varepsilon & \text{for } \lambda < 0. \end{cases}$$

Approximating $P(u)(x)$ by $s(u(x)) + f_\varepsilon(\lambda(u(x)))\mu(s(u(x)))$ and letting $\varepsilon \to 0$ shows that $P(u) \in W^{1,2}_{\mathrm{loc}}(\Omega, \bar{O})$, and for $u(x) \in O^\delta \setminus O$, we have $\nabla P(u)(x) = \nabla s \circ u(x)$, while for $u(x) \in O$, we have $\nabla P(u)(x) = \nabla u(x)$. Now Lemma 4.3 implies that $|\nabla P(u)| \le |\nabla u|$. Thus

$$\int_\Omega |\nabla P(u(x))|^2\,dx \le \int_\Omega |\nabla u(x)|^2\,dx. \tag{4.5}$$

Moreover, hypothesis (V_5) implies that

$$\int_\Omega V(x, P(u(x))\,dx \le \int_\Omega V(x, u(x))\,dx. \tag{4.6}$$

Then (4.5) and (4.6) show $J(P(u)) \le J(u)$. $\square$

Proof of Theorem 4.1. As a class of admissible functions, take

$$\Gamma(O^\delta) = \big\{u \in E \;\big|\; u(x) \in O^\delta \text{ for } x \in \Omega,\ \|u-a^\pm\|_{L^2(T_i,\mathbb{R}^m)} \to 0,\ i \to \pm\infty\big\}.$$

Define

$$c(O^\delta) = \inf_{u \in \Gamma(O^\delta)} J(u). \tag{4.7}$$

Since O is convex, it is readily seen that $\Gamma(O^\delta) \neq \emptyset$ and $0 \leq c(O^\delta) < \infty$. Let $\{u_k\} \subset \Gamma(O^\delta)$ be a minimizing sequence for (4.7). By the density of $C^1(\Omega, \mathbb{R}^m) \cap \Gamma(O^\delta)$ in $\Gamma(O^\delta)$, we may assume that $\{u_k\} \subset C^1(\Omega, \mathbb{R}^m) \cap \Gamma(O^\delta)$. Since P is a contraction on O^δ and is the identity map on O, for any $z \in O^\delta$ and $w \in O$, we have $|P(z) - w| \leq |z - w|$. Thus

$$\|P(u) - a^\pm\|_{L^2(T_i, \mathbb{R}^m)} \leq \|u - a^\pm\|_{L^2(T_i, \mathbb{R}^m)} \to 0, \quad i \to \pm\infty.$$

Hence Proposition 4.4 implies that $\{P(u_k)\}$ is also a minimizing sequence for (4.7) which is contained in $W^{1,2}_{\mathrm{loc}}(\Omega, \overline{O}) \cap \Gamma(O^\delta)$. The proof of Theorem 2.2 shows that there exists a $p(k) \in \mathbb{Z}$ such that a subsequence of $\{P(u_k(\cdot + (p(k), 0, \ldots, 0)))\}$ converges weakly in $W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$, strongly in $L^2_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ and pointwise a.e. to a minimizer $U \in W^{1,2}_{\mathrm{loc}}(\Omega, \overline{O}) \cap \Gamma(O^\delta)$ of (2.1). Since $U(x) \in \overline{O}$ for any $x \in \Omega$ and O is bounded, by Remark 2.37, U is a weak solution of (PDE) and (BC).

Following the argument in the Completion of the Proof of Theorem 3.1, we get that if $V_u \in C^1$ and $\partial\Omega \in C^3$, then $U \in C^{2,\alpha}(\Omega, O) \cap \Gamma$ and U is a classical (PDE) and (BC) with $\lim_{x_1 \to \pm\infty} U(x_1, \hat{x}) = a^\pm$ uniformly for $\hat{x} \in \mathcal{D}$. $\square$

For our second result, as earlier, let e_i be a unit vector in the positive x_i-direction, $1 \leq i \leq n$. Assume:

(Ω_1) $\Omega \subset \mathbb{R} \times \mathcal{D}$ for some bounded set $\mathcal{D} \subset \mathbb{R}^{n-1}$, $\partial\Omega$ is a C^3 manifold, and for all $x \in \Omega$, we have $x \pm e_1 \in \Omega$.

(Ω_2) Ω is a connected set.

Define the functional J as earlier with this new choice of Ω and for $i, j \in \mathbb{Z}$ with $i < j$, set $T_i = \{x \in \Omega \mid i < x_1 < i + 1\}$ and $T_i^j = \{x \in \Omega \mid i < x_1 < j\}$.

Then we have:

Theorem 4.8. *Suppose that V satisfies (V_1)–(V_3) and Ω satisfies (Ω_1), (Ω_2). Let*

$$\Gamma_1 = \left\{u \in E \;\middle|\; \|u - a^\pm\|_{L^2(T_i, \mathbb{R}^m)} \to 0,\ i \to \pm\infty\right\}.$$

Then there is a $U \in \Gamma_1$ such that

$$J(U) = \inf_{u \in \Gamma_1} J(u). \tag{4.9}$$

Proof. The proof of Theorem 2.2 uses Proposition 2.14 and Remark 2.19 to show that a minimizing sequence $\{u_k\}$ satisfying the normalization (2.13) and the bounds given by Proposition 2.24 has a subsequence which converges to a minimizer U of the functional J on Γ. Since Proposition 2.14 and Remark 2.19 can be proved in the same manner for a domain Ω satisfying (Ω_1) and (Ω_2), the proof carries over to the present setting provided that the bounds of Proposition 2.24 are also valid here; i.e., if $\{u_k\}$ is a minimizing sequence for (4.9), there is a constant $M > 0$ such that

$$\|u_k\|_{W^{1,2}(T_i, \mathbb{R}^m)} \leq M \tag{4.10}$$

for all $k \in \mathbb{N}$ and $i \in \mathbb{Z}$. We will show that this is the case. The proof uses the following result.

Lemma 4.11. *Assume that* (Ω_1) *and* (Ω_2) *hold. Then for any fixed* $k \geq 3$, *there exists a constant* $C > 0$, *independent of* $i \in \mathbb{Z}$, *such that for any* $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ *and* $j \in \{i+1, \dots, i+k-2\}$,

$$\|u\|_{L^2(T_j, \mathbb{R}^m)} \leq C \big(\|\nabla u\|_{L^2(T_i^{i+k}, \mathbb{R}^m)} + \|u\|_{W^{1,2}(T_i, \mathbb{R}^m)} + \|u\|_{L^2(T_{i+k-1}, \mathbb{R}^m)}\big).$$

Proof. By a translation in $\mathbb{Z}e_1$, it suffices to show that there exists a constant $C > 0$ such that for any $u \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ and $j \in \{1, \dots, k-2\}$,

$$\|u\|_{L^2(T_j, \mathbb{R}^m)} \leq C \big(\|\nabla u\|_{L^2(T_0^{k}, \mathbb{R}^m)} + \|u\|_{L^2(T_0, \mathbb{R}^m)} + \|u\|_{L^2(T_{k-1}, \mathbb{R}^m)}\big).$$

To the contrary, suppose that the inequality above does not hold. Then there is a sequence $\{w_l\} \subset W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ and $j \in \{1, \dots, k-2\}$ such that

$$\|w_l\|_{L^2(T_j, \mathbb{R}^m)} = 1 \tag{4.12}$$

and

$$\|\nabla w_l\|_{L^2(T_0^k, \mathbb{R}^m)} + \|w_l\|_{L^2(T_0, \mathbb{R}^m)} + \|u\|_{L^2(T_{k-1}, \mathbb{R}^m)} \to 0 \quad \text{as } l \to \infty. \tag{4.13}$$

Let Ω_0^j be a connected component of T_j and Ω^j a connected component of T_0^k containing Ω_0^j. If $\Omega^j \cap (T_0 \cup T_{k-1}) = \varnothing$, then Ω^j is an isolated connected component of Ω. But $k \geq 3$, so this contradicts the connectedness of Ω. Thus $\Omega^j \cap (T_0 \cup T_{k-1}) \neq \varnothing$. Assume that $\Omega^j \cap T_0 \neq \varnothing$. Then by the Poincaré inequality, there exists $c > 0$, independent of l, such that

$$\|w_l - [w_l]_j\|_{L^2(\Omega^j, \mathbb{R}^m)} \leq c \|\nabla w_l\|_{L^2(\Omega^j, \mathbb{R}^m)}, \tag{4.14}$$

where $[w_l]_j = (1/|\Omega_j|) \int_{\Omega_j} w_l \, dx$. Since $\lim_{l \to \infty} \|w_l\|_{L^2(T_0, \mathbb{R}^m)} = 0$ and

$$\begin{aligned}\|w_l - [w_l]_j\|_{L^2(\Omega^j \cap T_0, \mathbb{R}^m)} &\leq \|w_l - [w_l]_j\|_{L^2(\Omega^j, \mathbb{R}^m)} \leq c \|\nabla w_l\|_{L^2(\Omega^j, \mathbb{R}^m)} \\ &\leq c \|\nabla w_l\|_{L^2(T_0^k, \mathbb{R}^m)},\end{aligned}$$

(4.13) implies that $\lim_{l \to \infty} [w_l]_j = 0$. Then (4.14) shows that

$$\|w_l\|_{L^2(\Omega_0^j, \mathbb{R}^m)} \leq \|w_l\|_{L^2(\Omega^j, \mathbb{R}^m)} \to 0 \quad \text{as } l \to \infty.$$

If $\Omega^j \cap T_k \neq \varnothing$, we obtain the same conclusion. Thus, for each connected component Ω_0^j, we have $\lim_{l \to \infty} \|w_l\|_{L^2(\Omega_0^j, \mathbb{R}^m)} = 0$. This implies $\lim_{l \to \infty} \|w_l\|_{L^2(T_j, \mathbb{R}^m)} = 0$, contradicting (4.12) and completing the proof. □

Now, we argue as in the proof of Proposition 2.14. Since Proposition 2.14 and Remark 2.19 hold for a domain Ω satisfying (Ω_1) and (Ω_2), there exists $L \in \mathbb{N}$, independent of k, such that the number of elements of

$$A_k = \big\{ i \in \mathbb{N} \;\big|\; \|u_k - a^-\|_{L^2(T_i, \mathbb{R}^m)} \geq \rho, \ \|u_k - a^+\|_{L^2(T_i, \mathbb{R}^m)} \geq \rho \big\}$$

is bounded by L for each $k \in \mathbb{N}$. Note that if $i \notin A_k$,

$$\|u_k\|_{L^2(T_i, \mathbb{R}^m)} \leq \rho + \max\{|a^{-1}|, |a^+|\} |T_0|^{\frac{1}{2}}.$$

Then, applying Lemma 4.11, we get the boundedness (4.10). For the completion of the proof of Theorem 4.8, we follow exactly the same argument as in the Completion of the Proof of Theorem 2.2. Then, we get a minimizer $U \in \Gamma_1$ of J. $\square$

As a consequence of Theorem 4.8 and Theorem 3.1, we have:

Corollary 4.15. *If in addition to the hypotheses of Theorem 4.8, (V_4) is satisfied, $V_u \in C^1(\overline{\Omega} \times \mathbb{R}^m)$ and $\partial\Omega \in C^3$, then $U \in C^{2,\alpha}(\overline{\Omega}, \mathbb{R}^m)$ for any $\alpha \in (0,1)$ and U is a classical solution of (PDE) and (BC) with $\lim_{x_1 \to \pm\infty} U(x_1, \hat{x}) = a^{\pm}$ uniformly for $\hat{x} \in \mathcal{D}$. If V satisfies (V_5), then (PDE) and (BC) possess a solution $U \in C^{2,\alpha}(\Omega, \mathcal{O}) \cap \Gamma_1$.*

Theorem 3.1 and Corollary 4.15 require condition (V_4), which allows us to get an L^∞ bound for the solution. When $n = 1$, condition (V_4) is not required; conditions (V_1)–(V_3) suffice. Next an example will be given showing that a PDE perturbation of that case without any further conditions other than $V \in C^2$ gives classical solutions of (PDE) and (BC). Thus consider (PDE) and (BC) for $n = 1$. To better distinguish between the cases of $n = 1$ and the general case, set

$$\Gamma_0 = \left\{u \in W^{1,2}_{\text{loc}}(\mathbb{R}, \mathbb{R}^m) \;\middle|\; \|u - a_\pm\|_{L^2([i,i+1],\mathbb{R}^m)} \to 0,\ \pm i \to \infty\right\}$$

and

$$J_0(u) = \int_{\mathbb{R}} \left(\tfrac{1}{2}|u_{x_1}|^2 + V(x_1, u)\right) dx_1.$$

Then in Section 2, it was shown that

$$c_0 = \inf_{u \in \Gamma_0} J_0(u)$$

has a minimizer, $U_0 = U_0(x_1)$, which is a classical solution of (PDE). With the same choice of V, take any bounded domain $\mathcal{D} \subset \mathbb{R}^{n-1}$ with $\Omega = \mathbb{R} \times \mathcal{D}$ and Γ as in Section 2, J as in (1.2) and c as in (2.1). Note that $U_0 \in \Gamma$ so $J(U_0) = |\mathcal{D}| J_0(U_0) \geq c$.

