ANALYSIS & PDE

Volume 16 No. 9 2023

msp

Analysis & PDE

msp.org/apde

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

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber 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.

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. 16 (2023), No. 9, pp. 1989–2003
DOI: 10.2140/apde.2023.16.1989

OVERDETERMINED BOUNDARY PROBLEMS WITH NONCONSTANT DIRICHLET AND NEUMANN DATA

MIGUEL DOMÍNGUEZ-VÁZQUEZ, ALBERTO ENCISO AND DANIEL PERALTA-SALAS

We consider the overdetermined boundary problem for a general second-order semilinear elliptic equation on bounded domains of $\mathbb{R}^n$, where one prescribes both the Dirichlet and Neumann data of the solution. We are interested in the case where the data are not necessarily constant and where the coefficients of the equation can depend on the position, so that the overdetermined problem does not generally admit a radial solution. Our main result is that, nevertheless, under minor technical hypotheses nontrivial solutions to the overdetermined boundary problem always exist.

1. Introduction

The study of overdetermined boundary problems, that is, problems where one prescribes both Dirichlet and Neumann data, has grown into a major field of research in the theory of elliptic PDEs since its appearance in Lord Rayleigh's classic treatise [1877]. An outburst of activity started with the groundbreaking paper [Serrin 1971], where he combined an adaptation of Alexandrov's moving planes method with a subtle refinement of the maximum principle to prove a symmetry result for an overdetermined problem. More precisely, Serrin proved that, under mild technical hypotheses, positive solutions to elliptic equations of the form

$$\Delta u + F(u) = 0$$

inside a bounded domain $\Omega \subset \mathbb{R}^n$ satisfying the boundary conditions

$$u = 0 \quad \text{and} \quad \partial_\nu u = -c \quad \text{on } \partial\Omega, \tag{1-1}$$

where c is an unspecified constant that can be picked freely, only exist if Ω is a ball, in which case u is radial. The result remains true if F also depends on the norm of the gradient of u and if we replace the Laplacian by other position-independent operators of variational form [Cianchi and Salani 2009].

The influence of Serrin's result is such that the very considerable body of literature devoted to overdetermined boundary problems is mostly limited to proofs that solutions need to be radial in cases that can be handled using the method of moving planes. Without attempting to be comprehensive, some remarkable results about overdetermined boundary value problems include alternative approaches to radial symmetry results using P-functions [Garofalo and Lewis 1989; Kawohl 1998] or Pohozaev-type integral identities [Brandolini et al. 2008; Magnanini and Poggesi 2020a; 2020b], extensions of the moving plane method to the hyperbolic space and the hemisphere [Kumaresan and Prajapat 1998], to degenerate

MSC2020: primary 35N25; secondary 35J61.
Keywords: overdetermined boundary value problems, semilinear elliptic equations.

elliptic equations such as the p-Laplace equation [Damascelli et al. 1999], and to exterior [Aftalion and Busca 1998; Garofalo and Sartori 1999], unbounded [Farina and Valdinoci 2010] or nonsmooth domains [Prajapat 1998], and stability of symmetry [Aftalion et al. 1999]. Another direction of research that has attracted considerable recent attention is the study of connections with the theory of constant mean curvature surfaces and the construction of nontrivial solutions to Serrin-type problems in exterior domains [Traizet 2014; del Pino et al. 2015; Ros et al. 2020]. Nontrivial solutions for partially overdetermined problems or with degenerate ellipticity are also known to exist [Alessandrini and Garofalo 1989; Fragalà and Gazzola 2008; Fragalà et al. 2006; Farina and Valdinoci 2013].

In two surprising papers, Pacard and Sicbaldi [2009] and Delay and Sicbaldi [2015] proved the existence of extremal domains with small volume for the first eigenvalue of the Laplacian in any compact Riemannian manifold, that is, domains for which the overdetermined problem for the linear elliptic equation

$$\Delta_g u + \lambda u = 0$$

has a positive solution with zero Dirichlet data and constant Neumann data. Here Δ_g is the Laplacian operator associated with a Riemannian metric g on a compact manifold and the constant λ (which one eventually chooses as the first Dirichlet eigenvalue of the domain Ω) is not specified a priori. Very recently we managed to show the existence of nontrivial solutions, with the same overdetermined Dirichlet and Neumann conditions, for fairly general semilinear elliptic equations of second order with possibly nonconstant coefficients [Domínguez-Vázquez et al. 2019].

In all these results, the fact that one is imposing precisely the standard overdetermined boundary conditions (1-1) plays a crucial role. Roughly speaking, this is because one can relate the existence of overdetermined solutions with the critical points of certain functional via a variational argument. Therefore, the gist of the argument in these papers is that the overdetermined condition with constant data is connected with the local extrema for a natural energy functional, restricted to a specific class of functions labeled by points in the physical space. This ultimately permits one to derive the existence of solutions from the fact that a continuous function attains its maximum on a compact manifold. However, this strategy is successful only for constant boundary data. To our best knowledge, the only result in the literature which considers nonconstant (albeit special) Neumann data in relation to overdetermined boundary problems is [Bianchini et al. 2014].

In the recent paper [Domínguez-Vázquez et al. 2021], we have constructed new families of compactly supported stationary solutions to the three-dimensional Euler equation by proving that there are solutions to an associated overdetermined problem in two dimensions where one prescribes (modulo constants that can be picked freely) zero Dirichlet data and nonconstant Neumann data. The proof uses crucially that the space is two-dimensional, which ensures that the kernel and cokernel of a certain operator are one-dimensional, and does not work in higher dimensions.

Our objective in this paper is to prove the existence of solutions to overdetermined problems where one prescribes general Dirichlet and Neumann data (just as before, up to unspecified constants). For concreteness, we consider the model semilinear equation

$$Lu + \lambda F(x, u) = 0 \tag{1-2}$$

in a bounded domain $\Omega \subset \mathbb{R}^n$, with Dirichlet and Neumann boundary conditions

$$u = f_0(x), \quad \nu \cdot A(x)\nabla u = -c f_1(x) \quad \text{on } \partial\Omega. \tag{1-3}$$

Here f_0, f_1 are functions on $\mathbb{R}^n$, F is a function on $\mathbb{R}^n \times \mathbb{R}$, λ, c are unspecified positive constants, ν is the outer unit normal on $\partial\Omega$ and L is the second-order operator

$$Lu := a_{ij}(x)\,\partial_{ij}u + b_i(x)\,\partial_i u,$$

where $A(x) = (a_{ij}(x))$ is a (symmetric) matrix-valued function on $\mathbb{R}^n$ satisfying the (possibly nonuniform) ellipticity condition

$$\min_{|\xi|=1} \xi \cdot A(x)\xi > 0 \quad \text{for all } x \in \mathbb{R}^n.$$

Theorem 1.1. *Given any noninteger $s > 2$, let us take any functions F, f_0, f_1, b of class C^s and A of class C^{s+2}. Assume that the functions $F(\cdot\,, f_0(\cdot))$ and f_1 are positive and that the function f_0 has a nondegenerate critical point. Then there is a family of domains $\Omega_{\varepsilon,\bar\lambda}$ for which the overdetermined problem* (1-2)-(1-3) *admits a solution.*

More precisely, let $p \in \mathbb{R}^n$ be a nondegenerate critical point of f_0. Then, for any $\varepsilon \neq 0$ small enough and $\bar\lambda > 0$, the following statements hold:

(i) *The domain $\Omega_{\varepsilon,\bar\lambda}$ is a small deformation of the ball of radius ε centered at p, characterized by an equation of the form $|x - p|^2 < \varepsilon^2 + O(\varepsilon^3)$.*

(ii) *The dependence of λ and c on the parameter ε is of the form*

$$\lambda = \varepsilon^{-2}\bar\lambda, \quad c = \varepsilon^{-1}\bar c,$$

where $\bar c = \bar c(\varepsilon, \bar\lambda)$ is a positive constant of order 1.

Remark 1.2. In the case of the torsion problem, i.e., $\Delta u + \lambda = 0$ (i.e., $F(x, u) = 1$ in the previous notation), the condition that f_0 has a critical point can be relaxed: it is enough that the function $G_\kappa := f_0 + \kappa \log f_1$ has at least one nondegenerate critical point for some constant $\kappa > 0$. The statement then applies if p is a nondegenerate critical point of G_κ and taking $\bar\lambda := n\kappa > 0$ (not necessarily small).

Also, it is easy to obtain different variations on our main theorem following the same method of proof. In fact, one obtains new results even for the linear equation $\Delta u + b(x) \cdot \nabla u + \lambda f(x) = 0$ with standard overdetermined boundary data $f_0 := 0$, $f_1 := 1$; specifically, if p is a nondegenerate zero of the vector field $n\nabla f - f b$, then the statement still holds taking any $\bar\lambda > 0$. This does not follow from [Domínguez-Vázquez et al. 2019]. However, we shall not pursue these generalizations here.

Compared with [Domínguez-Vázquez et al. 2019], a major difference is that the theorem does not only ensure the existence of domains where the overdetermined problem under consideration admits a nontrivial solution, but also specifies the points around which those domains are located. This immediately permits one to translate this existence result to problems that are only defined in a subset of $\mathbb{R}^n$ or on a differentiable manifold. In view of the heuristic but fruitful connection between overdetermined boundary problems and the study of CMC hypersurfaces, a result that is somehow akin to our existence results for overdetermined boundary problems for semilinear equations is Ye's classical theorem [1991] on foliations by small CMC spheres on n-dimensional Riemannian manifolds.

The paper is organized as follows. We will start by setting up the problem in Section 2. For clarity of exposition, in Sections 2 to 4 we have chosen to assume that the matrix $A(x)$ is the identity and carry out the proof in this context. An essential ingredient of the proof is the computation of asymptotic expansions for the solution to the Dirichlet problem in small perturbations of a ball of radius $\varepsilon \ll 1$, when the constants λ and c scale with the radius as in Theorem 1.1. This computation is carried out in Section 3. These asymptotic estimates are put to use in Section 4, where we prove Theorem 1.1 in the particular case when $A(x) = I$. To obtain the general result, in Section 5 we show that the case of a general matrix-valued function $A(x)$ reduces to the study of the easiest case $A(x) = I$ subject to an inessential perturbation of order ε^2. Making this precise, however, involves using a heavier notation and geodesic-type normal coordinates adapted to the matrix $A(x)$ that might unnecessarily obscure the simple ideas the proof is based on. As a side remark, let us point out that the reason we ask for more regularity of the matrix A (which is of class C^{s+2} in contrast with the C^s regularity of the other functions) is precisely due to our use of geodesic coordinates.

2. Setting up the problem

For clarity of exposition, until Section 5 we will assume that $A(x) = I$. This assumption will enable us to obtain more compact expressions for the various quantities that appear in the problem and it will make it easier to point out the salient features of the proof.

Let us fix a point $p \in \mathbb{R}^n$ and introduce rescaled coordinates $z \in \mathbb{R}^n$ centered at p as

$$z := \frac{x - p}{\varepsilon},$$

where ε is a suitably small nonzero constant. We now consider spherical coordinates $(r, \omega) \in \mathbb{R}^+ \times \mathbb{S}$ for z, defined as

$$r := |z| = \left|\frac{x - p}{\varepsilon}\right|, \quad \omega := \frac{z}{|z|} = \frac{x - p}{|x - p|}.$$

Here and in what follows,

$$\mathbb{S} := \{\omega \in \mathbb{R}^n : |\omega| = 1\}$$

denotes the unit sphere of dimension $n - 1$. For simplicity of notation, we will notationally omit the dependence on the point p. Also, with some abuse of notation, we will denote the expression of the function $u(x)$ in these coordinates simply by $u(r, \omega)$.

Let us now consider a C^{s+1} function $B : \mathbb{S} \to \mathbb{R}$ and, for suitably small ε, let us describe the domain in terms of the above coordinates as

$$\Omega_{p,\varepsilon B} := \{r < 1 + \varepsilon B(\omega)\}. \tag{2-1}$$

We now consider (1-2) in the domain $\Omega_{p,\varepsilon B}$ and choose the constants λ, c as

$$\lambda =: \varepsilon^{-2}\bar{\lambda}, \quad c =: \varepsilon^{-1}\bar{c},$$

where we think of ε as a small constant and of $\bar{\lambda}$, $\bar{c}$ as positive constants of order 1. Equation (1-2) can then be rewritten in the rescaled coordinates as

$$\tilde{L}u + \bar{\lambda}\widetilde{F}(z, u) = 0, \tag{2-2}$$

where

$$\widetilde{F}(z, u) := F(p + \varepsilon z, u)$$

and $\tilde{L}$ is the differential operator

$$\tilde{L}u = \Delta u + \varepsilon \tilde{b}(z) \cdot \nabla u,$$

with $\tilde{b}_i(z) := b_i(p + \varepsilon z)$. We also denote the functions f_0 and f_1 in these coordinates as

$$\tilde{f}_0(z) := f_0(p + \varepsilon z), \quad \tilde{f}_1(z) := f_1(p + \varepsilon z).$$

Here and in what follows, Δ and ∇ denote the Laplacian and gradient operators in the rescaled coordinates z.

The Dirichlet boundary condition on $\partial\Omega_{p,\varepsilon B}$ can be simply written in rescaled hyperspherical coordinates as

$$u(1 + \varepsilon B(\omega), \omega) = \tilde{f}_0(1 + \varepsilon B(\omega), \omega) =: \tilde{f}_0(\varepsilon, \omega). \tag{2-3}$$

We notice that $\tilde{f}_0(0, \omega) = f_0(p)$. Analogously, the Neumann boundary condition reads as

$$\partial_\nu u(1 + \varepsilon B(\omega), \omega) = -\bar{c}\tilde{f}_1(1 + \varepsilon B(\omega), \omega),$$

where ν is the outwards normal unit vector on $\partial\Omega_{p,\varepsilon B}$.

We denote by $C^s_{\mathrm{Dir}}(\mathbb{B})$ the space of C^s functions on the unit n-dimensional ball $\mathbb{B} := \{|z| < 1\}$ with zero trace to the boundary. Also, $\mathcal{K} \subset C^\infty(\mathbb{S})$ denotes the restriction to the unit sphere of the space of linear functions on $\mathbb{R}^n$,

$$\mathcal{K} := \{V \cdot z : |z| = 1,\ V \in \mathbb{R}^n\}.$$

Equivalently, $\mathcal{K}$ is the eigenspace of the Laplacian $\Delta_{\mathbb{S}}$ of the unit sphere corresponding to the second eigenvalue, $n - 1$. Also, in what follows we will denote the partial derivatives of F (or $\widetilde{F}$) as

$$F'(x, u) := \partial_u F(x, u), \quad \nabla F(x, u) := \nabla_x F(x, u), \quad \partial_j F(x, u) := \partial_{x_j} F(x, u).$$

The following lemma is a reformulation of [Domínguez-Vázquez et al. 2019, Theorem 2.3 and Proposition 2.4]. Here $s > 2$ is assumed to be a noninteger real.

Lemma 2.1. *For each $p \in \mathbb{R}^n$, there is some $\bar{\lambda}_p > 0$ such that the following statements hold for all $\bar{\lambda} \in (0, \bar{\lambda}_p)$:*

(i) *There is a unique function $\phi_{p,\bar{\lambda}}(r)$ of class C^{s+2} satisfying the ODE*

$$\phi_{p,\bar{\lambda}}''(r) + \frac{n-1}{r}\phi_{p,\bar{\lambda}}'(r) + \bar{\lambda}F(p, f_0(p) + \phi_{p,\bar{\lambda}}(r)) = 0$$

and the boundary condition $\phi_{p,\bar{\lambda}}(1) = 0$ which is regular at $r = 0$. The function $\phi_{p,\bar{\lambda}}$ is well-defined for $r \in [0, 1 + \delta_p]$, with $\delta_p > 0$. Furthermore, $\phi_{p,\bar{\lambda}}(r) > 0$ for $r < 1$ and $\phi_{p,\bar{\lambda}}'(1) < 0$.

(ii) *The operator*

$$T_{p,\bar\lambda} v := \Delta v + \bar\lambda\, F'(p,\, f_0(p) + \phi_{p,\bar\lambda}(|z|))v$$

defines an invertible map $T_{p,\bar\lambda} : C^{s+1}_{\mathrm{Dir}}(\mathbb{B}) \to C^{s-1}(\mathbb{B})$.

(iii) *Consider the map* $H_{p,\bar\lambda}$ *defined for each function* ψ *on the boundary of the ball as*

$$H_{p,\bar\lambda}\psi := -{\phi_{p,\bar\lambda}}'(1)\,\partial_\nu v_\psi + {\phi_{p,\bar\lambda}}''(1) v_\psi,$$

where v_ψ *is the only solution to the problem* $T_{p,\bar\lambda} v_\psi = 0$ *on* $\mathbb{B}$, $v_\psi|_{\partial\mathbb{B}} = \psi$. *Then* $H_{p,\bar\lambda}$ *maps* $C^{s+1}(\mathbb{S}) \to C^s(\mathbb{S})$, *its kernel is* $\mathcal{K}$, *and its range is the set* $C^s(\mathbb{B}) \cap \mathcal{K}^\perp$ *of* C^s *functions orthogonal to* $\mathcal{K}$. *Furthermore,*

$$\|\psi\|_{C^{s+1}} \leqslant C_{p,\bar\lambda} \|H_{p,\bar\lambda}\psi\|_{C^s} \tag{2-4}$$

for all $\psi \in C^{s+1} \cap \mathcal{K}^\perp$.

(iv) *The function* $\phi_{p,\bar\lambda}$ *satisfies* $\|{\phi_{p,\bar\lambda}}'\|_{C^s((0,1+\delta_p))} \leqslant C\bar\lambda$ *and is of class* C^s *in* p *and* $\bar\lambda$.

Remark 2.2. When the equation is linear (that is, $F(x, u) = f(x)$), one can take $\bar\lambda_p$ arbitrarily large and

$$\phi_{p,\bar\lambda}(r) = -\frac{\bar\lambda}{2n} f(p)\,(r^2 - 1).$$

The operator $H_{p,\bar\lambda}$ is then

$$H_{p,\bar\lambda}\psi = \frac{\bar\lambda}{n} f(p)\,(\Lambda_0\psi - \psi),$$

where $\Lambda_0 := [(n/2-1)^2 - \Delta_{\mathbb{S}}]^{1/2} - n/2 + 1$ is the Dirichlet–Neumann map of the ball.

In what follows we shall always assume that $\bar\lambda < \bar\lambda_p$.

Proposition 2.3. *For any* ε *small enough and any function* $B \in C^{s+1}(\mathbb{S})$ *with* $\|B\|_{C^{s+1}} < 1$, *there is a unique function* $u = u_{p,\varepsilon,\bar\lambda,B}$ *in a small neighborhood of* $f_0(p) + \phi_{p,\bar\lambda}$ *in* $C^{s+1}(\Omega_{p,\varepsilon B})$ *that satisfies* (2-2) *and the Dirichlet boundary condition* (2-3).

Proof. Let $\chi_{p,\varepsilon B} : \mathbb{B} \to \Omega_{p,\varepsilon B}$ be the diffeomorphism defined in spherical coordinates as

$$(\rho, \omega) \mapsto ([1 + \varepsilon\chi(\rho)\, B(\omega)]\rho, \omega),$$

where $\chi(\rho)$ is a smooth cutoff function that is zero for $\rho < \frac14$ and 1 for $\rho > \frac12$. Then one can define a map

$$\mathcal{H}_{p,\bar\lambda,B} : (-\varepsilon_p, \varepsilon_p) \times C^{s+1}_{\mathrm{Dir}}(\mathbb{B}) \to C^{s-1}(\mathbb{B})$$

as

$$\mathcal{H}_{p,\bar\lambda,B}(\varepsilon, \phi) := [\tilde L(\phi \circ \chi^{-1}_{p,\varepsilon B})] \circ \chi_{p,\varepsilon B} + E \circ \chi_{p,\varepsilon B} + \bar\lambda[\widetilde F(\,\cdot\,, \tilde f_0 + \phi \circ \chi^{-1}_{p,\varepsilon B})] \circ \chi_{p,\varepsilon B},$$

with the function E defined as

$$E := \tilde L \tilde f_0. \tag{2-5}$$

Note that $\|E\|_{C^{s-1}(\Omega_{p,\varepsilon B})} \leqslant C\varepsilon^2$ because $\tilde f_0(z) := f_0(p + \varepsilon z)$. Clearly, $\mathcal{H}_{p,\bar\lambda,B}(\varepsilon, \phi) = 0$ if and only if $u := \tilde f_0 + \phi \circ \chi^{-1}_{p,\varepsilon B}$ solves the Dirichlet problem (2-2)-(2-3) in $\Omega_{p,\varepsilon B}$.

Note that, by definition and using (2-5), $\mathcal{H}_{p,\bar\lambda,B}(0,\phi_{p,\bar\lambda})=0$. Also, a short computation shows that the derivative of $\mathcal{H}_{p,\bar\lambda,B}(\varepsilon,\phi)$ with respect to ϕ satisfies

$$D_\phi\mathcal{H}_{p,\bar\lambda,B}(0,\phi_{p,\bar\lambda})=T_{p,\bar\lambda},$$

so it is an invertible map $C^{s+1}_{\mathrm{Dir}}(\mathbb{B})\to C^{s-1}(\mathbb{B})$; see Lemma 2.1. The implicit function theorem in Banach spaces then ensures that, for any ε close enough to 0, there is a unique function ϕ^ε in a small neighborhood of $\phi_{p,\bar\lambda}$ in $C^{s+1}_{\mathrm{Dir}}(\mathbb{B})$ satisfying

$$\mathcal{H}_{p,\bar\lambda,B}(\varepsilon,\phi^\varepsilon)=0.$$

Then $u_{p,\varepsilon,\bar\lambda,B}:=\tilde f_0+\phi^\varepsilon\circ\chi^{-1}_{p,\varepsilon B}$ is the desired solution to the Dirichlet problem in $\Omega_{p,\varepsilon B}$. □

We will henceforth denote by

$$\mathbb{P}_{p,\bar\lambda,\varepsilon B}:C^{s+1}(\mathbb{S})\to C^{s+1}(\Omega_{p,\varepsilon B})$$

the map $\psi\mapsto v_\psi$, where v_ψ is the only solution to the problem

$$T_{p,\bar\lambda}v_\psi=0\quad\text{in }\Omega_{p,\varepsilon B},$$

with the boundary condition

$$v_\psi(1+\varepsilon B(\omega),\omega)=\psi(\omega).$$

Note that the existence and uniqueness of v_ψ is an easy consequence of Lemma 2.1.

For future reference, let us record here the definition of the associated Dirichlet–Neumann operator $\Lambda_{p,\bar\lambda,\varepsilon B}:C^{s+1}(\mathbb{S})\to C^s(\mathbb{S})$,

$$\Lambda_{p,\bar\lambda,\varepsilon B}\psi(\omega):=\nu\cdot A\nabla\mathbb{P}_{p,\bar\lambda,\varepsilon B}\psi(1+\varepsilon B(\omega),\omega).$$

As $\Lambda_{p,\bar\lambda,\varepsilon B}$ reduces to the standard Dirichlet–Neumann map Λ_0 when $\varepsilon=\bar\lambda=0$, it is standard that

$$\|\Lambda_{p,\bar\lambda,\varepsilon B}-\Lambda_{p,\bar\lambda,0}\|_{C^{s+1}(\mathbb{S})\to C^s(\mathbb{S})}\leqslant C|\varepsilon|,\tag{2-6}$$

$$\|\Lambda_{p,\bar\lambda,\varepsilon B}-\Lambda_0\|_{C^{s+1}(\mathbb{S})\to C^s(\mathbb{S})}\leqslant C(|\varepsilon|+\bar\lambda).\tag{2-7}$$

3. Asymptotic expansions

In this section we compute asymptotic formulas for the solution to the Dirichlet problem in the domain (2-1) obtained in Proposition 2.3, valid for $|\varepsilon|\ll 1$. Let us begin with the estimates for the solutions to the Dirichlet problem:

Proposition 3.1. *The function $u_{p,\varepsilon,\bar\lambda,B}$ is of the form*

$$u_{p,\varepsilon,\bar\lambda,B}=f_0(p)+\phi_{p,\bar\lambda}(r)+\varepsilon\{W_{p,\bar\lambda}(r)\cdot z+\mathbb{P}_{p,\bar\lambda,\varepsilon B}[\nabla f_0(p)\cdot\omega-\phi_{p,\bar\lambda}{}'(1)\,B]\}+O(\varepsilon^2),$$

where $W_{p,\bar\lambda}:[0,1+\delta_p]\to\mathbb{R}^n$ is a function with $\|W_{p,\bar\lambda}\|_{C^{s+1}}\leqslant C\bar\lambda$.

Remark 3.2. In the case when $F(x,u)=f(x)$, the formula is slightly more explicit:

$$u_{p,\varepsilon,\bar\lambda,B}=f_0(p)-\frac{\bar\lambda}{2n}f(p)\,(r^2-1)\\ +\varepsilon\left\{\left[\nabla f_0(p)-\frac{\bar\lambda(r^2-1)}{2n+4}\left(\nabla f(p)-\frac{f(p)b(p)}{n}\right)\right]\cdot z+\frac{\bar\lambda\, f(p)}{n}\mathbb{P}_{\varepsilon B}B\right\}+O(\varepsilon^2).$$

Here we are using the notation $\mathbb{P}_{\varepsilon B}\equiv\mathbb{P}_{p,0,\varepsilon B}$, which does not depend on p because $F'=0$.

Proof. Note that $u_0:=f_0(p)+\phi_{p,\bar\lambda}(r)$ satisfies the equation

$$\Delta u_0+\bar\lambda F(p,u_0)=0,\quad u_0|_{r=1}=f_0(p).$$

Let us write $u_1:=(u_{p,\varepsilon,\bar\lambda,B}-u_0)/\varepsilon$ and observe that

$$\widetilde{F}(z,u_{p,\varepsilon,\bar\lambda,B})=F(p+\varepsilon z,u_0+\varepsilon u_1)=F(p,u_0)+\varepsilon[\nabla F(p,u_0)\cdot z+F'(p,u_0)u_1]+O(\varepsilon^2).$$

As $\tilde{L}u_{p,\varepsilon,\bar\lambda,B}+\bar\lambda\widetilde{F}(z,u_{p,\varepsilon,\bar\lambda,B})=0$ with the boundary condition

$$u_{p,\varepsilon,\bar\lambda,B}(1+\varepsilon B(\omega),\omega)=\tilde{f}_0(1+\varepsilon B(\omega),\omega)=f_0(p)+\varepsilon\nabla f_0(p)\cdot\omega+O(\varepsilon^2),$$

this ensures that u_1 satisfies an equation of the form

$$T_{p,\bar\lambda}u_1+\bar\lambda\nabla F(p,u_0)\cdot z+b(p)\cdot\frac{z}{r}\,\phi_{p,\bar\lambda}{}'(r)+O(\varepsilon)=0$$

in $\Omega_{p,\varepsilon B}$ and the boundary condition

$$u_1(1+\varepsilon B(\omega),\omega)=\nabla f_0(p)\cdot\omega-\phi_{p,\bar\lambda}{}'(1)\,B(\omega)+O(\varepsilon).$$

To analyze u_1, we start by noting that

$$u_1^*:=\mathbb{P}_{p,\bar\lambda,\varepsilon B}[\nabla f_0(p)\cdot\omega-\phi_{p,\bar\lambda}{}'(1)\,B(\omega)]$$

satisfies the equation $T_{p,\bar\lambda}u_1^*=0$ in $\Omega_{p,\varepsilon B}$ and the boundary condition

$$u_1^*(1+\varepsilon B(\omega),\omega)=\nabla f_0(p)\cdot\omega-\phi_{p,\bar\lambda}{}'(1)\,B(\omega).$$

It is an easy consequence of Lemma 2.1 that the equation

$$T_{p,\bar\lambda}w+\bar\lambda\nabla F(p,u_0(|z|))\cdot z+b(p)\cdot\frac{z}{r}\,u_0'(|z|)=0\quad\text{in }\mathbb{B},\quad w|_{\partial\mathbb{B}}=0,$$

has a unique solution w, which is then of the form $w=W_{p,\bar\lambda}(|z|)\cdot z$ for some $\mathbb{R}^n$-valued function $W_{p,\bar\lambda}$. Specifically, its j-th component $W_j(r):=W_{p,\bar\lambda}(r)\cdot e_j$ satisfies the ODE

$$W_j''(r)+\frac{n+1}{r}W_j'(r)+\bar\lambda F'(p,u_0(r))W_j(r)+\bar\lambda\,\partial_jF(p,u_0(r))+b_j(p)\,\frac{u_0'(r)}{r}=0,$$

with the boundary condition $W_j(1)=0$ and the requirement that W_j must be regular at 0. As $u_0(r)$ is well-defined up to $r=1+\delta_p$, so is $W_j(r)$. The function $W_{p,\bar\lambda}$ is obviously bounded as

$$\|W_{p,\bar\lambda}\|_{C^{s+1}((0,1+\delta_p))}\leqslant C\bar\lambda\|\partial_jF(p,u_0)\|_{C^{s-1}((0,1+\delta_p))}+C\left\|\frac{u_0'}{r}\right\|_{C^{s-1}((0,1+\delta_p))}.$$

Since $\|u_0'\|_{C^s((0,1+\delta_p))}\leqslant C\bar\lambda$ by Lemma 2.1, we infer that $\|W_{p,\bar\lambda}\|_{C^{s+1}}=O(\bar\lambda)$ as well.

By construction, we immediately obtain that $u_1 = u_1^* + w + O(\varepsilon)$, so the proposition follows. The expression of Remark 3.2 follows from the same argument taking into account the formula for $\phi_{p,\bar\lambda}$ provided in Remark 2.2. □

Next we obtain asymptotic formulas for the normal derivative of u:

Proposition 3.3. *The normal derivative of the function $u_{p,\varepsilon,\bar\lambda,B}$ satisfies*

$$\partial_\nu u_{p,\varepsilon,\bar\lambda,B} = {\phi_{p,\bar\lambda}}'(1) + \varepsilon\{H_{p,\bar\lambda}B + [\nabla f_0(p) + V_{p,\bar\lambda}]\cdot\omega\} + O(\varepsilon^2),$$

where the constant vector $V_{p,\bar\lambda} \in \mathbb{R}^n$ satisfies $|V_{p,\bar\lambda}| \leqslant C\bar\lambda$.

Remark 3.4. When $F(x,u) = f(x)$, one can obtain a more compact formula:

$$\partial_\nu u_{p,\varepsilon,\bar\lambda,B} = -\frac{\bar\lambda}{n} f(p) + \varepsilon\left\{-\frac{\bar\lambda}{n} f(p)\,(B - \Lambda_0 B) + \nabla f_0(p)\cdot\omega - \frac{\bar\lambda}{n+2}\left(\nabla f(p) - \frac{f(p)b(p)}{n}\right)\cdot\omega\right\} + O(\varepsilon^2). \tag{3-1}$$

Proof. Since the boundary of $\Omega_{p,\varepsilon B}$ is the zero set of the function $r - \varepsilon B(\omega) - 1$, it is clear that its unit normal vector at the point $(1+\varepsilon B(\omega), \omega)$ is

$$\nu = \left(\omega - \frac{\varepsilon}{1+\varepsilon B(\omega)}\nabla_{\mathbb{S}} B(\omega)\right)\left(1 + \frac{\varepsilon^2}{(1+\varepsilon B(\omega))^2}|\nabla_{\mathbb{S}} B(\omega)|^2\right)^{-1/2} = \omega - \varepsilon\nabla_{\mathbb{S}} B(\omega) + O(\varepsilon^2),$$

where $\nabla_{\mathbb{S}}$ denotes covariant differentiation on the unit sphere.

Using this formula, it follows from Proposition 3.1 that

$$\begin{aligned}\partial_\nu u_{p,\varepsilon,\bar\lambda,B} &= \nu\cdot\nabla u_{p,\varepsilon,\bar\lambda,B}(1+\varepsilon B(\omega),\omega)\\ &= {\phi_{p,\bar\lambda}}'(1+\varepsilon B(\omega)) + \varepsilon\{(rW_{p,\bar\lambda})'(1)\cdot\omega + \nu\cdot\nabla\mathbb{P}_{p,\bar\lambda,\varepsilon B}[\nabla f_0(p)\cdot\omega - {\phi_{p,\bar\lambda}}'(1)\,B]\} + O(\varepsilon^2).\end{aligned}$$

Since $\phi_{p,\bar\lambda}(r)$ is C^{s+1}-smooth for $r < 1+\delta_p$, let us now expand ${\phi_{p,\bar\lambda}}'$ and use the definition of the operator $\Lambda_{p,\bar\lambda,\varepsilon B}$ to write

$$\partial_\nu u_{p,\varepsilon,\bar\lambda,B} = {\phi_{p,\bar\lambda}}'(1) + \varepsilon\{{\phi_{p,\bar\lambda}}''(1)B - {\phi_{p,\bar\lambda}}'(1)\,\Lambda_{p,\bar\lambda,\varepsilon B}B + \Lambda_{p,\bar\lambda,\varepsilon B}(\nabla f_0(p)\cdot\omega) + W'_{p,\bar\lambda}(1)\cdot\omega\} + O(\varepsilon^2).$$

Let us now recall that $H_{p,\bar\lambda}B := {\phi_{p,\bar\lambda}}''(1)B - {\phi_{p,\bar\lambda}}'(1)\,\Lambda_{p,\bar\lambda,0}B$ (see Lemma 2.1) and that the usual Dirichlet–Neumann map of the ball satisfies $\Lambda_0(V\cdot\omega) = V\cdot\omega$ for all $V \in \mathbb{R}^n$. Therefore, we can use the bounds (2-6)-(2-7) and the estimate $|V_{p,\bar\lambda}| \leqslant C\bar\lambda$ with

$$V_{p,\bar\lambda} := W'_{p,\bar\lambda}(1),$$

proven in Proposition 3.1, to obtain the formula of the statement. The expression of Remark 3.4 follows from the above argument after taking into account the expression for $u_{p,\varepsilon,\bar\lambda,B}$ given in Remark 3.2. □

4. Proof of Theorem 1.1 when $A(x) = I$

For any given point $p \in \mathbb{R}^n$, let us now define a map

$$\mathcal{F}_{p,\bar\lambda} : (-\varepsilon_p, \varepsilon_p) \times X^1_{s+1} \to C^s(\mathbb{S}),$$

with $X_s^1 := \{b \in C^s(\mathbb{S}) : \|b\|_{C^s} < 1\}$, as

$$\mathcal{F}_{p,\bar\lambda}(\varepsilon, B) := \partial_\nu u_{p,\varepsilon,\bar\lambda,B} - \frac{\phi_{p,\bar\lambda}{}'(1)}{f_1(p)} \tilde f_1.$$

Roughly speaking, this map measures how far the Dirichlet solution $u_{p,\varepsilon,\bar\lambda,B}$ is from satisfying the Neumann condition in the domain $\Omega_{p,\varepsilon B}$ with a constant

$$\bar c := -\frac{\phi_{p,\bar\lambda}{}'(1)}{f_1(p)} > 0.$$

An immediate consequence of the asymptotic formulas for $\partial_\nu u_{p,\varepsilon,\bar\lambda,B}$ proved in Proposition 3.3 and the fact that

$$\tilde f_1(1+\varepsilon B(\omega), \omega) = f_1(p) + \varepsilon \nabla f_1(p)\cdot\omega + O(\varepsilon^2),$$

is the following:

Proposition 4.1. *For any fixed $p \in \mathbb{R}^n$, any $B \in X_{s+1}^1(\mathbb{S})$ and any $|\varepsilon| < \varepsilon_p$,*

$$\mathcal{F}_{p,\bar\lambda}(\varepsilon, B) = \varepsilon\left\{ H_{p,\bar\lambda} B + \left[\nabla f_0(p) - \frac{\phi_{p,\bar\lambda}{}'(1)}{f_1(p)}\nabla f_1(p) + V_{p,\bar\lambda}\right]\cdot\omega\right\} + O(\varepsilon^2).$$

Remark 4.2. When $F(x, u) = f(x)$, one can obtain a slightly more explicit formula:

$$\begin{aligned}\mathcal{F}_{p,\bar\lambda}(\varepsilon, B) = \varepsilon\bigg\{ -\frac{\bar\lambda}{n} f(p)\,(B - \Lambda_0 B) + \left[\nabla f_0(p) + \frac{\bar\lambda\, f(p)}{n\, f_1(p)}\nabla f_1(p)\right]\cdot\omega \\ - \frac{\bar\lambda}{n+2}\left[\nabla f(p) - \frac{f(p) b(p)}{n}\right]\cdot\omega\bigg\} + O(\varepsilon^2). \quad (4\text{-}1)\end{aligned}$$

It then follows that the function $\mathcal{F}_{p,\bar\lambda}(\varepsilon, B)/\varepsilon$ can be defined at $\varepsilon = 0$ by continuity. Furthermore, its derivative with respect to B involves the operator $H_{p,\bar\lambda}$, whose kernel was shown to be the space $\mathcal{K}$ in Lemma 2.1. Consequently, let us define the spaces

$$\mathcal{X}_s := \{b \in C^s(\mathbb{S}) : \mathcal{P}_{\mathcal{K}} b = 0\}, \quad \mathcal{X}_s^1 := \{b \in \mathcal{X}_s : \|b\|_{C^s} < 1\},$$

with $\mathcal{P}_{\mathcal{K}}$ being the orthogonal projector onto the subspace $\mathcal{K}$. We also define the operator

$$\mathcal{P} b := b - \mathcal{P}_{\mathcal{K}} b.$$

It is clear from these expressions that $\mathcal{P}$ maps each space $C^s(\mathbb{S})$ into itself and $\mathcal{X}_s^1 \subset X_s^1$.

By Proposition 4.1, we can now define a map

$$\mathcal{G}_{p,\bar\lambda} : (-\varepsilon_p, \varepsilon_p) \times \mathcal{X}_{s+1}^1 \to \mathcal{X}_s$$

as

$$\mathcal{G}_{p,\bar\lambda}(\varepsilon, B) := \frac{\mathcal{P}\mathcal{F}_{p,\bar\lambda}(\varepsilon, B)}{\varepsilon}.$$

Lemma 4.3. *Let $U \subset \mathbb{R}^n$ be any bounded domain. For any $\bar\lambda \in (0, \bar\lambda_U)$, with*

$$\bar\lambda_U := \inf_{p\in U} \bar\lambda_p > 0,$$

there exist some $\varepsilon_{U,\bar\lambda} > 0$ and a C^s function $Y_{\varepsilon,\bar\lambda} : U \to \mathbb{R}^n$ such that

$$\partial_\nu u_{p,\varepsilon,\bar\lambda,B_{\varepsilon,p,\bar\lambda}} - \frac{\phi_{p,\bar\lambda}{}'(1)}{f_1(p)} \tilde f_1 = Y_{\varepsilon,\bar\lambda}(p) \cdot \omega$$

for all $p \in U$ and all $|\varepsilon| < \varepsilon_{U,\bar\lambda}$. Here $Y_{\varepsilon,\bar\lambda}(p) := Y(\varepsilon, p, \bar\lambda)$ is of class C^s in all its arguments, and can be interpreted as a family of parametrized vector fields on U, and $B_{\varepsilon,p,\bar\lambda}$ is a certain function in $\mathcal{X}^1_{s+1}$.

Proof. Let us begin by showing that the Fréchet derivative $D_B\mathcal{G}_{p,\bar\lambda}(0, 0) : \mathcal{X}_{s+1} \to \mathcal{X}_s$ is one-to-one. To see this, note that Proposition 4.1 and the fact that $\mathcal{P}(A \cdot \omega) = 0$ for any $A \in \mathbb{R}^n$ imply that the derivative of $\mathcal{G}_{p,\bar\lambda}$ with respect to B is of the form

$$D_B\mathcal{G}_{p,\bar\lambda}(\varepsilon, 0) = H_{p,\bar\lambda} + \mathcal{E},$$

with $\|\mathcal{E}\|_{\mathcal{X}_{s+1}\to\mathcal{X}_s} \leqslant C|\varepsilon|$. Here we have used that, by Lemma 2.1, $\mathcal{P}H_{p,\bar\lambda} = H_{p,\bar\lambda}$ because the range of the elliptic first-order operator $H_{p,\bar\lambda}$ is contained in $\mathcal{K}^\perp$. The estimate (2-4) then ensures that $D_B\mathcal{G}_{p,\bar\lambda}(\varepsilon, 0)$ is an invertible map $\mathcal{X}_{s+1} \to \mathcal{X}_s$ provided that ε is small enough.

As $\mathcal{G}_{p,\bar\lambda}(0, 0) = 0$, the invertibility of $D_B\mathcal{G}_{p,\bar\lambda}(\varepsilon, 0)$ implies, via the implicit function theorem, that for any ε small enough, there is a unique function $B_{\varepsilon,p,\bar\lambda}$ in a small neighborhood of 0 such that

$$\mathcal{G}_{p,\bar\lambda}(\varepsilon, B_{\varepsilon,p,\bar\lambda}) = 0.$$

By the definition of $\mathcal{F}_{p,\bar\lambda}$ and the fact that $\mathcal{K} = \{Y\cdot\omega : Y \in \mathbb{R}^n\}$, this implies that there is some $Y(\varepsilon, p, \bar\lambda) \in \mathbb{R}^n$ such that

$$\partial_\nu u_{p,\varepsilon,\bar\lambda,B_{\varepsilon,p,\bar\lambda}} - \frac{\phi_{p,\bar\lambda}{}'(1)}{f_1(p)} \tilde f_1 = Y(\varepsilon, p, \bar\lambda) \cdot \omega.$$

Furthermore, $Y(\varepsilon, p, \bar\lambda)$ is a C^s-smooth function of its arguments because so is the left-hand side of this identity. □

Let us now note that the asymptotic expression of the vector field $Y_{\varepsilon,\bar\lambda}(p)$ can be read off Proposition 4.1:

Lemma 4.4. *The vector field $Y_{\varepsilon,\bar\lambda}$ is of the form*

$$Y_{\varepsilon,\bar\lambda}(p) = \varepsilon\left[\nabla f_0(p) - \frac{\phi_{p,\bar\lambda}{}'(1)}{f_1(p)}\nabla f_1(p) + V_{p,\bar\lambda}\right] + O(\varepsilon^2).$$

When $F(x, u) = f(x)$, one can write down the more precise expression

$$Y_{\varepsilon,\bar\lambda}(p) = \varepsilon\left\{\nabla f_0(p) + \frac{\bar\lambda\, f(p)}{n\, f_1(p)}\nabla f_1(p) - \frac{\bar\lambda}{n+2}\left[\nabla f(p) - \frac{f(p)b(p)}{n}\right]\right\} + O(\varepsilon^2).$$

Proof of Theorem 1.1 when $A(x) = I$ and of Remark 1.2. Let us suppose that p^* is a nondegenerate critical point of the function f_0. As $\phi_{p,\bar\lambda}{}'(1) = O(\bar\lambda)$ by Lemma 2.1, Lemma 4.4 implies that

$$\frac{Y_{\varepsilon,\bar\lambda}(p)}{\varepsilon} = \nabla f_0(p) + \mathcal{E},$$

with an error bounded as $\|\mathcal{E}\|_{C^1(U)} \leqslant C_U|\bar\lambda| + C_U|\varepsilon|$ for any bounded domain $U \ni p^*$. If $|\bar\lambda|$ and $|\varepsilon|$ are small enough, it is then standard that there is a unique point $p_{\varepsilon,\bar\lambda}$ in a small neighborhood of p^* such that

$$Y_{\varepsilon,\bar\lambda}(p_{\varepsilon,\bar\lambda}) = 0.$$

By Lemma 4.3, and setting $\bar c := -\phi'_{p_{\varepsilon,\bar\lambda},\bar\lambda}(1)/f_1(p_{\varepsilon,\bar\lambda})$, this ensures that

$$\partial_\nu u_{\varepsilon,p_{\varepsilon,\bar\lambda},\bar\lambda,B_{\varepsilon,p,\bar\lambda}} + \bar c\tilde f_1 \equiv 0,$$

which implies the claim of the theorem with the domain $\Omega_{p_{\varepsilon,\bar\lambda},\varepsilon B_{\varepsilon,p,\bar\lambda}}$.

To prove Remark 1.2 on overdetermined solutions for the torsion problem, let us assume that $F(x,u) = f(x) = 1$ and that p^* is a nondegenerate critical point of the function $f_0 + \kappa \log f_1$ for some constant $\kappa > 0$. In this case, since $f(x) = 1$ and $b(x) = 0$, Lemma 4.4 implies that

$$\frac{Y_{\varepsilon,\bar\lambda}(p)}{\varepsilon} = \nabla f_0(p) + \frac{\bar\lambda}{n}\nabla \log f_1(p) + \mathcal{E}',$$

with $\|\mathcal{E}'\|_{C^1(U)} \leqslant C_U\varepsilon$. As one can pick any positive value of $\bar\lambda$ by Remark 2.2, let us fix $\bar\lambda = \bar\lambda^* := n\kappa > 0$. The previous argument then allows us to conclude that, for any ε small enough, there exists some point p_ε close to p^* for which $Y_{\varepsilon,\bar\lambda}(p_\varepsilon) = 0$. Note that the condition that p^* is a nondegenerate critical point of $f_0 + \kappa \log f_1$ is crucially used to solve

$$\nabla f_0(p_\varepsilon) + \kappa\,\nabla \log f_1(p_\varepsilon) = -\mathcal{E}'$$

for small $\varepsilon > 0$ via the inverse function theorem. As above, this implies the existence of solutions to the overdetermined torsion problem. The case of $f_0 = 0$, $f_1 = 1$ and $F(x,u) = f(x)$ is handled similarly, so Remark 1.2 then follows. □

5. Introduction of a nonconstant matrix $A(x)$ and conclusion of the proof

In this section we will show why the proof of Theorem 1.1 carried out for the case when $A(x) = I$ remains valid, with only minor variations, in the case of a general matrix $A(x)$.

The key idea is that we are constructing domains that are small deformations of the ball of radius ε, with $\varepsilon \ll 1$. Over scales of order ε, the function $A(x)$ is essentially constant, so it stands to reason that one might be able to compensate for the effect of having a nonconstant matrix $A(x)$ (at least, to some orders when considering an asymptotic expansion in ε) by deforming the balls accordingly. More visually, this would correspond essentially to picking an ellipsoidal domain at each point x with axes determined by the matrix $A(x)$.

The way to implement this idea is through (a rescaling of) the normal coordinates associated with the matrix-valued function A, which we now regard as a Riemannian metric on $\mathbb{R}^n$ of class C^{s+2}. These are defined through the exponential map at a point $p \in \mathbb{R}^n$,

$$\exp_p^A : U_p \to \mathbb{R}^n,$$

which maps a certain domain $U_p \subset \mathbb{R}^n$ diffeomorphically onto its image. It is standard [DeTurck and Kazdan 1981] that $\exp_p^A(Z)$ is a C^{s+1} function of $Z \in U_p$ and of $p \in \mathbb{R}^n$. The normal coordinates at p are just the Cartesian coordinates $Z = (Z_1, \dots, Z_n)$ on $U_p \subset \mathbb{R}^n$. In these coordinates, the metric reads as $\hat{A}(Z) = I + O(|Z|^2)$. More precisely, $\hat{A}(Z) = (\hat{a}_{ij}(Z))$ is given by the pullback by the exponential map of the metric tensor, which is well known to be of the form

$$(\exp_p^A)^*[a_{ij}(x)\, dx_i\, dx_j] =: \hat{a}_{ij}(Z)\, dZ_i\, dZ_j,$$

with functions $\hat{a}_{ij}$ of class $C^s(U_p)$ such that

$$\hat{a}_{ij}(0) = \delta_{ij}, \quad \partial_{Z_k}\hat{a}_{ij}(0) = 0.$$

Therefore, normal coordinates enable us to write the matrix as the identity plus a C^s-smooth quadratic error. Incidentally, it is well known that the leading-order contribution of the error is determined by the curvature of the metric A at the point p.

We are now ready to reformulate the overdetermined problem with a general function A as a small perturbation of the case $A(x) = I$. For each function $B \in C^{s+1}(\mathbb{S})$ with $\|B\|_{C^{s+1}} < 1$ and each ε small enough, one can then define the domain $\Omega_{p,\varepsilon B} \subset \mathbb{R}^n$ (which will play the same role as (2-1)) as

$$\Omega_{p,\varepsilon B} := \{\exp_p^A(\varepsilon z) : |z| < 1 + \varepsilon B(z/|z|)\}.$$

Note that, in terms of the spherical coordinates associated with a point z,

$$r := |z| \in (0, \infty), \quad \omega := \frac{z}{|z|} \in \mathbb{S},$$

the above condition reads simply as $r < 1 + \varepsilon B(\omega)$. In the domain $\Omega_{p,\varepsilon B}$, (1-2) reads in the rescaled normal coordinates z at p as

$$\hat{L}u + \bar{\lambda}\widehat{F}(z, u) = 0,$$

where $\widehat{F}(z, u) := F(\exp_p^A(\varepsilon z), u)$ and now the linear operator $\hat{L}$ is of the form

$$\hat{L}u := \hat{a}_{ij}(\varepsilon z)\, \partial_{z_i z_j} u + \varepsilon \hat{b}_i(\varepsilon z)\, \partial_{z_i} u,$$

with $\hat{a}_{ij}(Z)$ as above and some functions $\hat{b}_i(Z)$ of class C^s.

Therefore,

$$\hat{L}u = \Delta u + \varepsilon \hat{b}_i(\varepsilon z)\, \partial_{z_i} u + \mathcal{E}u,$$

where the error term is bounded as $\|\mathcal{E}u\|_{C^{s-1}} \leqslant C\varepsilon^2 \|u\|_{C^{s+1}}$ and $\hat{L}u - \mathcal{E}u$ is just like the operator $\tilde{L}u$ introduced below (2-2). One can now go over the proof of Theorem 1.1 and readily see that all the arguments remain valid when one introduces an error of this form in the expressions. This is not surprising, as the proof only uses the formulas for the terms in the equations that are of zeroth and first order in ε. Since the nondegenerate critical points of f_0 do not depend on the coordinate system, Theorem 1.1 is then proven for a general matrix-valued function A.

Acknowledgements

Domínguez-Vázquez is supported by the grants PID2019-105138GB-C21 (AEI, Spain) and ED431C 2019/10, ED431F 2020/04 (Xunta de Galicia, Spain), and by the Ramón y Cajal program of the Spanish Ministry of Science. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme through the Consolidator Grant agreement 862342 (Enciso). It is also supported by the grants CEX2019-000904-S, RED2022-134301-T, and PID2022-136795NB-I00 (Enciso and Peralta-Salas) funded by MCIN/ AEI.

References

[Aftalion and Busca 1998] A. Aftalion and J. Busca, "Radial symmetry of overdetermined boundary-value problems in exterior domains", *Arch. Ration. Mech. Anal.* **143**:2 (1998), 195–206. MR Zbl

[Aftalion et al. 1999] A. Aftalion, J. Busca, and W. Reichel, "Approximate radial symmetry for overdetermined boundary value problems", *Adv. Differential Equations* **4**:6 (1999), 907–932. MR Zbl

[Alessandrini and Garofalo 1989] G. Alessandrini and N. Garofalo, "Symmetry for degenerate parabolic equations", *Arch. Ration. Mech. Anal.* **108**:2 (1989), 161–174. MR Zbl

[Bianchini et al. 2014] C. Bianchini, A. Henrot, and P. Salani, "An overdetermined problem with non-constant boundary condition", *Interfaces Free Bound.* **16**:2 (2014), 215–241. MR Zbl

[Brandolini et al. 2008] B. Brandolini, C. Nitsch, P. Salani, and C. Trombetti, "Serrin-type overdetermined problems: an alternative proof", *Arch. Ration. Mech. Anal.* **190**:2 (2008), 267–280. MR Zbl

[Cianchi and Salani 2009] A. Cianchi and P. Salani, "Overdetermined anisotropic elliptic problems", *Math. Ann.* **345**:4 (2009), 859–881. MR Zbl

[Damascelli et al. 1999] L. Damascelli, F. Pacella, and M. Ramaswamy, "Symmetry of ground states of p-Laplace equations via the moving plane method", *Arch. Ration. Mech. Anal.* **148**:4 (1999), 291–308. MR Zbl

[Delay and Sicbaldi 2015] E. Delay and P. Sicbaldi, "Extremal domains for the first eigenvalue in a general compact Riemannian manifold", *Discrete Contin. Dyn. Syst.* **35**:12 (2015), 5799–5825. MR Zbl

[DeTurck and Kazdan 1981] D. M. DeTurck and J. L. Kazdan, "Some regularity theorems in Riemannian geometry", *Ann. Sci. École Norm. Sup.* (4) **14**:3 (1981), 249–260. MR Zbl

[Domínguez-Vázquez et al. 2019] M. Domínguez-Vázquez, A. Enciso, and D. Peralta-Salas, "Solutions to the overdetermined boundary problem for semilinear equations with position-dependent nonlinearities", *Adv. Math.* **351** (2019), 718–760. MR Zbl

[Domínguez-Vázquez et al. 2021] M. Domínguez-Vázquez, A. Enciso, and D. Peralta-Salas, "Piecewise smooth stationary Euler flows with compact support via overdetermined boundary problems", *Arch. Ration. Mech. Anal.* **239**:3 (2021), 1327–1347. MR Zbl

[Farina and Valdinoci 2010] A. Farina and E. Valdinoci, "Flattening results for elliptic PDEs in unbounded domains with applications to overdetermined problems", *Arch. Ration. Mech. Anal.* **195**:3 (2010), 1025–1058. MR Zbl

[Farina and Valdinoci 2013] A. Farina and E. Valdinoci, "On partially and globally overdetermined problems of elliptic type", *Amer. J. Math.* **135**:6 (2013), 1699–1726. MR Zbl

[Fragalà and Gazzola 2008] I. Fragalà and F. Gazzola, "Partially overdetermined elliptic boundary value problems", *J. Differential Equations* **245**:5 (2008), 1299–1322. MR Zbl

[Fragalà et al. 2006] I. Fragalà, F. Gazzola, and B. Kawohl, "Overdetermined problems with possibly degenerate ellipticity: a geometric approach", *Math. Z.* **254**:1 (2006), 117–132. MR Zbl

[Garofalo and Lewis 1989] N. Garofalo and J. L. Lewis, "A symmetry result related to some overdetermined boundary value problems", *Amer. J. Math.* **111**:1 (1989), 9–33. MR Zbl

[Garofalo and Sartori 1999] N. Garofalo and E. Sartori, "Symmetry in exterior boundary value problems for quasilinear elliptic equations via blow-up and a priori estimates", *Adv. Differential Equations* **4**:2 (1999), 137–161. MR Zbl

[Kawohl 1998] B. Kawohl, “Symmetry or not?”, *Math. Intelligencer* **20**:2 (1998), 16–22. MR Zbl

[Kumaresan and Prajapat 1998] S. Kumaresan and J. Prajapat, “Serrin’s result for hyperbolic space and sphere”, *Duke Math. J.* **91**:1 (1998), 17–28. MR Zbl

[Magnanini and Poggesi 2020a] R. Magnanini and G. Poggesi, “Nearly optimal stability for Serrin’s problem and the soap bubble theorem”, *Calc. Var. Partial Differential Equations* **59**:1 (2020), art. id. 35. MR Zbl

[Magnanini and Poggesi 2020b] R. Magnanini and G. Poggesi, “Serrin’s problem and Alexandrov’s soap bubble theorem: enhanced stability via integral identities”, *Indiana Univ. Math. J.* **69**:4 (2020), 1181–1205. MR Zbl

[Pacard and Sicbaldi 2009] F. Pacard and P. Sicbaldi, “Extremal domains for the first eigenvalue of the Laplace–Beltrami operator”, *Ann. Inst. Fourier* (*Grenoble*) **59**:2 (2009), 515–542. MR Zbl

[del Pino et al. 2015] M. del Pino, F. Pacard, and J. Wei, “Serrin’s overdetermined problem and constant mean curvature surfaces”, *Duke Math. J.* **164**:14 (2015), 2643–2722. MR Zbl

[Prajapat 1998] J. Prajapat, “Serrin’s result for domains with a corner or cusp”, *Duke Math. J.* **91**:1 (1998), 29–31. MR Zbl

[Rayleigh 1877] J. W. Strutt, 3rd baron Rayleigh, *The theory of sound, I*, Macmillan, London, 1877.

[Ros et al. 2020] A. Ros, D. Ruiz, and P. Sicbaldi, “Solutions to overdetermined elliptic problems in nontrivial exterior domains”, *J. Eur. Math. Soc.* **22**:1 (2020), 253–281. MR Zbl

[Serrin 1971] J. Serrin, “A symmetry problem in potential theory”, *Arch. Ration. Mech. Anal.* **43** (1971), 304–318. MR Zbl

[Traizet 2014] M. Traizet, “Classification of the solutions to an overdetermined elliptic problem in the plane”, *Geom. Funct. Anal.* **24**:2 (2014), 690–720. MR Zbl

[Ye 1991] R. Ye, “Foliation by constant mean curvature spheres”, *Pacific J. Math.* **147**:2 (1991), 381–396. MR Zbl

Received 18 Aug 2020. Revised 23 Mar 2022. Accepted 14 Apr 2022.

MIGUEL DOMÍNGUEZ-VÁZQUEZ: miguel.dominguez@usc.es
CITMAga, Department of Mathematics, Universidade de Santiago de Compostela, Spain

ALBERTO ENCISO: aenciso@icmat.es
Instituto de Ciencias Matemáticas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

DANIEL PERALTA-SALAS: dperalta@icmat.es
Instituto de Ciencias Matemáticas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2005–2040
DOI: 10.2140/apde.2023.16.2005

MONGE–AMPÈRE GRAVITATION AS A Γ-LIMIT OF GOOD RATE FUNCTIONS

LUIGI AMBROSIO, AYMERIC BARADAT AND YANN BRENIER

Monge–Ampère gravitation is a modification of the classical Newtonian gravitation where the linear Poisson equation is replaced by the nonlinear Monge–Ampère equation. This paper is concerned with the rigorous derivation of Monge–Ampère gravitation for a finite number of particles from the stochastic model of a Brownian point cloud, following the formal ideas of a recent work by Brenier (*Bull. Inst. Math. Acad. Sin.* **11**:1(2016), 23–41). This is done in two steps. First, we compute the good rate function corresponding to a large deviation problem related to the Brownian point cloud at fixed positive diffusivity. Second, we study the Γ-convergence of this good rate function, as the diffusivity tends to zero, toward a (nonsmooth) Lagrangian encoding the Monge–Ampère dynamic. Surprisingly, the singularities of the limiting Lagrangian correspond to dissipative phenomena. As an illustration, we show that they lead to sticky collisions in one space dimension.

1. Introduction

Monge–Ampère gravitation. On a periodic domain such as $\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$, Newtonian gravitation is commonly described in terms of the density of probability $f(t, x, \xi)$ to find gravitating matter at time t, position $x \in \mathbb{T}^d$ and velocity $\xi \in \mathbb{R}^d$, subject to the Vlasov–Poisson equation

$$\begin{gathered}\partial_t f(t,x,\xi) + \operatorname{div}_x(\xi f(t,x,\xi)) - \operatorname{div}_\xi(\nabla\varphi(t,x) f(t,x,\xi)) = 0,\\ \Delta\varphi(t,x) = \int_{\mathbb{R}^d} f(t,x,\xi)\,\mathrm{d}\xi - 1, \quad (t,x,\xi)\in\mathbb{R}\times\mathbb{T}^d\times\mathbb{R}^d,\end{gathered}$$

where φ is the gravitational potential. Notice that the averaged density, say 1, has been subtracted out from the right-hand side of the Poisson equation, due to the periodicity of the spatial domain. This is a common feature of computational cosmology and it lets the uniform density be a stationary solution. The Vlasov–Poisson system can be seen as an "approximation" to the more nonlinear Vlasov–Monge–Ampère (VMA) system

$$\partial_t f(t,x,\xi) + \operatorname{div}_x(\xi f(t,x,\xi)) - \operatorname{div}_\xi(\nabla\varphi(t,x) f(t,x,\xi)) = 0, \tag{1}$$

$$\det(\mathbb{I} + \mathrm{D}^2\varphi(t,x)) = \int_{\mathbb{R}^d} f(t,x,\xi)\,\mathrm{d}\xi, \quad (t,x,\xi)\in\mathbb{R}\times\mathbb{T}^d\times\mathbb{R}^d, \tag{2}$$

where the fully nonlinear Monge–Ampère equation substitutes for the linear Poisson equation of Newtonian gravitation. Indeed, for "weak" gravitational potentials, by expanding the determinant around the identity matrix $\mathbb{I}$, we get

$$\det(\mathbb{I} + \mathrm{D}^2\varphi(t,x)) \sim 1 + \operatorname{tr}(\mathrm{D}^2\varphi(t,x)) = 1 + \Delta\varphi(t,x)$$

MSC2020: 49J52, 60F10, 70F40, 70B05.
Keywords: Monge–Ampère gravitation, large deviations, Γ-convergence, Lagrangian mechanics, interacting particle systems.

and recover the Newtonian model approximately (and exactly as $d = 1$). In this paper, we will speak of "Monge–Ampère gravitation" ("MAG" in short). The Vlasov–Monge–Ampère system was introduced and related to the Vlasov–Poisson system in [Brenier and Loeper 2004], and studied as an ODE on the Wasserstein space in [Ambrosio and Gangbo 2008]. It can also be solved numerically thanks to efficient Monge–Ampère solvers recently designed by Mérigot [2011]. It was argued in [Brenier 2011] that MAG may also be seen as an approximation of Newtonian gravitation for which the "Zeldovich approximation" [1970] (see [Frisch et al. 2002; Brenier et al. 2003]), popular in computational cosmology, becomes exact.

Derivation of a discrete model of MAG. In what follows, we will not be directly interested in the aforementioned system, but rather in its discrete version, i.e., when the number of particles is finite. As is well known in optimal transport theory [Brenier 1987; 1991; Villani 2003], the Monge–Ampère equation (2) is solved by the unique function φ such that the map $\mathrm{Id}+\nabla\varphi$ realizes the optimal transport with quadratic cost from the density $\int f \,\mathrm{d}\xi$ to the Lebesgue measure. Then, the kinetic equation (1) is known to be the continuous version of the Newton equations of classical mechanics in a potential given by φ.