Proposition 4.16. *$J(U_0) = c$ and any minimizer $U \in C^{2,\alpha}(\overline{\Omega}) \cap L^\infty(\Omega)$ of (2.1) depends only on x_1.*

Proof. Let $\{u_k\}$ be a minimizing sequence for (2.1). Write $x = (x_1, \hat{x})$ for $x \in \mathbb{R}^n$ and fix $k \in \mathbb{N}$. Then by Fubini's theorem, there exists a set $A_k \subset \mathcal{D}$ with $|A_k| = |\mathcal{D}|$ such that for any $\hat{x} \in A_k$,

$$\int_{\Omega} \left(\tfrac{1}{2}|\nabla u_k(x_1, \hat{x})|^2 + V(x_1, u_k(x_1, \hat{x}))\right) dx_1 < \infty. \tag{4.17}$$

Therefore by Proposition 2.3, there exist $e_k^{\pm}(\hat{x}) \in \{a^-, a^+\}$ such that

$$\lim_{i \to \pm\infty} \|u_k(\cdot, \hat{x}) - e_k^{\pm}(\hat{x})\|_{L^2([i,i+1],\mathbb{R}^m)} = 0. \tag{4.18}$$

We claim that $e_k^{\pm}(\hat{x}) = a^{\pm}$ for all $\hat{x} \in A_k$. Indeed for each $i \in \mathbb{Z}$, set

$$f_i^{\pm}(\hat{x}) = \int_{[i,i+1]} |u_k(x_1, \hat{x}) - a^{\pm}|^2 \, dx_1.$$

Then each function $f_i^{\pm}$ is measurable on $\mathcal{D}$ and by Fubini's theorem again,

$$\lim_{i\to\pm\infty}\int_{\mathcal{D}} f_i^{\pm}(\hat{x})\,d\hat{x} = \lim_{i\to\pm\infty}\int_{\mathcal{D}}\int_{[i,i+1]} |u_k(x_1,\hat{x})-a^{\pm}|^2\,dx_1\,d\hat{x}$$
$$= \lim_{i\to\pm\infty}\|u_k-a^{\pm}\|_{L^2(T_i)} = 0$$

since $u_k\in\Gamma$. But $f_i^{\pm}(\hat{x})\geq 0$, so $f_i^{\pm}\to 0$ in $L^1(\mathcal{D})$ as $i\to\pm\infty$. Hence there exist subsequences $i_j^{\pm}\to\pm\infty$ such that

$$f_{i_j^{\pm}}^{\pm}(\hat{x})\to 0 \quad \text{for a.e. } x\in\mathcal{D}. \tag{4.19}$$

Comparing (4.19) to (4.18) shows the existence of a set $B_k\subset A_k$ with $|B_k|=|A_k|=|\mathcal{D}|$ and $e_k^{\pm}(\hat{x})=a^{\pm}$ for all $\hat{x}\in B_k$. Defining $B=\bigcap_k B_k$, we have $|B|=|\mathcal{D}|$, and for any $\hat{x}\in B$ and $k\in\mathbb{N}$, we have

$$u_k(\,\cdot\,,\hat{x})\in W^{1,2}_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^m) \quad\text{and}\quad \lim_{i\to\pm\infty}\|u_k(\,\cdot\,,\hat{x})-a^{\pm}\|_{L^2([i,i+1],\mathbb{R}^m)}\to 0.$$

This implies that for each $\hat{x}\in B$, we have $u_k(x_1,\hat{x})\in\Gamma_0$. Therefore, for each $\hat{x}\in B$,

$$J_0(u_k(\,\cdot\,,\hat{x}))\geq J_0(U_0). \tag{4.20}$$

Integrating (4.20) over $\mathcal{D}$ then shows $J(u_k)\geq J(U_0)$, which implies $J(U_0)=c$, yielding the first part of the proposition.

For the second part, suppose that c is attained by $U\in C^{2,\alpha}(\overline{\Omega})\cap L^{\infty}(\Omega)$. As in (4.20), for a.e. $\hat{x}\in\mathcal{D}$,

$$J_0(U(\,\cdot\,,\hat{x}))\geq J_0(U_0).$$

Since $J(U)=c$, this implies that for a.e. $\hat{x}\in\mathcal{D}$,

$$J_0(U(\,\cdot\,,\hat{x}))= J_0(U_0).$$

Then, for a.e. $\hat{x}\in\mathcal{D}$,

$$\frac{\partial^2 U(x_1,\hat{x})}{\partial x_1^2}-V_u(x_1,U(x_1,\hat{x}))=0.$$

This implies that

$$\Delta_{\hat{x}}U\equiv\Delta U-\frac{\partial^2 U(x_1,\hat{x})}{\partial x_1^2}=0 \quad \text{for any } x_1\in\mathbb{R};$$

i.e., $U(x_1,\hat{x})$ as a function of $\hat{x}$ is harmonic. Thus using the boundary condition (BC) shows that $U(x_1,\hat{x})$ does not depend on $\hat{x}\in\mathcal{D}$. This completes the proof. $\square$

Now the perturbation result can be formulated. Suppose:

(V_0) For some $\varepsilon_0>0$, there exists a function $W\in C^1((-\varepsilon,\varepsilon)\times\overline{\Omega}\times\mathbb{R}^m)$ such that for each $\varepsilon\in(-\varepsilon_0,\varepsilon_0)$, $W(\varepsilon,\cdot)$ satisfies (V_1)–(V_3) and $W(0,x,u)=V(x_1,u)$.

For $|\varepsilon| < \varepsilon_0$, consider the family of equations

$$-\Delta u + W_u(\varepsilon, x, u) = 0, \quad x \in \Omega, \tag{4.21}$$

with boundary conditions

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial\Omega. \tag{4.22}$$

Then we have:

Theorem 4.23. *Suppose (V_0) is satisfied and $\partial\Omega \in C^3$. Then there is an $\varepsilon_1 \in (0, \varepsilon_0)$ such that the problem (4.21)–(4.22) has a classical solution U_ε for each $|\varepsilon| \le \varepsilon_1$.*

Proof. Let u_0 be any minimizer of J_0 on Γ_0. Then (2.43) provides an upper bound K for $\|u_0\|_{L^\infty(\mathbb{R},\mathbb{R}^m)}$ and any such u_0. To obtain the solutions U_ε, the family of functions $W(\varepsilon,\cdot)$ will be truncated. Let $W_K \in C((-\varepsilon,\varepsilon)\times\overline{\Omega}\times\mathbb{R}^m)$ satisfy (V_0) with

(1) $W_K(0,x,u)$ independent of $\hat{x}$,
(2) $W_K(\varepsilon,x,u) = W(\varepsilon,x,u)$ for $|u| \le 2K$,
(3) $|(W_K)_u(\varepsilon,x,u)| \le K_1$ for some constant K_1,
(4) $\liminf_{|u|\to\infty} W_K(\varepsilon,x,u) \ge \underline{V} > 0$ uniformly for $x \in \Omega$ and $|\varepsilon| \le \varepsilon_0$,
(5) $\inf\{W_K(0,x,u) \mid t \in \mathbb{R},\ \min\{|u-a^-|,|u-a^+|\} \ge \rho\} \ge \beta(\rho)$,

where $\beta(\rho) = \inf\{V(x_1,u) \mid x_1 \in \mathbb{R},\ \min\{|u-a^-|,|u-a^+|\} \ge \rho\}$. It is straightforward to construct such a family of functions. By (V_0) and Theorem 2.2, the functional

$$J_{\varepsilon,K}(u) \equiv \int_\Omega \left(\tfrac{1}{2}|\nabla u|^2 + W_K(\varepsilon,x,u)\right)dx \equiv \int_\Omega L_{\varepsilon,K}(u)\,dx,$$

corresponding to (4.21) with W replaced by W_K has a minimizer $U_{\varepsilon,K} \in \Gamma$ for each $|\varepsilon| \le \varepsilon_0$. By (3) of the properties of W_K and Theorem 3.1(1), there is a constant $\overline{M}_1$ that is independent of ε but depends on K such that

$$\|U_{\varepsilon,K}\|_{L^\infty(\Omega,\mathbb{R}^m)} \le \overline{M}_1.$$

Thus by Theorem 3.1(2), each component of $U_{\varepsilon,K}$ can be viewed as a $C^{2,\alpha}(\Omega,\mathbb{R})$ solution of a linear elliptic equation of the form

$$-\Delta v = f, \quad x \in \Omega,$$

with $\partial v/\partial\nu = 0$, $x \in \partial\Omega$ and $f \in L^\infty(\Omega,\mathbb{R})$. Applying the $W^{2,q}_{\mathrm{loc}}$ estimates for such equations gives a constant $\overline{M}_2 > 0$ that is independent of ε but depends on K such that

$$\|U_{\varepsilon,K}\|_{W^{2,q}(\Omega,\mathbb{R})} \le \overline{M}_2.$$

Taking $q > n$ and applying the Schauder estimates for each component yields a constant $\overline{M} > 0$ that is independent of ε but depends on K such that

$$\|U_{\varepsilon,K}\|_{C^{2,\alpha}(\Omega,\mathbb{R}^m)} \le \overline{M}. \tag{4.24}$$

Thus, $U_{\varepsilon,K}$ is a classical solution of (4.21)–(4.22). It can be assumed that the functions $U_{\varepsilon,K}$ are normalized as in (2.13). We claim there is an $\varepsilon_1 \in (0,\varepsilon_0]$ such that for $|\varepsilon| \le \varepsilon_1$, we have $\|U_{\varepsilon,K}\|_{L^\infty(\Omega,\mathbb{R}^m)} < 2K$. If so, $W_K(\varepsilon,x,U_{\varepsilon,K}(x)) = W(\varepsilon,x,U_{\varepsilon,K}(x))$ and $U_{\varepsilon,K}$ is the desired solution of (4.21)–(4.22) for $|\varepsilon| \le \varepsilon_1$. To show that such an ε_1 exists, suppose that there exists $\varepsilon_l \in (-\varepsilon_0,\varepsilon_0)$ with $\lim_{l\to\infty}\varepsilon_l = 0$ such that

$$\limsup_{l\to\infty} \|U_{\varepsilon_l,K}\|_{L^\infty(\Omega,\mathbb{R}^m)} \ge 2K. \tag{4.25}$$

By (4.24), it can be assumed that $U_{\varepsilon_l,K}$ converges in $C^2_{\mathrm{loc}}(\overline{\Omega},\mathbb{R}^m)$ to a solution U^* of (4.21)–(4.22) for $\varepsilon = 0$. Due to equations (4.21)–(4.22) again, the convergence is in $C^{2,\alpha}_{\mathrm{loc}}(\Omega,\mathbb{R}^m)$ so by (4.24),

$$\|U^*\|_{C^{2,\alpha}(\Omega,\mathbb{R}^m)} \le \overline{M}. \tag{4.26}$$

Suppose for the moment that U^* minimizes $J_{0,K}$ on Γ. Then by Remark 2.44, $\|U^*\|_{L^\infty(\Omega,\mathbb{R}^m)} \le K$ and (4.26) is in contradiction to (4.25). Hence ε_1 exists and the theorem is proved.