In the discrete setting, the stationary Lebesgue measure is replaced by a family $A = (a_1, \dots, a_N) \in (\mathbb{R}^d)^N$ of $N \geq 1$ points in $\mathbb{R}^d$ (here we make the presentation in $\mathbb{R}^d$ instead of $\mathbb{T}^d$ for the sake of simplicity). One can for instance think of a regular lattice approximating in some region a constant density, even though in the sequel the particular choice of $(a_1, \dots, a_N)$ will play no role. We will consider the evolution of a cloud $X = (x_1, \dots, x_N)$ of N particles $x_1, \dots, x_N$ in $\mathbb{R}^d$ whose dynamic is ruled by the $(-1/N)$-convex function induced by the discrete optimal transport problem

$$F(X) := -\min_{\sigma\in\mathfrak{S}_N} \frac{1}{2N}\sum_{i=1}^N |x_i - a_{\sigma(i)}|^2 = -\frac{1}{2}W_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{a_i}, \frac{1}{N}\sum_{i=1}^N \delta_{x_i}\right), \tag{3}$$

where W_2 is the so-called Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$, the set of Borel probability measures on $\mathbb{R}^d$ having a finite second-order moment. At least in the case where the optimization problem in (3) admits a unique minimizer $\sigma_{\mathrm{opt}} = \sigma^X_{\mathrm{opt}}$, the analogue of (1), (2) in this framework is easily seen to be formally,

$$\text{for all } i = 1, \dots, N, \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2} x_i(t) = x_i(t) - a_{\sigma_{\mathrm{opt}}(i)}, \tag{4}$$

which can be rewritten as, letting $\mathcal{X}_t := (x_1(t), \dots, x_N(t))$,

$$\frac{1}{N}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{X}_t = -\nabla F(\mathcal{X}_t). \tag{5}$$

Following the ideas of the recent paper [Brenier 2016], we will derive this discrete dynamic from the very elementary stochastic model of a Brownian point cloud. However, in [Brenier 2016], the derivation was obtained by applying two successive large deviation principles (LDP), through a purely formal use of the Freidlin–Wentzell theory [1998]. The main purpose of the present paper is to explain how such a derivation can be made rigorous by substituting for one of the applications of the LDP a PDE method inspired by the famous concept of "onde pilote" introduced by Louis de Broglie [1927] at the early stage of quantum mechanics.

Dealing with the singularities. Due to the lack of uniqueness in the discrete optimal transport problem, solutions of (4) are not always well-defined a priori. Otherwise stated, F is singular, and therefore ∇F in (5) is not everywhere meaningful. A standard choice to give sense to (5) is to restate it as

$$-\frac{1}{N}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{X}_t \in \partial F(\mathcal{X}_t), \tag{6}$$

where $\partial F(\mathcal{X}_t)$ is the subdifferential of F at $\mathcal{X}_t$, or

$$-\frac{1}{N}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{X}_t = \overline{\nabla} F(\mathcal{X}_t),$$

where $\overline{\nabla} F(\mathcal{X}_t)$ is the element of $\partial F(\mathcal{X}_t)$ with minimal Euclidean norm (see Definition 8 below). In these formulations, existence results are available even in the nondiscrete case [Ambrosio and Gangbo 2008].

This is not what we do: our approach selects minimizers of actions appearing as Γ-limits of good rate functions associated with some LDP, under endpoint constraints. These curves do solve (4) in the case where σ_{opt} is unique, but this time, the relaxation is made at the level of the Lagrangian formulation, and not at the level of the Hamiltonian one. In view of (5), we would expect to find the action

$$\int_{t_0}^{t_1}\left\{\frac{|\dot{\mathcal{X}}_t|^2}{2} - NF(\mathcal{X}_t)\right\}\mathrm{d}t, \tag{7}$$

where t_0, t_1 are some prescribed initial and final times. Instead, our derivation ends up with the smaller action

$$\int_{t_0}^{t_1}\left\{\frac{|\dot{\mathcal{X}}_t|^2}{2} + \frac{|\mathcal{X}_t - \overline{\nabla} f(\mathcal{X}_t)|^2}{2}\right\}\mathrm{d}t, \tag{8}$$

$$f(X) := \max_{\sigma\in\mathfrak{S}_N}\sum_{i=1}^N x_i a_{\sigma(i)} = \sum_{i=1}^N x_i a_{\sigma_{\mathrm{opt}}^X(i)}, \qquad X = (x_1,\dots,x_N)\in(\mathbb{R}^d)^N.$$

Note that these two actions coincide on curves $\mathcal{X}$ such that, for almost every t, $\sigma_{\mathrm{opt}}^{\mathcal{X}_t}$ is unique (see Section 2.7 for more details). Unexpectedly, this action is *exactly* the one previously suggested by the third author in [Brenier 2011] in order to include dissipative phenomena (such as sticky collisions in one space dimension) in the Monge–Ampère gravitational model!

The classical theory for sticky particles vs. our approach. Systems of particles moving along the line and that stick together when they meet have been studied for a long time, for instance because they were suggested to model the formation of large structures in the universe [Zeldovich 1970]. On the mathematical side, a lot of works have been devoted to studying the limit of this kind of system when the number of particles tends to infinity (see for instance [E et al. 1996; Brenier and Grenier 1998]), and the most recent works generally build on a connection with the theory of optimal transport (see [Natile and Savaré 2009; Brenier et al. 2013; Hynd 2020]). An example illustrating this link, which is one of the main theorems in [Natile and Savaré 2009], is that up to a change of time, the one-dimensional pressureless Euler system with sticky collisions is the gradient flow in the Wasserstein space of $-\frac{1}{2}W_2^2(\mathsf{m},\cdot)$, where $\mathsf{m}\in\mathcal{P}_2(\mathbb{R})$ is a reference probability measure on the line. In plain English, in these models, particles are only allowed to stick when they meet, and it corresponds to the optimal way of decreasing a certain functional.

Our approach is different. In fact, our model is a least action principle, and therefore is *conservative* and *time-reversible*. In this context, sticky collisions happen due to the presence of an internal energy, corresponding to the discontinuities of the potential energy $X \mapsto -\frac{1}{2}|X - \overline{\nabla} f(X)|^2$ (see formula (51)), and which grows when particles aggregate. Kinetic energy can hence be transferred into internal energy through perfectly inelastic shocks. An output of these considerations is that in our case, particles are not only allowed to stick together; they can also separate.

Outline. In Section 2 we show how to derive MAG starting from a finite number of Brownian particles. This is done in several steps, the main one being the Γ-convergence towards the "effective" singular functional (8) of the good rate functions associated with the large deviations of the solutions of a family of SDEs (up to a change of time). This is stated in Theorem 9, which is our main result. Section 3 is devoted to the proof of Theorem 9. The purpose of Section 4 is to show that in one space dimension, the dissipative phenomena induced by this functional lead to sticky collisions.

Notation. We will work with N particles in $\mathbb{R}^d$, and hence in $(\mathbb{R}^d)^N$. Points of $(\mathbb{R}^d)^N$ will be denoted with capital letters, mainly X, Y or Z. Curves with values in $(\mathbb{R}^d)^N$ will be denoted with calligraphic letters $\mathcal{X}$, $\mathcal{Y}$ or $\mathcal{Z}$. The positions of $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ at time $t \in \mathbb{R}$ will be denoted by $\mathcal{X}_t$, $\mathcal{Y}_t$ and $\mathcal{Z}_t$ respectively.

In order to avoid heavy notation, in most cases, the laws of the processes that we will consider will be continuously parametrized. In these cases, we will use abuses of notation: for instance, we will say that the family of laws $(\mu_\eta)_{\eta>0}$ is tight whenever it is tight for sufficiently small values of η. This is equivalent to $(\mu_{\eta_n})_{n\in\mathbb{N}}$ being tight for all $(\eta_n) \in (\mathbb{R}_+^*)^{\mathbb{N}}$ decreasing to 0.

2. Derivation of the discrete model

2.1. ***The stochastic model of a lattice with Brownian motion.*** Take $A = (a_1, \dots, a_N) \in (\mathbb{R}^d)^N$ to be a family of $N > 1$ points in $\mathbb{R}^d$. We assume each point of this lattice to be subject to Brownian motion for times $t \geq 0$. At time t, the position of point i is

$$a_i + \sqrt{\varepsilon} B_t^i,$$

where $(B^i)_{i=1,\dots,d}$ is a family of N independent normalized Brownian curves and ε monitors the (common) level of noise. As a consequence, at time $t > 0$, the density of probability $\rho_\varepsilon(t, X)$ for the point cloud

$$(a_1 + \sqrt{\varepsilon} B_t^1, \dots, a_N + \sqrt{\varepsilon} B_t^N)$$

to be observed at location $X = (x_1, \dots, x_d) \in (\mathbb{R}^d)^N$, up to a permutation $\sigma \in \mathfrak{S}_N$ of the labels, is easy to compute. We find

$$\rho_\varepsilon(t, X) = \frac{1}{N!\,\sqrt{2\pi\varepsilon t}^{dN}} \sum_{\sigma\in\mathfrak{S}_N} \prod_{i=1}^{N} \exp\biggl(-\frac{|x_i - a_{\sigma(i)}|^2}{2\varepsilon t}\biggr),$$

or, in short,

$$\frac{1}{N!\,\sqrt{2\pi\varepsilon t}^{Nd}} \sum_{\sigma\in\mathfrak{S}_N} \exp\biggl(-\frac{|X - A^\sigma|^2}{2\varepsilon t}\biggr),$$

where $|\cdot|$ denotes the euclidean norm in $\mathbb{R}^d$ or $(\mathbb{R}^d)^N$ depending on the context, and where, for all $X = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$, we let

$$X^\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(N)}).$$

This was the starting point of the discussion made in [Brenier 2016], using a double large deviation principle.

In the present paper, we rather turn to a PDE viewpoint, where ρ_ε is the solution of the heat equation in $(\mathbb{R}^d)^N$,

$$\frac{\partial \rho_\varepsilon}{\partial t}(t, X) = \frac{\varepsilon}{2} \Delta \rho_\varepsilon(t, X), \tag{9}$$

with, as initial condition, the delta measure located at $A = (a_1, \ldots, a_N) \in (\mathbb{R}^d)^N$ and symmetrized with respect to $\sigma \in \mathfrak{S}_N$, namely

$$\rho_\varepsilon(0, X) = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \delta_{A^\sigma}. \tag{10}$$

In some sense, we have solved the heat equation in the space of "point clouds" $(\mathbb{R}^d)^N / \mathfrak{S}_N$, with initial position A, defined up to a permutation $\sigma \in \mathfrak{S}_N$ of the labels $i = 1, \ldots, N$.

2.2. ***"Surfing" the "heat wave".*** After solving the heat equation (9)–(10), in the space of "clouds" $(\mathbb{R}^d)^N / \mathfrak{S}_N$, we introduce the companion ODE in the space $(\mathbb{R}^d)^N$:

$$\frac{\mathrm{d}\mathcal{X}_t^\varepsilon}{\mathrm{d}t} = v_\varepsilon(t, \mathcal{X}_t^\varepsilon), \quad v_\varepsilon(t, X) = -\frac{\varepsilon}{2} \nabla \log \rho_\varepsilon(t, X), \tag{11}$$

or, more explicitly

$$\begin{aligned} v_\varepsilon(t, X) &= \frac{1}{2t} \frac{\sum_{\sigma \in \mathfrak{S}_N} (X - A^\sigma) \exp(-|X - A^\sigma|^2/(2\varepsilon t))}{\sum_{\sigma \in \mathfrak{S}_N} \exp(-|X - A^\sigma|^2/(2\varepsilon t))} \\ &= \frac{1}{2t} \left(X - \frac{\sum_{\sigma \in \mathfrak{S}_N} A^\sigma \exp((X \cdot A^\sigma)/(\varepsilon t))}{\sum_{\sigma \in \mathfrak{S}_N} \exp((X \cdot A^\sigma)/(\varepsilon t))} \right), \end{aligned} \tag{12}$$

where if U and V are in $(\mathbb{R}^d)^N$, then $U \cdot V$ denotes the inner product between U and V. This velocity is chosen so that

$$\frac{\partial \rho_\varepsilon}{\partial t}(t, X) + \operatorname{div}(\rho_\varepsilon(t, X) v_\varepsilon(t, X)) = 0,$$

i.e., for the density ρ_ε to be transported by the velocity field v_ε. We may solve this ODE for arbitrarily chosen position $\mathcal{X}_{t_0} \in (\mathbb{R}^d)^N$ (up to reordering) and initial time $t_0 > 0$.

Put another way, we consider the characteristics corresponding to the heat equation (9)–(10), interpreted as a continuity equation, associated to our Brownian point cloud.

Remark 1. By doing that change of perspective, we just mimic the idea of quantum particles driven by the "onde pilote", as imagined by Louis de Broglie [1927; 1959] at the early stage of quantum mechanics.

Indeed, in our case, the velocity $v^\varepsilon = \nabla\varphi^\varepsilon$ is the gradient of the scalar function $\varphi^\varepsilon := (-\varepsilon/2)\log\rho^\varepsilon$, and the pair $(\rho^\varepsilon, \varphi^\varepsilon)$ is easily seen to solve the system

$$\begin{cases} \partial_t \rho^\varepsilon + \operatorname{div}(\rho^\varepsilon \nabla\varphi^\varepsilon) = 0, \\ \partial_t \varphi^\varepsilon + \dfrac{1}{2}|\nabla\varphi^\varepsilon|^2 = -\dfrac{\varepsilon^2}{2}\dfrac{\Delta\sqrt{\rho^\varepsilon}}{\sqrt{\rho^\varepsilon}}, \end{cases} \tag{13}$$

that is, the characteristics follow the trajectories of Newton's law in a potential induced by ρ^ε.

In the quantum case, something very similar can be found with the help of the Madelung transform [1927]. Namely, if the complex function Ψ^ε solves the Schrödinger equation

$$i\partial_t\Psi^\varepsilon + \frac{\varepsilon}{2}\Delta\Psi^\varepsilon = 0,$$

writing $\Psi^\varepsilon = \sqrt{\rho^\varepsilon}e^{i\varphi^\varepsilon/\varepsilon}$ for a pair $(\rho^\varepsilon, \varphi^\varepsilon)$ of real functions, then this pair is shown to formally solve the very similar system

$$\begin{cases} \partial_t \rho^\varepsilon + \operatorname{div}(\rho^\varepsilon \nabla\varphi^\varepsilon) = 0, \\ \partial_t \varphi^\varepsilon + \dfrac{1}{2}|\nabla\varphi^\varepsilon|^2 = \dfrac{\varepsilon^2}{2}\dfrac{\Delta\sqrt{\rho^\varepsilon}}{\sqrt{\rho^\varepsilon}}, \end{cases} \tag{14}$$

and this observation was the starting point of de Broglie's interpretation of quantum mechanics. In this case, the potential in the right-hand side of the second equation is called the Bohm quantum potential. However, the analysis of (14) is substantially more difficult than the one of (13), due to the possible vanishing of the wave function Ψ^ε during the evolution.

This analogy is not a coincidence. Indeed, it is known [von Renesse 2012] that the Schrödinger equation in its Madelung formulation (14) is formally the Hamiltonian flow corresponding to the Lagrangian

$$\mathcal{L}^\varepsilon_{\text{quantum}}(\rho, \nabla\varphi) := \frac{1}{2}\int \left\{ |\nabla\varphi|^2 - \left|\frac{\varepsilon}{2}\nabla\log\rho\right|^2 \right\}\rho,$$

in the geometry of optimal transport, while system (13), which admits solutions of the heat equations as particular solutions, is *rigorously* the Hamiltonian flow corresponding to the Lagrangian

$$\mathcal{L}^\varepsilon_{\text{heat}}(\rho, \nabla\varphi) := \frac{1}{2}\int \left\{ |\nabla\varphi|^2 + \left|\frac{\varepsilon}{2}\nabla\log\rho\right|^2 \right\}\rho,$$

in the geometry of optimal transport [Conforti 2019]. The latter Lagrangian appears naturally in the theory of entropic optimal transport; see [Gentil et al. 2017; Gigli and Tamanini 2020].

2.3. ***Large deviations of the "heat wave" ODE.*** Let us now add to the ODE of the previous subsection a noise of the form

$$\mathrm{d}\mathcal{X}^{\varepsilon,\eta}_t = v_\varepsilon(t, \mathcal{X}^{\varepsilon,\eta}_t)\,\mathrm{d}t + \sqrt{\frac{\eta}{t}}\,\mathrm{d}\mathcal{W}_t, \tag{15}$$

where η is a positive number, $(\mathcal{W}_t)$ is a standard Brownian motion in $(\mathbb{R}^d)^N$, and where the scaling prefactor $1/\sqrt{t}$ has been chosen to recover MAG at Section 2.6. That is, we include a second time-dependent level of noise to the model: we perturb the characteristics that were already generated, through the heat equation, by the Brownian motion of our original point cloud.

Our main finding is that when η and ε are small and up to a change of time, the trajectories charged by the solution of this SDE starting from $P \in (\mathbb{R}^d)^N$ at time $t_0 > 0$ and which happen to be close to $Q \in (\mathbb{R}^d)^N$ at time $t_1 > t_0$ are well-approximated by MAG.

The purpose of the rest of this section will be to make this rough statement precise. When we say that some random trajectories are well-approximated by MAG, we mean that they are close in the uniform topology to minimizers of the action (8), with large probability. Justifying this fact will require several steps and intermediate functionals. As the times t_0 and t_1, as well as the endpoints P and Q, will be fixed in what follows, we decided not to refer to them in the notation for the different functionals and laws that will appear.

Since for fixed $\varepsilon > 0$ and $t \geq t_0 > 0$, v_ε is a smooth velocity field, the existence of a strong solution and pathwise uniqueness for (15) is standard once fixed a law for the initial position $\mathcal{X}_{t_0}^{\varepsilon,\eta}$ at some $t_0 > 0$. Since we want to consider indistinguishable particles, a relevant choice of initial law consists in taking $\mathcal{X}_{t_0}^{\varepsilon,\eta} = P^\sigma$ with probability $1/(N!)$, given some $P \in (\mathbb{R}^d)^N$ and $\sigma \in \mathfrak{S}_N$. For convenience, from now on, we denote by $\{P^\sigma\}$ the set $\{P^\sigma : \sigma \in \mathfrak{S}_N\}$. The law just described is nothing but the uniform law on $\{P^\sigma\}$.

Remark 2. Actually, at this stage, it would be possible to reintroduce distinguishability: Theorem 3, Corollary 4, Proposition 7 and Theorem 9 below could easily be written for distinguishable particles, that is, with constraints on the endpoints of the curves, and not on these endpoints up to reordering. We decided to keep on working on clouds of indistinguishable particles in order to avoid crossings of trajectories in Section 4.

The first step consists in using classical Freidlin–Wentzell theory [1998] (see also [Dembo and Zeitouni 1998]) in order to pass to the limit $\eta \to 0$, while $\varepsilon > 0$ is kept fixed, in the sense of large deviations (we omit the proof since it consists in adapting in a straightforward way [Dembo and Zeitouni 1998, Theorem 5.6.3] to time-dependent entries and more general initial law for the SDE).

Theorem 3. *Let us fix two positive times $0 < t_0 < t_1$ and $P \in (\mathbb{R}^d)^N$. For fixed $\varepsilon > 0$ and as $\eta \to 0$, the family of laws $(\mu_{\varepsilon,\eta})$ of the solution of* (15) *between times t_0 and t_1 and starting from the uniform law on $\{P^\sigma\}$ satisfies the LDP on $C^0([t_0, t_1]; (\mathbb{R}^d)^N)$ with good rate function L_ε^0 defined for all $\mathcal{X} = (\mathcal{X}_t)_{t\in[t_0,t_1]}$ by*

$$L_\varepsilon^0(\mathcal{X}) = \begin{cases} \frac{1}{2}\int_{t_0}^{t_1} |\dot{\mathcal{X}}_t - v_\varepsilon(t, \mathcal{X}_t)|^2 \times t \,\mathrm{d}t & \text{if } \mathcal{X} \in H^1([t_0, t_1]; (\mathbb{R}^d)^N) \text{ and } \mathcal{X}_{t_0} \in \{P^\sigma\}, \\ +\infty, & \text{else}. \end{cases} \tag{16}$$

In the rest of the article, we will call these kind of functionals "actions", instead of the usual terminology "good rate function".

An outcome of this result is that with large probability, when η is small, $\mathcal{X}_{t_1}^{\eta,\varepsilon}$ is close to the position at time t_1 of the solution of the ODE (11) starting from P, up to reordering. But now, we want to use Theorem 3 in order to describe the behavior of the solutions of the SDE (15) when η is small, under the large deviation assumption that its final position $\mathcal{X}_{t_1}^{\varepsilon,\eta}$ is far from this expected value.

For this, we take $Q \in (\mathbb{R}^d)^N$, and we assume that we observe $\mathcal{X}_{t_1}^{\varepsilon,\eta}$ to be close to Q, up to reordering. To quantify this closeness, we consider a new small parameter $\delta > 0$, and we work with the laws $(\mu_{\varepsilon,\eta})$ from Theorem 3, conditioned to the event $\{\mathcal{X}_{t_1}^{\varepsilon,\eta} \in \bigcup_{\sigma\in\mathfrak{S}_N} \bar{B}(Q^\sigma, \delta)\}$, where for a given $X \in (\mathbb{R}^d)^N$, $\bar{B}(X, \delta)$ stands for the closed ball of center X and radius δ. MAG will be obtained by studying the limit of these conditional laws when $\eta \to 0$, then $\delta \to 0$ and finally $\varepsilon \to 0$. We refer to Remark 12 for a discussion about the order in which we let the different parameters tend to 0.

Concerning the limit $\eta \to 0$, Theorem 3 implies the following.

Corollary 4. *Let us fix $\varepsilon, \delta > 0$, and call $\mathcal{E}^\delta$ the closed subset of $C^0([t_0, t_1]; (\mathbb{R}^d)^N)$ defined by*

$$\mathcal{E}^\delta := \Big\{ \mathcal{X} \in C^0([t_0, t_1]; (\mathbb{R}^d)^N) : \mathcal{X}_{t_1} \in \bigcup_{\sigma \in \mathfrak{S}_N} \bar{B}(Q^\sigma, \delta) \Big\}.$$

The family of conditional laws $(\mu^\delta_{\varepsilon,\eta} := \mu_{\varepsilon,\eta}(\,\cdot : \mathcal{E}^\delta))_{\eta > 0}$ is tight. Moreover, its limit points for the topology of narrow convergence as $\eta \to 0$ only charge minimizers of the functional

$$L^\delta_\varepsilon(\mathcal{X}) = \begin{cases} \frac{1}{2}\int_{t_0}^{t_1} |\dot{\mathcal{X}}_t - v_\varepsilon(t, \mathcal{X}_t)|^2 \times t \,\mathrm{d}t & \text{if } \mathcal{X} \in H^1([t_0, t_1]; (\mathbb{R}^d)^N),\ \mathcal{X}_{t_0} \in \{P^\sigma\} \text{ and } \mathcal{X}_{t_1} \in \bigcup_{\sigma \in \mathfrak{S}_N} \bar{B}(Q^\sigma, \delta), \\ +\infty, & \text{else.} \end{cases} \tag{17}$$

Proof. Let us first prove the tightness property. Let $\mathcal{X}$ be a curve in the interior of $\mathcal{E}^\delta$. As it satisfies an LDP associated with a good rate function in a Polish space, by virtue of [Dembo and Zeitouni 1998, Exercise 4.1.10], for fixed $\varepsilon > 0$, the family of laws $(\mu_{\varepsilon,\eta})_{\eta > 0}$ is exponentially tight. Hence, there is a compact K (we call K^c its complement in $C^0([t_0, t_1]; (\mathbb{R}^d)^N)$) such that

$$\limsup_{\eta \to 0} \eta \log \mu_{\varepsilon,\eta}(K^c) \le -L^0_\varepsilon(\mathcal{X}) - 1.$$

Therefore, we find

$$\begin{aligned} \limsup_{\eta \to 0} \eta \log \mu^\delta_{\varepsilon,\eta}(K^c) &= \limsup_{\eta \to 0} \{\eta \log \mu_{\varepsilon,\eta}(K^c \cap \mathcal{E}^\delta) - \eta \log \mu_{\varepsilon,\eta}(\mathcal{E}^\delta)\} \\ &\le \limsup_{\eta \to 0} \eta \log \mu_{\varepsilon,\eta}(K^c) - \liminf_{\eta \to 0} \eta \log \mu_{\varepsilon,\eta}(\mathring{\mathcal{E}}^\delta) \\ &\le -L^0_\varepsilon(\mathcal{X}) - 1 + L^0_\varepsilon(\mathcal{X}) \le -1. \end{aligned}$$

The tightness follows.

Now, let us consider μ a limit point of $(\mu^\delta_{\varepsilon,\eta})$ as $\eta \to 0$, and (η_n) a sequence of positive numbers decreasing to 0, with $\mu^\delta_{\varepsilon,\eta_n} \to \mu$ as $n \to +\infty$. We will argue that whenever $\mathcal{X}$ is not a minimizer of L^δ_ε, then $\mathcal{X}$ is not in the support of μ. First, for all $\eta > 0$, the support of $\mu^\delta_{\varepsilon,\eta}$ is a subset of $\mathcal{E}^\delta$. As the latter is closed, this is also the case for the support of μ. So let us take $\mathcal{X} \in \mathcal{E}^\delta$, which is not a minimizer of L^δ_ε. In particular, $L^0_\varepsilon(\mathcal{X}) > \inf_{\mathcal{E}^\delta} L^0_\varepsilon$. As L^0_ε is lower semicontinuous, there exists an open set U of $C^0([t_0, t_1]; (\mathbb{R}^d)^N)$ containing $\mathcal{X}$ such that $\inf_{\bar{U}} L^0_\varepsilon > \inf_{\mathcal{E}^\delta} L^0_\varepsilon$. Let us show that $\mu(U) = 0$.

By the Portmanteau theorem, we have

$$\mu(U) \le \liminf_{n \to +\infty} \mu^\delta_{\varepsilon,\eta_n}(U).$$

By the definition of $(\mu^\delta_{\varepsilon,\eta})$, we have

$$\eta_n \log \mu^\delta_{\varepsilon,\eta_n}(U) = \eta_n \log \mu_{\varepsilon,\eta_n}(U \cap \mathcal{E}^\delta) - \eta_n \log \mu_{\varepsilon,\eta_n}(\mathcal{E}^\delta) \le \eta_n \log \mu_{\varepsilon,\eta_n}(\bar{U}) - \eta_n \log \mu_{\varepsilon,\eta_n}(\mathring{\mathcal{E}}^\delta).$$

The large deviation principle of Theorem 3 lets us estimate the lim sup of this quantity by

$$\limsup_{n \to +\infty} \eta_n \log \mu^\delta_{\varepsilon,\eta_n}(U) \le \inf_{\mathring{\mathcal{E}}^\delta} L^0_\varepsilon - \inf_{\bar{U}} L^0_\varepsilon.$$

To conclude that this quantity is negative, and therefore that $\mu^{\delta}_{\varepsilon,\eta_n}(U)$ tends to 0 as $n \to +\infty$, it suffices to notice that $\inf_{\mathring{\mathcal{E}}^\delta} L^0_\varepsilon = \inf_{\mathcal{E}^\delta} L^0_\varepsilon$ (for instance by the easy fact that the infimum of L^δ_ε is continuous with respect to δ), and to use the definition of U. The result follows. □

2.4. ***From the Γ-convergence of the actions to the narrow convergence of the laws.*** In the previous subsection, we justified why the conditional laws $(\mu^\delta_{\varepsilon,\eta})$ from Corollary 4 are well-described by the action L^δ_ε defined by formula (17) as $\eta \to 0$: in this limit, these laws mainly charge small neighborhoods of minimizers of that action. Now, we want to argue that in order to study these laws when not only η is small, but also δ and ε, we have to study the action L^δ_ε in that regime, *in the sense of* Γ*-convergence*.

This assertion relies on the two following lemmas:

Lemma 5. *Let* (Ω, d) *be a metric space, and* $(\mathcal{L}_n)_{n\in\mathbb{N}}$ *be a sequence of lower semicontinuous functionals from* Ω *to* $\mathbb{R}_+ \cup \{+\infty\}$ *having compact sublevels, uniformly in* $n \in \mathbb{N}$. *Assume that* $(\mathcal{L}_n)$ *has a* Γ*-limit* $\mathcal{L}$. *Assume furthermore that* $\mathcal{L}$ *is not uniformly equal to* $+\infty$. *Finally, consider* $(\mu_n) \in \mathcal{P}(\Omega)^{\mathbb{N}}$ *a sequence of Borel probability measures on* Ω, *such that, for all* n, μ_n *only charges minimizers of* $\mathcal{L}_n$. *Then,* (μ_n) *is tight, and any of its limit points in the narrow topology only charges minimizers of* $\mathcal{L}$.

Lemma 6. *The family of actions* (L^δ_ε) *defined in* (17) *have compact sublevels in* $C^0([t_0, t_1]; (\mathbb{R}^d)^N)$, *uniformly in* $\varepsilon, \delta > 0$.

Using these lemmas, we see that if we manage to identify a Γ-limit L for L^δ_ε as $\varepsilon, \delta \to 0$, then in this limit, any family (μ^δ_ε) of limits of $(\mu^\delta_{\varepsilon,\eta})$ as $\eta \to 0$ will mainly charge small neighborhoods of minimizers of the limiting L. Before doing so in the next subsection, let us prove our two lemmas.

Proof of Lemma 5. Let x be a minimizer of $\mathcal{L}$, and (x_n) be an associated recovery sequence, that is, $x_n \to x$ as $n \to +\infty$, and $\limsup_{n\to+\infty} \mathcal{L}_n(x_n) \le \mathcal{L}(x) = \inf \mathcal{L}$. Up to forgetting the first terms, we can assume that $\mathcal{L}_n(x_n)$ is finite for all $n \in \mathbb{N}$. Now, call $M := \sup_{n\in\mathbb{N}} \mathcal{L}_n(x_n)$. By assumption, the set

$$K := \overline{\bigcup_{n\in\mathbb{N}} \{z \in \Omega : \mathcal{L}_n(z) \le M\}}$$

is compact, and by definition of M it contains all the minimizers of all the functionals $\mathcal{L}_n$, $n \in \mathbb{N}$. Therefore, for all $n \in \mathbb{N}$, $\mu_n(K) = 1$, and the tightness follows.

Let μ be a limit point of (μ_n) for the topology of narrow convergence. Up to considering a subsequence, we assume that $\mu_n \to \mu$. Let x be in the support of μ. It is easy to see that there exists a sequence (x_n) such that $x_n \to x$ as $n \to +\infty$, and, for all $n \in \mathbb{N}$, x_n is in the support of μ_n. But then by assumption, for all n, x_n is a minimizer of $\mathcal{L}_n$, and therefore, by standard considerations about Γ-convergence, x is a minimizer of $\mathcal{L}$. □

Proof of Lemma 6. For all $\varepsilon, \delta > 0$, the action L^δ_ε coincides with L^0_ε (defined in (16)) inside of the closed set $\mathcal{E}^\delta$ and is $+\infty$ outside of this closed set. Therefore, we just need to prove that L^0_ε has compact sublevels, uniformly in $\varepsilon > 0$. Actually, precompacity suffices by lower semicontinuity of L^0_ε. To do so, we will use the following bound, which holds as a consequence of (12) for all $\varepsilon > 0$, $t \in [t_0, t_1]$ and $X \in (\mathbb{R}^d)^N$:

$$|v_\varepsilon(t, X)| \le \frac{|A| + |X|}{2t_0}. \tag{18}$$

We will prove that, for all $M > 0$, there exists $M' > 0$ (uniform in ε) such that, for all $\varepsilon > 0$ and $\mathcal{X} \in C^0([t_0, t_1]; (\mathbb{R}^d)^N)$, whenever $L^0_\varepsilon(\mathcal{X}) \le M$, we have

$$\frac{1}{2}\int_{t_0}^{t_1} |\dot{\mathcal{X}}_t|^2 \,\mathrm{d}t \le M'.$$

This is enough to conclude since it is well-known that the set $H^1([t_0, t_1]; (\mathbb{R}^d)^N)$ is compactly embedded in $C^0([t_0, t_1]; (\mathbb{R}^d)^N)$.