It remains to verify that U^* minimizes $J_{0,K}$ on Γ. As a first step, let $w \in C^1(\mathbb{R},\mathbb{R}^m)$ with $w(t) = a^-$ for $t \le -1$ and $w(t) = a^+$ for $t \ge 1$. We define $\tilde{w}(x_1,\hat{x}) = w(x_1)$. Then there is a constant M_1 independent of ε but depending on K such that

$$J_{\varepsilon,K}(\hat{w}) \le M_1. \tag{4.27}$$

Thus $J_{\varepsilon,K}(U_{\varepsilon,K}) \le M_1$ for $|\varepsilon| \le \varepsilon_0$. Now for any $R > 0$, due to the C^1_{loc} convergence of $U_{\varepsilon,K}$,

$$\int_{[-R,R]\times\mathcal{D}} L_{0,K}(U^*)\,dx = \lim_{\varepsilon\to 0}\int_{[-R,R]\times\mathcal{D}} L_{\varepsilon,K}(U_{\varepsilon,K})\,dx \le M_1.$$

Thus letting $R \to \infty$ shows

$$J_{0,K}(U^*) \le M_1. \tag{4.28}$$

By (4.28), as $|i| \to \infty$,

$$\int_{T_i} L_{0,K}(U^*)\,dx \to 0. \tag{4.29}$$

Due to the bounds (4.26) and the Poincaré inequality,

$$\|U^* - a^\pm\|_{W^{1,2}(T_i,\mathbb{R}^m)} \to 0, \quad |i| \to \infty. \tag{4.30}$$

Employing the bounds, (4.26) again with (4.30) and an interpolation inequality shows

$$\|U^* - a^\pm\|_{C^1(T_i,\mathbb{R}^m)} \to 0, \quad i \to \pm\infty. \tag{4.31}$$

The estimate (4.31) also holds for any u_0 minimizing $J_{0,K}$ on Γ. Let $\sigma > 0$. By (4.31), there is a $q = q(\sigma) \in \mathbb{N}$ such that for $u = U^*$ or $u = u_0$,

$$\|u - a^\pm\|_{C^1(T_i,\mathbb{R}^m)} \le \sigma, \quad \pm i \ge q. \tag{4.32}$$

By (4.28) again, by taking q larger if need be, it can be assumed that

$$\int_{\{|x_1|\ge q+1\}\times\mathcal{D}} L_{0,K}(U^*)\,dx \le \sigma \quad \text{and} \quad \int_{\{|x_1|\ge q+1\}\times\mathcal{D}} L_{0,K}(u_0)\,dx \le \sigma. \tag{4.33}$$

Next observe that U^*, being a limit of minimizers, possesses a minimality property. Indeed since $U_{\varepsilon,K}$ minimizes $J_{\varepsilon,K}$ over Γ, for any $\varphi \in W^{1,2}(\Omega,\mathbb{R}^m)$ having compact support,

$$\int_\Omega \big(L_{\varepsilon,K}(U_{\varepsilon,K}+\varphi)-L_{\varepsilon,K}(U_{\varepsilon,K})\big)\,dx = \int_{\operatorname{supp}\varphi} \big(L_{\varepsilon,K}(U_{\varepsilon,K}+\varphi)-L_{\varepsilon,K}(U_{\varepsilon,K})\big)\,dx \geq 0. \tag{4.34}$$

Thus taking $\varepsilon \to 0$ in (4.34) yields

$$\int_{\operatorname{supp}\varphi} \big(L_{0,K}(U^*+\varphi)-L_{0,K}(U^*)\big)\,dx \geq 0. \tag{4.35}$$

Taking $q = q(\sigma)$, choose $\varphi = f_q$, where

$$f_q(x) = \begin{cases} u_0 - U^* & \text{for } |x_1| \leq q, \\ (x_1-q-1)(U^*-u_0) & \text{for } q \leq x_1 \leq q+1, \\ (-q-1-x_1)(U^*-u_0) & \text{for } -q-1 \leq x_1 \leq -q, \\ 0 & \text{for } |x_1| \geq q+1. \end{cases}$$

With this choice of φ, (4.35) becomes

$$\int_{[-q,q]\times\mathcal{D}} L_{0,K}(u_0)\,dx + \int_{T_{-q-1}\cup T_q} L_{0,K}(U^*+f_q)\,dx \geq \int_{[-q-1,q+1]\times\mathcal{D}} L_{0,K}(U^*)\,dx. \tag{4.36}$$

The choice of f_q and (4.32) show

$$\int_{T_{-q-1}\cup T_q} L_{0,K}(U^*+f_q)\,dx \leq \gamma(\sigma),$$

where $\gamma(\sigma) \to 0$ as $\sigma \to 0$. Recall that by Remark 2.44 and Proposition 4.16,

$$c_{0,K} \equiv \inf_{u\in\Gamma} J_{0,K}(u) = \inf_{u\in\Gamma} J(u) \equiv c.$$

Consequently, letting $\sigma \to 0$, $q \to \infty$ and (4.36) implies

$$c = J(u_0) \geq J_{0,K}(U^*) \geq c_{0,K}$$

and Theorem 4.23 is proved. □

Remark 4.37. One can also allow for perturbations of the domain in the setting of Theorem 4.23. For example, with a condition like:

(Ω_0) For some $\varepsilon_0 > 0$ and each $|\varepsilon| \leq \varepsilon_0$, there is a domain $\Omega_\varepsilon \subset \mathbb{R}^n$, where Ω_ε satisfies (Ω_1)–(Ω_2), the map $\varepsilon \to \Omega_\varepsilon$ is continuous, and $\Omega_0 = \mathbb{R}\times\mathcal{D}$.

To conclude this section, we will briefly mention the case of (V_2) replaced by:

(V_2') There are points $a^i \in \mathbb{R}^m$ such that $V(x,a^i)=0$, $1 \leq i \leq s$, for all $x \in \Omega$, and $V(x,u) > 0$ otherwise,

i.e., V is a multiwell potential. Existence and multiplicity results for such multiwell potentials and even infinite well potentials have been studied, e.g., in [Montecchiari and Rabinowitz 2016]. Using the methods of this paper, such treatments can readily be extended to the current setting. For example, suppose that V is an s-well potential and set

$$A = \{a^1, \dots, a^s\}.$$

Then it is straightforward to show:

Theorem 4.38. *Suppose that V satisfies (V_1), (V_2'), (V_3), (V_4), $V_u \in C^1$, $\Omega = \mathbb{R} \times \mathcal{D}$ with $\mathcal{D} \subset \mathbb{R}^{n-1}$ a bounded open set and $\partial\mathcal{D}$ a C^3 manifold. Then*:

(1) *For any $a^i \in A$, there exists an $a^j \in A$ with $i \neq j$ and corresponding classical solution $U_{i,j}$ of (PDE) and (BC) such that U_{ij} is heteroclinic in x_1 from a^i to a^j and $U_{i,j}$ minimizes J over the set*

$$\Big\{u \in E \,\Big|\, \lim_{k\to-\infty} \|u - a^i\|_{L^2(T_k,\mathbb{R}^m)} = \lim_{k\to\infty} \|u - a^j\|_{L^2(T_k,\mathbb{R}^m)} = 0 \text{ for some } j \neq i\Big\}.$$

(2) *For any $a^i, a^j \in A$, with $i \neq j$, there exists a (minimal) heteroclinic chain of solutions U_{i,p_1}, $U_{p_1,p_2}, \dots, U_{p_t,j}$ of (PDE) and (BC), where $U_{k,l}$ are as in (1) and the integers $i, p_1, \dots, p_t, j$ are distinct. Moreover, if*

$$c_{i,j} = \inf_{u\in\Gamma_{i,j}} J(u),$$

where

$$\Gamma_{i,j} = \Big\{u \in E \,\Big|\, \|u - a^i\|_{L^2(T_k,\mathbb{R}^m)} \to 0,\ k \to -\infty;\ \|u - a^j\|_{L^2(T_k,\mathbb{R}^m)} \to 0,\ k \to \infty\Big\},$$

then

$$c_{i,j} = J(U_{i,p_1}) + \cdots + J(U_{p_t,j}).$$

5. Multitransition solutions

In this section, it will be shown how the approach of [Montecchiari and Rabinowitz 2016] can be mirrored to construct multitransition homoclinic and heteroclinic solutions of (PDE). More precisely, we seek solutions of (PDE) that as a function of x_1 make multiple transitions between small neighborhoods of a^- and a^+. In order to find such solutions, we need a mild nondegeneracy condition on the set of minimizing heteroclinics given by Theorem 2.2. To make this precise, we replace Γ by $\Gamma(a^-, a^+)$ and c by $c(a^-, a^+)$. Thus interchanging the roles of a^- and a^+ gives us $\Gamma(a^+, a^-)$ and $c(a^+, a^-)$. For $\xi \in \{a^+, a^-\}$, and $\eta \in \{a^+, a^-\} \setminus \{\xi\}$, set

$$\mathcal{M}(\xi, \eta) \equiv \{u \in \Gamma(\xi, \eta) \mid J(u) = c(\xi, \eta)\}.$$

Define

$$\mathcal{S}(\xi, \eta) \equiv \{u|_{T_0} \mid u \in \mathcal{M}(\xi, \eta)\}$$

and put the $W^{1,2}(T_0, \mathbb{R}^m)$ topology on this set. Then we have:

Proposition 5.1. *Suppose V satisfies (V_1)–(V_4), $V_u \in C^1(\bar{\Omega} \times \mathbb{R}^m)$ and $\partial\Omega \in C^3$. Then*

(1) $\bar{\mathcal{S}}(\xi, \eta) = \mathcal{S}(\xi, \eta) \cup \{\xi\} \cup \{\eta\}$,

(2) $\bar{\mathcal{S}}(\xi, \eta)$ *is compact.*

Proof. Due to the asymptotic behavior of the members u of $\mathcal{M}(\xi,\eta)$, we know $u(x_1+j,x_2,\dots,x_n)$ converges in $L^2(T_0,\mathbb{R}^m)$ to η as $j\to\infty$ and to ξ as $j\to-\infty$. Then, by the L^∞ uniform boundedness of minimizers $U\in S(\xi,\eta)$ in Proposition 6.2 and elliptic estimates, we see that $\{\xi\}\cup\{\eta\}\in\bar{\mathcal{S}}(\xi,\eta)$.