So let us consider $M, \varepsilon > 0$, and a curve $\mathcal{X}$ such that $L^0_\varepsilon(\mathcal{X}) \le M$. Note that in particular, $\mathcal{X}_{t_0} \in \{P^\sigma\}$. We have, for all $t \in [t_0, t_1]$,

$$\begin{aligned}
\frac{1}{2}\int_{t_0}^{t} |\dot{\mathcal{X}}_s|^2 \,\mathrm{d}s &= \int_{t_0}^{t} \frac{|\dot{\mathcal{X}}_s - v_\varepsilon(s, \mathcal{X}_s) + v_\varepsilon(s, \mathcal{X}_s)|^2}{2} \,\mathrm{d}s \\
&\le \int_{t_0}^{t_1} |\dot{\mathcal{X}}_s - v_\varepsilon(s, \mathcal{X}_s)|^2 \,\mathrm{d}s + \int_{t_0}^{t} |v_\varepsilon(s, \mathcal{X}_s)|^2 \,\mathrm{d}s \\
&\le \frac{1}{t_0}\int_{t_0}^{t_1} |\dot{\mathcal{X}}_s - v_\varepsilon(s, \mathcal{X}_s)|^2 \times s \,\mathrm{d}s + \frac{1}{4t_0^2}\int_{t_0}^{t} (|A| + |\mathcal{X}_s|)^2 \,\mathrm{d}s \\
&\le \frac{2M}{t_0} + \frac{(t_1 - t_0)|A|^2}{2t_0^2} + \frac{1}{2t_0^2}\int_{t_0}^{t} \left| \mathcal{X}_{t_0} + \int_{t_0}^{s} \dot{\mathcal{X}}_\tau \,\mathrm{d}\tau \right|^2 \mathrm{d}s \\
&\le \frac{2M}{t_0} + \frac{t_1 - t_0}{t_0^2}\left\{ \frac{|A|^2}{2} + |P|^2 + \int_{t_0}^{t}\int_{t_0}^{s} |\dot{\mathcal{X}}_\tau|^2 \,\mathrm{d}\tau \,\mathrm{d}s \right\},
\end{aligned}$$

where we used (18) to get the third line. We deduce our claim from the Grönwall lemma. □

2.5. *The convergence results.* As already explained, understanding the behavior of families (μ^δ_ε) of limit points of $(\mu^\delta_{\varepsilon,\eta})$ as $\eta \to 0$ when ε and δ are small amounts to understanding the behavior of the family of actions (L^δ_ε) in the Γ-convergence sense. This is what we propose to do now. More specifically, we will see that (L^δ_ε) has a Γ-limit, when first $\delta \to 0$, and then $\varepsilon \to 0$. Doing so, we ensure that limit points of the family (μ^δ_ε) in the relevant asymptotic only charge minimizers of the corresponding actions; see Corollary 11 below. We discuss the question of swapping these limits in Remark 12.

Thanks to the smoothness of v^ε, the first Γ-limit, as $\delta \to 0$, is very simple and we omit the proof.

Proposition 7. *Let $\varepsilon > 0$. As δ tends to zero, the family of actions (L^δ_ε) Γ-converges to*

$$L_\varepsilon(\mathcal{X}) = \begin{cases} \frac{1}{2}\int_{t_0}^{t_1} |\dot{\mathcal{X}}_t - v_\varepsilon(t, \mathcal{X}_t)|^2 \times t \,\mathrm{d}t & \text{if } \mathcal{X} \in H^1([t_0, t_1]; (\mathbb{R}^d)^N),\ \mathcal{X}_{t_0} \in \{P^\sigma\} \text{ and } \mathcal{X}_{t_1} \in \{Q^\sigma\}, \\ +\infty, & \text{else.} \end{cases}$$

The second Γ-convergence, as $\varepsilon \to 0$, is more intricate and can be seen as the main result of this paper, because it involves the singular limit of the vector fields (v^ε) as $\varepsilon \to 0$. Before stating it, we need to introduce a few objects.

Define the following smooth functions, which are convex in X:

$$\text{for all } \varepsilon > 0,\ t > 0,\ X \in (\mathbb{R}^d)^N, \quad f_\varepsilon(t, X) := \varepsilon t \log\left[\frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \exp\left(\frac{X \cdot A^\sigma}{t\varepsilon} \right) \right]. \tag{19}$$

It has the property that, for all $\varepsilon > 0$, $t > 0$, and $X \in (\mathbb{R}^d)^N$,

$$v_\varepsilon(t, X) = \frac{X - \nabla f_\varepsilon(t, X)}{2t}.$$

As a consequence, we can rewrite L_ε for all $\varepsilon > 0$ as

$$L_\varepsilon(\mathcal{X}) = \begin{cases} \frac{1}{2}\int_{t_0}^{t_1} |\dot{\mathcal{X}}_t - (\mathcal{X}_t - \nabla f_\varepsilon(t, \mathcal{X}_t))/(2t)|^2 \times t \,\mathrm{d}t & \text{if } \mathcal{X} \in H^1([t_0, t_1]; (\mathbb{R}^d)^N),\ \mathcal{X}_{t_0} \in \{P^\sigma\} \text{ and } \mathcal{X}_{t_1} \in \{Q^\sigma\}, \\ +\infty & \text{else.} \end{cases}$$

When ε tends to zero, by virtue of the so-called Laplace's principle, we have the pointwise convergence

$$\lim_{\varepsilon \to 0} f_\varepsilon(t, X) = \max_{\sigma \in \mathfrak{S}_N} X \cdot A^\sigma =: f(X). \tag{20}$$

Notice that f is linked to the function F defined in (3) by the formula,

$$\text{for all } X \in (\mathbb{R}^d)^N, \quad f(X) = \frac{|A|^2 + |X|^2}{2} + N F(X). \tag{21}$$

The function f no longer depends on the time variable, and it is a convex function with finite values. As a consequence, for each $X \in (\mathbb{R}^d)^N$, the subdifferential $\partial f(X)$ of f at X is nonempty. We will consider the extended gradient $\overline{\nabla} f(X)$ of f at X defined as:

Definition 8 (extended gradient). We define the *extended gradient* of a real-valued convex function h at X, denoted by $\overline{\nabla} h(X)$, to be the element of $\partial h(X)$ with minimal Euclidean norm.

We are now ready to state our result concerning the limit $\varepsilon \to 0$.

Theorem 9. *As ε tends to 0, the family of actions $(L_\varepsilon)_{\varepsilon>0}$ Γ-converges to*

$$L(\mathcal{X}) = \begin{cases} \frac{1}{2}\int_{t_0}^{t_1} |\dot{\mathcal{X}}_t - (\mathcal{X}_t - \overline{\nabla} f(\mathcal{X}_t))/(2t)|^2 \times t \,\mathrm{d}t & \text{if } \mathcal{X} \in H^1([t_0, t_1]; (\mathbb{R}^d)^N),\ \mathcal{X}_{t_0} \in \{P^\sigma\} \text{ and } \mathcal{X}_{t_1} \in \{Q^\sigma\}, \\ +\infty & \text{else} \end{cases} \tag{22}$$

for the topology of uniform convergence of $C^0([t_0, t_1]; (\mathbb{R}^d)^N)$.

Remark 10. It is relevant to wonder what exactly in the convergence $f_\varepsilon \to f$ implies Theorem 9. It is not so simple to answer due to the dependence in t of f_ε and because the proof involves several manipulations of formula (22). However, the main step of the proof is Lemma 15 below. Now, at least in the autonomous case, several works that are posterior to the first version of the present paper study results similar to Lemma 15 in greater generality, namely in Hilbert spaces [Ambrosio et al. 2021] or in measured metric spaces [Monsaingeon et al. 2023]. In [Ambrosio et al. 2021], the good notion of convergence for $f_\varepsilon \to f$ is *Mosco convergence*. We give more details on this in Remark 16.

As a consequence of Lemmas 5 and 6, this theorem clearly implies the following.

Corollary 11. *Consider the family of laws $(\mu^\delta_{\varepsilon,\eta})$ defined in Corollary 4, and three sequences $(\eta_n)_{n\in\mathbb{N}}$, $(\delta_m)_{m\in\mathbb{N}}$ and $(\varepsilon_p)_{p\in\mathbb{N}}$ decreasing to 0. Then, there exist subsequences $(\eta'_n)_{n\in\mathbb{N}}$, $(\delta'_m)_{m\in\mathbb{N}}$ and $(\varepsilon'_p)_{p\in\mathbb{N}}$ such that the triple limit*

$$\lim_{p\to+\infty} \lim_{m\to+\infty} \lim_{n\to+\infty} \mu^{\delta_m}_{\varepsilon_p,\eta_n}$$

exists in the topology of narrow convergence and only charge minimizers of L as defined by (22).

In particular, if L admits a unique minimizer $\mathcal{X}$, the whole family converges:

$$\lim_{\varepsilon\to 0}\lim_{\delta\to 0}\lim_{\eta\to 0}\mu^{\delta}_{\varepsilon,\eta}=\delta_{\mathcal{X}}.$$

Let us now comment on the order in which these limits are taken.

Remark 12. Up to potentially considering subsequences, we are studying the behavior of the conditioned laws $(\mu^{\delta}_{\varepsilon,\eta})$ in the limit $\lim_{\varepsilon\to 0}\lim_{\delta\to 0}\lim_{\eta\to 0}$, and one could wonder whether these limits could be swapped. We recall that ε stands for the level of noise of the original point cloud, that η stands for the level of perturbation of the companion ODE, and that δ is the precision of the observation at the final time.

- Swapping $\lim_{\varepsilon\to 0}$ and $\lim_{\delta\to 0}$ is easy: it amounts to studying the dependence of the limiting action (22) when Q varies. Essentially, this swap would be a consequence of the fact that v^{ε} is bounded on compact sets, uniformly in time and ε.

- Swapping $\lim_{\delta\to 0}$ and $\lim_{\eta\to 0}$ would be more delicate, but doable as well. We would first need to prove that the family $(\mu^{\delta}_{\varepsilon,\eta})$ from Corollary 4 converges when $\delta\to 0$, with fixed ε and η, as classically done in the theory of bridges of processes, and then write a large deviation principle for these bridges in place of Theorem 3.

- Finally, not taking into consideration the limit in δ because of the two previous points, the question of how to swap $\lim_{\varepsilon\to 0}$ with $\lim_{\eta\to 0}$ relates to the question of building solutions to SDEs with singular coefficients, and lies beyond the scope of this article. A related question that we also do not want to address is the question of quantifying how small η needs to be with respect to ε to be able to take a simultaneous limit in ε and η. To answer it, we would need to study the dependence in ε of the rates of convergence in the large deviation principle, which is probably a very delicate question, once again because of the singularities of v^{ε} appearing as $\varepsilon\to 0$.

We will prove Theorem 9 in Section 3 below, but before doing so, let us show that up to changing time, we recover MAG. Notice L has compact sublevels as a consequence of the Γ-convergence and Lemma 6. Hence, the existence of global minimizers for L (and hence for all the forthcoming functionals) follows from the direct method of calculus of variations.

2.6. ***A change of time leading to Monge–Ampère gravitation.*** Through the change of variable

$$t=\exp(2\theta),\qquad \mathcal{Z}_\theta=\mathcal{X}_{\exp(2\theta)},\qquad \theta_0=\tfrac{1}{2}\log t_0,\qquad \theta_1=\tfrac{1}{2}\log t_1,$$

we observe that, for all $\mathcal{X}\in C^0([t_0,t_1];(\mathbb{R}^d)^N)$, $L(\mathcal{X})=\frac{1}{2}\Lambda(\mathcal{Z})$, with

$$\Lambda(\mathcal{Z})=\begin{cases}\frac{1}{2}\int_{\theta_0}^{\theta_1}|\dot{\mathcal{Z}}_\theta-(\mathcal{Z}_\theta-\overline{\nabla}f(\mathcal{Z}_\theta))|^2\,\mathrm{d}\theta & \text{if } \mathcal{Z}\in H^1([\theta_0,\theta_1];(\mathbb{R}^d)^N),\ \mathcal{Z}_{\theta_0}\in\{P^\sigma\} \text{ and } \mathcal{Z}_{\theta_1}\in\{Q^\sigma\},\\ +\infty & \text{else.}\end{cases}$$

(Recall the definition (20) of f.)

It turns out to be equivalent to the following one (in which we recognize (8)):

$$\Lambda'(\mathcal{Z})=\begin{cases}\int_{\theta_0}^{\theta_1}\left\{\frac{1}{2}|\dot{\mathcal{Z}}_\theta|^2+\frac{1}{2}|\mathcal{Z}_\theta-\overline{\nabla}f(\mathcal{Z}_\theta)|^2\right\}\mathrm{d}\theta & \text{if } \mathcal{Z}\in H^1([\theta_0,\theta_1];(\mathbb{R}^d)^N),\ \mathcal{Z}_{\theta_0}\in\{P^\sigma\} \text{ and } \mathcal{Z}_{\theta_1}\in\{Q^\sigma\},\\ +\infty & \text{else.}\end{cases}$$

To see this, it suffices to expand the square and to remark that the mixed product is an exact time derivative, so that its integral only involves the endpoints P and Q. This is done in a slightly different context in the proof of Lemma 14 below.

2.7. *Application of the least action principle.* We observe that the points Z where f is differentiable are those for which the maximum in the definition (20) of f is reached by a unique permutation σ_{opt} so that $\nabla f(Z)$ is nothing but $A^{\sigma_{\mathrm{opt}}}$. For such points Z, we get

$$\tfrac{1}{2}|Z-\nabla f(Z)|^2 = \tfrac{1}{2}|Z-A^{\sigma_{\mathrm{opt}}}|^2 = -NF(Z)$$

(by definition (3) of F), while, on the set $\mathcal{N}$ of nondifferentiability of f, we rather have

$$\tfrac{1}{2}|Z-\overline{\nabla} f(Z)|^2 < -NF(Z);$$

see for instance Proposition 27 below in the case of dimension 1. So the action we have obtained in the previous section, namely Λ', bounds from below

$$\Lambda^+(\mathcal{Z}) = \begin{cases} \int_{\theta_0}^{\theta_1} \big\{\tfrac{1}{2}|\dot{\mathcal{Z}}_\theta|^2 - NF(\mathcal{Z}_\theta)\big\}\,\mathrm{d}\theta & \text{if } \mathcal{Z} \in H^1([\theta_0,\theta_1];(\mathbb{R}^d)^N),\ \mathcal{Z}_{\theta_0} \in \{P^\sigma\} \text{ and } \mathcal{Z}_{\theta_1} \in \{Q^\sigma\}, \\ +\infty & \text{else.} \end{cases}$$

This second action, already announced in (7), is definitely strictly larger than the first one for those curves $\theta \to \mathcal{Z}_\theta$ which take values in $\mathcal{N}$ (where f and F are not differentiable) on a set of times $\theta \in [\theta_0,\theta_1]$ with positive Lebesgue measure. So the least action principle may provide different optimal curves, depending on the action we choose. However, if a curve is optimal for Λ' and almost surely takes value outside of $\mathcal{N}$, then it must also be optimal for Λ^+. Clearly, it is much easier to get the optimality equation for such a curve, by working with Λ^+ rather than with Λ'. By varying action Λ^+, we get (6) as optimality equation. Therefore, the optimal curves of our functional Λ' taking value in $\mathcal{N}$ for a negligible set of times solve (4) (in a distributional sense), which is the MAG discrete model announced in the Introduction.

Of course, these equations have to be suitably modified for those curves which are optimal for the action Λ' but not for Λ^+ because they take values in $\mathcal{N}$ for a nonnegligible amount of time. At this stage, we do not know how to do it. However, at least in the one-dimensional case $d=1$, such modifications are tractable and correspond to sticky collisions as $x_i(t)=x_j(t)$ occurs for different "particles" of labels $i \neq j$ and during intervals of times of strictly positive Lebesgue measure; see Section 4.

3. Proof of the Γ-convergence

The purpose of this section is to prove Theorem 9.

3.1. *The proof as a consequence of three lemmas.* As we will see, Theorem 9 will be a consequence of three lemmas that we state below. Lemmas 14 and 15 both involve a family of smooth functions $(g_\varepsilon)_{\varepsilon>0}$ on $[\theta_0,\theta_1]\times\mathbb{R}^p$ for some $\theta_0<\theta_1$ and $p\in\mathbb{N}$, pointwise converging to a L^1_{loc} function g. On these functions, we will assume the following:

Assumptions 13. (H1) For all $\varepsilon>0$ and $\theta\in[\theta_0,\theta_1]$, $g_\varepsilon(\theta,0)=0$.

(H2) For all $\varepsilon>0$ and $\theta\in[\theta_0,\theta_1]$, $g_\varepsilon(\theta,\cdot)$ is convex. Therefore, $g(\theta,\cdot)$ is convex as well.

(H3) The maps ∇g_ε are uniformly bounded, that is,

$$L := \sup_{\varepsilon>0} \sup_{\theta\in[\theta_0,\theta_1]} \sup_{Y\in\mathbb{R}^p} |\nabla g_\varepsilon(\theta, Y)| < +\infty. \tag{23}$$

Therefore, we also have

$$\sup_{\theta\in[\theta_0,\theta_1]} \sup_{Y\in\mathbb{R}^p} |\overline{\nabla} g(\theta, Y)| \le L.$$

(H4) The distributional derivative $\partial_\theta g$ is $L^2([\theta_0,\theta_1]; L^\infty_{\mathrm{loc}}(\mathbb{R}^d)^N)$, and, for all $\mathcal{Y} \in H^1([\theta_0,\theta_1]; (\mathbb{R}^d)^N)$, the map $\theta \mapsto g(\theta, \mathcal{Y}_\theta)$ is also H^1, with, for almost all $\theta \in [\theta_0,\theta_1]$,

$$\frac{\mathrm{d}}{\mathrm{d}\theta} g(\theta, \mathcal{Y}_\theta) = \partial_\theta g(\theta, \mathcal{Y}_\theta) + \overline{\nabla} g(\theta, \mathcal{Y}_\theta) \cdot \dot{\mathcal{Y}}_\theta. \tag{24}$$

(H5) The maps $\partial_\theta \nabla g_\varepsilon$ are uniformly bounded, that is,

$$M := \sup_{\varepsilon>0} \sup_{\theta\in[\theta_0,\theta_1]} \sup_{Y\in\mathbb{R}^p} |\partial_\theta \nabla g_\varepsilon(\theta, Y)| < +\infty. \tag{25}$$

In order to keep the proofs simple, we did not try to optimize these assumptions for Lemmas 14 and 15, which are probably true in a far more general context (see Remark 16 in the case of Lemma 15). However, as we will see in the proof of Theorem 9, it suffices to check these assumptions for the family $(f_\varepsilon)_{\varepsilon>0}$ after suitable change of temporal and spatial scale. This is done in Lemma 17.

Lemma 14. *Let us consider $\theta_0 < \theta_1 \in \mathbb{R}$, $\eta \in C^\infty([\theta_0,\theta_1]; \mathbb{R}^*_+)$ and a family $(g_\varepsilon)_{\varepsilon>0}$ of smooth functions from $[\theta_0,\theta_1] \times \mathbb{R}^p$ to $\mathbb{R}$ pointwise converging to a function g, which satisfy* (H1), (H3), (H4) *and* (H5) *from Assumptions 13. If a family of curves $(\mathcal{Y}^\varepsilon)_{\varepsilon>0}$ in $H^1([\theta_0,\theta_1]; \mathbb{R}^p)$ uniformly converges to a curve $\mathcal{Y} \in H^1([\theta_0,\theta_1]; \mathbb{R}^p)$, then*

$$\int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}^\varepsilon_\theta \cdot \nabla g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta)\eta(\theta)\,\mathrm{d}\theta \xrightarrow[\varepsilon\to 0]{} \int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}_\theta \cdot \overline{\nabla} g(\theta, \mathcal{Y}_\theta)\eta(\theta)\,\mathrm{d}\theta.$$

Lemma 15. *Let us consider $\theta_0 < \theta_1 \in \mathbb{R}$, $\eta \in C^\infty([\theta_0,\theta_1]; \mathbb{R}^*_+)$ and a family $(g_\varepsilon)_{\varepsilon>0}$ of smooth functions from $[\theta_0,\theta_1] \times \mathbb{R}^p$ to $\mathbb{R}$ pointwise converging to a function g, and satisfying* (H2), (H3)*and* (H5) *from Assumptions 13. Let us fix $R, S \in \mathbb{R}^p$ and define for $\varepsilon > 0$ and $\mathcal{Y} \in C^0([\theta_0,\theta_1]; \mathbb{R}^p)$*

$$K_\varepsilon(\mathcal{Y}) := \begin{cases} \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta & \text{if } \mathcal{Y} \in H^1([\theta_1,\theta_1]; \mathbb{R}^p),\ \mathcal{Y}_{\theta_0} = R \text{ and } \mathcal{Y}_{\theta_1} = S, \\ +\infty & \text{else,} \end{cases}$$

$$K(\mathcal{Y}) := \begin{cases} \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta|^2 + |\overline{\nabla} g(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta & \text{if } \mathcal{Y} \in H^1([\theta_1,\theta_1]; \mathbb{R}^p),\ \mathcal{Y}_{\theta_0} = R \text{ and } \mathcal{Y}_{\theta_1} = S, \\ +\infty, & \text{else.} \end{cases}$$

Then $(K_\varepsilon)_{\varepsilon>0}$ Γ-converges to K for the topology of uniform convergence of $C^0([\theta_0,\theta_1]; \mathbb{R}^p)$.

Remark 16. This lemma is the keystone of the proof, and one may wonder how it can be generalized and what is really necessary among our assumptions. In [Ambrosio et al. 2021], we show that at least when (g_ε) and g have no dependence on θ and $\eta \equiv 1$, the result holds true, even in Hilbert spaces, whenever (g_ε) is a family of proper lower semicontinuous uniformly λ-convex functions Mosco converging towards g, plus some uniform Lipschitz conditions at the extreme points.

Lemma 17. *With the notation of Theorem 9, let us define* $\theta_0 := \log t_0/2,\ \theta_1 := \log t_1/2,\ p = dN$, *and for* $\theta \in [\theta_0, \theta_1]$, $\varepsilon > 0$ *and* $Y \in (\mathbb{R}^d)^N$,

$$g_\varepsilon(\theta, Y) := \frac{f_\varepsilon(\exp(2\theta), \exp(\theta)Y)}{\exp(2\theta)} \quad \textit{and} \quad g(\theta, Y) := \frac{f(\exp(\theta)Y)}{\exp(2\theta)}. \tag{26}$$

Then $(g_\varepsilon)_{\varepsilon>0}$ *converges pointwise to* g, *and satisfies* (H1), (H2), (H3), (H4) *and* (H5) *from Assumptions 13.*

In the next subsections, we will prove these three lemmas one by one. The most involved one is undoubtedly Lemma 15, which can be seen as the main step in the proof of Theorem 9. Let us start by proving Theorem 9 using Lemmas 14, 15 and 17.

Proof of Theorem 9. In this proof, the notation $\mathcal{X} = \mathcal{X}_t$ will stand for a generic curve from $[t_0, t_1]$ to $(\mathbb{R}^d)^N$. Associated with $\mathcal{X}$, we define by $\mathcal{Y} = \mathcal{Y}_\theta$ the curve from $[\theta_0, \theta_1]$ to $(\mathbb{R}^d)^N$, where $\theta_0 := \log t_0/2$, $\theta_1 := \log t_1/2$, and, for all $\theta \in [\theta_0, \theta_1]$, $\mathcal{Y}_\theta := \mathcal{X}_{\exp(2\theta)}/\exp(\theta)$. Note that $\mathcal{X}$ is H^1 if and only if $\mathcal{Y}$ is H^1. If $(\mathcal{X}^\varepsilon)_{\varepsilon>0}$ is a family of curves from $[t_0, t_1]$ to $(\mathbb{R}^d)^N$, we define in the same way the family of corresponding curves $(\mathcal{Y}^\varepsilon)_{\varepsilon>0}$ from $[\theta_0, \theta_1]$ to $(\mathbb{R}^d)^N$.

A quick computation shows that, for all $\mathcal{X} \in H^1([\theta_0, \theta_1]; (\mathbb{R}^d)^N)$, considering $\eta(\theta) := \exp(2\theta)$ and $(g_\varepsilon)_{\varepsilon>0}$, g as defined in Lemma 17, we have

$$L_\varepsilon(\mathcal{X}) = \frac{1}{2}\int_{t_0}^{t_1} \left| \dot{\mathcal{X}}_t - \frac{\mathcal{X}_t - \nabla f_\varepsilon(t, \mathcal{X}_t)}{2t} \right|^2 \frac{\mathrm{d}t}{t} = \frac{1}{4}\int_{\theta_0}^{\theta_1} |\dot{\mathcal{Y}}_\theta + \nabla g_\varepsilon(\theta, \mathcal{Y}_\theta)|^2 \eta(\theta)\,\mathrm{d}\theta \tag{27}$$

$$= \frac{1}{4}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta + \frac{1}{2}\int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}_\theta \cdot \nabla g_\varepsilon(\theta, \mathcal{Y}_\theta)\eta(\theta)\,\mathrm{d}\theta \tag{28}$$

and

$$L(\mathcal{X}) = \frac{1}{2}\int_{t_0}^{t_1} \left| \dot{\mathcal{X}}_t - \frac{\mathcal{X}_t - \overline{\nabla} f(\mathcal{X}_t)}{2t} \right|^2 \frac{\mathrm{d}t}{t} = \frac{1}{4}\int_{\theta_0}^{\theta_1} |\dot{\mathcal{Y}}_\theta + \overline{\nabla} g(\theta, \mathcal{Y}_\theta)|^2 \eta(\theta)\,\mathrm{d}\theta$$

$$= \frac{1}{4}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta|^2 + |\overline{\nabla} g(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta + \frac{1}{2}\int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}_\theta \cdot \overline{\nabla} g(\theta, \mathcal{Y}_\theta)\eta(\theta)\,\mathrm{d}\theta. \tag{29}$$

(Note that due to Lemma 17, g is convex with respect to the space variable, and so $\overline{\nabla} g$ is well-defined.)

<u>Proof of the Γ-lim inf</u>: Let $\mathcal{X}^\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathcal{X}$ for the topology of uniform convergence. Of course, we also have $\mathcal{Y}^\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathcal{Y}$. Without loss of generality, we can suppose

$$\liminf_{\varepsilon\to 0} L_\varepsilon(\mathcal{X}^\varepsilon) < +\infty.$$

Indeed, if it is not the case, there is nothing to prove. Let us take $(\varepsilon_n)_{n\in\mathbb{N}}$ to be a sequence tending to 0 along which the lim inf is achieved.

As $\nabla g_\varepsilon(\theta, Y)$ is bounded uniformly in ε, θ, Y (this is (H3)), we easily deduce with (27)

$$\limsup_{n\to+\infty} \int_{\theta_0}^{\theta_1} |\dot{\mathcal{Y}}^{\varepsilon_n}_\theta|^2\,\mathrm{d}\theta < +\infty.$$

In particular, by the lower semicontinuity of this H^1 seminorm with respect to uniform convergence, $\mathcal{Y}$ is in $H^1([\theta_0, \theta_1]; (\mathbb{R}^d)^N)$. Applying Lemma 14, thanks to Lemma 17, we have

$$\int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}^{\varepsilon_n}_\theta \cdot \nabla g_{\varepsilon_n}(\theta, \mathcal{Y}^{\varepsilon_n}_\theta)\eta(\theta)\,\mathrm{d}\theta \xrightarrow[n\to+\infty]{} \int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}_\theta \cdot \overline{\nabla} g(\theta, \mathcal{Y}_\theta)\eta(\theta)\,\mathrm{d}\theta. \tag{30}$$

On the other hand, for n sufficiently large, $L_{\varepsilon_n}(\mathcal{X}^{\varepsilon_n}) < +\infty$. So the endpoints of $\mathcal{X}^{\varepsilon_n}$ belong to a finite set, and because of the convergence $\mathcal{X}^{\varepsilon_n} \to \mathcal{X}$, for even larger n the endpoints of $\mathcal{X}^{\varepsilon_n}$ are independent of n. In other terms, $\mathcal{X}^{\varepsilon_n}_{t_0} = P^{\sigma_0}$ and $\mathcal{X}^{\varepsilon_n}_{t_1} = Q^{\sigma_1}$ with σ_0, σ_1 independent of n. Hence, for such n, $\mathcal{Y}^{\varepsilon_n}$ satisfies the endpoint constraint for K_{ε_n} with $R := P^{\sigma_0}/\sqrt{t_0}$ and $S := Q^{\sigma_1}/\sqrt{t_1}$. Hence, applying Lemma 15 thanks to Lemma 17, we have

$$\begin{aligned}\frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta|^2 + |\overline{\nabla} g(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta = K(\mathcal{Y}) &\le \liminf_{n\to+\infty} K_{\varepsilon_n}(\mathcal{Y}^{\varepsilon_n}) \\ &= \liminf_{n\to+\infty} \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}^{\varepsilon_n}_\theta|^2 + |\nabla g_{\varepsilon_n}(\theta, \mathcal{Y}^{\varepsilon_n}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta. \end{aligned} \tag{31}$$

The result follows easily by gathering (28), (30), (31) and (29).

<u>Proof of the Γ-lim sup</u>: Let $\mathcal{X} \in C^0([t_0, t_1]; (\mathbb{R}^d)^N)$. Without loss of generality, we can suppose that $\mathcal{X} \in H^1([t_0, t_1]; (\mathbb{R}^d)^N)$ and that it satisfies the endpoint constraint for L. In particular, $\mathcal{Y}$ belongs to $H^1([\theta_0, \theta_1]; (\mathbb{R}^d)^N)$ and satisfies the endpoint constraint for K with $R := \mathcal{X}_{t_0}/\sqrt{t_0}$ and $S := \mathcal{X}_{t_0}/\sqrt{t_1}$. Lemmas 15 and 17 let us find a family $(\mathcal{Y}^\varepsilon)_{\varepsilon>0}$ converging to the corresponding $\mathcal{Y}$ such that

$$\limsup_{\varepsilon\to 0} K_\varepsilon(\mathcal{Y}^\varepsilon) \le K(\mathcal{Y}). \tag{32}$$

In particular $\mathcal{Y}^\varepsilon$ is in H^1 for sufficiently small ε, and by Lemmas 14 and 17,

$$\int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}^\varepsilon_\theta \cdot \nabla g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta)\eta(\theta)\,\mathrm{d}\theta \xrightarrow[\varepsilon\to 0]{} \int_{\theta_0}^{\theta_1} \dot{\mathcal{Y}}_\theta \cdot \overline{\nabla} g(\theta, \mathcal{Y}_\theta)\eta(\theta)\,\mathrm{d}\theta. \tag{33}$$

The result follows easily from (28), (32), (33) and (29), by noticing, that because of (32), $\mathcal{Y}^\varepsilon$ satisfies the endpoint constraint for K_ε. Hence, for such ε, $\mathcal{X}^\varepsilon$ satisfies the endpoint constraint for L_ε. □

3.2. ***Proof of Lemma 14.*** The proof of Lemma 14 just consists in integrating by parts and using the convergence properties of $(g_\varepsilon)_{\varepsilon>0}$.