Let $\{w_j\}$ be a sequence in $\mathcal{S}(\xi,\eta)$. Then the proof of (1)–(2) is complete if a subsequence of $\{w_j\}$ converges to a member of $\mathcal{S}(\xi,\eta)\cup\{\xi\}\cup\{\eta\}$. If a subsequence of w_j converges to ξ or η, we are done. Thus suppose this is not the case. For any j, we have $w_j=W_j|_{T_0}$, where $W_j\in\mathcal{M}(\xi,\eta)$, so $J(W_j)=c(\xi,\eta)$. By Proposition 6.2 and elliptic estimates, there exists $K>0$ such that $\|W_j\|_{C^{2,\alpha}(\Omega,\mathbb{R}^m)}\le K$. Then, a subsequence of W_j converges in $C^2_{\mathrm{loc}}(\Omega,\mathbb{R}^m)$ to a function $W\in E\cap C^2(\Omega,\mathbb{R}^m)$ and W is a classical solution of (PDE). In particular $w_j\to w=W|_{T_0}\ne\xi,\eta$. Since for each $p<q\in\mathbb{Z}$,

$$\sum_{i=p}^{q}\int_{T_i}L(W)\,dx=\lim_{j\to\infty}\sum_{i=p}^{q}\int_{T_i}L(W_j)\,dx\le J(W_j)=c(\xi,\eta),$$

letting $q,-p\to\infty$ shows

$$\sum_{i\in\mathbb{Z}}\int_{T_i}L(W)\,dx=J(W)\le c(\xi,\eta).\tag{5.2}$$

Equation (5.2) and (V_2) imply there are points $\xi^\pm\in\{\xi,\eta\}$ such that $\|W-\xi^\pm\|_{L^2(T_i,\mathbb{R}^m)}\to0$ as $i\to\pm\infty$, respectively. We must show $\xi^-=\xi$ and $\xi^+=\eta$. Arguing indirectly, suppose that $\xi^-\ne\xi$, so $\xi^-=\eta$. Let $\varepsilon>0$. Then there is a negative integer $i_0=i_0(\varepsilon)\in\mathbb{Z}$ such that $\|W-\eta\|_{W^{1,2}(T_i,\mathbb{R}^m)}\le\varepsilon$ for all $i\le i_0+2$. For large $k=k(i_0)$ and $i\in\{i_0-1,\dots,i_0+2\}$, we have $\|W_k-\eta\|_{W^{1,2}(T_i,\mathbb{R}^m)}\le2\varepsilon$. Define $f_k\in\Gamma(\xi,\eta)$ by

$$f_k=\begin{cases}W_k & \text{for } x_1\le i_0-1,\\ (x_1-i_0+1)\eta+(i_0-x_1)W_k & \text{for } i_0-1\le x_1\le i_0,\\ \eta & \text{for } i_0\le x_1\le i_0+1,\\ (x_1-i_0-1)W_k+(2+i_0-x_1)\eta & \text{for } i_0+1\le x_1\le i_0+2,\\ W_k & \text{for } i_0+2\le x_1.\end{cases}\tag{5.3}$$

Note that

$$|J(W_k)-J(f_k)|=\left|\int_{\bigcup_{i=i_0-1}^{i_0+2}T_i}(L(W_k)-L(f_k))\,dx\right|\le\kappa(\varepsilon),\tag{5.4}$$

where $\kappa(\varepsilon)\to0$ as $\varepsilon\to0$. Therefore by (5.4),

$$c(\xi,\eta)=J(W_k)\ge J(f_k)-\kappa(\varepsilon).\tag{5.5}$$

Further define functions $g_k\in\Gamma(\xi,\eta)$ and h_k via

$$g_k=\begin{cases}f_k & \text{for } x_1\le i_0,\\ \eta & \text{for } x_1\ge i_0,\end{cases}\tag{5.6}$$

$$h_k=\begin{cases}\eta & \text{for } x_1\le i_0+1,\\ f_k & \text{for } x_1\ge i_0+1,\end{cases}\tag{5.7}$$

so by construction,

$$J(f_k)=J(g_k)+J(h_k).\tag{5.8}$$

By (5.5) and (5.8),

$$c(\xi,\eta) \geq J(g_k) + J(h_k) - \kappa(\varepsilon) \geq c(\xi,\eta) + J(h_k) - \kappa(\varepsilon).$$

Thus, we get

$$\kappa(\varepsilon) \geq J(h_k) \geq \sum_{i=i_0+1}^{\infty} \int_{T_i} L(f_k)\,dx \geq \sum_{i=0}^{\infty} \int_{T_i} L(f_k)\,dx, \tag{5.9}$$

where the last inequality follows since i_0 is negative. But on T_0, we have $f_k = w_k \to w$ in $W^{1,2}(T_0,\mathbb{R}^m)$ as $k \to \infty$ and $w \neq a_\pm$. Therefore $\int_{T_0} L(f_k)\,dx \geq \omega > 0$ for all large k. Since the left-hand side of (5.9) goes to 0 as $\varepsilon \to 0$, we have a contradiction. Thus $\xi^- = \xi$. Similarly, $\xi^+ = \eta$ and the proposition is proved. □

Next define $\mathcal{C}_\xi(\xi,\eta)$ to be the connected component of $\bar{\mathcal{S}}(\xi,\eta)$ to which ξ belongs and define $\mathcal{C}_\eta(\xi,\eta)$ similarly. Then the following alternative holds.

Proposition 5.10. *One of the following items holds*:

(i) $\mathcal{C}_\xi(\xi,\eta) = \mathcal{C}_\eta(\xi,\eta)$;

(ii) $\mathcal{C}_\xi(\xi,\eta) = \{\xi\}$ *and* $\mathcal{C}_\eta(\xi,\eta) = \{\eta\}$.

If (ii) *holds, there exist nonempty disjoint compact sets* $K_\xi(\xi,\eta),\ K_\eta(\xi,\eta) \subset \bar{\mathcal{S}}(\xi,\eta)$ *such that*

(a) $\xi \in K_\xi(\xi,\eta),\quad \eta \in K_\eta(\xi,\eta)$,

(b) $\bar{\mathcal{S}}(\xi,\eta) = K_\xi(\xi,\eta) \cup K_\eta(\xi,\eta)$,

(c) $\operatorname{dist}(K_\xi(\xi,\eta), K_\eta(\xi,\eta)) \equiv 5r(\xi,\eta) > 0$.

Proof. The proofs of these statements are exactly the same as their counterparts in Proposition 2.43 of [Montecchiari and Rabinowitz 2016]. □

Remark 5.11. Note that Proposition 5.10(i) occurs if V is independent of x_1.

To continue, we assume that the nondegeneracy condition, alternative (ii) of Proposition 5.10, holds for both $\mathcal{C}_\xi(\xi,\eta)$ and $\mathcal{C}_\xi(\eta,\xi)$. Since the arguments are very close to those of [Montecchiari and Rabinowitz 2016], we will give the proof for the simplest case of two transition solutions and merely set up the variational problem that finds the multitransition solutions as local minima of J, referring to [Montecchiari and Rabinowitz 2016] for further results and details.

Recalling the definition of ρ given after (2.12), by Proposition 5.10,

$$\bar{r} = \min\big(\rho, r(a_-,a_+), r(a_+,a_-)\big) > 0.$$

Define the set

$$\Lambda(\xi,\eta) = \big\{u \in \Gamma(\xi,\eta) \;\big|\; \|u - K_\xi(\xi,\eta)\|_{W^{1,2}(T_0,\mathbb{R}^m)} = \bar{r} \text{ or } \|u - K_\eta(\xi,\eta)\|_{W^{1,2}(T_0,\mathbb{R}^m)} = \bar{r}\big\}$$

and

$$d(\xi,\eta) = \inf_{u \in \Lambda(\xi,\eta)} J(u). \tag{5.12}$$

Arguing as in the proof of Proposition 2.47 of [Montecchiari and Rabinowitz 2016] shows

$$d(\xi,\eta) > c(\xi,\eta). \tag{5.13}$$

To set up the variational framework to find the simplest two transition solutions of (PDE) and (BC), following [Montecchiari and Rabinowitz 2016], let $\boldsymbol{m} = (m_1,\dots,m_4) \in \mathbb{Z}^4$ and $l \in \mathbb{N}$ be such that

$$m_1 + 2l < m_2 - 2l < m_2 + 2l < m_3 - 2l < m_3 + 2l < m_4 - 2l.$$

Finally define

$$\mathcal{A}_2 = \mathcal{A}_2(\boldsymbol{m}, l) = \{u \in E \mid u \text{ satisfies (5.14)}\},$$

where

$$u(\cdot + je_1)|_{T_0} \in \begin{cases} N_r(K_{a_-}(a_-,a_+)), & j < m_1 + l, \\ N_r(K_{a_+}(a_-,a_+)), & m_2 - l \le j < m_2 + l, \\ N_r(K_{a_+}(a_+,a_-)), & m_3 - l \le j < m_3 + l, \\ N_r(K_{a_-}(a_+,a_-)), & m_4 - l \le j. \end{cases} \tag{5.14}$$

Here $N_r(A) \equiv \{u \in W^{1,2}(T_0,\mathbb{R}^m) \mid \operatorname{dist}_{W^{1,2}(T_0,\mathbb{R}^m)}(u,A) \le r\}$ for any $A \subset W^{1,2}(T_0\mathbb{R}^m)$.

We seek 2-transition solutions as minima of J on $\mathcal{A}_2$. Define

$$b_2 = b_2(\boldsymbol{m}, l) = \inf_{u \in \mathcal{A}_2} J(u). \tag{5.15}$$

Theorem 5.16. *Suppose* (V_1)–(V_4) *are satisfied and that Proposition 5.10*(ii) *holds for* $\mathcal{C}_\xi(\xi,\eta)$ *whenever* $\xi \ne \eta \in \{a_-,a_+\}$. *There exists an* $m_0 \in \mathbb{N}$ *such that if* $l \ge m_0$ *and* $m_{i+1} - m_i - 6l \ge m_0$ *for* $i = 1,2,3$, *then*

$$\mathcal{M}(b_2) \equiv \{u \in \mathcal{A}_2 \mid J(u) = b_2\} \ne \varnothing.$$

Moreover, any $U \in \mathcal{M}(b_2)$ *is a classical solution of* (PDE) *satisfying* (BC) *and* $\|U - a_-\|_{W^{1,2}(T_p,\mathbb{R}^m)} \to 0$ *as* $p \to \pm\infty$.

Proof. Let $\{u_k\} \subset \mathcal{A}_2$ be such that $J(u_k) \to b_2$. Arguments similar to the ones used to prove Propositions 2.14 and 2.24 show that $\{\|u_k\|_{W^{1,2}(T_i,\mathbb{R}^m)}\}_{i\in\mathbb{Z},k\in\mathbb{N}}$ is bounded. Then, along a subsequence (denoted again by $\{u_k\}$), $u_k \to U$ weakly in E. Since $\mathcal{A}_2$ is weakly closed, we have $U \in \mathcal{A}_2$ and J is weakly lower semicontinuous, so $J(U) = b_2$. Since $J(U) < +\infty$,

$$\operatorname{dist}_{W^{1,2}(T_p,\mathbb{R}^m)}(U,\{a_-,a_+\}) \to 0 \quad \text{as } p \to \pm\infty,$$

and by the definition of $\mathcal{A}_2$, it follows that $\lim_{\pm\to\infty} \|U - a_-\|_{W^{1,2}(T_p,\mathbb{R}^m)} = 0$. To show that U is a classical solution of (PDE) satisfying (BC), the arguments of Section 3 can be applied here once we have verified that U is a weak solution of (PDE), i.e.,

$$\int_\Omega \nabla U \cdot \nabla\varphi + V_u(x,U)\cdot\varphi\,dx = 0 \quad \text{for any } \varphi \in C_0^\infty(\mathbb{R}^n,\mathbb{R}^m). \tag{5.17}$$

To verify (5.17), it suffices to show that if m_0 is large enough, U satisfies the inequalities defining $\mathcal{A}_2$ with strict inequalities. Towards this end, define $K_1 = K_{a_-}(a_-,a_+)$, $K_2 = K_{a_+}(a_-,a_+)$, $K_3 = K_{a_+}(a_+,a_-)$, $K_4 = K_{a_-}(a_+,a_-)$, and $a_1 = a_4 = a_-$, $a_2 = a_3 = a_+$. If U does not satisfy one of the inequalities defining $\mathcal{A}_2$ with strict inequality, then

(∗) there exist integers $j \in \{1, \dots, 4\}$ and

$$p_j \in \begin{cases} (-\infty, m_1 + l] \cap \mathbb{Z} & \text{if } j = 1, \\ [m_j - l, m_j + l) \cap \mathbb{Z} & \text{if } 1 < j < 4, \\ [m_4 - l, +\infty) \cap \mathbb{Z} & \text{if } j = 4, \end{cases}$$

for which

$$\bar{r} = \operatorname{dist}_{W^{1,2}(T_0, \mathbb{R}^m)}(U(\cdot + p_j e_1)|_{T_0}, K_j).$$

We show here below how (∗) is not possible if m_0 is large enough. The arguments are slightly different depending on whether $j = 1, 4$ (the "boundary" case) or $j = 2, 3$ (the "interior" case). We will show here how to get a contradiction only for the interior case, the other being very similar (and simpler).

Assume that for some $\bar{j} \in \{2, 3\}$ there exists a $p \in [m_{\bar{j}} - l, m_{\bar{j}} + l)$ such that

$$\bar{r} = \operatorname{dist}_{W^{1,2}(T_0, \mathbb{R}^m)}(U(\cdot + p e_1)|_{T_0}, K_{\bar{j}}). \tag{5.18}$$

Let $\varepsilon \in (0, \bar{r})$. First note that if m_0 is sufficiently large then

$$b_2 < c(a_-, a_+) + c(a_+, a_-) + 2 \tag{5.19}$$

independently of the choice on $\boldsymbol{m}$ — see the simple argument at the beginning of the proof of Proposition 3.29 in [Montecchiari and Rabinowitz 2016]. Moreover, taking m_0 larger if need be, observe that for any $j \in \{1, \dots, 4\}$ there exists $\iota_j \in [m_j - l + 2, m_j + l - 2) \cap \mathbb{Z}$ such that

$$\|U - a_j\|_{W^{1,2}(T_i, \mathbb{R}^m)} < \varepsilon \quad \text{for } i \in [\iota_j - 2, \iota_j + 2] \cap \mathbb{Z}. \tag{5.20}$$

Indeed, suppose for every $X_i \equiv \bigcup_{k=i-2}^{i+2} T_k \subset [m_j - l, m_j + l] \times \mathcal{D}$, there exists $T_j \subset X_i$ such that $\|U - a_j\|_{W^{1,2}(T_j, \mathbb{R}^m)} \geq \varepsilon$. Since $U \in \mathcal{A}_2$, it follows that $\operatorname{dist}_{W^{1,2}(T_j, \mathbb{R}^m)}(U, \{a_-, a_+\}) \geq \varepsilon$. Then, the argument in Proposition 2.14 shows $\int_{T_j} L(U)\, dx \geq \beta(\varepsilon) > 0$. Therefore

$$b_2 = J_1(U) \geq \tfrac{1}{5}(2l + 1)\beta(\varepsilon) \geq \tfrac{2}{5} m_0 \beta(\varepsilon),$$

which is in contradiction with (5.19) for large values of m_0.

By (5.20), there are integers $i_- \in (m_{\bar{j}-1} - l + 2, m_{\bar{j}-1} + l - 2)$ and $i_+ \in (m_{\bar{j}+1} - l + 2, m_{\bar{j}+1} + l - 2)$ and corresponding regions X_{i_-} and X_{i_+} such that if $T_l \subset X_{i_-}$ and $T_k \subset X_{i_+}$, then

$$\|U - a_{\bar{j}-1}\|_{W^{1,2}(T_l, \mathbb{R}^m)} < \varepsilon \quad \text{and} \quad \|U - a_{\bar{j}+1}\|_{W^{1,2}(T_k, \mathbb{R}^m)} < \varepsilon. \tag{5.21}$$

Define

$$f = \begin{cases} a_{\bar{j}-1} & \text{for } x_1 \leq i_-, \\ U & \text{for } i_- + 1 \leq x_1 \leq i_+ - 1, \\ a_{\bar{j}+1} & \text{for } i_+ \leq x_1, \end{cases} \tag{5.22}$$

with interpolations as in (5.3) in the other regions.

By construction, $f \in \Gamma(a_{\bar{j}-1}, a_{\bar{j}+1})$ and since $f = U$ on T_p, by (5.18) we have $f \in \Lambda(a_{\bar{j}-1}, a_{\bar{j}+1})$. Then, by (5.13) and (5.22),

$$\begin{aligned} d(a_{\bar{j}-1}, a_{\bar{j}+1}) &\le J(f) \\ &\le \int_{\bigcup_{i_-+1}^{i_+-2} T_i} L(U)\,dx + \int_{T_{i_-}} L(f)\,dx + \int_{T_{i_+-1}} L(f)\,dx \\ &\le \int_{\bigcup_{i_-+1}^{i_+-2} T_i} L(U)\,dx + 2\kappa(\varepsilon). \end{aligned} \tag{5.23}$$

If m_0 is large enough, there exists $u \in \mathcal{M}(a_{\bar{j}-1}, a_{\bar{j}+1})$ such that

$$\begin{aligned} \|u - a_{\bar{j}-1}\|_{W^{1,2}(T_q,\mathbb{R}^m)} &\le \varepsilon \quad \text{for any } q \le m_{\bar{j}-1} + l, \\ \|u - a_{\bar{j}+1}\|_{W^{1,2}(T_q,\mathbb{R}^m)} &\le \varepsilon \quad \text{for any } q \ge m_{\bar{j}+1} - l. \end{aligned}$$

Define

$$\Phi = \begin{cases} U & \text{for } x_1 \le i_- - 2, \\ a_{\bar{j}-1} & \text{for } i_- - 1 \le x_1 \le i_-, \\ u & \text{for } i_- + 1 \le x_1 \le i_+ - 1, \\ a_{\bar{j}+1} & \text{for } i_+ \le x_1 \le i_+ + 1, \\ U & \text{for } x_1 \ge i_+ + 2, \end{cases} \tag{5.24}$$

making the usual interpolations in the remaining regions. Observe that $\Phi \in \mathcal{A}_2$. Consequently, with the aid of (5.23), we obtain

$$\begin{aligned} 0 \le J(\Phi) - J(U) &= \int_{\bigcup_{i=i_--2}^{i_++1} T_i} L(\Phi) - L(U)\,dx \\ &\le \int_{\bigcup_{i=i_-+1}^{i_+-1} T_i} L(u)\,dx + 2\kappa(\varepsilon) - \int_{\bigcup_{i=i_--2}^{i_++1} T_i} L(U)\,dx \\ &\le c(a_{\bar{j}-1}, a_{\bar{j}+1}) - d(a_{\bar{j}-1}, a_{\bar{j}+1}) + 4\kappa(\varepsilon), \end{aligned}$$

a contradiction to (5.13) if $4\kappa(\varepsilon) < d(a_{\bar{j}-1}, a_{\bar{j}+1}) - c(a_{\bar{j}-1}, a_{\bar{j}+1})$. An analogous argument leads to a contradiction in the boundary case. Thus $(*)$ cannot occur and the theorem is proved. □

Remark 5.25. Varying the values of $\boldsymbol{m}$, Theorem 5.16 provides the existence of infinitely many 2-transition solutions of (PDE) homoclinic to a_-. Reversing the roles of a_- and a_+, an analogous result is obtained giving infinitely many solutions homoclinic to a_+.

As in [Montecchiari and Rabinowitz 2016], Theorem 5.16 can be generalized also to the case of k-transition and infinite transition solutions. We state here the case of k-transition solutions referring to [Montecchiari and Rabinowitz 2016] for more details.

For $k \in \mathbb{N}$, let $\{a_1, \dots, a_{2k}\} \in \{a_-, a_+\}^{2k}$ be such that

$$a_1 \ne a_2 = a_3 \ne \cdots \ne a_{2k-2} = a_{2k-1} \ne a_{2k}.$$

Consider also the family of sets $\{K_1, \dots, K_{2k}\}$ defined as

$$K_{2j-1} = K_{a_{2j-1}}(a_{2j-1}, a_{2j}) \quad \text{and} \quad K_{2j} = K_{a_{2j}}(a_{2j-1}, a_{2j}), \quad j = 1, \dots, k.$$

Given $l \in \mathbb{N}$ and $\boldsymbol{m} = (m_1, \ldots, m_{2k}) \in \mathbb{Z}^{2k}$ with $m_j - m_{j-1} > 2l$ for $j = 2, \ldots, 2k$, consider the set

$$\mathcal{A}(k, \boldsymbol{m}, l) = \{u \in E \mid u \text{ satisfies (5.26)}\},$$

where

$$u(\cdot + pe_1)|_{T_0} \in \begin{cases} N_{\bar{r}}(K_1), & p \in (-\infty, m_1 + l) \cap \mathbb{Z}, \\ N_{\bar{r}}(K_j), & p \in [m_j - l, m_j + l) \cap \mathbb{Z},\ 2 \leq j \leq 2k-1, \\ N_{\bar{r}}(K_{2k}), & p \in [m_{2k} - l, +\infty) \cap \mathbb{Z}, \end{cases} \tag{5.26}$$

and let

$$b_k = b(k, \boldsymbol{m}, l) = \inf_{u \in \mathcal{A}(k, \boldsymbol{m}, l)} J(u). \tag{5.27}$$

Theorem 5.28. *Under the hypotheses of Theorem 5.16, there is an $m_0 \in \mathbb{N}$ for which if $k \in \mathbb{N}$, $l \geq m_0$ and $m_{i+1} - m_i - 6l \geq m_0$ for $i = 1, \ldots, 2k-1$, then*

$$\mathcal{M}(b_k) \equiv \big\{u \in \mathcal{A}(k, \boldsymbol{m}, l) \;\big|\; J(u) = b(k, \boldsymbol{m}, l)\big\} \neq \varnothing.$$

Moreover, any $U \in \mathcal{M}(b_k)$ is a classical solution of (PDE) satisfying (BC).

6. Proof of Theorem 3.1

In this section the proof of Theorem 3.1 will be carried out. It is similar to the proof of the corresponding scalar case. The proof consists of several steps. First note that since (V_1)–(V_3) are satisfied and $\partial\Omega = \mathbb{R} \times \partial\mathcal{D} \in C^1$, by Theorem 2.2, there exists a minimizer $U \in \Gamma$ of (2.1). For any $\varphi \in W^{1,2}_{\text{loc}}(\Omega)$ with compact support in $\bar{\Omega}$ and $t \in \mathbb{R}$, we see that $U + t\varphi \in \Gamma$. Since $V(x, \cdot) \in C^1(\mathbb{R}^m)$ for each $x \in \Omega$ and (V_4) holds, $\lim_{t\to 0}(J(U + t\varphi) - J(U))/t$ exists. Since $J(U) \leq J(U + t\varphi)$ for any $t \in \mathbb{R}$, we see that

$$\lim_{t \to 0} \frac{J(U + t\varphi) - J(U)}{t} = \int_\Omega \nabla U \cdot \nabla\varphi + V_u(x, U)\varphi \, dx = 0.$$

This implies that U is a weak solution of (PDE) and (BC).

Now two rather technical steps are required and will be stated as separate propositions. The first provides an L^∞ bound for any weak solution U of (PDE) and (BC). When $m = 1$, such results are well known; see, e.g., [Gilbarg and Trudinger 1983]. In general, they are not true for systems, but we will show that due to the semilinear structure of (PDE) and (V_4), variants of arguments in [Gilbarg and Trudinger 1983], that in turn go back to work of Moser, can be modified to treat the current setting.

Note that for any $U \in E$, there is a constant $M_4 > 0$ depending on U such that

$$\int_\Omega |\nabla U|^2 \, dx + \sup_{i \in \mathbb{Z}} \int_{T_i} |U|^2 \, dx \leq M_4. \tag{6.1}$$

Proposition 6.2. *Suppose V satisfies (V_1)–(V_4), and that $\partial\Omega = \mathbb{R} \times \partial\mathcal{D} \in C^1$. Then for any weak solution $U \in E$ of (PDE) and (BC), there exists a constant $M_5 > 0$ depending on U such that*

$$\|U\|_{L^\infty(\Omega, \mathbb{R}^m)} \leq M_5.$$

If U is a minimizer of (2.1), M_4 and M_5 are independent of U.