Proof of Lemma 14. <u>Integration by parts</u>: First, notice that as soon as $\mathcal{Y} \in H^1([\theta_0, \theta_1]; \mathbb{R}^p)$ and $\varepsilon > 0$, then $\theta \mapsto g_\varepsilon(\theta, \mathcal{Y}_\theta)$ and $\theta \mapsto g(\theta, \mathcal{Y}_\theta)$ are also in H^1, with, for almost every θ,

$$\frac{\mathrm{d}}{\mathrm{d}\theta} g_\varepsilon(\theta, \mathcal{Y}_\theta) = \partial_\theta g_\varepsilon(\theta, \mathcal{Y}_\theta) + \nabla g_\varepsilon(\theta, \mathcal{Y}_\theta) \cdot \dot{\mathcal{Y}}_\theta \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\theta} g(\theta, \mathcal{Y}_\theta) = \partial_\theta g(\theta, \mathcal{Y}_\theta) + \overline{\nabla} g(\theta, \mathcal{Y}_\theta) \cdot \dot{\mathcal{Y}}_\theta.$$

It is clear in the case of g_ε because g_ε is smooth, and it is the assumption (H4) in the case of g. As a consequence, by an integration by parts, it suffices to prove that whenever $(\mathcal{Y}^\varepsilon)_{\varepsilon>0}$ converges to $\mathcal{Y}$ as $\varepsilon \to 0$ for the topology of uniform convergence,

$$\begin{aligned} &g_\varepsilon(\theta_1, \mathcal{Y}^\varepsilon_{\theta_1})\eta(\theta_1) - g_\varepsilon(\theta_0, \mathcal{Y}^\varepsilon_{\theta_0})\eta(\theta_0) - \int_{\theta_0}^{\theta_1} g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta)\eta'(\theta)\,\mathrm{d}\theta - \int_{\theta_0}^{\theta_1} \partial_\theta g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta)\eta(\theta)\,\mathrm{d}\theta \\ &\quad\xrightarrow[\varepsilon\to 0]{} g(\theta_1, \mathcal{Y}_{\theta_1})\eta(\theta_1) - g(\theta_0, \mathcal{Y}_{\theta_0})\eta(\theta_0) - \int_{\theta_0}^{\theta_1} g(\theta, \mathcal{Y}_\theta)\eta'(\theta)\,\mathrm{d}\theta - \int_{\theta_0}^{\theta_1} \partial_\theta g(\theta, \mathcal{Y}_\theta)\eta(\theta)\,\mathrm{d}\theta. \end{aligned}$$

Convergence term by term: The convergence

$$g_\varepsilon(\theta_1, \mathcal{Y}^\varepsilon_{\theta_1})\eta(\theta_1) - g_\varepsilon(\theta_0, \mathcal{Y}^\varepsilon_{\theta_0})\eta(\theta_0) \xrightarrow[\varepsilon\to 0]{} g(\theta_1, \mathcal{Y}_{\theta_1})\eta(\theta_1) - g(\theta_0, \mathcal{Y}_{\theta_0})\eta(\theta_0)$$

is an easy consequence of the pointwise convergence and of the uniform Lipschitz bound (H3).

For the same reason, we have, for all $\theta \in [\theta_0, \theta_1]$, $g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta) \xrightarrow[\varepsilon\to 0]{} g(\theta, \mathcal{Y}_\theta)$. But on the other hand, because of (H1) and (H3), g_ε is locally bounded, uniformly in ε. Hence,

$$\int_{\theta_0}^{\theta_1} g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta)\eta'(\theta)\,\mathrm{d}\theta \xrightarrow[\varepsilon\to 0]{} \int_{\theta_0}^{\theta_1} g(\theta, \mathcal{Y}_\theta)\eta'(\theta)\,\mathrm{d}\theta$$

is a consequence of the dominated convergence theorem.

Because of (H1) and (H5), for all θ, $(\partial_\theta g_\varepsilon(\theta, \cdot))_{\varepsilon>0}$ is compact for the topology of local uniform convergence. But its only possible limit point is the distributional derivative $\partial_\theta g$. As a consequence, $(\partial_\theta g_\varepsilon)_{\varepsilon>0}$ converges pointwise to $\partial_\theta g$, and because of the uniform bound (H5), for all θ, $\partial_\theta g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta) \xrightarrow[\varepsilon\to 0]{}$ $\partial_\theta g(\theta, \mathcal{Y}_\theta)$. Because of (H1) and (H5), $\partial_\theta g_\varepsilon$ is locally bounded, uniformly in ε, and so

$$\int_{\theta_0}^{\theta_1} \partial_\theta g_\varepsilon(\theta, \mathcal{Y}^\varepsilon_\theta)\eta(\theta)\,\mathrm{d}\theta \xrightarrow[\varepsilon\to 0]{} \int_{\theta_0}^{\theta_1} \partial_\theta g(\theta, \mathcal{Y}_\theta)\eta(\theta)\,\mathrm{d}\theta$$

is also a consequence of the dominated convergence theorem. □

3.3. ***Proof of Lemma 15.*** Before entering the proof of Lemma 15, we need to state a few standard results concerning the extended gradient $\overline{\nabla}$ as defined in Definition 8, and its links with the so-called *resolvent map*. These tools could even be set in the infinite-dimensional setting, that is, in Hilbert spaces [Strömberg 1996], or in metric spaces [Ambrosio et al. 2005].

The following proposition is a lower semicontinuity property of the slope with respect to both convergence of the function and of the evaluation point.

Proposition 18. *Consider $h : \mathbb{R}^p \to \mathbb{R}$ a convex function with finite values. Let $(h_\varepsilon)_{\varepsilon>0}$ be a family of convex functions on $\mathbb{R}^p$ pointwise converging to h, and let $(X^\varepsilon)_{\varepsilon>0}$ be a family of points in $\mathbb{R}^p$ converging to X. Then*

$$|\overline{\nabla} h(X)| \le \liminf_{\varepsilon\to 0} |\overline{\nabla} h_\varepsilon(X^\varepsilon)|.$$

Proof. As all these functions are convex and h has finite values, standard arguments show that the convergence of $h_\varepsilon \to h$ is also locally uniform. First of all, if

$$\liminf_{\varepsilon\to 0} |\overline{\nabla} h_\varepsilon(X^\varepsilon)| = +\infty,$$

there is nothing to prove. Else, up to considering a subsequence, there exists $D \in \mathbb{R}^p$ such that

$$\lim_{\varepsilon\to 0} \overline{\nabla} h_\varepsilon(X^\varepsilon) = D.$$

But sending $\varepsilon \to 0$ in the inequality,

$$\text{for all } Y \in \mathbb{R}^p, \quad h_\varepsilon(Y) \ge h_\varepsilon(X^\varepsilon) + \langle \overline{\nabla} h_\varepsilon(X^\varepsilon), Y - X^\varepsilon\rangle,$$

and using the local uniformity of the convergence, we see that $D \in \partial h(X)$ (that is, the subdifferential is upper semicontinuous). So $|D| \ge |\overline{\nabla} h(X)|$, and the result follows. □

For $\tau > 0$ and $X \in \mathbb{R}^p$, define the resolvent operator by

$$J_{\tau,h}(X) := \operatorname*{argmin}_{Y \in \mathbb{R}^p} h(Y) + \frac{|Y-X|^2}{2\tau}.$$

Once again, the following proposition is standard. It is an application in the very simple case of convex functions in finite dimension of the so-called maximal monotone operators theory in Hilbert spaces, for which we refer for instance to [Brézis 1973] (see in particular Section 2.4 for the properties of the resolvent in a general setting).

Proposition 19. (1) *We have for all $X \in \mathbb{R}^p$ and $\tau > 0$,*

$$|\overline{\nabla} h(J_{\tau,h}(X))| \leq \left| \frac{X - J_{\tau,h}(X)}{\tau} \right| \leq |\overline{\nabla} h(X)|. \tag{34}$$

(2) *If h is differentiable at $J_{\tau,h}(X)$ for some $X \in \mathbb{R}^p$, then the following first-order condition holds*:

$$\frac{X - J_{\tau,h}(X)}{\tau} = \nabla h(J_{\tau,h}(X)).$$

(3) *If $(h_\varepsilon)_{\varepsilon>0}$ is a family of convex functions on $\mathbb{R}^p$ pointwise converging to h, then, for all $\tau > 0$ and $X \in \mathbb{R}^p$,*

$$J_{\tau,h_\varepsilon}(X) \xrightarrow[\varepsilon \to 0]{} J_{\tau,h}(X). \tag{35}$$

Proof. By [Brézis 1973, Lemma 2.1], we have

$$\frac{X - J_{\tau,h}(X)}{\tau} \in \partial h(J_{\tau,h}(X)). \tag{36}$$

The first inequality in (34) and the second point of the statement follow.

To get the second inequality in (34), apply the monotone inequality of [Brézis 1973, Definition 2.1] to the maximal monotone operator ∂h (see [Brézis 1973, Example 2.1.4]), with $x_1 = X$, $x_2 = J_{\tau,h}(X)$, $y_1 = \overline{\nabla} h(X) \in \partial h(X)$ and $(X - J_{\tau,h}(X))/\tau \in \partial h(J_{\tau,h}(X))$, thanks to (36). We find

$$\left\langle \overline{\nabla} h(X) - \frac{X - J_{\tau,h}(X)}{\tau}, X - J_{\tau,h}(X) \right\rangle \geq 0,$$

which can be rewritten as

$$\left| \frac{X - J_{\tau,h}(X)}{\tau} \right|^2 \leq \left\langle \frac{X - J_{\tau,h}(X)}{\tau}, \overline{\nabla} h(X) \right\rangle.$$

Therefore, the result follows from the Cauchy–Schwarz inequality.

Let us now focus on the third point. Let us fix $\tau > 0$ and $X \in \mathbb{R}^p$, and set,

$$\text{for all } \varepsilon > 0,\ Y \in \mathbb{R}^p, \quad f_\varepsilon(Y) := h_\varepsilon(Y) + \frac{|Y-X|^2}{2\tau} \quad \text{and} \quad f(Y) := h(Y) + \frac{|Y-X|^2}{2\tau}.$$

The family $(f_\varepsilon)_{\varepsilon>0}$ converges pointwise to f, but by convexity and finiteness of the limit, as before, this convergence is also locally uniform. As a consequence, the only thing to prove is that for sufficiently small

$\varepsilon_0 > 0$, the set $\{J_{\tau,h_\varepsilon}(X) : 0 < \varepsilon \le \varepsilon_0\}$ is bounded. Indeed, if it is the case, by local uniform convergence, any limit point Z of $J_{\tau,h_\varepsilon}(X)$ as ε tends to 0 would satisfy

$$f(Z) \le \limsup_{\varepsilon\to 0} f_\varepsilon(J_{\tau,h_\varepsilon}(X)) \le \lim_{\varepsilon\to 0} f_\varepsilon(J_{\tau,h}(X)) = f(J_{\tau,h}(X)),$$

so that, by the definition of $J_{\tau,h}(X)$, $Z = J_{\tau,h}(X)$, which lets us conclude.

Call B the open ball of center $J_{\tau,h}(X)$ and radius 1. We have by the strict convexity of f and minimality of $J_{\tau,h}(X)$

$$f(J_{\tau,h}(X)) < \inf_{Y\in\partial B} f(Y),$$

and this property is open for the topology of local uniform convergence. Hence, we can find ε_0 sufficiently small so that for all $\varepsilon \le \varepsilon_0$

$$f_\varepsilon(J_{\tau,h}(X)) < \inf_{Y\in\partial B} f_\varepsilon(Y).$$

Then, if $Y \notin B$, we call $\overline{Y}$ the projection of Y on ∂B and $\lambda := 1/|Y - J_{\tau,h}(X)| \le 1$, so that $\overline{Y} = (1-\lambda)J_{\tau,h}(X) + \lambda Y$. As soon as $\varepsilon \le \varepsilon_0$, $f_\varepsilon(\overline{Y}) > f_\varepsilon(J_{\tau,h}(X))$. By using the convexity inequality

$$f_\varepsilon(\overline{Y}) \le (1-\lambda) f_\varepsilon(J_{\tau,h}(X)) + \lambda f_\varepsilon(Y),$$

we find $f_\varepsilon(Y) > f_\varepsilon(J_{\tau,h}(X))$. As a consequence, $\{J_{\tau,h_\varepsilon}(X) : 0 < \varepsilon \le \varepsilon_0\} \subset B$ and the result follows. □

We are now ready for the proof of Lemma 15.

Proof of Lemma 15. Proof of the Γ-lim inf: It is straightforward using Fatou's lemma, Proposition 18 and the lower semicontinuity of $\mathcal{Y} \mapsto \int_{\theta_0}^{\theta_1} |\dot{\mathcal{Y}}_\theta|^2 \,\mathrm{d}\theta$ with respect to the topology of uniform convergence.

Proof of the Γ-lim sup: Let us consider a curve $\mathcal{Y} \in H^1([\theta_0, \theta_1]; \mathbb{R}^p)$ with $\mathcal{Y}_{\theta_0} = R$ and $\mathcal{Y}_{\theta_1} = S$ (else there is nothing to prove). For all $\varepsilon > 0$ and $\tau > 0$, we define

$$\mathcal{Y}^{\tau,\varepsilon} : \theta \mapsto J_{\tau, g_\varepsilon(\theta,\cdot)}(\mathcal{Y}_\theta),$$

and correspondingly

$$\mathcal{Y}^{\tau} : \theta \mapsto J_{\tau, g(\theta,\cdot)}(\mathcal{Y}_\theta).$$

First, we prove

$$\limsup_{\tau\to 0}\limsup_{\varepsilon\to 0} \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta^{\tau,\varepsilon}|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Y}_\theta^{\tau,\varepsilon})|^2\}\eta(\theta)\,\mathrm{d}\theta \le \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta|^2 + |\overline{\nabla} g(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta. \quad (37)$$

We will then choose τ as a function of ε and show how to fix the endpoints.

Proof of (37): By the second point of Proposition 19, for all $\varepsilon, \tau, \theta$, we have

$$\mathcal{Y}_\theta = \mathcal{Y}_\theta^{\tau,\varepsilon} + \tau \nabla g_\varepsilon(\theta, \mathcal{Y}_\theta^{\tau,\varepsilon}).$$

For all θ, $g_\varepsilon(\theta,\cdot)$ is convex, so $Y \mapsto Y + \nabla g_\varepsilon(\theta, Y)$ is invertible and its inverse is 1-Lipschitz. In addition, the smoothness of $g_\varepsilon = g_\varepsilon(\theta, Y)$ with respect to θ lets us deduce from $\mathcal{Y} \in H^1$ that $\mathcal{Y}^{\tau,\varepsilon}$ is in H^1, and that, for almost all θ,

$$\dot{\mathcal{Y}}_\theta = (\mathbb{I} + \tau \mathrm{D}^2 g_\varepsilon(\theta, \mathcal{Y}_\theta^{\tau,\varepsilon})) \cdot \dot{\mathcal{Y}}_\theta^{\tau,\varepsilon} + \tau \partial_\theta \nabla g_\varepsilon(\theta, \mathcal{Y}_\theta^{\tau,\varepsilon}).$$

By the convexity of g_ε, we have $\mathbb{I} \leq \mathbb{I} + \tau \mathrm{D}^2 g_\varepsilon$ in the sense of symmetric matrices, and hence

$$|\dot{\mathcal{Y}}^{\tau,\varepsilon}_\theta| \leq |\dot{\mathcal{Y}}_\theta - \tau \partial_\theta \nabla g_\varepsilon(\theta, \mathcal{Y}^{\tau,\varepsilon}_\theta)| \leq |\dot{\mathcal{Y}}_\theta| + \tau M. \tag{38}$$

Recall that M was defined in the uniform integrability assumption (25) on $\partial_\theta \nabla g_\varepsilon$. (In the case when $\partial_\theta \nabla g_\varepsilon = 0$, we recover the known fact that for h independent of time, $J_{\tau,h}$ is contractive.) Then, we deduce

$$\begin{aligned}\limsup_{\varepsilon\to 0} \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}^{\tau,\varepsilon}_\theta|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Y}^{\tau,\varepsilon}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta &\overset{(34),(38)}{\leq} \limsup_{\varepsilon\to 0} \frac{1}{2}\int_{\theta_0}^{\theta_1} \left\{(|\dot{\mathcal{Y}}_\theta| + \tau M)^2 + \left|\frac{\mathcal{Y}_\theta - \mathcal{Y}^{\tau,\varepsilon}_\theta}{\tau}\right|^2\right\}\eta(\theta)\,\mathrm{d}\theta \\ &\overset{(35)}{\leq} \frac{1}{2}\int_{\theta_0}^{\theta_1} \left\{(|\dot{\mathcal{Y}}_\theta| + \tau M)^2 + \left|\frac{\mathcal{Y}_\theta - \mathcal{Y}^{\tau}_\theta}{\tau}\right|^2\right\}\eta(\theta)\,\mathrm{d}\theta \\ &\overset{(34)}{\leq} \frac{1}{2}\int_{\theta_0}^{\theta_1} \{(|\dot{\mathcal{Y}}_\theta| + \tau M)^2 + |\overline{\nabla} g(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta.\end{aligned}$$

Formula (37) follows.

<u>Choice of $\tau = \tau(\varepsilon)$</u>: Because of (37), and because,

$$\text{for all } \varepsilon > 0, \quad \mathcal{Y}^{\tau,\varepsilon}_{\theta_0} \xrightarrow[\tau\to 0]{} R \quad \text{and} \quad \mathcal{Y}^{\tau,\varepsilon}_{\theta_1} \xrightarrow[\tau\to 0]{} S,$$

it is possible to find a nonincreasing function $\tau = \tau(\varepsilon)$ converging sufficiently slowly to 0 so that

$$\limsup_{\varepsilon\to 0} \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}^{\tau(\varepsilon),\varepsilon}_\theta|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Y}^{\tau(\varepsilon),\varepsilon}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta \leq \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta|^2 + |\overline{\nabla} g(\theta, \mathcal{Y}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta, \tag{39}$$

$$\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0} \xrightarrow[\varepsilon\to 0]{} R \quad \text{and} \quad \mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_1} \xrightarrow[\tau\to 0]{} S. \tag{40}$$

<u>Fixing the endpoints</u>: For fixed ε and small $\delta > 0$, we will define $\mathcal{Z}^{\delta,\varepsilon}$ as a slight modification of the curve $\mathcal{Y}^{\varepsilon,\tau(\varepsilon)}$ in such a way that $\mathcal{Z}^{\delta,\varepsilon}$ joins R to S. For this, we just set for $\theta \in [\theta_0, \theta_1]$

$$\mathcal{Z}^{\delta,\varepsilon}_\theta = \begin{cases} R + ((\theta - \theta_0)/\delta)(\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\delta_0+\delta} - R) & \text{if } \theta \in [\theta_0, \theta_0 + \delta], \\ \mathcal{Y}^{\tau(\varepsilon),\varepsilon}_\theta & \text{if } \theta \in [\theta_0 + \delta, \theta_1 - \delta], \\ S + ((\theta_1 - \theta)/\delta)(\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\delta_1-\delta} - S) & \text{if } \theta \in [\theta_1 - \delta, \theta_1]. \end{cases}$$

A quick computation shows

$$\begin{aligned}&\frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Z}}^{\delta,\varepsilon}_\theta|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Z}^{\delta,\varepsilon}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta \\ &\leq \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}^{\tau(\varepsilon),\varepsilon}_\theta|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Y}^{\tau(\varepsilon),\varepsilon}_\theta)|^2\}\eta(\theta)\,\mathrm{d}\theta + \|\eta\|_\infty \left(\frac{|\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0+\delta} - R|^2}{2\delta} + \frac{|\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_1-\delta} - S|^2}{2\delta} + \delta L^2\right),\end{aligned} \tag{41}$$

where L is defined in the uniform Lipschitz assumption (23) for g_ε.

Let us estimate $|\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0+\delta} - R|^2/2\delta$. We have

$$\frac{|\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0+\delta} - R|^2}{2\delta} \leq \frac{|\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0} - R|^2}{\delta} + \frac{|\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0+\delta} - \mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0}|^2}{\delta} \leq \frac{|\mathcal{Y}^{\tau(\varepsilon),\varepsilon}_{\theta_0} - R|^2}{\delta} + \int_{\theta_0}^{\theta_0+\delta} |\dot{\mathcal{Y}}^{\tau(\varepsilon),\varepsilon}_\theta|^2\,\mathrm{d}\theta.$$

Because of (38), (25) and $\mathcal{Y} \in H^1$, the integral $\int_{\theta_0}^{\theta_0+\delta} |\dot{\mathcal{Y}}_\theta^{\tau(\varepsilon),\varepsilon}|^2 \, \mathrm{d}\theta$ tends to 0 as $\delta \to 0$, uniformly in ε: we bound it by a function $v_i = v_i(\delta)$ tending to 0 as $\delta \to 0$. In the same way,

$$\frac{|\mathcal{Y}_{\theta_1-\delta}^{\tau(\varepsilon),\varepsilon} - S|^2}{2\delta} \le \frac{|\mathcal{Y}_{\theta_1}^{\tau(\varepsilon),\varepsilon} - S|^2}{\delta} + v_f(\delta),$$

where $v_f(\delta) \to 0$ as $\delta \to 0$.

Plugging these bounds into (41), we get

$$\frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Z}}_\theta^{\delta,\varepsilon}|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Z}_\theta^{\delta,\varepsilon})|^2\}\eta(\theta)\,\mathrm{d}\theta \\ \le \frac{1}{2}\int_{\theta_0}^{\theta_1} \{|\dot{\mathcal{Y}}_\theta^{\tau(\varepsilon),\varepsilon}|^2 + |\nabla g_\varepsilon(\theta, \mathcal{Y}_\theta^{\tau(\varepsilon),\varepsilon})|^2\}\eta(\theta)\,\mathrm{d}\theta + \|\eta\|_\infty\left(\frac{u(\varepsilon)}{\delta} + v(\delta)\right),$$

where $u(\varepsilon) := |\mathcal{Y}_{\theta_0}^{\tau(\varepsilon),\varepsilon} - R|^2 + |\mathcal{Y}_{\theta_1}^{\tau(\varepsilon),\varepsilon} - S|^2 \to 0$ as $\varepsilon \to 0$ by (40), and $v(\delta) := v_i(\delta) + v_f(\delta) + \delta L^2 \to 0$ as $\delta \to 0$. Hence, choosing $\delta(\varepsilon) := \sqrt{u(\varepsilon)}$, we find with the help of (39) that $\mathcal{Z}^{\delta(\varepsilon),\varepsilon}$ is a recovery sequence for the $\Gamma - \limsup$ of K_ε towards K. □

3.4. *Proof of Lemma 17.* The proof is straightforward, and relies on explicit computations.

Proof of Lemma 17. Let us define for $X \in (\mathbb{R}^d)^N$

$$h(X) := \log\left[\frac{1}{N!}\sum_{\sigma\in\mathfrak{S}_N} \exp(X \cdot A^\sigma)\right]. \tag{42}$$

For $\varepsilon > 0$, $\theta \in [\theta_0, \theta_1]$ and $Y \in (\mathbb{R}^d)^N$, we have by the definition of f_ε and g_ε (formulas (19) and (26) respectively)

$$g_\varepsilon(\theta, Y) = \varepsilon h\left(\frac{Y}{\varepsilon \exp(\theta)}\right). \tag{43}$$

Proof of (H1): It is obvious.

Proof of (H2): By (43), it suffices to check that h is convex. Differentiating (42) twice, we get for all $X \in (\mathbb{R}^d)^N$

$$\mathrm{D}^2 h(X) = \langle A^\sigma \otimes A^\sigma\rangle_X - \langle A^\sigma\rangle_X \otimes \langle A^\sigma\rangle_X = \langle A^\sigma - \langle A^\sigma\rangle_X\rangle_X \otimes \langle A^\sigma - \langle A^\sigma\rangle_X\rangle_X, \tag{44}$$

where if a is a function of σ, then $\langle a(\sigma)\rangle_X$ stands for

$$\langle a(\sigma)\rangle_X := \frac{\sum_{\sigma\in\mathfrak{S}_N} a(\sigma)\exp(X \cdot A^\sigma)}{\sum_{\sigma\in\mathfrak{S}_N} \exp(X \cdot A^\sigma)}.$$

It follows that $\mathrm{D}^2 h(X)$ is a nonnegative symmetric matrix.

Proof of (H3): In view of (43) and as $\theta_0 > -\infty$, it suffices to check that ∇h is bounded. Differentiating (42) at $X \in (\mathbb{R}^d)^N$ leads to

$$\nabla h(X) = \langle A^\sigma\rangle_X,$$

which is clearly bounded by $|A|$.

Proof of (H4): By the definitions (20) of f and (26) of g, we have for all $\theta \in [\theta_0, \theta_1]$ and $Y \in (\mathbb{R}^d)^N$

$$g(\theta, Y) = \frac{f(Y)}{\exp(\theta)}.$$

The integrability property of $\partial_\theta g$ is clear; let us check (24). Let us consider $\mathcal{Y} \in H^1([\theta_0, \theta_1]; (\mathbb{R}^d)^N)$. The function g is locally Lipschitz both in θ and Y. As a consequence, the map $G : \theta \mapsto g(\theta, \mathcal{Y}_\theta)$ is also H^1.

Now, instead of proving (24), we will prove that for all curves $\mathcal{D} = \mathcal{D}_\theta$ such that, for almost all $\theta \in [\theta_0, \theta_1]$, $\mathcal{D}_\theta$ belongs to the subdifferential of $g(\theta, \cdot)$ at $Y = \mathcal{Y}_\theta$, we have for almost all $\theta \in [\theta_0, \theta_1]$

$$\frac{\mathrm{d}}{\mathrm{d}\theta} g(\theta, \mathcal{Y}_\theta) = \partial_\theta g(\theta, \mathcal{Y}_\theta) + \mathcal{D}_\theta \cdot \dot{\mathcal{Y}}_\theta,$$

so that (24) is an application of this property to $\mathcal{D}_\theta := \overline{\nabla} g(\theta, \mathcal{Y}_\theta)$. Notice that this property implies that up to negligible sets, $\mathcal{D}_\theta \cdot \dot{\mathcal{Y}}_\theta$ does not depend on the choice of $\mathcal{D}$. Let us give ourselves such a curve $\mathcal{D}$.

Let us take a point $\theta \in (\theta_0, \theta_1)$ where both $\mathcal{Y}$ and G are differentiable (this happens for almost every θ). We have

$$\begin{aligned} G'(\theta) &= \lim_{\delta \downarrow 0} \frac{1}{\delta} \left\{ \frac{f(\mathcal{Y}_{\theta+\delta})}{\exp(\theta+\delta)} - \frac{f(\mathcal{Y}_\theta)}{\exp(\theta)} \right\} = -\frac{f(\mathcal{Y}_\theta)}{\exp(\theta)} + \lim_{\delta \downarrow 0} \frac{g(\theta, \mathcal{Y}_{\theta+\delta}) - g(\theta, \mathcal{Y}_\theta)}{\delta} \\ &\geq -\frac{f(\mathcal{Y}_\theta)}{\exp(\theta)} + \limsup_{\delta \downarrow 0} \mathcal{D}_\theta \cdot \frac{\mathcal{Y}_{\theta+\delta} - \mathcal{Y}_\theta}{\delta} = \partial_\theta g(\theta, \mathcal{Y}_\theta) + \mathcal{D}_\theta \cdot \dot{\mathcal{Y}}_\theta, \end{aligned}$$

where we used $g(\theta, \mathcal{Y}_{\theta+\delta}) \geq g(\theta, \mathcal{Y}_\theta) + \mathcal{D}_\theta \cdot (\mathcal{Y}_{\theta+\delta} - \mathcal{Y}_\theta)$ to get the second line.

In the same way, we have

$$G'(\theta) = \lim_{\delta \downarrow 0} \frac{1}{\delta} \left\{ \frac{f(\mathcal{Y}_\theta)}{\exp(\theta)} - \frac{f(\mathcal{Y}_{\theta-\delta})}{\exp(\theta-\delta)} \right\} \leq \partial_\theta g(\theta, \mathcal{Y}_\theta) + \mathcal{D}_\theta \cdot \dot{\mathcal{Y}}_\theta.$$

The result follows from gathering these two inequalities.

Proof of (H5): Using (43), we get for all $\varepsilon > 0$, $\theta \in [\theta_0, \theta_1]$ and $Y \in (\mathbb{R}^d)^N$,

$$\partial_\theta \nabla g_\varepsilon(\theta, Y) = -\frac{1}{\exp(\theta)} \left(\nabla h\left(\frac{Y}{\varepsilon \exp(\theta)} \right) + \mathrm{D}^2 h\left(\frac{Y}{\varepsilon \exp(\theta)} \right) \cdot \frac{Y}{\varepsilon \exp(\theta)} \right).$$

As we already saw in (H3) that ∇h is bounded, it suffices to prove that $X \mapsto \mathrm{D}^2 h(X) \cdot X$ is bounded. Let us expand everything in (44) and apply X to the right. We get

$$\mathrm{D}^2 h(X) \cdot X = \frac{\sum_{\sigma, \eta \in \mathfrak{S}_N} X \cdot (A^\sigma - A^\eta) A^\sigma \exp(X \cdot (A^\sigma + A^\eta))}{\sum_{\sigma, \eta \in \mathfrak{S}_N} \exp(X \cdot (A^\sigma + A^\eta))}.$$

As a consequence, it suffices to show that, for each $\sigma, \eta \in \mathfrak{S}_N$,

$$T(\sigma, \eta, X) := \frac{X \cdot (A^\sigma - A^\eta) \exp(X \cdot (A^\sigma + A^\eta))}{\sum_{\sigma', \eta' \in \mathfrak{S}_N} \exp(X \cdot (A^{\sigma'} + A^{\eta'}))}$$

is bounded, uniformly in X. First, if $\eta = \sigma$, then $T(\sigma, \sigma, X) = 0$. Else, let us use the bound

$$\sum_{\sigma', \eta' \in \mathfrak{S}_N} \exp(X \cdot (A^{\sigma'} + A^{\eta'})) \leq \exp(2X \cdot A^\sigma) + \exp(2X \cdot A^\eta),$$

obtained by only keeping the terms corresponding to $\sigma' = \eta' = \sigma$ and $\sigma' = \eta' = \eta$ in the sum. This leads to

$$|T(\sigma, \eta, X)| \leq \frac{|X \cdot (A^\sigma - A^\eta)| \exp(X \cdot (A^\sigma + A^\eta))}{\exp(2X \cdot A^\sigma) + \exp(2X \cdot A^\eta)} = \frac{|X \cdot (A^\sigma - A^\eta)|}{\exp(-|X \cdot (A^\sigma - A^\eta)|) + \exp(|X \cdot (A^\sigma - A^\eta)|)},$$

which is clearly bounded uniformly in X. The result follows. $\square$

4. The case of dimension 1: sticky collisions

In this section, we will study the global minimizers of the functional Λ' obtained in Section 2.6, in dimension $d = 1$. If we call t the time variable and if we replace θ_0 and θ_1 by 0 and T respectively, due to the invariance of the functional through translation in time, Λ' reads

$$\Lambda'(\mathcal{Z}) = \begin{cases} \int_0^T \left\{\frac{1}{2}|\dot{\mathcal{Z}}_t|^2 + \frac{1}{2}|\mathcal{Z}_t - \overline{\nabla} f(\mathcal{Z}_t)|^2\right\} \mathrm{d}t & \text{if } \mathcal{Z} \in H^1([0,T];\mathbb{R}^N), \mathcal{Z}_0 \in \{P^\sigma\} \text{ and } \mathcal{Z}_T \in \{Q^\sigma\}, \\ +\infty & \text{else,} \end{cases} \tag{45}$$

where

$$f(X) = \max_{\sigma \in \mathfrak{S}_N} X \cdot A^\sigma, \quad X \in \mathbb{R}^N. \tag{46}$$

Here, we chose a *strictly ordered* $A = (a_1, \dots, a_N)$, that is, such that $a_1 < \cdots < a_N$, $P, Q \in \mathbb{R}^N$ and $T > 0$. Once again, when $X = (x_1, \dots, x_N) \in \mathbb{R}^N$ and $\sigma \in \mathfrak{S}_N$, we let $X^\sigma := (x_{\sigma(1)}, \dots, x_{\sigma(N)})$, and $\{P^\sigma\}$ and $\{Q^\sigma\}$ refer to $\{P^\sigma : \sigma \in \mathfrak{S}_N\}$ and $\{Q^\sigma : \sigma \in \mathfrak{S}_N\}$ respectively. Of course $P = (p_1, \dots, p_N)$ and $Q = (q_1, \dots, q_N)$ can be supposed to be ordered, that is, $p_1 \le \cdots \le p_N$ and $q_1 \le \cdots \le q_N$. We recall that we defined the *extended gradient* $\overline{\nabla} f$ in Definition 8. As already noticed in Section 2.5, the existence of global minimizers for Λ' follows from the direct method of calculus of variations. Uniqueness does not hold in general, even up to permutations.

The purpose of the section is two-fold. On the one hand, we will show that the model has nice regularity properties: any global minimizer of Λ' is smooth except on a finite number of "sticking" or "separation" times.[1] On the other hand, we will justify as claimed in Section 2 that Λ' describes a model with sticky collisions in the sense that a minimizer $\mathcal{Z} = (z_1(t), \dots, z_N(t))$ of Λ' will typically exhibit some sticking effects as $z_i(t) = z_j(t)$ for $i \neq j$ on nontrivial intervals.

To describe the sticking effect, it is convenient to introduce the following definition:

Definition 20 (partition of $[\![1, N]\!]$). Let $X \in \mathbb{R}^N$. We say that X is divided according to $\pi(X)$ when $\pi(X)$ is the partition of $[\![1, N]\!]$ induced by the relation,

$$\text{for all } (i, j) \in [\![1, N]\!]^2, \quad i \sim j \iff x_i = x_j.$$

We call $C(X, i)$ the class of $i \in [\![1, N]\!]$ in $\pi(X)$, namely, $C(X, i) = \{j : x_j = x_i\}$.

The main result of the section is the following:

Theorem 21 (regularity of the optimal trajectories). *For given $A, P, Q \in \mathbb{R}^N$ and $T > 0$ as before, let $\mathcal{Z}$ be a global minimizer of Λ' defined in* (45). *Then $\mathcal{Z}$ is continuous and there exist*

$$0 = t_0 < t_1 < \cdots < t_p = T,$$

a family of times such that, for each $i = 1, \dots, p$, $\mathcal{Z}$ is smooth on $[t_{i-1}, t_i]$, and $\pi(\mathcal{Z})$ is constant on (t_{i-1}, t_i).

It will be quite clear from the proof that sticking effects do occur. This exactly means that there exist trajectories $\mathcal{Z}$ for which, with the notation of Section 2.7, $\Lambda'(\mathcal{Z}) < \Lambda^+(\mathcal{Z})$. For such trajectories, $\mathcal{Z}_t$ is

[1]Notice that Λ' is invariant under time inversion, so that if particles are allowed to stick, they are also allowed to separate.

located on the set where f is not differentiable for a set of times of positive Lebesgue measure. But in dimension 1, this set is exactly the set where at least two particles are located at the same place. That is, the set of times when $\pi(\mathcal{Z}) \neq \{\{1\}, \dots, \{N\}\}$ is typically of positive Lebesgue measure. As a consequence of Theorem 21, it is even a finite union of intervals.

Still it might be convenient to illustrate the sticking effects included in the model by the following easy proposition. It asserts that the set of times when all the particles are stuck is an interval: if all the particles are stuck at two different times, the cheapest behavior between these two times is to remain stuck. It also shows that this phenomenon occurs: if all the particles are sufficiently close at the initial and final time, then they necessarily stick together during a nontrivial interval along the evolution.

Proposition 22 (intervals of full degeneration). (1) *For given $A, P, Q \in \mathbb{R}^N$ and $T > 0$ as before, let $\mathcal{Z} = (z_1(t), \dots, z_N(t))$ be a global minimizer of Λ'. Suppose there exist two times $0 \leq t_1 < t_2 \leq T$ such that*

$$z_1(t_1) = \cdots = z_N(t_1) \quad \textit{and} \quad z_1(t_2) = \cdots = z_N(t_2).$$

Then, for all $t \in [t_1, t_2]$, we have $z_1(t) = \cdots = z_N(t)$.

(2) *For given $A \in \mathbb{R}^N$ and $T > 0$ as before, the set $\mathcal{U}$ of endpoints $P, Q \in \mathbb{R}^N$ with the property that, for all minimizers $\mathcal{Z} = (z_1(t), \dots, z_N(t))$ of Λ', the set of times*

$$\{t \in [0, T] : z_1(t) = \cdots = z_N(t)\}$$

is a nontrivial interval, is a neighborhood of $\{P, Q \in \mathbb{R}^N : p_1 = \cdots = p_N \text{ and } q_1 = \cdots = q_N\}$.

The proof of Proposition 22 uses almost nothing and is given in Section 4.2. Except for that, the whole section is dedicated to the proof of Theorem 21. For this we take once for all $A, P, Q \in \mathbb{R}^N$ and $T > 0$, A being strictly ordered and P, Q being ordered.

Even if all the arguments are elementary, we will need a certain number of steps, including:

- The explicit computation of the potential $|X - \overline{\nabla} f(X)|^2$ (Section 4.1 and 4.4).
- The justification of a priori knowledge on the optimal trajectories: they can be supposed to be ordered at all times (Section 4.3).
- The conservation of energy and momentum holds during shocks[2] (Section 4.5).

Then, the main ingredient in the proof of Theorem 21 is an estimate given in Section 4.6: during a nonpathological shock (pathological shocks are excluded a posteriori), at least one particle has a lower-bounded jump in its velocity (Proposition 31). We finally provide the proof of Theorem 21 in Section 4.7.

Throughout the section, we will work with several types of finite sets: the partitions of type $\pi(X)$ and the class of particles of type $C(X, i)$. Some of the arguments or computations will deal with their cardinality. Thus, if $\mathcal{F}$ is a finite set, we will denote by $\#\mathcal{F}$ its cardinality.

[2]We say that $\mathcal{Z}$ presents a shock at time t if t is a discontinuity point of $\pi(\mathcal{Z})$; see Definition 30.

4.1. ***Properties of the extended gradient.*** In Lemma 24, we gather easy properties of $\overline{\nabla} f$ that will be needed in the following. Before doing so, let us introduce some notation.

Definition 23. Let π be a partition of $[\![1, N]\!]$. We call E_π the linear subspace of $\mathbb{R}^N$ of all X such that π is a refinement of $\pi(X)$, that is,

$$E_\pi := \bigcap_{C\in\pi} \bigcap_{i,j\in C} \{X = (x_1, \dots, x_N) \in \mathbb{R}^N : x_i = x_j\}.$$

Lemma 24 (properties of $\overline{\nabla} f$). (1) *The extended gradient $\overline{\nabla} f$ has the following symmetry*:

$$\text{for all } X \in \mathbb{R}^N,\ \sigma \in \mathfrak{S}_N, \quad \overline{\nabla} f(X^\sigma) = (\overline{\nabla} f(X))^\sigma. \tag{47}$$

(2) *The function $X \mapsto |X - \overline{\nabla} f(X)|$ is symmetric*:

$$\text{for all } X \in \mathbb{R}^N,\ \sigma \in \mathfrak{S}_N, \quad |X^\sigma - \overline{\nabla} f(X^\sigma)|^2 = |X - \overline{\nabla} f(X)|^2. \tag{48}$$

(3) *If X is ordered*, *then $\overline{\nabla} f(X)$ is the orthogonal projection of A on $E_{\pi(X)}$.*

(4) *If X is ordered and $i \in \{1, \dots, N\}$*,

$$(\overline{\nabla} f(X))_i = \frac{1}{\#C(X,i)} \sum_{j\in C(X,i)} a_j. \tag{49}$$

(*Recall that $C(X, i)$ is defined in Definition 20.*)

Remark 25. The extended gradient $\overline{\nabla} f$ is completely characterized by points (1) and (3) (or (4)) of Lemma 24.

Proof. (1) Let $\sigma \in \mathfrak{S}_N$. By the definition (46) of f, for all $X \in \mathbb{R}^N$, $f(X^\sigma) = f(X)$. Letting $I^\sigma : X \mapsto X^\sigma$, we easily deduce that at the level of subdifferentials: $\partial f(X^\sigma) = I^\sigma(\partial f(X))$. We conclude by the fact that I^σ is orthogonal.

(2) It is a direct consequence of point (1).

(3) Let $X = (x_1, \dots x_N) \in \mathbb{R}^N$ be an ordered vector. Considering the definition (46) of f and noticing that the maximum is achieved exactly for those σ such that $X^\sigma = X$, it appears that $\overline{\nabla} f(X)$ belongs to the convex hull:

$$\mathrm{Conv}(\{A^\sigma : \sigma \in \mathfrak{S}_N \text{ such that } X^\sigma = X\}).$$

For a given $i \in \{1, \dots, N\}$, we call $V^i \in \mathbb{R}^N$ the vector whose j-th coordinate is 1 if $j \in C(X, i)$ and 0 otherwise. On the one hand, we have $E_{\pi(X)} = \mathrm{Span}\{V^i : i = 1, \dots, N\}$, and on the other hand, for all i, the scalar product $V^i \cdot Y$ is constant on the above-mentioned convex hull. So we deduce

$$A - \overline{\nabla} f(X) \in (E_{\pi(X)})^\perp.$$

Hence, we just have to prove that $\overline{\nabla} f(X) \in E_{\pi(X)}$. If $i, j \in \{1, \dots, N\}$ are such that $x_i = x_j$, let us apply formula (47) to the permutation $\sigma := (i, j)$:

$$(\overline{\nabla} f(X))_i = ((\overline{\nabla} f(X))^\sigma)_j = (\overline{\nabla} f(X^\sigma))_j = (\overline{\nabla} f(X))_j.$$

The result follows.

(4) Let X be ordered and $i \in \{1, \dots, N\}$. As $\overline{\nabla} f(X) \in E_{\pi(X)}$, with the notation of the proof of (3),

$$\begin{aligned}(\overline{\nabla} f(X))_i &= \frac{1}{\#C(X,i)} \sum_{j \in C(X,i)} (\overline{\nabla} f(X))_j = \frac{1}{\#C(X,i)} \overline{\nabla} f(X) \cdot V^i \\ &= \frac{1}{\#C(X,i)} A \cdot V^i = \frac{1}{\#C(X,i)} \sum_{j \in C(X,i)} a_j,\end{aligned}$$

where we used $A - \overline{\nabla} f(X) \perp V^i$ to get the first identity in the second line. □

The three next subsections will be dedicated to consequences of this lemma:

- A proof of Proposition 22.
- When proving Theorem 21, it is enough to consider ordered trajectories (Proposition 26).
- For ordered trajectories, the potential in Λ' can be decomposed as sum of a smooth "external" potential and an "internal" energy only depending on $\pi(X)$ (Proposition 27).

4.2. *Proof of Proposition 22.* With the help of Lemma 24, we are ready to prove Proposition 22.

Proof of Proposition 22. (1) Without loss of generality, we can suppose $t_1 = 0$ and $t_2 = T$, that is, $P = (p_1, \dots, p_N)$ and $Q = (q_1, \dots, q_N)$ are such that $p_1 = \dots = p_N$ and $q_1 = \dots = q_N$.

Call Ψ the orthogonal projection on the line $E_{[\![1,N]\!]} := \{X = (x_, \dots, x_N) \in \mathbb{R}^N \mid x_1 = \dots = x_N\}$. It suffices to prove that when $\mathcal{Z}$ is a continuous trajectory joining P to Q, then $\Lambda'(\Psi(\mathcal{Z})) \le \Lambda'(\mathcal{Z})$, and with equality if and only if $\mathcal{Z} = \Psi(\mathcal{Z})$. As Ψ is 1-Lipschitz, it reduces the kinetic part of Λ'. For the potential part, we remark that, for all $X \in \mathbb{R}^N$, $E_{\pi(\Psi(X))} = E_{[\![1,N]\!]} \subset E_{\pi(X)}$. As a consequence, by point (3) of Lemma 24, we have as soon as X is ordered $\overline{\nabla} f(\Psi(X)) = \Psi(\overline{\nabla} f(X))$. Hence

$$|\Psi(X) - \overline{\nabla} f(\Psi(X))|^2 = |\Psi(X - \overline{\nabla} f(X))|^2 \le |X - \overline{\nabla} f(X)|^2,$$

with equality if and only if $X \in E_{[\![1,N]\!]}$, i.e., if and only if $\Psi(X) = X$. This property is extended to nonordered X using (48), and the result follows.

(2) The function $\overline{\Lambda'} = \overline{\Lambda'}(P, Q)$, defined for all $P, Q \in \mathbb{R}^N$ as the minimal value of Λ', is continuous. Indeed, if $P, P', Q, Q' \in \mathbb{R}^N$ are chosen so that $|P' - P| + |Q' - Q| \ll 1$ and if $\mathcal{Z}$ is a trajectory joining P to Q, we can find a trajectory $\widetilde{\mathcal{Z}}$ joining P' to Q' with[3]

$$\Lambda'(\widetilde{\mathcal{Z}}) \le \Lambda'(\mathcal{Z}) + \underset{(P',Q') \to (P,Q)}{o}(1). \tag{50}$$

To do so, it suffices to choose $\tau \sim |P' - P| + |Q' - Q|$, and to define $\widetilde{\mathcal{Z}}$ as the trajectory joining P' to P in straight line between times 0 and τ, joining P to Q between times τ and $T - \tau$ by following $\mathcal{Z}$ with a proper affine change of time, and finally joining Q to Q' in straight line between times $T - \tau$ and T. This shows that $\overline{\Lambda'}$ is lower semicontinuous, but the continuity is obtained by noticing that the o in (50) is locally uniform on $P, Q \in \mathbb{R}^N$. The argument is easily adapted to show that $\widetilde{\Lambda'} = \widetilde{\Lambda'}(P, Q)$, defined

[3]With a slight abuse of notation, we do not refer explicitly to the dependence of Λ' on P, Q.

for $P, Q \in \mathbb{R}^N$ by

$$\widetilde{\Lambda'}(P,Q) := \inf\{\Lambda'(\mathcal{Z}) : \mathcal{Z} \text{ whose set of } t \text{ such that } \mathcal{Z}_t \in E_{[\![1,N]\!]} \text{ is negligible}\},$$

is also continuous. Additionally, the set $\mathcal{U}$ defined in the statement clearly satisfies

$$\mathcal{V} := \{P, Q \in \mathbb{R}^N : \overline{\Lambda'}(P,Q) < \widetilde{\Lambda'}(P,Q)\} \subset \mathcal{U}.$$

By the continuity of $\overline{\Lambda'}$ and $\widetilde{\Lambda'}$, $\mathcal{V}$ is an open set. Hence it remains to prove that

$$\{P, Q \in \mathbb{R}^N : p_1 = \cdots = p_N \text{ and } q_1 = \cdots = q_N\} = E_{[\![1,N]\!]} \times E_{[\![1,N]\!]} \subset \mathcal{V}.$$

To do so, we take $P, Q \in E_{[\![1,N]\!]}$ and $\mathcal{Z}$ a curve joining P to Q such that $\{t : \mathcal{Z}_t \in E_{[\![1,N]\!]}\}$ is negligible, we still call Ψ the orthogonal projection on $E_{[\![1,N]\!]}$, and we prove that

$$\Lambda'(\mathcal{Z}) \geq \Lambda'(\Psi(\mathcal{Z})) + a,$$

where $a > 0$ does not depend on $\mathcal{Z}$. Let us call $\Phi := \mathrm{Id} - \Psi$ the orthogonal projection on the orthogonal of $E_{[\![1,N]\!]}$. As in the proof of the first point, $\overline{\nabla} f \circ \Psi = \Psi \circ \overline{\nabla} f$. As a consequence

$$\begin{aligned}\Lambda'(\mathcal{Z}) &= \int_0^T \{|\Psi(\dot{\mathcal{Z}}_t)|^2 + |\Psi(\mathcal{Z}_t) - \Psi(\overline{\nabla} f(\mathcal{Z}_t))|^2\}\,\mathrm{d}t + \int_0^T \{|\Phi(\dot{\mathcal{Z}}_t)|^2 + |\Phi(\mathcal{Z}_t) - \Phi(\overline{\nabla} f(\mathcal{Z}_t))|^2\}\,\mathrm{d}t \\ &= \Lambda'(\Psi(\mathcal{Z})) + \int_0^T \{|\dot{\mathcal{Z}}_t^\perp|^2 + |\mathcal{Z}_t^\perp - \Phi(\overline{\nabla} f(\mathcal{Z}_t))|^2\}\,\mathrm{d}t,\end{aligned}$$

where $\mathcal{Z}^\perp = \mathcal{Z}_t^\perp := \Phi(\mathcal{Z}_t)$ is a curve joining 0 to 0. But for almost all t, $\mathcal{Z}_t \notin E_{[\![1,N]\!]}$, so as we saw in the proof of the first point, $\overline{\nabla} f(\mathcal{Z}_t) \notin E_{[\![1,N]\!]}$. As $\overline{\nabla} f$ only takes a finite number of values (see Lemma 24), for almost all t, $\Phi(\overline{\nabla} f(\mathcal{Z}_t))$ belongs to some finite set, say $\mathcal{G}$, which does not contain 0. Hence,

$$\int_0^T \{|\dot{\mathcal{Z}}_t^\perp|^2 + |\mathcal{Z}_t^\perp - \Phi(\overline{\nabla} f(\mathcal{Z}_t))|^2\}\,\mathrm{d}t \geq \int_0^T \{|\dot{\mathcal{Z}}_t^\perp|^2 + \mathrm{dist}(\mathcal{Z}_t^\perp, \mathcal{G})^2\}\,\mathrm{d}t,$$

where $\mathrm{dist}(Z, \mathcal{G})$ denotes the distance from Z to $\mathcal{G}$. Because $\mathcal{Z}^\perp$ joins 0 to 0 and $\mathcal{G}$ does not contain 0, this last integral is easily seen to be bounded below away from 0 independently of $\mathcal{Z}$, and the result follows. □

4.3. ***Ordering of the particles.*** The purpose of this subsection is to show that when proving Theorem 21, we can restrict ourselves to study trajectories that remain ordered (see Figure 1). This is due to the following proposition.

Proposition 26. *Let $\mathcal{Z} = \mathcal{Z}_t$ be a global minimizer of Λ'. We call $\widetilde{\mathcal{Z}} = \widetilde{\mathcal{Z}}_t$ the trajectory obtained by reordering the coordinates of $\mathcal{Z}$ in increasing order. Then $\widetilde{\mathcal{Z}}$ is also a global minimizer of Λ'.*

Moreover, $\mathcal{Z}$ has the regularity stated in Theorem 21 if and only if $\widetilde{\mathcal{Z}}$ does.

In particular, Λ' always admits an ordered minimizer, and it is enough to prove Theorem 21 for such minimizers.

Thanks to this proposition, from now on, we only work with ordered minimizers of Λ'. These minimizers $\mathcal{Z} = \mathcal{Z}_t$ satisfy in particular $\mathcal{Z}_0 = P$ and $\mathcal{Z}_T = Q$ (as we chose them to be ordered in the first place).

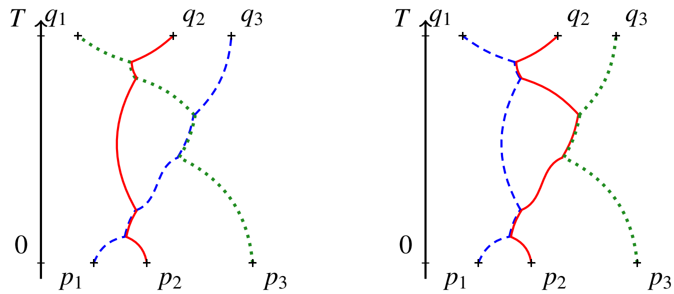

Figure 1. These two trajectories share their initial and final positions up to ordering and their actions. But to the right, the order is preserved, while to the left, this is not the case.

Proof. Let $\mathcal{Z}$ and $\widetilde{\mathcal{Z}}$ be as in the statement of the proposition. Point (2) of Lemma 24 implies

$$\int_0^T |\dot{\widetilde{\mathcal{Z}}}_t - \overline{\nabla} f(\widetilde{\mathcal{Z}}_t)|^2 \,\mathrm{d}t = \int_0^T |\dot{\mathcal{Z}}_t - \overline{\nabla} f(\mathcal{Z}_t)|^2 \,\mathrm{d}t.$$

We call $\Psi : \mathbb{R}^N \to \mathbb{R}^N$ the operator that reorders the coordinates of a vector in increasing order, so that in particular, for all t, $\widetilde{\mathcal{Z}}_t = \Psi(\mathcal{Z}_t)$. A simple application of the rearrangement inequality shows that Ψ is 1-Lipschitz. In particular, it reduces the action of curves

$$\int_0^T |\dot{\widetilde{\mathcal{Z}}}_t|^2 \,\mathrm{d}t \le \int_0^T |\dot{\mathcal{Z}}_t|^2 \,\mathrm{d}t.$$

By adding the two last formulas, and by noticing that the endpoint constraint is fulfilled, we get $\Lambda'(\widetilde{\mathcal{Z}}) \le \Lambda'(\mathcal{Z})$. As $\mathcal{Z}$ is a minimizer, this inequality is in fact an equality, and $\widetilde{\mathcal{Z}}$ is also a minimizer.

Note that both $\mathcal{Z}$ and $\widetilde{\mathcal{Z}}$ are continuous because they have finite action. Hence, the second claim of the proposition is a consequence of the two following facts:

- For all $t \in [0, T]$, $\#\pi(\widetilde{\mathcal{Z}}_t) = \#\pi(\mathcal{Z}_t)$.
- For any continuous trajectory $t \in I \mapsto \mathcal{X}_t \in \mathbb{R}^N$, where I is an interval, $t \mapsto \pi(\mathcal{X}_t)$ is constant if and only if $t \mapsto \#\pi(\mathcal{X}_t)$ is constant.

Indeed in that case, $t \mapsto \pi(\mathcal{Z}_t)$ and $t \mapsto \pi(\widetilde{\mathcal{Z}}_t)$ are constant on the same intervals, and the result follows.

The first point and the "only if" part of the second point are trivial.

For the "if" part of the second one, we reason by contraposition. Suppose $s \mapsto \pi(\mathcal{X}_s)$ has a discontinuity at time t and we prove that $s \mapsto \#\pi(\mathcal{X}_s)$ also does. If $s \mapsto \pi(\mathcal{X}_s)$ has a discontinuity at time t, we can find two distinct accumulation points π_1 and π_2 of $s \mapsto \pi(\mathcal{X}_s)$ at time t. As the set E_π is closed for all π, we find that $\mathcal{X}_t$ belongs to $E_{\pi_1} \cap E_{\pi_2}$. But this set is nothing but $E_{\bar{\pi}}$, where $\bar{\pi}$ is the finest partition of which π_1 and π_2 are refinements, that is, the partition corresponding to the relation

$$i \sim j \quad \iff \quad \text{there exists } C \in \pi_1 \cup \pi_2 \text{ such that } \{i, j\} \subset C.$$

In particular, $\pi(\mathcal{X}_t)$ is a refinement of $\bar{\pi}$ and as $\pi_1 \neq \pi_2$,

$$\#\pi(\mathcal{X}_t) \le \#\bar{\pi} < \max(\#\pi_1, \#\pi_2).$$

So $s \mapsto \#\pi(\mathcal{X}_s)$ has a discontinuity at time t, and the result follows. $\square$

4.4. ***Decomposition of the potential.*** Here, we compute explicitly the values of the potential $X \mapsto |X-\overline{\nabla} f(X)|^2$ on ordered vectors $X \in \mathbb{R}^N$. Notice that, for such vectors X, $\pi(X)$ has an additional structure: if $C \in \pi(X)$, then C is an interval of integers. We say that such partitions are *ordered*. We prove the following:

Proposition 27. *For all ordered $X \in \mathbb{R}^N$,*

$$|X-\overline{\nabla} f(X)|^2 = |X-A|^2 + h(\pi(X)) - |A|^2, \tag{51}$$

where h is defined on a partition π of $[\![1, N]\!]$ by

$$h(\pi) := \sum_{C\in\pi} \frac{1}{\#C} \Big|\sum_{j\in C} a_j\Big|^2. \tag{52}$$

In particular, h has the following monotonicity property: if π and π' are two ordered partitions and if π' is a strict refinement of π, then $h(\pi) < h(\pi')$.

The more particles are stuck together, the lower h is. This is the reason for which Λ' favors the sticking of particles. The function $-h/2$ can be understood as the internal energy of the system.

Dropping the constant term $|A|^2/2$ in (51) and defining Λ'' on a trajectory $\mathcal{Z}$ by

$$\Lambda''(\mathcal{Z}) = \begin{cases} \frac{1}{2}\int_0^T \{|\dot{\mathcal{Z}}_t|^2 + |\mathcal{Z}_t - A|^2 + h(\pi(\mathcal{Z}_t))\}\,\mathrm{d}t & \text{if } \mathcal{Z} \in H^1([0,T];\mathbb{R}^N),\ \mathcal{Z}_0 = P \text{ and } \mathcal{Z}_T = Q, \\ +\infty & \text{else,} \end{cases} \tag{53}$$

it is clear that Λ' and Λ'' have the same minimizers in the class of ordered trajectories. Hence, as a consequence of Proposition 26, it suffices to prove the conclusion of Theorem 21 for the minimizers of Λ'' in the class of ordered trajectories.

Proof of Proposition 27. Let $X \in \mathbb{R}^N$ be an ordered vector. By point (3) of Lemma 24, we have $A-\overline{\nabla} f(X) \in (E_{\pi(X)})^{\perp}$ and both X and $\overline{\nabla} f(X) \in E_{\pi(X)}$. So using the Pythagorean theorem twice, we get

$$|X-\overline{\nabla} f(X)|^2 = |X-A|^2 - |A-\overline{\nabla} f(X)|^2 = |X-A|^2 + |\overline{\nabla} f(X)|^2 - |A|^2.$$

The identities (51) and (52) are obtained by computing $|\overline{\nabla} f(X)|^2$ using (49).

If we recap, $h(\pi)$ is the squared norm of the orthogonal projection of A on E_π. But if π' is a refinement of π, $E_\pi \subset E_{\pi'}$, and hence $h(\pi) \le h(\pi')$. The strict inequality is obtained by noticing with the help of (49) and using the strict ordering of A that if in addition π and π' are ordered and $\pi' \neq \pi$, then the projection of A on $E_{\pi'}$ does not belong to E_π. □

4.5. ***Conserved quantities.*** In this subsection, we discuss two simple and yet structural properties of the dynamic prescribed by the functionals Λ', Λ'': the Hamiltonian of the system is conserved (Proposition 28), and its center of mass draws a smooth curve (Proposition 29). In particular, the momentum of the system is conserved during shocks.

Proposition 28. *Let $\mathcal{Z}$ be an ordered minimizer of Λ''. Then*

$$\mathcal{E} = \mathcal{E}(t) := \tfrac{1}{2}\{|\dot{\mathcal{Z}}_t|^2 - |\mathcal{Z}_t - A|^2 - h(\pi(\mathcal{Z}_t))\} \tag{54}$$

is constant in the sense of distributions.

Proof. The proof is completely standard and is done by comparing the value of Λ'' on $\mathcal{Z}$ and $t \mapsto \mathcal{Z}_{t+\varepsilon\varphi(t)}$ for small ε and functions φ that are smooth and compactly supported in $(0, T)$. ☐

Proposition 29. *Let $\mathcal{Z} = (z_1(t), \dots, z_N(t))$ be an ordered minimizer of Λ''. Call $a := (a_1 + \dots + a_N)/N$ and for $t \in [0, T]$*

$$\mathcal{M}(t) := \frac{1}{N}\sum_{i=1}^{N} z_i(t) \quad \textit{and} \quad \mathcal{P}(t) := \frac{1}{N}\sum_{i=1}^{N} \dot{z}_i(t).$$

($\mathcal{M}$ is well-defined for all t, and $\mathcal{P}$ for almost all t.) Then $\mathcal{M}$, $\mathcal{P}$ solve distributionally

$$\dot{\mathcal{M}}(t) = \mathcal{P}(t), \quad \dot{\mathcal{P}}(t) = \mathcal{M}(t) - a.$$

In particular, $\mathcal{M}$ is smooth and $\mathcal{P}$ coincide almost surely with a smooth function.

Proof. Here the proof consists in comparing the value of Λ'' on $\mathcal{Z}$ and $t \mapsto \mathcal{Z}_t + \varepsilon\varphi(t)V$ for small ε, smooth and compactly supported φ, and where we call $V = (1, \dots, 1)$. The only somehow unusual thing to remark is that π and hence $h \circ \pi$ are invariant under translations in the direction of V. ☐

4.6. ***Shocks, isolated shocks and minimal deviation.*** This subsection contains the main estimate that allows us to prove Theorem 21. Roughly speaking, if at time t some of the particles stick or separate, there is a lower bound on the change of the velocity of at least one particle. The proof of Theorem 21 will then consist in showing that this cannot happen an infinite number of times.

Let us first define as "shocks" these sticking and separating behaviors:

Definition 30 (shocks). Let $\mathcal{X} = \mathcal{X}_t = (x_1(t), \dots, x_N(t))$ be a continuous trajectory on $\mathbb{R}^N$.

(1) We call a shock of $\mathcal{X}$ a triplet (t, q, C) with $t \in [0, T]$, $q \in \mathbb{R}$ and $C \subset [\![1, N]\!]$ such that
   - $C \in \pi(\mathcal{X}_t)$,
   - for all $i \in C$, $x_i(t) = q$,
   - for all $\tau > 0$, there exists $s \in (t - \tau, t + \tau)$ such that $C \notin \pi(\mathcal{X}_s)$.

(2) If (t, q, C) is a shock of $\mathcal{X}$, we say that it is isolated if (t, q) is isolated in

$$\{(t', q') : \text{there exists } C' \subset [\![1, N]\!] \text{ such that } (t', q', C') \text{ is a shock}\},$$

i.e., if there is no other shock than (t, q, C) in the neighborhood of $(t, q) \in [0, T] \times \mathbb{R}$.

We provide in Figure 2 a picture of a shock which does not seem to be isolated. The following result is the main step in the proof of Theorem 21.

Proposition 31. *Let $\mathcal{Z} = (z_1(t), \dots, z_N(t))$ be an ordered minimizer of Λ' (or equivalently a minimizer of Λ'' in the class of ordered trajectories), and let $t \in [0, T]$.*

(1) *If particle i is not involved in a shock at time t, then for s in the neighborhood of t, $C := C(\mathcal{Z}_s, i)$ is constant and z_i is a smooth solution of*

$$\ddot{z}_i(s) = z_i(s) - \frac{1}{\#C}\sum_{j \in C} a_j. \tag{55}$$

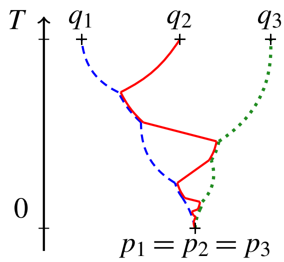

Figure 2. A shock with three particles which does not seem to be isolated. We will see later on that this kind of shock cannot occur in our model.

In particular, if i is involved in an isolated shock at time t, then z_i admits left and right derivatives at time t, denoted by $\dot{z}_i(t-)$ and $\dot{z}_i(t+)$ respectively.

(2) *There is $\alpha = \alpha(N, A) > 0$ such that for any isolated shock (t, q, C), calling $i := \min C$,*

$$\dot{z}_i(t-) - \dot{z}_i(t+) \geq \alpha. \tag{56}$$

(Note that the quantity $\dot{z}_i(t-) - \dot{z}_i(t+)$ is not affected by time inversion. In particular, this lower bound is coherent with the invariance of the Lagrangian through time inversion.)

Proof. (1) If particle i is not involved in a shock at time t, by the definition of a shock, it means that $C := C(\mathcal{Z}_t, i) \in \pi(\mathcal{Z}_s)$ for all s in a neighborhood of t. In particular, for all $j \in C$ and s sufficiently close to t, by (49),

$$(\overline{\nabla} f(\mathcal{Z}_s))_j = \frac{1}{\#C} \sum_{k \in C} a_k.$$

On the other hand, it is easy to find a neighborhood U of $(t, z_i(t))$ in $[0, T] \times \mathbb{R}$ such that, for all $j \in \{1, \dots, N\}$ and all $s \in [0, T]$, $(s, z_j(s)) \in U$ implies $j \in C$.