Proof. First observe that by Proposition 2.24, M_4 can be chosen independently of U if U is a minimizer of (2.1). Let $\eta \in C^1(\mathbb{R}, [0,1])$ have compact support. Then η extends to a C^1-function on Ω by defining $\eta(x_1, \ldots, x_n) = \eta(x_1)$. For each $\sigma > 0$ and $i = 1, \ldots, m$, define a function U_i^σ by $U_i^\sigma(x) = U_i(x)$ if $|U_i(x)| < \sigma$, by $U_i^\sigma(x) = \sigma$ if $U_i(x) \geq \sigma$ and by $U_i^\sigma(x) = -\sigma$ if $U_i(x) \leq -\sigma$. If $U = (U_1, \ldots, U_m)$, set $U^\sigma = (U_1^\sigma, \ldots, U_m^\sigma)$. Let $\beta > 0$ and take $\varphi_j = \eta^2 U_j |U_j^\sigma|^{2\beta}$, $1 \leq j \leq m$. Then, taking $\varphi = \varphi_j e_j$, with e_j the j-th unit vector in $\mathbb{R}^m$, we see that $\varphi \in W^{1,2}_{\mathrm{loc}}(\Omega)$ and the support of φ is compact. Thus, (1.1) implies that for $1 \leq j \leq m$,

$$\int_\Omega \nabla U_j \cdot \nabla(\eta^2 U_j |U_j^\sigma|^{2\beta}) + V_{u_j}(x, U)\eta^2 U_j |U_j^\sigma|^{2\beta}\, dx = 0. \tag{6.3}$$

Note that

$$\begin{aligned}\nabla U_j &\cdot \nabla(\eta^2 U_j |U_j^\sigma|^{2\beta}) \\ &= \eta^2 |U_j^\sigma|^{2\beta} |\nabla U_j|^2 + 2\beta\eta^2 U_j |U_j^\sigma|^{2\beta-1} \nabla U_j \cdot \nabla |U_j^\sigma| + 2\eta U_j |U_j^\sigma|^{2\beta} \nabla U_j \cdot \nabla\eta. \end{aligned}\tag{6.4}$$

Observing that the middle term on the right in (6.4) satisfies

$$2\beta\eta^2 U_j |U_j^\sigma|^{2\beta-1} \nabla U_j \cdot \nabla |U_j^\sigma| \geq 0, \tag{6.5}$$

substituting (6.4)–(6.5) in (6.3) and using (V_4) shows for some constant $C_2 > 0$, independent of σ, j, β,

$$\int_\Omega \eta^2 |U_j^\sigma|^{2\beta} |\nabla U_j|^2\, dx \leq 2\int_\Omega \eta |U_j| |U_j^\sigma|^{2\beta} |\nabla U_j| |\nabla\eta|\, dx + C_2 \int_\Omega \eta^2 |U_j| |U_j^\sigma|^{2\beta}(1 + |U|^p)\, dx. \tag{6.6}$$

Simplifying the right-hand side of (6.6) gives

$$\begin{aligned}\int_\Omega &\eta^2 |U_j^\sigma|^{2\beta} |\nabla U_j|^2\, dx \\ &\leq \frac{1}{2}\int_\Omega \eta^2 |U_j^\sigma|^{2\beta} |\nabla U_j|^2\, dx + 8\int_\Omega |U_j|^2 |U_j^\sigma|^{2\beta} |\nabla\eta|^2\, dx + C_2 \int_\Omega \eta^2 |U_j^\sigma|^{2\beta}(|U_j| + |U|^{p+1})\, dx.\end{aligned}$$

Hence there is a constant $C_3 > 0$, independent of σ, j, β, such that

$$\begin{aligned}\int_\Omega &\eta^2 |U_j^\sigma|^{2\beta} |\nabla U_j|^2\, dx \\ &\leq C_3 \int_{\Omega\cap\operatorname{supp}(\eta)} \left[(1 + |\nabla\eta|_{L^\infty})^2 (|U_j| + |U_j|^2)|U_j^\sigma|^{2\beta} + \eta^2 |U_j^\sigma|^{2\beta} |U|^{p+1}\right] dx.\end{aligned}\tag{6.7}$$

Since

$$\left|\nabla(\eta U_j |U_j^\sigma|^\beta)\right|^2 \leq 2(\beta+1)^2 \eta^2 |U_j^\sigma|^{2\beta} |\nabla U_j|^2 + 2(U_j)^2 |U_j^\sigma|^{2\beta} |\nabla\eta|^2,$$

using this estimate in (6.7) shows there is a constant $C_4 > 0$, independent of σ, j, β, such that

$$\begin{aligned}\int_\Omega &|\nabla(\eta U_j |U_j^\sigma|^\beta)|^2 + \eta^2 (U_j)^2 |U_j^\sigma|^{2\beta}\, dx \\ &\leq C_4(\beta+1)^2 \int_{\Omega\cap\operatorname{supp}(\eta)} \left[(1 + |\nabla\eta|_{L^\infty})^2 (|U_j| + |U_j|^2)|U_j^\sigma|^{2\beta} + \eta^2 |U_j^\sigma|^{2\beta} |U|^{p+1}\right] dx.\end{aligned}\tag{6.8}$$

Due to the Sobolev inequality and (6.8), there exists a constant $C_5 > 0$, independent of σ, j, β, such that

$$\begin{aligned}\left(\int_\Omega (\eta |U_j^\sigma|^{\beta+1})^{\frac{2n}{n-2}}\,dx\right)^{\frac{n-2}{n}} &\\ \le C_5(\beta+1)^2 &\int_{\Omega\cap \operatorname{supp}(\eta)} \left[(1+|\nabla\eta|_{L^\infty})^2 |U_j|^2 |U_j^\sigma|^{2\beta} + \eta^2 |U_j^\sigma|^{2\beta}|U|^{p+1}\right] dx \\ \equiv M(j,\sigma,\beta). &\end{aligned} \tag{6.9}$$

Suppose for the moment that $n \ge 3$. Define β_1 by

$$2(\beta_1+1)+p-1 = \frac{2n}{n-2}.$$

The restriction on p in (V_4) implies $\beta_1 > 0$. By this choice of β_1, (6.1) and the Sobolev inequality, $M(j,\sigma,\beta_1)$ is bounded independently of σ. Consequently letting $\sigma \to \infty$ and choosing η so that $\eta(x_1) = 1$ for $|x_1 - i| \le l$, (6.9) shows there is a constant $K = K(\beta_1, l)$ such that for each $l \ge 1$,

$$\int_{\Omega\cap\{|x_1-i|\le l\}} (|U|^{\beta_1+1})^{\frac{2n}{n-2}}\,dx \le K(\beta_1,l) \tag{6.10}$$

independently of i. For $t \in \mathbb{N}$ with $t \ge 2$, define β_t via

$$2(\beta_t+1)+p-1 = 2(\beta_{t-1}+1)\frac{n}{n-2}$$

and repeat the above argument, obtaining

$$\int_{\Omega\cap\{|x_1-i|\le l\}} (|U|^{\beta_t+1})^{\frac{2n}{n-2}}\,dx \le K(\beta_{t-1}, l+t-1) \tag{6.11}$$

independently of i. Since $\beta_{t+1} - \beta_t = n/(n-2)(\beta_t - \beta_{t-1})$ and $\beta_2 - \beta_1 > 0$, it follows that $\beta_t \to \infty$ as $t \to \infty$. Thus for each fixed $q > 0$ and $l \ge 1$,

$$\int_{\Omega\cap\{|x_1-i|\le l\}} |U|^q\,dx$$

is bounded independently of i, the bound depending on q, n, M_4 and the constants in (V_4).

Now the Moser iteration argument will be used to get the L^∞ bound of the proposition. Returning to (6.9), our above observations show there is a constant $C_6 > 0$, independent of η and $\beta > 0$, such that

$$\begin{aligned}\left(\int_\Omega (\eta^2 |U|^{(2\beta+2)})^{\frac{n}{n-2}}\,dx\right)^{\frac{n-2}{n}} &\\ \le C_6(\beta+1)^2 &\int_{\Omega\cap\operatorname{supp}(\eta)} \left[(1+|\nabla\eta|_{L^\infty})^2(1+|U|)|U|^{2\beta+1} + \eta^2|U|^{2\beta+2+p-1}\right] dx.\end{aligned} \tag{6.12}$$

Consider the last term in (6.12). Let $h > 0$. Note that

$$\int_{\Omega\cap\operatorname{supp}(\eta)} \eta^2 |U|^{2\beta+2+p-1}\,dx = \int_{R_1} \eta^2|U|^{2\beta+2+p-1}\,dx + \int_{R_2}\eta^2|U|^{2\beta+2+p-1}\,dx \equiv I_1 + I_2, \tag{6.13}$$

where $R_1 = \{x \in \Omega \cap \mathrm{supp}(\eta) \mid |U| \le h\}$ and $R_2 = \{x \in \Omega \cap \mathrm{supp}(\eta) \mid |U| > h\}$. Then

$$I_1 \le h^{p-1} \int_\Omega \eta^2 |U|^{2\beta+2}\, dx. \tag{6.14}$$

By Hölder's inequality,

$$I_2 \le \left(\int_{R_2} |U|^{n(p-1)/2}\, dx \right)^{\frac{2}{n}} \left(\int_\Omega (\eta^2 |U|^{2\beta+2})^{\frac{n}{n-2}}\, dx \right)^{\frac{n-2}{n}} \equiv I_3 I_4. \tag{6.15}$$

Noting that $\frac{1}{2}n(p-1) < 2n/(n-2)$ and setting $d = \frac{1}{4}(p-1)(n-2)$, another application of Hölder's inequality implies

$$I_3^{\frac{n}{2}} \le |R_2|^{1-d} \left(\int_{R_2} |U|^{2n/(n-2)}\, dx \right)^d. \tag{6.16}$$

Since

$$|R_2| = \left| \{x \in \Omega \cap \mathrm{supp}(\eta) \mid |U| > h\} \right| \le h^{-2n/(n-2)} \int_{\Omega \cap \mathrm{supp}(\eta)} |U|^{2n/(n-2)}\, dx,$$

(6.16) can be rewritten as

$$I_3 \le h^{-\frac{4(1-d)}{n-2}} \left(\int_{\Omega \cap \mathrm{supp}(\eta)} |U|^{\frac{2n}{n-2}}\, dx \right)^{\frac{2}{n}}. \tag{6.17}$$

Combining (6.13)–(6.17), (6.12) becomes

$$\begin{aligned} &\left(\int_\Omega (\eta^2 |U|^{2\beta+2})^{\frac{n}{n-2}}\, dx \right)^{\frac{n-2}{n}} \\ &\quad \le C_6(\beta+1)^2 \int_{\Omega \cap \mathrm{supp}(\eta)} (1 + |\nabla \eta|_{L^\infty})^2 (|U|^{2\beta+1} + |U|^{2\beta+2})\, dx \\ &\qquad + C_6(\beta+1)^2 \Bigg[h^{p-1} \int_\Omega \eta^2 |U|^{2\beta+2}\, dx \\ &\qquad\qquad + h^{-\frac{4(1-d)}{n-2}} \left(\int_{\Omega \cap \mathrm{supp}(\eta)} |U|^{\frac{2n}{n-2}}\, dx \right)^{\frac{2}{n}} \left(\int_\Omega (\eta^2 |U|^{2\beta+2})^{\frac{n}{n-2}}\, dx \right)^{\frac{n-2}{n}} \Bigg]. \end{aligned} \tag{6.18}$$

Using the freedom in the choice of h, we require that

$$C_6(\beta+1)^2 h^{-\frac{4(1-d)}{n-2}} \left(\int_{\Omega \cap \mathrm{supp}(\eta)} |U|^{\frac{2n}{n-2}}\, dx \right)^{\frac{2}{n}} = \frac{1}{2}$$

or equivalently

$$h = \left(2C_6(\beta+1)^2 \left(\int_{\Omega \cap \mathrm{supp}(\eta)} |U|^{\frac{2n}{n-2}}\, dx \right)^{\frac{2}{n}} \right)^{\frac{n-2}{4(1-d)}}.$$

This makes the coefficient of the last integral term in (6.18) equal to $\frac{1}{2}$ so it can be absorbed on the left-hand side of the inequality (6.18). Thus (6.18) becomes

$$\begin{aligned}&\left(\int_\Omega (\eta^2|U|^{(2\beta+2)})^{\frac{n}{n-2}}\,dx\right)^{\frac{n-2}{n}}\\&\quad\le 2C_6(\beta+1)^2\int_{\Omega\cap\operatorname{supp}(\eta)}(1+\|\nabla\eta\|_{L^\infty})^2(|U|^{2\beta+1}+|U|^{2\beta+2})\,dx\\&\qquad+(2C_6(\beta+1)^2)^{1+\frac{(p-1)(n-2)}{4(1-d)}}\left(\int_{\Omega\cap\operatorname{supp}(\eta)}|U|^{\frac{2n}{n-2}}\,dx\right)^{\frac{2(n-2)(p-1)}{4n(1-d)}}\int_\Omega \eta^2|U|^{2\beta+2}\,dx.\end{aligned}\tag{6.19}$$