As a consequence, if $\xi : [0, T] \to \mathbb{R}$ is smooth and compactly supported in a sufficiently small neighborhood of t, and if ε is sufficiently small, by defining $\widetilde{\mathcal{Z}} = (\tilde{z}_1(s), \dots, \tilde{z}_N(s))$ for any $j \in \{1, \dots, N\}$ and $s \in [0, T]$ by

$$\tilde{z}_j(s) := \begin{cases} z_j(s) + \varepsilon\xi(s) & \text{if } j \in C, \\ z_j(s) & \text{else,} \end{cases}$$

then $\pi(\mathcal{Z})$ and $\pi(\widetilde{\mathcal{Z}})$ (and hence $\overline{\nabla} f(\mathcal{Z})$ and $\overline{\nabla} f(\widetilde{\mathcal{Z}})$) coincide at all time. The ODE follows from comparing the values of Λ' on $\mathcal{Z}$ and trajectories of type $\widetilde{\mathcal{Z}}$.

In particular, by boundedness of $\mathcal{Z}$, if particle i is not involved in a shock at time t, $|\ddot{z}_i|$ is bounded by a constant that is not depending on t. The existence of $\dot{z}_i(t-)$ and $\dot{z}_i(t+)$ at the times of isolated shocks follows easily.

(2) This is the heart of our study of the dynamical system, and maybe the less standard part of Section 4. But still the idea is very easy: with the notation of the statement, if $\dot{z}_i(t-) - \dot{z}_i(t+)$ is too small, then it is cheaper to stick particle i with other particles, as shown in Figure 3. The proof goes as follows.

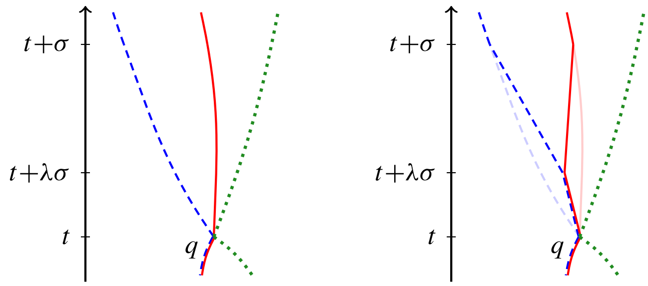

Figure 3. To the left, a piece of the trajectory $\mathcal{Z}$, and to the right, the competitor $\mathcal{Z}^{\sigma,\lambda}$ that we describe in the proof.

Step 1: Definition of a competitor. Let us consider (t, q, C) an isolated shock. Because it is isolated, we can find $\tau > 0$ such that the particles of C are not involved in another shock between times $t - \tau$ and $t + \tau$. By the definition of a shock, we cannot have $C \in \pi(\mathcal{Z}_s)$ for all $s \in (t - \tau, t + \tau)$, so either, for all $s \in (t - \tau, t)$, $C \notin \pi(\mathcal{Z}_s)$ or, for all $s \in (t, t + \tau)$, $C \notin \pi(\mathcal{Z}_s)$. Without loss of generality, we suppose that the second one holds: the particles of C are not all stuck right after the shock. Moreover, by our choice of τ, for all $C' \subset C$, the assertion $C' \in \pi(\mathcal{Z}_s)$ is either true or false independently of $s \in (t, t + \tau)$. Then, for $s \in (t, t + \tau)$, the following definitions of $C_1, C_2 \in \pi(\mathcal{Z}_s)$ do not depend on s:

$$C_1 := C(\mathcal{Z}_s, i) \quad \text{for } i = \min C \qquad \text{and} \qquad C_2 := C(\mathcal{Z}_s, i) \quad \text{for } i = \min C \backslash C_1.$$

(The classes C_1 and C_2 are the two leftmost packs of particles of C right after the shock.) Let us define for $j = 1, 2$

$$k_j := \#C_j, \qquad v_j := \dot{z}_i(t+) \quad \text{for } i \in C_j, \qquad \text{and} \qquad p := \frac{k_1 v_1 + k_2 v_2}{k_1 + k_2}. \tag{57}$$

For $0 \le \sigma < \tau$ and $\lambda \in [0, 1)$, we define a competitor $\mathcal{Z}^{\sigma,\lambda} = (z_1^{\sigma,\lambda}(s), \dots, z_N^{\sigma,\lambda}(s))$ by setting for all $i = \{1, \dots, N\}$ and $s \in [0, T]$

$$z_i^{\sigma,\lambda}(s) = \begin{cases} z_i(s) & \text{if } i \notin C_1 \cup C_2 \text{ or } s \notin (t, t+\sigma), \\ q + (s - t)p & \text{if } i \in C_1 \cup C_2 \text{ and } s \in (t, t + \lambda\sigma), \\ \dfrac{t+\sigma-s}{(1-\lambda)\sigma}(q + \lambda\sigma p) + \dfrac{s-(t+\lambda\sigma)}{(1-\lambda)\sigma} z_i(t+\sigma) & \text{if } i \in C_1 \cup C_2 \text{ and } s \in (t + \lambda\sigma, t + \sigma). \end{cases}$$

(See Figure 3 for an illustration of this competitor.) We will get a lower bound on $v_2 - v_1$ by comparing the value of Λ'' on $\mathcal{Z}$ and $\mathcal{Z}^{\sigma,\lambda}$, and by differentiating the corresponding inequality first with respect to σ at $\sigma = 0$ (we zoom so that the particles of $\mathcal{Z}$ only travel along straight lines), and then with respect to λ at $\lambda = 0$ (we compute the first variation of the action when we let the particles stick together).

Step 2: A lower bound on $v_2 - v_1$. The partitions $\pi(\mathcal{Z}_s^{\sigma,\lambda})$ and $\pi(\mathcal{Z}_s)$ coincide at all times except between t and $t + \lambda\sigma$, when $\pi(\mathcal{Z}_s)$ is a strict refinement of $\pi(\mathcal{Z}_s^{\sigma,\lambda})$. Hence, letting

$$\delta = \delta(N, A) := \min\{h(\pi) - h(\pi') : (\pi, \pi') \text{ ordered partition of } [\![1, N]\!],\ \pi \text{ strict refinement of } \pi'\} > 0,$$

we have, for all $s \in (t, t+\lambda\sigma)$,

$$h(\pi(\mathcal{Z}_s^{\lambda,\sigma})) + \delta \le h(\pi(\mathcal{Z}_s)). \tag{58}$$

As $\mathcal{Z}^\sigma$ coincide with $\mathcal{Z}$ for times outside $(t, t+\sigma)$ and for coordinates that are not in $C_1 \cup C_2$, by definition (53) of Λ'', we have

$$\begin{aligned}\Lambda''(\mathcal{Z}^{\sigma,\lambda}) - \Lambda''(\mathcal{Z}) &= \sum_{i\in C_1\cup C_2} \int_t^{t+\sigma} \{|\dot z_i^{\sigma,\lambda}(s)|^2 + |z_i^{\sigma,\lambda}(s) - a_i|^2 - |\dot z_i(s)|^2 - |z_i(s)-a_i|^2\}\,\mathrm{d}s \\ &\qquad + \int_t^{t+\lambda\sigma} \{h(\pi(\mathcal{Z}_s^{\sigma,\lambda})) - h(\pi(\mathcal{Z}_s))\}\,\mathrm{d}s \\ &\le \sum_{i\in C_1\cup C_2} \int_t^{t+\sigma} \{|\dot z_i^{\tau,\sigma}(s)|^2 - |\dot z_i(s)|^2\}\,\mathrm{d}s - \delta\lambda\sigma + \underset{\sigma\to 0}{o}(\sigma),\end{aligned} \tag{59}$$

where to obtain the second line, we used (58) and the fact that between times t and $t+\sigma$, both z_i and $z_i^{\sigma,\lambda}$ remain at a distance of order σ of q.

Let us consider $i \in C_j$ for $j = 1, 2$. On one hand, as z_i admits v_j as a right derivative at time t, we have

$$\int_t^{t+\sigma} |\dot z_i(s)|^2\,\mathrm{d}s = v_j^2\sigma + \underset{\sigma\to 0}{o}(\sigma). \tag{60}$$

On the other hand, we can compute explicitly

$$\begin{aligned}\int_t^{t+\sigma} |\dot z_i^{\sigma,\lambda}(s)|^2\,\mathrm{d}s &= \lambda p^2\sigma + (1-\lambda)\sigma\left(\frac{z_i(t+\sigma) - (q+\lambda p\sigma)}{(1-\lambda)\sigma}\right)^2 \\ &= \lambda p^2\sigma + \frac{1}{(1-\lambda)\sigma}\Big(q + v_j\sigma + \underset{\sigma\to 0}{o}(\sigma) - q - \lambda p\sigma\Big)^2 \\ &= \lambda p^2\sigma + (v_j - \lambda p)^2\frac{\sigma}{1-\lambda} + \underset{\sigma\to 0}{o}(\sigma).\end{aligned} \tag{61}$$

By plugging (60) and (61) into (59) and by using the definition (57) of k_1, k_2 and p, we get

$$\begin{aligned}\Lambda''(\mathcal{Z}^{\sigma,\lambda}) - \Lambda''(\mathcal{Z}) &\le \left\{(k_1+k_2)\lambda p^2 + \frac{k_1(v_1-\lambda p)^2 + k_2(v_2-\lambda p)^2}{1-\lambda} - k_1v_1^2 - k_2v_2^2 - \delta\lambda\right\}\sigma + \underset{\sigma\to 0}{o}(\sigma) \\ &= \{(k_1+k_2)p^2 + k_1v_1^2 + k_2v_2^2 - 2p(k_1v_1+k_2v_2) - \delta(1-\lambda)\}\frac{\lambda}{1-\lambda}\sigma + \underset{\sigma\to 0}{o}(\sigma) \\ &= \left\{k_1v_1^2 + k_2v_2^2 - \frac{(k_1v_1+k_2v_2)^2}{k_1+k_2} - \delta(1-\lambda)\right\}\frac{\lambda}{1-\lambda}\sigma + \underset{\sigma\to 0}{o}(\sigma) \\ &= \left\{\frac{k_1k_2}{k_1+k_2}(v_2-v_1)^2 - \delta(1-\lambda)\right\}\frac{\lambda}{1-\lambda}\sigma + \underset{\sigma\to 0}{o}(\sigma).\end{aligned}$$

By the minimality of $\Lambda''(\mathcal{Z})$, this quantity must be nonnegative. If we divide it by $\lambda\sigma$, and if we let σ and then λ go to zero, we end up with

$$\frac{k_1k_2}{k_1+k_2}(v_2-v_1)^2 \ge \delta. \tag{62}$$

Step 3: Conservation of momentum during an isolated shock and conclusion. Because (t, q, C) is isolated, it is easy to justify that we can replace V by the vector V^C whose j-th coordinate is 1 if $j \in C$ and 0 otherwise in the proof of Proposition 29. Doing so, we obtain the "local" conservation of momentum

$$\frac{1}{\#C}\sum_{i\in C}\dot z_i(t-) = \frac{1}{\#C}\sum_{i\in C}\dot z_i(t+) =: \mathcal{P}^C(t).$$

By ordering of the particles, we have, for $i = \min C$,

$$\dot z_i(t-) \geq \mathcal{P}^C(t) = \frac{1}{\#C}\sum_{i\in C}\dot z_i(t+) \geq \frac{k_1}{\#C}v_1 + \frac{\#C - k_1}{\#C}v_2.$$

(Indeed, $j \in C \mapsto \dot z_j(t-)$ and $j \in C \mapsto \dot z_j(t+)$ are clearly nonincreasing and nondecreasing respectively.) By recalling that $v_1 = \dot z_i(t+)$ and using (62), we get

$$\dot z_i(t-) - \dot z_i(t+) \geq \frac{\#C - k_1}{\#C}(v_2 - v_1) \geq \frac{\#C - k_1}{\#C}\sqrt{\frac{k_1+k_2}{k_1k_2}\delta}.$$

The minimal right-hand side's value is $\sqrt{\delta/(\#C^2 - \#C)}$, obtained for $k_1 = \#C - 1$ and $k_2 = 1$. Hence, we get the result by choosing $\alpha = \sqrt{\delta/(N^2 - N)}$. □

4.7. *Conclusion*: *proof of Theorem 21.* We are now ready to give the proof of Theorem 21. We give ourselves $\mathcal{Z}$ a global minimizer of Λ'. Thanks to Proposition 26, we can suppose that $\mathcal{Z}$ is ordered, and thanks to Proposition 27, we can consider Λ'' instead of Λ'.

Because of Proposition 31, it suffices to prove that there is a finite number of shocks. Indeed, in that case one can take for $0 = t_0 < t_1 < \cdots < t_p = T$ the moments of these shocks (and the endpoints of $[0, T]$). The smoothness of $\mathcal{Z}$ on each $[t_{i-1}, t_i]$, $i = 1, \dots, p$, follows directly from Proposition 31. Then $\pi(\mathcal{Z})$ is constant on each (t_{i-1}, t_i), $i = 1, \dots, p$, because by Definition 30 of a shock, at each time of discontinuity of $\pi(\mathcal{Z})$, there is at least one shock.

The set

$$\{(t', q') : \text{there exists } C' \subset [\![1, N]\!] \text{ such that } (t', q', C') \text{ is a shock}\}$$

is easily seen to be compact. So if it is not finite, it admits at least one accumulation point. That is, if there is an infinite number of shocks, then there is at least one shock which is not isolated. Let us consider such a shock (t, q, C) with minimal number of particles involved, i.e., with minimal $\#C$. The rest of the proof is dedicated to showing that the existence of (t, q, C) leads to a contradiction.

Step 1: The velocities are bounded. As $\mathcal{Z}$ is continuous on $[0, T]$, it is bounded. On the other hand, by definition, $h \leq |A|^2$. Now if $i \in \{1, \dots, N\}$ and $t \in [0, T]$ is such that $\mathcal{Z}$ is differentiable at t (which is true for almost any t), recalling the definition (54) of $\mathcal{E}$,

$$\dot z_i(t)^2 \leq |\dot{\mathcal{Z}}_t|^2 \leq 2\mathcal{E} + |\mathcal{Z}_t - A|^2 + h(\pi(\mathcal{Z}_t)),$$

which is bounded uniformly in t.

Step 2: All the shocks in the neighborhood of (t, q) that are distinct from (t, q) are isolated. Let U be a neighborhood of (t, q) in $[0, T] \times \mathbb{R}$ such that, for all $s \in [0, T]$ and $i \in \{1, \dots, N\}$, $(s, z_i(s)) \in U$ implies $i \in C$. This is possible since $\mathcal{Z}$ is continuous and, for all $j \notin C$, $z_j(t) \neq q$ by Definition 30 of a shock. Let us consider (t', q', C') a shock with $(t', q') \in U$. If $\#C' < \#C$, then (t', q', C') is isolated by the minimality of $\#C$. If $\#C' = \#C$, then $C' = C$ by the definition of U. But then it is easy to adapt the proof of point (1) of Proposition 22 to prove that $C \in \pi(\mathcal{Z}_s)$ for all s between t and t', and so there is no shock in U between t and t'. Hence there exists at most one such shock in U: either one before t or one after t, but not both because else (t, q, C) would contradict the third point of the definition of a shock. Up to reducing U, we can then exclude (t', q', C').

Step 3: Conclusion using Proposition 31. As (t, q, C) is not isolated, there is an infinite number of (isolated) shocks in U. Without loss of generality, we can assume that there is an infinite number of shocks in U after time t. Call $i \in C$ the smallest index such that particle i is involved in an infinite number of shocks in U after time t. When $i \neq \min C$, up to reducing U and by the minimality of i, we can assume that no particle $j \in C$ with $j < i$ is involved in a shock in U after time t.

As the shocks in U involving i after time t are isolated (Step 2), we can enumerate their times in decreasing order $(t_p)_{p\in\mathbb{N}}$. The boundedness of $\mathcal{Z}$ together with (55) allows us to take M as an upper bound for $\ddot{z}_i$ between the times of shocks. For all $p \in \mathbb{N}$ and $s \in (t_{p+1}, t_p)$, taking α as in (56), we have

$$\begin{aligned}\dot{z}_i(s) &= \dot{z}_i(t_0-) + \sum_{k=1}^{p}\left\{\dot{z}_i(t_k-) - \dot{z}_i(t_k+) - \int_{t_{k-1}}^{t_k} \ddot{z}_i(\tau)\,\mathrm{d}\tau\right\} - \int_{t_p}^{s} \ddot{z}_i(\tau)\,\mathrm{d}\tau \\ &\geq \dot{z}_i(t_0-) + p\alpha - M(t_0 - t),\end{aligned}$$

which contradicts Step 1 as soon as p is sufficiently large. □

Acknowledgment

Ambrosio acknowledges the support of the PRIN 2017 project "Gradient Flows, Optimal Transport and Metric Measure Structures".

References

[Ambrosio and Gangbo 2008] L. Ambrosio and W. Gangbo, "Hamiltonian ODEs in the Wasserstein space of probability measures", *Comm. Pure Appl. Math.* **61**:1 (2008), 18–53. MR Zbl

[Ambrosio et al. 2005] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Birkhäuser, Basel, 2005. MR Zbl

[Ambrosio et al. 2021] L. Ambrosio, A. Baradat, and Y. Brenier, "Γ-convergence for a class of action functionals induced by gradients of convex functions", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **32**:1 (2021), 97–108. MR Zbl

[Brenier 1987] Y. Brenier, "Décomposition polaire et réarrangement monotone des champs de vecteurs", *C. R. Acad. Sci. Paris Sér. I Math.* **305**:19 (1987), 805–808. MR Zbl

[Brenier 1991] Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions", *Comm. Pure Appl. Math.* **44**:4 (1991), 375–417. MR Zbl

[Brenier 2011] Y. Brenier, "A modified least action principle allowing mass concentrations for the early universe reconstruction problem", *Confluentes Math.* **3**:3 (2011), 361–385. MR Zbl

[Brenier 2016] Y. Brenier, "A double large deviation principle for Monge–Ampère gravitation", *Bull. Inst. Math. Acad. Sin. (N.S.)* **11**:1 (2016), 23–41. MR Zbl

[Brenier and Grenier 1998] Y. Brenier and E. Grenier, "Sticky particles and scalar conservation laws", *SIAM J. Numer. Anal.* **35**:6 (1998), 2317–2328. MR Zbl

[Brenier and Loeper 2004] Y. Brenier and G. Loeper, "A geometric approximation to the Euler equations: the Vlasov–Monge–Ampère system", *Geom. Funct. Anal.* **14**:6 (2004), 1182–1218. MR Zbl

[Brenier et al. 2003] Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, R. Mohayaee, and A. Sobolevskiĭ, “Reconstruction of the early universe as a convex optimization problem”, *Monthly Notices Roy. Astronom. Soc.* **346**:2 (2003), 501–524.

[Brenier et al. 2013] Y. Brenier, W. Gangbo, G. Savaré, and M. Westdickenberg, “Sticky particle dynamics with interactions”, *J. Math. Pures Appl.* (9) **99**:5 (2013), 577–617. MR Zbl

[Brézis 1973] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Math. Stud. **5**, North-Holland, Amsterdam, 1973. MR Zbl

[de Broglie 1927] L. de Broglie, “La mécanique ondulatoire et la structure atomique de la matière et du rayonnement”, *J. Phys. Radium* **8** (1927), 225–241. Zbl

[de Broglie 1959] L. de Broglie, “L’interprétation de la mécanique ondulatoire”, *J. Phys. Radium* (8) **20** (1959), 963–979. MR Zbl

[Conforti 2019] G. Conforti, “A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost”, *Probab. Theory Related Fields* **174**:1-2 (2019), 1–47. MR Zbl

[Dembo and Zeitouni 1998] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd ed., Appl. Math. (New York) **38**, Springer, 1998. MR Zbl

[E et al. 1996] W. E, Y. G. Rykov, and Y. G. Sinai, “Generalized variational principles, global weak solutions and behavior with random initial data for systems of conservation laws arising in adhesion particle dynamics”, *Comm. Math. Phys.* **177**:2 (1996), 349–380. MR Zbl

[Freidlin and Wentzell 1998] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, 2nd ed., Grundlehren der Math. Wissenschaften **260**, Springer, 1998. MR Zbl

[Frisch et al. 2002] U. Frisch, S. Matarrese, R. Mohayaee, and A. Sobolevski, “A reconstruction of the initial conditions of the universe by optimal mass transportation”, *Nature* **417** (2002), 260–262.

[Gentil et al. 2017] I. Gentil, C. Léonard, and L. Ripani, “About the analogy between optimal transport and minimal entropy”, *Ann. Fac. Sci. Toulouse Math.* (6) **26**:3 (2017), 569–601. MR Zbl

[Gigli and Tamanini 2020] N. Gigli and L. Tamanini, “Benamou–Brenier and duality formulas for the entropic cost on $\mathrm{RCD}^*(K, N)$ spaces”, *Probab. Theory Related Fields* **176**:1-2 (2020), 1–34. MR Zbl

[Hynd 2020] R. Hynd, “A trajectory map for the pressureless Euler equations”, *Trans. Amer. Math. Soc.* **373**:10 (2020), 6777–6815. MR Zbl

[Madelung 1927] E. Madelung, “Quantentheorie in hydrodynamischer Form”, *Z. Phys.* **40** (1927), 322–326. Zbl

[Mérigot 2011] Q. Mérigot, “A multiscale approach to optimal transport”, *Comput. Graphics Forum* **30**:5 (2011), 1583–1592.

[Monsaingeon et al. 2023] L. Monsaingeon, L. Tamanini, and D. Vorotnikov, “The dynamical Schrödinger problem in abstract metric spaces”, *Adv. Math.* **426** (2023), art. id. 109100. MR Zbl

[Natile and Savaré 2009] L. Natile and G. Savaré, “A Wasserstein approach to the one-dimensional sticky particle system”, *SIAM J. Math. Anal.* **41**:4 (2009), 1340–1365. MR Zbl

[von Renesse 2012] M.-K. von Renesse, “An optimal transport view of Schrödinger’s equation”, *Canad. Math. Bull.* **55**:4 (2012), 858–869. MR Zbl

[Strömberg 1996] T. Strömberg, *The operation of infimal convolution*, Dissertationes Math. **352**, Polish Acad. Sci., Warsaw, 1996. MR Zbl

[Villani 2003] C. Villani, *Topics in optimal transportation*, Grad. Stud. in Math. **58**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Zeldovich 1970] Y. B. Zeldovich, “Gravitational instability: an approximate theory for large density perturbations”, *Astron. Astrophys.* **5** (1970), 84–89.

Received 8 Sep 2020. Revised 17 Jan 2022. Accepted 9 Apr 2022.

LUIGI AMBROSIO: luigi.ambrosio@sns.it
Scuola Normale Superiore, Pisa, Italy

AYMERIC BARADAT: baradat@math.univ-lyon1.fr
Université Lyon, CNRS, Université Claude Bernard Lyon 1, UMR5208, Institut Camille Jordan, Villeurbanne, France

YANN BRENIER: brenier@dma.ens.fr
École Normale Supérieure, Paris, France

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2041–2077
DOI: 10.2140/apde.2023.16.2041

IDA AND HANKEL OPERATORS ON FOCK SPACES

ZHANGJIAN HU AND JANI A. VIRTANEN

We introduce a new space IDA of locally integrable functions whose integral distance to holomorphic functions is finite, and use it to completely characterize boundedness and compactness of Hankel operators on weighted Fock spaces. As an application, for bounded symbols, we show that the Hankel operator H_f is compact if and only if $H_{\bar{f}}$ is compact, which complements the classical compactness result of Berger and Coburn. Motivated by recent work of Bauer, Coburn, and Hagger, we also apply our results to the Berezin–Toeplitz quantization.

1. Introduction

Denote by L^2 the Hilbert space of all Gaussian square-integrable functions f on $\mathbb{C}^n$, that is,

$$\int_{\mathbb{C}^n} |f(z)|^2 e^{-|z|^2}\, dv(z) < \infty,$$

where v is the standard Lebesgue measure on $\mathbb{C}^n$. The Fock space F^2 (aka Segal–Bargmann space) consists of all holomorphic functions in L^2. The orthogonal projection of L^2 onto F^2 is denoted by P and called the Bergman projection. For a suitable function $f : \mathbb{C}^n \to \mathbb{C}$, the Hankel operator H_f and the Toeplitz operator T_f are defined on F^2 by

$$H_f = (I - P)M_f \quad \text{and} \quad T_f = PM_f.$$

The function f is referred to as the symbol of H_f and T_f. Since P is a bounded operator, it follows that both H_f and T_f are well-defined and bounded on F^2 if f is a bounded function. For unbounded symbols, despite considerable efforts, see, e.g., [Bauer 2005; Berger and Coburn 1994; Coburn et al. 2021; Hu and Wang 2018], characterization of boundedness or compactness of these operators has remained an open problem for more than 20 years.

In this paper, as a natural evolution from BMO (see [John and Nirenberg 1961; Zhu 2012]), we introduce a notion of integral distance to holomorphic (aka analytic) functions IDA and use it to completely characterize boundedness and compactness of Hankel operators on Fock spaces. Recently, in [Hu and Virtanen 2022], which continues our present work, we used IDA in the Hilbert space setting to characterize the Schatten class properties of Hankel operators. Indeed, the space IDA is broad in scope, and should have more applications, which we hope to demonstrate in future work in connection with Toeplitz operators.

All our results are proved for weighted Fock spaces $F^p(\varphi)$ consisting of holomorphic functions for which

$$\int_{\mathbb{C}^n} |f(z)|^p e^{-p\varphi(z)}\, dv(z) < \infty,$$

MSC2020: primary 47B35; secondary 32A25, 32A37, 81S10.
Keywords: Fock space, Hankel operator, boundedness, compactness, quantization, $\bar{\partial}$-equation.

where $0 < p < \infty$ and φ is a suitable weight function (see Section 2 for further details). Obviously, with $p = 2$ and $\varphi(z) = (\alpha/2)|z|^2$, we obtain the weighted Fock space F^2_α. The study of L^p-type Fock spaces was initiated in [Janson et al. 1987] and has since grown considerably, as seen in [Zhu 2012].

We also revisit and complement a surprising result due to [Berger and Coburn 1987], which states that for bounded symbols

$$H_f : F^2 \to L^2 \text{ is compact if and only if } H_{\bar{f}} \text{ is compact.}$$

In particular, we give a new proof and show that this phenomenon remains true for Hankel operators from $F^p(\varphi)$ to $L^q(\varphi)$ for general weights. What also makes this result striking is that it is not true for Hankel operators acting on other important function spaces, such as Hardy or Bergman spaces.

As an application, we will apply our results to the Berezin–Toeplitz quantization, which complements the results in [Bauer et al. 2018].

1A. ***Main results.*** We introduce the following new function spaces to characterize bounded and compact Hankel operators. Let $0 < s \le \infty$ and $0 < q < \infty$. For $f \in L^q_{\text{loc}}$, set

$$(G_{q,r}(f)(z))^q = \inf_{h \in H(B(z,r))} \frac{1}{|B(z,r)|} \int_{B(z,r)} |f - h|^q \, dv, \quad z \in \mathbb{C}^n,$$

where $H(B(z,r))$ stands for the set of holomorphic functions in the ball $B(z,r)$. We say that $f \in L^q_{\text{loc}}$ is in $\text{IDA}^{s,q}$ if

$$\|f\|_{\text{IDA}^{s,q}} = \|G_{q,1}(f)\|_{L^s} < \infty.$$

We further write BDA^q for $\text{IDA}^{\infty,q}$ and say that $f \in \text{VDA}^q$ if

$$\lim_{z \to \infty} G_{q,1}(f)(z) = 0.$$

The properties of these spaces will be studied in Section 3.

We denote by $\mathcal{S}$ the set of all measurable functions f that satisfy the condition in (2-7), which ensures that the Hankel operator H_f is densely defined on $F^p(\varphi)$ provided that $0 < p < \infty$ and φ is a suitable weight. Notice that the symbol class $\mathcal{S}$ contains all bounded functions. Further, we write $\text{Hess}_{\mathbb{R}}\,\varphi$ for the Hessian of φ and E for the $2n \times 2n$ identity matrix — these concepts will be discussed in more detail in Section 2. It is important to notice that the condition $\text{Hess}_{\mathbb{R}}\,\varphi \simeq \text{E}$ in the following theorems is satisfied by the classical Fock space F^2, the Fock spaces F^2_α generated by *standard weights* $\varphi(z) = (\alpha/2)|z|^2$, $\alpha > 0$, Fock–Sobolev spaces, and a large class of nonradial weights.

Theorem 1.1. *Let $f \in \mathcal{S}$ and suppose that* $\text{Hess}_{\mathbb{R}}\,\varphi \simeq \text{E}$ *as in* (2-1).

(a) *For $0 < p \le q < \infty$ and $q \ge 1$, $H_f : F^p(\varphi) \to L^q(\varphi)$ is bounded if and only if $f \in \text{BDA}^q$, and H_f is compact if and only if $f \in \text{VDA}^q$. For the operator norm of H_f, we have the estimate*

$$\|H_f\| \simeq \|f\|_{\text{BDA}^q}. \tag{1-1}$$

(b) *For* $1 \le q < p < \infty$, $H_f : F^p(\varphi) \to L^q(\varphi)$ *is bounded if and only if it is compact, which is equivalent to* $f \in \mathrm{IDA}^{s,q}$, *where* $s = pq/(p-q)$, *and*

$$\|H_f\| \simeq \|f\|_{\mathrm{IDA}^{s,q}}. \tag{1-2}$$

(c) *For* $0 < p \le q \le 1$ *and* $f \in L^\infty$, $H_f : F^p(\varphi) \to L^q(\varphi)$ *is bounded with*

$$\|H_f\| \le C\|f\|_{L^\infty} \tag{1-3}$$

and compact if and only if $f \in \mathrm{VDA}^q$.

We first note that Theorem 1.1 is new even for Hankel operators acting from F^2 to L^2. Previously only characterizations for H_f and $H_{\bar{f}}$ to be *simultaneously* bounded (or *simultaneously* compact) were known. These were given in terms of the bounded (or vanishing) mean oscillation of f in [Bauer 2005] for F^2 and in [Hu and Wang 2018] for Hankel operators from F^p_α to L^q_α. In Theorem 7.1 of Section 7, we obtain these results as a simple consequence of Theorem 1.1. We also mention our recent work [Hu and Virtanen 2022], which gives a complete characterization of Schatten class Hankel operators.

Theorem 1.1 should also be compared with the results for Hankel operators on Bergman spaces A^p. Indeed, characterizations for boundedness and compactness can be found in [Axler 1986] for antianalytic symbols, in [Hagger and Virtanen 2021] for bounded symbols, and in [Hu and Lu 2019; Li 1994; Luecking 1992; Pau et al. 2016] for unbounded symbols. These two cases are different to study because of properties such as $F^p \subset F^q$ for $p \le q$ (as opposed to $A^q \subset A^p$) and certain nice geometry on the boundary of these bounded domains, which in turn helps with the treatment of the $\bar{\partial}$-problem.

What is very different about the results on Hankel operators acting on these two types of spaces is that our next result is only true in Fock spaces (see [Hagger and Virtanen 2021] for an interesting counterexample for the Bergman space).