Now for each $i\in\mathbb{Z}$ and $j\in\mathbb{N}$, choose $\eta_j\in C^\infty(\mathbb{R})$ such that $\eta_j(x_1)=1$ for $|x_1-i|\le 1+2^{-j-1}$, $\eta_j(x_1)=0$ for $|x_1-i|\ge 1+2^{-j}$ and $\|\nabla\eta_j\|_{L^\infty}\le 2^{j+2}$. By (6.1) and the Sobolev inequality, there exists a constant $M'>0$, independent of $i\in\mathbb{Z}$ and $j\in\mathbb{N}$, such that

$$\int_{\Omega\cap\operatorname{supp}(\eta_j)}|U|^{\frac{2n}{n-2}}\,dx\le M'.$$

Set $\theta=n/(n-2)$ and for each $j\in\mathbb{N}$, define γ_j by $\gamma_j=2\theta^j$. Then, taking $\eta=\eta_j$ and $2\beta+2=\gamma_j$ in (6.19), simple estimates show there is a constant $C_7\ge 1$, independent of $j\in\mathbb{N}$ and $i\in\mathbb{Z}$, so that

$$\begin{aligned}&\left(\int_{\{x\in\Omega\,|\,|x_1-i|\le 1+2^{-j-1}\}}|U|^{2\theta^{j+1}}\,dx\right)^{\frac{n-2}{n}}\\&\quad\le C_7\Big((2\theta)^{2j}+\theta^{2j\left(1+\frac{(p-1)(n-2)}{4(1-d)}\right)}\Big)\int_{\{x\in\Omega\,|\,|x_1-i|\le 1+2^{-j}\}}(|U|^{2\theta^j-1}+|U|^{2\theta^j})\,dx\\&\quad\le C_7\Big((2\theta)+\theta^{1+\frac{(p-1)(n-2)}{4(1-d)}}\Big)^{2j}\int_{\{x\in\Omega\,|\,|x_1-i|\le 1+2^{-j}\}}(|U|^{2\theta^j-1}+|U|^{2\theta^j})\,dx.\end{aligned}\tag{6.20}$$

Thus setting

$$A_j\equiv\left(\int_{\{x\in\Omega\,|\,|x_1-i|\le 1+2^{-j}\}}|U|^{2\theta^j}\,dx\right)^{\frac{1}{2\theta^j}},$$

by the Hölder inequality,

$$\int_{\{x\in\Omega\,|\,|x_1-i|\le 1+2^{-j}\}}|U|^{2\theta^j-1}\,dx\le\Big|\{x\in\Omega\mid |x_1-i|\le 2\}\Big|^{\frac{1}{2\theta^j}}(A_j)^{2\theta^j-1}.\tag{6.21}$$

With the aid of (6.21), there is a constant $C_8\ge 1$, independent of $j\in\mathbb{N}$ and $i\in\mathbb{Z}$, such that (6.20) yields the simpler inequality

$$A_{j+1}\le (C_8)^{\frac{j}{2\theta^j}}\left(1+\frac{1}{A_j}\right)^{\frac{1}{2\theta^j}}A_j.\tag{6.22}$$

Since $\left(1+\frac{1}{A_j}\right)^{\frac{1}{2\theta^j}}\le 1+\frac{1}{2\theta^j}\frac{1}{A_j}$, (6.22) implies

$$A_{j+1}\le (C_8)^{\frac{j}{2\theta^j}}\left(A_j+\frac{1}{2\theta^j}\right).\tag{6.23}$$

This inequality can be further rewritten as

$$A_{j+1} \le (C_8)^{\frac{j}{2\theta^j}} A_j + \frac{C_9}{2\theta^j}, \tag{6.24}$$

where the constant $C_9 \ge 1$ is independent of $j \in \mathbb{N}$ and $i \in \mathbb{Z}$. Since $D \equiv \sum_{j\in\mathbb{N}} j/(2\theta^j) < \infty$, by (6.24),

$$\begin{aligned} A_{j+1} \le (C_8)^{\frac{j}{2\theta^j}} A_j + \frac{C_9}{2\theta^j} &\le (C_8)^{\frac{j}{2\theta^j}+\frac{j-1}{2\theta^{j-1}}} A_{j-1} + (C_8)^{\frac{j}{2\theta^j}} \frac{C_9}{2\theta^{j-1}} + \frac{C_9}{2\theta^j} \\ &\le \cdots \le (C_8)^D \Big(A_1 + C_9 \sum_{j=1}^{\infty} \theta^{-j} \Big) < \infty. \end{aligned} \tag{6.25}$$

Consequently there exists a constant $M_5 > 0$, independent of $i \in \mathbb{Z}$, such that for any weak solution U of (1.1),

$$\|U\|_{L^\infty(\{x\in\Omega \mid |x_1-i|\le 1\})} = \lim_{j\to\infty} A_j \le M_5$$

as claimed. Note that M_5 depends on n, M_4, and the constants in (V_4).

When $n = 2$, by the Sobolev inequality, for any $q > 2$, there exists a constant, C_{10}, depending on q but independent of i, such that

$$\|U\|_{L^q(\{x\in\Omega \mid |x_1-i|\le 2\})} \le C_{10} \|U\|_{W^{1,2}(\{x\in\Omega \mid |x_1-i|\le 2\})}. \tag{6.26}$$

Therefore the case of $n \ge 3$ can be simplified and modified by, e.g., replacing our earlier φ_j by $\eta^2 U_j |U_j|^{2(\beta+1)}$. This leads to a simpler version of (6.8). Then employing (6.26) in going from (6.8) to (6.9) leads to the following variant of (6.12)with q replacing $n/(n-2)$:

$$\Big(\int_\Omega (\eta |U_j|^{2(\beta+1)})^q \, dx \Big)^{\frac{1}{q}} \le C_{11} (\beta+1)^2 \int_{\Omega\cap \mathrm{supp}(\eta)} \big[(1+|\nabla\eta|_{L^\infty})^2 |U_j|^2 |U_j|^{2\beta} + \eta^2 |U_j|^{2\beta} |U|^{p+1} \big] \, dx$$

for any $q > 2$, where C_{11} depends on q. Then continuing as earlier completes the proof for this case. □

As the next step in the proof of Theorem 3.1, we have:

Proposition 6.27. *Suppose that V satisfies (V_1)–(V_4) and $\partial\Omega \in C^1$. If $U \in E$ is a weak solution of* (PDE) *and* (BC), *then*:

1. *For any $\Omega' \subset\subset \Omega$, we have $U \in W^{2,2}(\Omega')$.*
2. *If $V_u \in C^1(\Omega \times \mathbb{R}^m, \mathbb{R}^m)$, then $U \in C^{2,\alpha}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ for any $\alpha \in (0,1)$ and satisfies* (PDE) *in Ω.*
3. *If $\partial\Omega \in C^2$, then $U \in W^{2,2}_{\mathrm{loc}}(\overline{\Omega})$ and U is a strong solution of* (PDE) *and* (BC).

Proof. First since U is a weak solution of (PDE) and by Theorem 3.1, $V(x, U) \in L^2(\Omega', \mathbb{R}^m)$, (1) follows from Theorem 8.8 of [Gilbarg and Trudinger 1983]. Moreover, this additional differentiability shows U is a strong solution of (PDE). Next by Theorem 3.1 again, $V_u(\cdot, U) \in L^q(\Omega', \mathbb{R}^m)$ for any $q > 1$ so by Theorem 9.11 of [Gilbarg and Trudinger 1983], $U \in W^{2,q}(\Omega', \mathbb{R}^m)$. The Sobolev inequality then implies $U \in C^{1,\alpha}(\Omega', \mathbb{R}^m)$ for any $\alpha \in (0,1)$. Then, since $V_u(x,U) \in C^1(\Omega')$, invoking the linear Schauder theory then gives $U \in C^{2,\alpha}(\Omega', \mathbb{R}^m)$ and (2) holds. Lastly the proof of Theorem 4 in §6.3.2 of [Evans

1998] with the modification that $U \in W^{1,2}$ rather than $U \in W_0^{1,2}$ yields the first part of (3). For the second, taking any $\varphi \in W^{1,2}_{\mathrm{loc}}(\Omega)$ with compact support in $\overline{\Omega}$, by (2.36) and integration by parts due to the fact $U \in W^{2,2}_{\mathrm{loc}}(\overline{\Omega}) \hookrightarrow W^{1,2}_{\mathrm{loc}}(\partial\Omega)$, we get

$$\int_\Omega (-\Delta U + V_u(x,U)) \cdot \varphi \, dx - \int_{\partial\Omega} \frac{\partial U}{\partial \nu} \cdot \varphi \, dS = 0.$$

Thus since U satisfies (PDE),

$$\int_{\partial\Omega} \frac{\partial U}{\partial \nu} \cdot \varphi \, dS = 0 \tag{6.28}$$

for all $\varphi \in W^{1,2}_{\mathrm{loc}}(\Omega) \hookrightarrow L^2_{\mathrm{loc}}(\partial\Omega)$ having compact support and (6.28) implies (BC). Thus, U is a strong solution of (PDE) and (BC). □

Completion of proof of Theorem 3.1. It remains to show the regularity of U in a neighborhood of $\partial\Omega$ when $\partial\Omega \in C^3$. Since in Section 4 we consider a more general domain than $\Omega = \mathbb{R} \times \mathcal{D}$, the special nature of Ω will be suppressed here so that our argument also adapts easily to the case treated elsewhere. Let $\mathbb{R}^n_+ \equiv \{x \in \mathbb{R}^n \mid x_n > 0\}$ and $z \in \partial\Omega = \mathbb{R} \times \partial\mathcal{D}$. Slightly modifying the proof of Theorem 8.12 of [Gilbarg and Trudinger 1983], there exists a C^3 diffeomorphism Ψ defined on $B_R(z)$ such that $\Psi(B_R(z) \cap \Omega) \subset \mathbb{R}^n_+$, $\Psi(B_R(z) \cap \partial\Omega) \subset \partial\mathbb{R}^n_+$. Choose $\sigma < R$ and set $B^+ = B_\sigma(z) \cap \Omega$, $D' = \Psi(B_\sigma(z))$, and $D^+ = \Psi(B^+)$. Then setting $\Phi = \Psi^{-1}$ and $w \equiv U \circ \Phi$, (PDE) in B^+ is transformed into the equation

$$-\sum_{1 \le i,j \le n} a_{ij}(y) \frac{\partial^2 w}{\partial y_i \partial y_j} + \sum_{j=1}^n b_i(y) \frac{\partial w}{\partial y_i} + V_u(\Phi(y), w) = 0 \quad \text{in } D', \tag{6.29}$$

where

$$\begin{aligned} a_{ij}(y) &= \sum_{l=1}^n \frac{\partial \Psi_i}{\partial x_l}(\Phi(y)) \frac{\partial \Psi_j}{\partial x_l}(\Phi(y)), & 1 \le i, j \le n, \\ b_i(y) &= (\Delta \Psi_i)(\Phi(y)), & i = 1, \dots, n. \end{aligned}$$

Moreover, since $U \in W^{1,2}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$, we know $w \in W^{1,2}(D^+, \mathbb{R}^m)$.

Next we will show that an appropriate choice of Ψ, or equivalently of Φ, allows us to get the regularity of U near z and satisfy (BC). Translate and rotate variables for convenience so that z becomes 0 and ∂R^n_+ is the tangent space to $\partial\Omega$ at $z = 0$. Since $\partial\Omega$ is a C^3 manifold, for r small, there is a C^3 $\mathbb{R}$-valued map ϕ defined on $B_r(0) \cap \partial\mathbb{R}^n_+$ with $\phi(0) = 0 = |\nabla\phi(0)|$ and such that near 0, the boundary $\partial\Omega$ is given by

$$\{(y', \phi(y')) \mid y' \in B_r(0) \cap \partial\mathbb{R}^n_+\}.$$

Then, for $y = (y', y_n) = (y_1, \dots, y_n)$, extend ϕ to $\Phi = (\Phi_1, \dots, \Phi_n) : B_r(0) \to \mathbb{R}^n$ with $\Phi : B_r(0) \cap \overline{\mathbb{R}^n_+} \to \overline{\Omega}$ via

$$\Phi_j(y) = \begin{cases} y_j - y_n \dfrac{\partial \phi}{\partial y_j}(y') & \text{for } j = 1, \dots, n-1, \\ y_n + \phi(y') & \text{for } j = n. \end{cases}$$

This extension of ϕ makes Φ a C^2 function with $\Phi(0)=0$ and $\Phi'(0)=I$, the identity matrix. Thus Φ is a diffeomorphism in $B_r(0)$ for r small and

$$\frac{\partial \Phi}{\partial y_n}(y',0)=\left(-\frac{\partial \psi}{\partial y_1}(y'),\dots,-\frac{\partial \psi}{\partial y_{n-1}}(y'),1\right)$$

is the inward normal to $\partial\Omega$. Hence for small $r>0$ and an open neighborhood N of 0 in $\mathbb{R}^n$, the map $\Phi:\overline{\mathbb{R}^n_+}\cap B_r(0)\to N\cap\overline{\Omega}$ is a diffeomorphism. Let $\Psi:N\cap\overline{\Omega}\to\overline{\mathbb{R}^n_+}\cap B_r(0)$ be the inverse of the map Φ. Note that $\nabla\Psi_n(\Phi(y',0))$ is orthogonal to the surface $\{(y',\psi(y'))\mid y'\in B_r(0)\cap\partial\mathbb{R}^n_+\}\subset\partial\Omega$ since Ψ_n vanishes on the surface $\{(y',\psi(y'))\mid y'\in B_r(0)\cap\partial\mathbb{R}^n_+\}$. For $i\in\{1,\dots,n-1\}$, we know $\nabla\Psi_i(\Phi(y',0))$ is orthogonal to the surface

$$\left\{(y'-y_n\nabla\phi(y'),y_n+\phi(y'))\;\middle|\;y'\in B_r(0)\cap\partial\mathbb{R}^n_+,\ y_n\ge 0 \text{ and fixed } y_i\right\}$$

at $(y',\phi(y'))$. This implies that $\nabla\Psi_i(\Phi(y',0))$ is in the tangent space of $\partial\Omega$ at $\Phi(y',0)$. Thus for $i\in\{1,\dots,n-1\}$,

$$\nabla\Psi_n(\Phi(y',0))\cdot\nabla\Psi_i(\Phi(y',0))=0.$$

Hence $a_{in}=a_{ni}=0$ when $y_n=0$ and $i=1,\dots,n-1$. Now for $(y',y_n)\in B_r(0)$, we define

$$\begin{aligned}
\bar a_{ij}(y',y_n)&=a_{ij}(y',|y_n|) &&\text{if } i,j\le n-1,\\
\bar a_{in}(y',y_n)&=\frac{y_n}{|y_n|}a_{ij}(y',|y_n|) &&\text{if } 1\le i\le n-1,\\
\bar a_{nn}(y',y_n)&=a_{nn}(y',|y_n|).
\end{aligned}$$

We also define $b_i(y',y_n)=b_i(y',|y_n|)$ for $i=1,\dots,n-1$, and $b_n(y',y_n)=-b_n(y',y_n)$. For the solution w, we define a function $\bar w$ on $B_r(0)$ by $\bar w(y',y_n)=w(y',|y_n|)$. Then, we see that $\bar w$ is a strong solution of

$$-\sum_{1\le i,j\le n}\overline{a_{ij}}(y)\frac{\partial^2\bar w}{\partial y_i\partial y_j}+\sum_{j=1}^{n}\overline{b_i}(y)\frac{\partial \bar w}{\partial y_i}+V_u(\bar\Phi(y),\bar w)=0,\qquad y\in B_r(0).$$

Since $\overline{a_{ij}}$ is continuous and $\overline{b_i},V_u\in L^\infty$, Theorem 9.11 in [Gilbarg and Trudinger 1983] shows first that $\bar w\in W^{2,p}(B_{\frac{r}{2}}(0))$ for any $p>1$, and then $\bar w\in C^{1,\alpha}(B_{\frac{r}{2}}(0))$. This implies that $U\in C^{2,\alpha}(\Omega)\cap C^{1,\alpha}(\overline{\Omega})$.

Now, returning to the original equation (PDE) and applying Theorem 6.31 of [Gilbarg and Trudinger 1983], we get the regularity $U\in C^{2,\alpha}(\overline{\Omega})$.

It remains to show

$$\lim_{x_1\to\pm\infty}U(x_1,\hat x)=a^\pm\quad\text{uniformly for } \hat x\in\mathcal{D}. \tag{6.30}$$

From the proof of Theorem 2.2, we have

$$\lim_{i\to\pm\infty}\left(\|\nabla(U-a^\pm)\|_{L^2(T_i,\mathbb{R}^m)}+\|U-a^\pm\|_{L^2(T_i,\mathbb{R}^m)}\right)=0.$$

By Theorem 3.1, Proposition 6.2, and (PDE), there is a constant, $M^*>0$ such that $\|U\|_{C^{2,\alpha}(\Omega,\mathbb{R}^m)}\le M^*$. Therefore standard interpolation inequalities imply $\|U-a^\pm\|_{L^\infty(T_i)}\to 0$, $i\to\pm\infty$, which gives (6.30) and completes the proof of Theorem 3.1. $\square$

Remark 6.31. The arguments we have given to establish the regularity of solutions of (PDE) and (BC), in particular Proposition 6.2 obtaining an L^∞ bound for the solution, Proposition 6.27 giving interior regularity, and the final arguments establishing regularity up to the boundary, work equally well for any divergence structure semilinear elliptic system of PDEs satisfying (V_4) provided the coefficients are sufficiently smooth.

Acknowledgments

Byeon was supported by Mid-career Researcher Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (NRF-2013R1A2A2A05006371).

References

[Alessio 2013] F. Alessio, "Stationary layered solutions for a system of Allen–Cahn type equations", *Indiana Univ. Math. J.* **62**:5 (2013), 1535–1564. MR 3188554 Zbl 1300.35035

[Alessio and Montecchiari 2014] F. Alessio and P. Montecchiari, "Multiplicity of layered solutions for Allen–Cahn systems with symmetric double well potential", *J. Differential Equations* **257**:12 (2014), 4572–4599. MR 3268736 Zbl 1301.35033

[Alikakos 2012] N. D. Alikakos, "A new proof for the existence of an equivariant entire solution connecting the minima of the potential for the system $\Delta u - W_u(u) = 0$", *Comm. Partial Differential Equations* **37**:12 (2012), 2093–2115. MR 3005537 Zbl 1268.35051

[Alikakos 2013] N. D. Alikakos, "On the structure of phase transition maps for three or more coexisting phases", pp. 1–31 in *Geometric partial differential equations*, CRM Series **15**, Ed. Norm., Pisa, 2013. MR 3156886 Zbl 1298.35054

[Alikakos and Fusco 2008] N. D. Alikakos and G. Fusco, "On the connection problem for potentials with several global minima", *Indiana Univ. Math. J.* **57**:4 (2008), 1871–1906. MR 2440884 Zbl 1162.65060

[Alikakos and Fusco 2009] N. D. Alikakos and G. Fusco, "Entire solutions to nonconvex variational elliptic systems in the presence of a finite symmetry group", pp. 1–26 in *Singularities in nonlinear evolution phenomena and applications*, vol. 9, CRM Series, Ed. Norm., Pisa, 2009. MR 2528696 Zbl 1193.35034

[Alikakos and Fusco 2011] N. D. Alikakos and G. Fusco, "Entire solutions to equivariant elliptic systems with variational structure", *Arch. Ration. Mech. Anal.* **202**:2 (2011), 567–597. MR 2847535 Zbl 1266.35055

[Alikakos and Fusco 2015] N. D. Alikakos and G. Fusco, "A maximum principle for systems with variational structure and an application to standing waves", *J. Eur. Math. Soc.* (*JEMS*) **17**:7 (2015), 1547–1567. MR 3361722 Zbl 1331.35124

[Alikakos and Smyrnelis 2012] N. D. Alikakos and P. Smyrnelis, "Existence of lattice solutions to semilinear elliptic systems with periodic potential", *Electron. J. Differential Equations* (2012), art. id. 15, 1–15. MR 2889621 Zbl 1239.35043

[Bolotin 1978] S. V. Bolotin, "Libration motions of natural dynamical systems", *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* 6 (1978), 72–77. MR 524544 Zbl 0403.34053

[Bronsard and Reitich 1993] L. Bronsard and F. Reitich, "On three-phase boundary motion and the singular limit of a vector-valued Ginzburg–Landau equation", *Arch. Rational Mech. Anal.* **124**:4 (1993), 355–379. MR 1240580 Zbl 0785.76085

[Evans 1998] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics **19**, American Mathematical Society, Providence, RI, 1998. MR 1625845 Zbl 0902.35002

[Gilbarg and Trudinger 1983] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, 2nd ed., Grundlehren der Mathematischen Wissenschaften **224**, Springer, Berlin, 1983. MR 737190 Zbl 0562.35001

[Gui and Schatzman 2008] C. Gui and M. Schatzman, "Symmetric quadruple phase transitions", *Indiana Univ. Math. J.* **57**:2 (2008), 781–836. MR 2414335 Zbl 1185.35080

[Kozlov 1985] V. V. Kozlov, "Calculus of variations in the large and classical mechanics", *Uspekhi Mat. Nauk* **40**:2(242) (1985), 33–60, 237. In Russian; translated in *Russian Mathematical Surveys* **40**:2 (1985), 37–71. MR 786086 Zbl 0579.70020

[Montecchiari and Rabinowitz 2016] P. Montecchiari and P. H. Rabinowitz, "On the existence of multi-transition solutions for a class of elliptic systems", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **33**:1 (2016), 199–219. MR 3436431 Zbl 1332.35113

[Rabinowitz 1989] P. H. Rabinowitz, "Periodic and heteroclinic orbits for a periodic Hamiltonian system", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6**:5 (1989), 331–346. MR 1030854 Zbl 0701.58023

[Rabinowitz 1993] P. H. Rabinowitz, "Homoclinic and heteroclinic orbits for a class of Hamiltonian systems", *Calc. Var. Partial Differential Equations* **1**:1 (1993), 1–36. MR 1261715 Zbl 0791.34042

[Rabinowitz 1994] P. H. Rabinowitz, "Solutions of heteroclinic type for some classes of semilinear elliptic partial differential equations", *J. Math. Sci. Univ. Tokyo* **1**:3 (1994), 525–550. MR 1322690 Zbl 0823.35062

[Rabinowitz 2002] P. H. Rabinowitz, "Spatially heteroclinic solutions for a semilinear elliptic P.D.E.", *ESAIM Control Optim. Calc. Var.* **8** (2002), 915–931. MR 1932980 Zbl 1092.35026

[Rabinowitz 2004] P. H. Rabinowitz, "A new variational characterization of spatially heteroclinic solutions of a semilinear elliptic PDE", *Discrete Contin. Dyn. Syst.* **10**:1–2 (2004), 507–515. MR 2026208 Zbl 1174.35341

[Rabinowitz 2012] P. H. Rabinowitz, "On a class of reversible elliptic systems", *Netw. Heterog. Media* **7**:4 (2012), 927–939. MR 3004692 Zbl 1266.35046

[Schatzman 2002] M. Schatzman, "Asymmetric heteroclinic double layers", *ESAIM Control Optim. Calc. Var.* **8** (2002), 965–1005. MR 1932983 Zbl 1092.35030

[Sternberg 1991] P. Sternberg, "Vector-valued local minimizers of nonconvex variational problems", *Rocky Mountain J. Math.* **21**:2 (1991), 799–807. MR 1121542 Zbl 0737.49009

Received 23 Mar 2016. Revised 26 Jun 2016. Accepted 30 Jul 2016.

JAEYOUNG BYEON: byeon@kaist.ac.kr
Department of Mathematical Sciences, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, South Korea

PIERO MONTECCHIARI: p.montecchiari@univpm.it
Dipartimento di Ingegneria Civile, Edile e Architettura, Università Politecnica delle Marche, Via brecce bianche, Ancona I-60131, Italy

PAUL H. RABINOWITZ: rabinowi@math.wisc.edu
Department of Mathematics, University of Wisconsin–Madison, Madison, WI 53706, United States

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 9 No. 7 2016