Theorem 1.2. *Let* $f \in L^\infty$ *and suppose that* $\mathrm{Hess}_{\mathbb{R}}\,\varphi \simeq \mathrm{E}$ *as in* (2-1). *If* $0 < p \le q < \infty$ *or* $1 \le q < p < \infty$, *then* $H_f : F^p(\varphi) \to L^q(\varphi)$ *is compact if and only if* $H_{\bar{f}}$ *is compact.*

For Hankel operators on the Fock space F^2, Theorem 1.2 was proved in [Berger and Coburn 1987] using C^*-algebra and Hilbert space techniques and in [Stroethoff 1992] using elementary methods. More recently in [Hagger and Virtanen 2021], limit operator techniques were used to treat the reflexive Fock spaces F^p_α. However, our result is new even in the Hilbert space case because of the more general weights that we consider. As a natural continuation of our present work, in [Hu and Virtanen 2022], we prove that, for $f \in L^\infty$, the Hankel operator H_f is in the Schatten class S_p if and only if $H_{\bar{f}}$ is in the Schatten class S_p provided that $1 < p < \infty$.

As an application and further generalization of our results, in Section 6, we provide a complete characterization of those $f \in L^\infty$ for which

$$\lim_{t\to 0} \|T^{(t)}_f T^{(t)}_g - T^{(t)}_{fg}\|_t = 0 \tag{1-4}$$

for all $g \in L^\infty$, where $T^{(t)}_f = P^{(t)} M_f : F^2_t(\varphi) \to F^2_t(\varphi)$ and $P^{(t)}$ is the orthogonal projection of $L^2_t(\varphi)$ onto $F^2_t(\varphi)$. Here $L^2_t = L^2(\mathbb{C}^n, d\mu_t)$ and

$$d\mu_t(z) = \frac{1}{t^n}\exp\left\{-2\varphi\left(\frac{z}{\sqrt{t}}\right)\right\} dv(z).$$

The importance of the semiclassical limit in (1-4) stems from the fact that it is one of the essential ingredients of the deformation quantization of [Rieffel 1989; 1990] in mathematical physics. Our conclusion related to (1-4) extends and complements the main result in [Bauer et al. 2018].

1B. ***Approach.*** A careful inspection shows that the methods and techniques used in [Berger and Coburn 1986; 1987; Hagger and Virtanen 2021; Perälä et al. 2014; Stroethoff 1992] depend heavily upon the following three aspects. First, the explicit representation of the Bergman kernel $K(z, w)$ for standard weights $\varphi(z) = (\alpha/2)|z|^2$ has the property that

$$K(z, w)e^{-(\alpha/2)|z|^2-(\alpha/2)|w|^2} = e^{(\alpha/2)|z-w|^2}. \tag{1-5}$$

However, for the class of weights we consider, this quadratic decay is known not to hold (even in dimension $n = 1$) and is expected to be very rare [Christ 1991]. The second aspect involves the Weyl unitary operator W_a defined as

$$W_a f = f \circ \tau_a k_a,$$

where τ_a is the translation by a and k_a is the normalized reproducing kernel. As a unitary operator on F^p_α (or on L^p_α), W_a plays a very important role in the theory of the Fock spaces F^p_α (see [Zhu 2012]). Unfortunately, no analogue of Weyl operators is currently available for $F^p(\varphi)$ when $\varphi \neq (\alpha/2)|w|^2$. The third aspect we mention is Banach (or Hilbert) space techniques, such as the adjoint (for example, H_f^*) and the duality. However, when $0 < p < 1$, $F^p(\varphi)$ is only an F-space (in the sense of [Rudin 1973]) and the usual Banach space techniques can no longer be applied.

To overcome the three difficulties mentioned above, we introduce function spaces IDA, BDA and VDA, and develop their theory, which we use to characterize those symbols f such that H_f are bounded (or compact) from $F^p(\varphi)$ to $L^q(\varphi)$. Our characterization of the boundedness of H_f extends the main results of [Bauer 2005; Hu and Wang 2018; Perälä et al. 2014]. It is also worth noting that as a natural generalization of BMO, the space IDA will have its own interest and will likely be useful to study other (related) operators (such as Toeplitz operators).

In our analysis, we appeal to the $\bar\partial$-techniques several times. As the canonical solution to $\bar\partial u = g\partial f$, $H_f g$ is naturally connected with the $\bar\partial$-theory. Hörmander's theory provides us with the L^2-estimate, but less is known about L^p-estimates on $\mathbb{C}^n$ when $p \neq 2$. With the help of a certain auxiliary integral operator, we obtain L^p-estimates of the Berndtsson–Anderson solution [1982] to the $\bar\partial$-equation. Our approach to handling weights whose curvature is uniformly comparable to the Euclidean metric form is similar to the treatment in [Schuster and Varolin 2012] which was initiated in [Berndtsson and Ortega Cerdà 1995], and a number of the techniques we use here were inspired by this approach. Although the work in [Berndtsson and Ortega Cerdà 1995] is restricted to $n = 1$, some of the results were extended to higher dimensions in [Lindholm 2001], and the others are easy to modify.

The outline of the paper is as follows. In Section 2 we study preliminary results on the Bergman kernel which are needed throughout the paper, and we also establish estimates for the $\bar\partial$-solution developed in [Berndtsson and Andersson 1982]. In Section 3, a notion of function spaces $\mathrm{IDA}^{s,q}$ is introduced. We obtain a useful decomposition for functions in $\mathrm{IDA}^{s,q}$ (compare with the decompositions of BMO and VMO). Using this decomposition, we obtain the completeness of $\mathrm{IDA}^{s,q}/H(\mathbb{C}^n)$ in $\|\cdot\|_{\mathrm{IDA}^{s,q}}$. In Sections 4 and 5 we prove Theorems 1.1 and 1.2, respectively. For the latter theorem, we also appeal to the Calderón–Zygmund theory of singular integrals, and in particular employ the Ahlfors–Beurling operator to obtain certain estimates on ∂- and $\bar{\partial}$-derivatives. In Section 6, we present an application of our results to quantization. In the last section, we give further remarks together with two conjectures.

Throughout the paper, C stands for positive constants which may change from line to line, but does not depend on functions being considered. Two quantities A and B are called equivalent, denoted by $A \simeq B$, if there exists some C such that $C^{-1}A \le B \le CA$.

2. Preliminaries

Let $\mathbb{C}^n = \mathbb{R}^{2n}$ be the n-dimensional complex Euclidean space and denote by v the Lebesgue measure on $\mathbb{C}^n$. For $z = (z_1, \dots, z_n)$ and $w = (w_1, \dots, w_n)$ in $\mathbb{C}^n$, we write $z \cdot \bar{w} = z_1\bar{w}_1 + \cdots + z_n\bar{w}_n$ and $|z| = \sqrt{z \cdot \bar{z}}$. Let $H(\mathbb{C}^n)$ be the family of all holomorphic functions on $\mathbb{C}^n$. Given a domain Ω in $\mathbb{C}^n$ and a positive Borel measure μ on Ω, we denote by $L^p(\Omega, d\mu)$ the space of all Lebesgue measurable functions f on Ω for which

$$\|f\|_{L^p(\Omega,d\mu)} = \left\{\int_\Omega |f|^p \, d\mu\right\}^{1/p} < \infty \quad \text{for } 0 < p < \infty$$

and $\|f\|_{L^\infty(\Omega,dv)} = \operatorname{ess\,sup}_{z\in\Omega} |f(z)| < \infty$ for $p = \infty$. For ease of notation, we simply write L^p for the space $L^p(\mathbb{C}^n, dv)$.

2A. ***Weighted Fock spaces.*** For a real-valued weight $\varphi \in C^2(\mathbb{C}^n)$ and $0 < p < \infty$, denote by $L^p(\varphi)$ the space $L^p(\mathbb{C}^n, e^{-p\varphi}dv)$ with norm $\|\cdot\|_{p,\varphi} = \|\cdot\|_{L^p(\mathbb{C}^n,e^{-p\varphi}dv)}$. Then the Fock space $F^p(\varphi)$ is defined as

$$\begin{aligned} F^p(\varphi) &= L^p(\varphi) \cap H(\mathbb{C}^n), \\ F^\infty(\varphi) &= \left\{f \in H(\mathbb{C}^n) : \|f\|_{\infty,\varphi} = \sup\nolimits_{z\in\mathbb{C}^n} |f(z)|e^{-\varphi(z)} < \infty\right\}. \end{aligned}$$

For $1 \le p \le \infty$, $F^p(\varphi)$ is a Banach space in the norm $\|\cdot\|_{p,\varphi}$ and $F^2(\varphi)$ is a Hilbert space. For $0 < p < 1$, $F^p(\varphi)$ is an F-space with metric given by $d(f, g) = \|f - g\|_{p,\varphi}^p$.

Other related and widely studied holomorphic function spaces include the Bergman spaces $A^p_\alpha(\mathbb{B}^n)$ of the unit ball $\mathbb{B}^n$ consisting of all holomorphic functions f in $L^p(\mathbb{B}^n, dv_\alpha)$, where $0 < p < \infty$, $dv_\alpha(z) = (1 - |z|^2)^\alpha \, dv(z)$ and $\alpha > -1$.

In this paper we are interested in Fock spaces $F^p(\varphi)$ with certain uniformly convex weights φ. More precisely, suppose $\varphi = \varphi(x_1, x_2, \dots, x_{2n}) \in C^2(\mathbb{R}^{2n})$ is real-valued, and there are positive constants m and M such that $\mathrm{Hess}_\mathbb{R}\, \varphi$, the real Hessian, satisfies

$$m\mathrm{E} \le \mathrm{Hess}_\mathbb{R}\, \varphi(x) = \left(\frac{\partial^2\varphi(x)}{\partial x_j \partial x_k}\right)_{j,k=1}^{2n} \le M\mathrm{E}, \tag{2-1}$$

where E is the $2n \times 2n$ identity matrix; above, for symmetric matrices A and B, we used the convention that $A \le B$ if $B - A$ is positive semidefinite. When (2-1) is satisfied, we write $\mathrm{Hess}_\mathbb{R}\, \varphi \simeq \mathrm{E}$. A typical model of such weights is given by $\varphi(z) = (\alpha/2)|z|^2$ for $z = (z_1, z_2, \dots, z_n)$ with $z_j = x_{2j-1} + \mathrm{i}x_{2j}$, which induces the weighted Fock space F^p_α studied by many authors (see, e.g., [Zhu 2012]). Another popular example is $\varphi(z) = |z|^2 - \frac{1}{2}\log(1+|z|^2)$, which gives the so-called Fock–Sobolev spaces studied for example in [Cho and Zhu 2012]. Notice that the weights φ satisfying (2-1) are not only radial functions, as the example $\varphi(z) = |z|^2 + \sin[(z_1 + \bar{z}_1)/2]$ clearly shows.

For $x = (x_1, x_2, \dots, x_{2n})$, $t = (t_1, t_2, \dots, t_{2n}) \in \mathbb{R}^{2n}$, write $z_j = x_{2j-1} + \mathrm{i}x_{2j}$, $\xi_j = t_{2j-1} + \mathrm{i}t_{2j}$ and $\xi = (\xi_1, \xi_2, \dots, \xi_n)$. An elementary calculation similar to that on page 125 of [Krantz 1992] shows

$$\operatorname{Re}\sum_{j,k=1}^{n} \frac{\partial^2\varphi}{\partial z_j \partial z_k}(z)\xi_j\xi_k + \sum_{j,k=1}^{n} \frac{\partial^2\varphi}{\partial z_j \partial \bar{z}_k}(z)\xi_j\bar{\xi}_k = \frac{1}{2}\sum_{j,k=1}^{2n} \frac{\partial^2\varphi}{\partial x_j \partial x_k}(x)t_jt_k \ge \frac{1}{2}m|\xi|^2.$$

Replacing ξ with $\mathrm{i}\xi$ in the above inequality gives

$$-\operatorname{Re}\sum_{j,k=1}^{n} \frac{\partial^2\varphi}{\partial z_j \partial z_k}(z)\xi_j\xi_k + \sum_{j,k=1}^{n} \frac{\partial^2\varphi}{\partial z_j \partial \bar{z}_k}(z)\xi_j\bar{\xi}_k \ge \frac{1}{2}m|\xi|^2.$$

Thus,

$$\sum_{j,k=1}^{n} \frac{\partial^2\varphi}{\partial z_j \partial \bar{z}_k}(z)\xi_j\bar{\xi}_k \ge \frac{1}{2}m|\xi|^2.$$

Similarly, we have an upper bound for the complex Hessian of φ. Therefore, $m\omega_0 \le dd^c\varphi \le M\omega_0$, where $\omega_0 = dd^c|z|^2$ is the Euclidean Kähler form on $\mathbb{C}^n$ and $d^c = \frac{1}{4}\sqrt{-1}(\bar{\partial} - \partial)$. This implies that the theory in [Schuster and Varolin 2012; Hu and Lv 2014] is applicable in the present setting.

For $z \in \mathbb{C}^n$ and $r > 0$, let $B(z,r) = \{w \in \mathbb{C}^n : |w - z| < r\}$ be the ball with center at z with radius r. For the proof of the following weighted Bergman inequality, we refer to Proposition 2.3 of [Schuster and Varolin 2012].

Lemma 2.1. *Suppose $0 < p \le \infty$. For each $r > 0$ there is some $C > 0$ such that if $f \in F^p(\varphi)$ then*

$$|f(z)e^{-\varphi(z)}|^p \le C\int_{B(z,r)} |f(\xi)e^{-\varphi(\xi)}|^p \, dv(\xi).$$

It follows from the preceding lemma that $\|f\|_{q,\varphi} \le C\|f\|_{p,\varphi}$ and

$$F^p(\varphi) \subseteq F^q(\varphi) \quad \text{for } 0 < p \le q \le \infty. \tag{2-2}$$

This inclusion is completely different from that of the Bergman spaces.

Lemma 2.2. *There exist positive constants θ and C_1, depending only on n, m and M, such that*

$$|K(z,w)| \le C_1 e^{\varphi(z)+\varphi(w)} e^{-\theta|z-w|} \quad \text{for all } z, w \in \mathbb{C}^n, \tag{2-3}$$

and there exist positive constants C_2 and r_0 such that

$$|K(z,w)| \ge C_2 e^{\varphi(z)+\varphi(w)} \tag{2-4}$$

for $z \in \mathbb{C}^n$ and $w \in B(z, r_0)$.

The estimate (2-3) appeared in [Christ 1991] for $n = 1$ and in [Delin 1998] for $n \ge 2$, while the inequality (2-4) can be found in [Schuster and Varolin 2012].

For $z \in \mathbb{C}^n$, write

$$k_z(\cdot) = \frac{K(\cdot, z)}{\sqrt{K(z,z)}}$$

for the normalized Bergman kernel. Then Lemma 2.2 implies

$$\frac{1}{C}e^{\varphi(z)} \le \|K(\cdot, z)\|_{p,\varphi} \le Ce^{\varphi(z)} \quad \text{and} \quad \frac{1}{C} \le \|k_z\|_{p,\varphi} \le C \quad \text{for } z \in \mathbb{C}^n, \tag{2-5}$$

and $\lim_{|z|\to\infty} k_z(\xi) = 0$ uniformly in ξ on compact subsets of $\mathbb{C}^n$.

2B. *The Bergman projection.* For Fock spaces, we denote by P the orthogonal projection of $L^2(\varphi)$ onto $F^2(\varphi)$, and refer to it as the Bergman projection. It is well known that P can be represented as an integral operator

$$Pf(z) = \int_{\mathbb{C}^n} K(z, w) f(w) e^{-2\varphi(w)}\, dv(w) \tag{2-6}$$

for $z \in \mathbb{C}^n$, where $K(\cdot, \cdot)$ is the Bergman (reproducing) kernel of $F^2(\varphi)$.

As a consequence of Lemma 2.2, it follows that the Bergman projection P is bounded on $L^p(\varphi)$ for $1 \le p \le \infty$, and $P|_{F^p(\varphi)} = \mathrm{I}$ for $0 < p \le \infty$; for further details, see Proposition 3.4 and Corollary 3.7 of [Schuster and Varolin 2012].

2C. *Hankel operators.* To define Hankel operators with unbounded symbols, consider

$$\Gamma = \left\{\textstyle\sum_{j=1}^N a_j K(\cdot, z_j) : N \in \mathbb{N},\ a_j \in \mathbb{C},\ z_j \in \mathbb{C}^n \text{ for } 1 \le j \le N\right\},$$

and the symbol class

$$\mathcal{S} = \{f \text{ measurable on } \mathbb{C}^n : fg \in L^1(\varphi) \text{ for } g \in \Gamma\}. \tag{2-7}$$

Given $f \in \mathcal{S}$, the Hankel operator $H_f = (\mathrm{I} - P)M_f$ with symbol f is well-defined on Γ. According to Proposition 2.5 of [Hu and Virtanen 2020], for $0 < p < \infty$, the set Γ is dense in $F^p(\varphi)$, and hence the Hankel operator H_f is densely defined on $F^p(\varphi)$.

2D. *Lattices in $\mathbb{C}^n$.* Given $r > 0$, a sequence $\{a_k\}_{k=1}^\infty$ in $\mathbb{C}^n$ is called an r-lattice if the balls $\{B(a_k, r)\}_{k=1}^\infty$ cover $\mathbb{C}^n$ and $\{B(a_k, r/(2\sqrt{n}))\}_{k=1}^\infty$ are pairwise disjoint. A typical model of an r-lattice is the sequence

$$\left\{\frac{r}{\sqrt{n}}(m_1 + k_1\mathrm{i}, m_2 + k_2\mathrm{i}, \dots, m_n + k_n\mathrm{i}) \in \mathbb{C}^n : m_j, k_j \in \mathbb{Z},\ j = 1, 2, \dots, n\right\}. \tag{2-8}$$

Notice that there exists an integer N depending only on the dimension of $\mathbb{C}^n$ such that, for any r-lattice $\{a_k\}_{k=1}^\infty$,

$$1 \le \sum_{k=1}^\infty \chi_{B(a_k, 2r)}(z) \le N \tag{2-9}$$

for $z \in \mathbb{C}^n$, where χ_E is the characteristic function of $E \subset \mathbb{C}^n$. These well-known facts are explained in [Zhu 2012] when $n = 1$ and they can be easily generalized to any $n \in \mathbb{N}$.

2E. *Fock Carleson measures.* In the theory of Bergman spaces, Carleson measures provide an essential tool for treating various problems, especially in connection with bounded operators, functions of bounded mean oscillation, and their applications; see, e.g., [Zhu 2005]. In Fock spaces, Carleson measures play a similar role; see [Zhu 2012] for the Fock spaces F^p_α. Carleson measures for Fock–Sobolev spaces were described in [Cho and Zhu 2012]. In [Schuster and Varolin 2012], Carleson measures for generalized Fock spaces (which include the weights considered in the present work) were used to study bounded and compact Toeplitz operators. Finally, their generalization to (p, q)-Fock Carleson measures was carried out in [Hu and Lv 2014], which is indispensable to the study of operators between distinct Banach spaces and will be applied to analyze Hankel operators acting from $F^p(\varphi)$ to $L^q(\varphi)$ in our work.

We recall the basic theory of these measures. Let $0 < p, q < \infty$ and let $\mu \geq 0$ be a positive Borel measure on $\mathbb{C}^n$. We call μ a (p, q)-*Fock Carleson measure* if the embedding $\mathrm{I} : F^p(\varphi) \to L^q(\mathbb{C}^n, e^{-q\varphi} d\mu)$ is bounded. Further, the measure μ is referred to as a vanishing (p, q)-Fock Carleson measure if in addition

$$\lim_{j\to\infty} \int_{\mathbb{C}^n} |f_j(z)e^{-\varphi(z)}|^q \, d\mu(z) = 0$$

whenever $\{f_j\}_{j=1}^{\infty}$ is bounded in $F^p(\varphi)$ and converges to 0 uniformly on any compact subset of $\mathbb{C}^n$ as $j \to \infty$. Fock Carleson measures were completely characterized in [Hu and Lv 2014] and we only add the following simple result, which is trivial for Banach spaces and can be easily proved in the other cases.

Proposition 2.3. *Let $0 < p, q < \infty$ and μ be a positive Borel measure on $\mathbb{C}^n$. Then μ is a vanishing (p, q)-Fock Carleson measure if and only if the inclusion map I is compact from $F^p(\varphi) \to L^q(\mathbb{C}^n, d\mu)$.*

Proof. It is not difficult to show that the image of the unit ball of $F^p(\varphi)$ under the inclusion is relatively compact in $L^q(\mathbb{C}^n, e^{q\varphi} d\mu)$. We leave out the details. □

2F. *Differential forms and an auxiliary integral operator.* As in [Krantz 1992], given two nonnegative integers $s, t \leq n$, we write

$$\omega = \sum_{|\alpha|=s,|\beta|=t} \omega_{\alpha,\beta} \, dz^{\alpha} \wedge d\bar{z}^{\beta} \tag{2-10}$$

for a differential form of type (s, t). We denote by $L_{s,t}$ the family of all (s, t)-forms ω as in (2-10) with coefficients $\omega_{\alpha,\beta}$ measurable on $\mathbb{C}^n$ and set

$$|\omega| = \sum_{|\alpha|=s,|\beta|=t} |\omega_{\alpha,\beta}| \quad \text{and} \quad \|\omega\|_{p,\varphi} = \| |\omega| \|_{p,\varphi}. \tag{2-11}$$

Given a weight function φ satisfying (2-1), we define an integral operator A_φ as

$$A_\varphi(\omega)(z) = \int_{\mathbb{C}^n} e^{\langle 2\partial\varphi, z-\xi\rangle} \sum_{j<n} \omega(\xi) \wedge \frac{\partial|\xi - z|^2 \wedge (2\bar{\partial}\partial\varphi(\xi))^j \wedge (\bar{\partial}\partial|\xi - z|^2)^{n-1-j}}{j!\,|\xi - z|^{2n-2j}} \tag{2-12}$$

for $\omega \in L_{0,1}$, where

$$\langle \partial\varphi(\xi), z - \xi\rangle = \sum_{j=1}^{n} \frac{\partial\varphi}{\partial\xi_j}(\xi)(z_j - \xi_j)$$

as denoted on page 92 in [Berndtsson and Andersson 1982].

For an (s_1, t_1)-form ω_A and an (s_2, t_2)-form ω_B with $s_1 + s_2 \leq n,\ t_1 + t_2 \leq n$, it is easy to verify that $|\omega_A \wedge \omega_B| \leq |\omega_A| \, |\omega_B|$. Therefore, for the (n, n)-form inside the integral of the right-hand side of (2-12), we obtain

$$\left|\omega(\xi)\wedge\frac{\partial|\xi-z|^2\wedge(2\bar\partial\partial\varphi)^j\wedge(\bar\partial\partial|\xi-z|^2)^{n-1-j}}{j!\,|\xi-z|^{2n-2j}}\right|\le C\frac{|\omega(\xi)|}{|\xi-z|^{2n-2j-1}}$$

because $\mathrm{i}\partial\bar\partial\,\varphi(\xi)\simeq\mathrm{i}\partial\bar\partial\,|\xi|^2$.

Recall that

$$\Gamma=\left\{\textstyle\sum_{j=1}^N a_jK_{z_j}:N\in\mathbb{N},\ a_j\in\mathbb{C},\ z_j\in\mathbb{C}^n\text{ for }1\le j\le N\right\}$$

is dense in $F^p(\varphi)$ for all $0<p<\infty$.

Lemma 2.4. *Suppose* $1\le p\le\infty$.

(I) *There is a constant C such that* $\|A_\varphi(\omega)\|_{p,\varphi}\le C\|\omega\|_{p,\varphi}$ *for* $\omega\in L_{0,1}$.

(II) *For $g\in\Gamma$ and $f\in C^2(\mathbb{C}^n)$ satisfying $|\bar\partial f|\in L^p$, it holds that $\bar\partial A_\varphi(g\bar\partial f)=g\bar\partial f$.*

Proof. Let $z\in\mathbb{C}^n$. By (2-1), using Taylor expansion of φ at ξ, we get

$$\varphi(z)-\varphi(\xi)\ge 2\,\mathrm{Re}\sum_{j=1}^n\frac{\partial\varphi(\xi)}{\partial\xi_j}(z_j-\xi_j)+m|z-\xi|^2.$$

Then (2-12) gives

$$|A_\varphi(\omega)(z)e^{-\varphi(z)}|\le C\int_{\mathbb{C}^n}|\omega(\xi)|e^{-\varphi(\xi)}\left\{\frac{1}{|\xi-z|}+\frac{1}{|\xi-z|^{2n-1}}\right\}e^{-m|\xi-z|^2}\,dv(\xi).\tag{2-13}$$

For $l<2n$ fixed, define another integral operator $\mathcal{A}_l$ as

$$\mathcal{A}_l:h\mapsto\int_{\mathbb{C}^n}h(\xi)\frac{e^{-m|\xi-z|^2}}{|\xi-z|^l}\,dv(\xi).$$

It is easy to verify, by interpolation, that $\mathcal{A}_l$ is bounded on L^p for $1\le p\le\infty$. Therefore,

$$\begin{aligned}\|A_\varphi(\omega)\|_{p,\varphi}&\le C\|(\mathcal{A}_1+\mathcal{A}_{2n-1})(|\omega|e^{-\varphi})\|_{L^p}\\&\le C(\|\mathcal{A}_1\|_{L^p\to L^p}+\mathcal{A}_{2n-1}\|_{L^p\to L^p})\|\omega\|_{p,\varphi},\end{aligned}$$

which completes the proof of part (A).

Notice that the convexity assumption in (2-1) yields $dd^c\varphi\simeq\omega_0$, which in turn means that $|\partial\bar\partial\varphi(\xi)|\simeq 1$. We use p' to denote the conjugate of p, $1/p+1/p'=1$. Now, for $f\in C^2(\mathbb{C}^n)$ satisfying $|\bar\partial f|\in L^p$, and $z,z_0\in\mathbb{C}^n$, we have

$$\begin{aligned}&\int_{\mathbb{C}^n}|K(\xi,z_0)\,\bar\partial f(\xi)|\sum_{j=0}^{n-1}\frac{e^{-\varphi(\xi)}|\bar\partial\partial\varphi(\xi)|^j}{|\xi-z|^{2n-2j-1}}\,dv(\xi)\\&\qquad\le C\left\{\sup_{\xi\in B(z,1)}|K(\xi,z_0)\,\bar\partial f(\xi)e^{-\varphi(\xi)}|+\int_{\mathbb{C}^n\setminus B(z,1)}|K(\xi,z_0)\,\bar\partial f(\xi)|e^{-\varphi(\xi)}\,dv(\xi)\right\}\\&\qquad\le Ce^{\varphi(z_0)}\Big\{\sup_{\xi\in B(z,1)}|\bar\partial f(\xi)|+\|\bar\partial f\|_{L^p}\|K(\,\cdot\,,z_0)\|_{p',\varphi}\Big\}<\infty.\end{aligned}$$

Hence, for $g \in \Gamma$ and $z \in \mathbb{C}^n$, it holds that

$$\int_{\mathbb{C}^n} |g(\xi)\bar\partial f(\xi)| \sum_{j=0}^{n-1} \frac{e^{-\varphi(\xi)}|\bar\partial\partial\varphi(\xi)|^j}{|\xi - z|^{2n-2j-1}}\, dv(\xi) < \infty.$$

From Proposition 10 of [Berndtsson and Andersson 1982], we get (B) (pay attention to the mistake in the last line of that result where f is left out on the right-hand side). □

Corollary 2.5. *Suppose $f \in \mathcal{S} \cap C^1(\mathbb{C}^n)$ and $|\bar\partial f| \in L^s$ with some $1 \le s \le \infty$. For $g \in \Gamma$, it holds that*

$$H_f(g) = A_\varphi(g\bar\partial f) - P(A_\varphi(g\bar\partial f)). \tag{2-14}$$

Proof. Given $f \in \mathcal{S} \cap C^1(\mathbb{C}^n)$ with $|\bar\partial f| \in L^s$ and $g \in \Gamma$, we have $\|g\bar\partial f\|_{1,\varphi} \le \|g\|_{s',\varphi}\|\bar\partial f\|_{L^s} < \infty$, where s' is the conjugate of s. Lemma 2.4 implies that $u = A_\varphi(g\bar\partial f) \in L^1(\varphi)$ and $\bar\partial u = g\bar\partial f$. Then $fg - u \in L^1(\varphi)$. Notice that $\bar\partial(fg-u) = g\bar\partial f - \bar\partial u = 0$, and so $fg - u \in F^1(\varphi)$. Since $P|_{F^1_\varphi} = \mathrm{I}$, we have

$$fg - u = P(fg-u) = P(fg) - P(u).$$

This shows that $H_f(g) = u - P(u)$. □

3. The space IDA

We now introduce a new space to characterize boundedness and compactness of Hankel operators. The space IDA is related to the space of bounded mean oscillation BMO (see, e.g., [John and Nirenberg 1961; Zhu 2012]), which has played an important role in many branches of analysis and their applications for decades. We find that IDA is also broad in scope and should have more applications in operator theory and related areas.

3A. ***Definitions and preliminary lemmas.*** Let $0 < q < \infty$ and $r > 0$. For $f \in L^q_{\mathrm{loc}}$ (the collection of q-th locally Lebesgue integrable functions on $\mathbb{C}^n$), following [Luecking 1992], we define $G_{q,r}(f)$ as

$$G_{q,r}(f)(z) = \inf\left\{\left(\frac{1}{|B(z,r)|}\int_{B(z,r)} |f-h|^q\, dv\right)^{1/q} : h \in H(B(z,r))\right\} \tag{3-1}$$

for $z \in \mathbb{C}^n$.

Definition 3.1. Suppose $0 < s \le \infty$ and $0 < q < \infty$. The space $\mathrm{IDA}^{s,q}$ (integral distance to holomorphic functions) consists of all $f \in L^q_{\mathrm{loc}}$ such that

$$\|f\|_{\mathrm{IDA}^{s,q}} = \|G_{q,1}(f)\|_{L^s} < \infty.$$

The space $\mathrm{IDA}^{\infty,q}$ is also denoted by BDA^q. The space VDA^q consists of all $f \in \mathrm{BDA}^q$ such that

$$\lim_{z\to\infty} G_{q,1}(f)(z) = 0.$$

We will see in Section 6 that $\mathrm{IDA}^{s,q}$ is an extension of the space $\mathrm{IMO}^{s,q}$ introduced in [Hu and Wang 2018].

Notice that the space BDA^2 was first introduced in the context of the Bergman spaces of the unit disk in [Luecking 1992], where it is called the space of functions with bounded distance to *analytic* functions (BDA).

Remark 3.2. As is the case with the classical BMO^q and VMO^q spaces, we have

$$\mathrm{BDA}^{q_2} \subset \mathrm{BDA}^{q_1} \quad \text{and} \quad \mathrm{VDA}^{q_2} \subset \mathrm{VDA}^{q_1}$$

properly for $0 < q_1 < q_2 < \infty$.

Let $0 < q < \infty$. For $z \in \mathbb{C}^n$, $f \in L^q(B(z,r), dv)$ and $r > 0$, we define the q-th mean of $|f|$ over $B(z,r)$ by setting

$$M_{q,r}(f)(z) = \left(\frac{1}{|B(z,r)|} \int_{B(z,r)} |f|^q \, dv \right)^{1/q}.$$

For $\omega \in L_{0,1}$, we set $M_{q,r}(\omega)(z) = M_{q,r}(|\omega|)(z)$.

Lemma 3.3. *Suppose $0 < q < \infty$. Then for $f \in L^q_{\mathrm{loc}}$, $z \in \mathbb{C}^n$ and $r > 0$, there is some $h \in H(B(z,r))$ such that*

$$M_{q,r}(f-h)(z) = G_{q,r}(f)(z) \tag{3-2}$$

and

$$\sup_{w \in B(z,r/2)} |h(w)| \le C \|f\|_{L^q(B(z,r),dv)}, \tag{3-3}$$

where the constant C is independent of f and r.

Proof. Let $f \in L^q_{\mathrm{loc}}$, $z \in \mathbb{C}^n$ and $r > 0$. Taking $h = 0$ in the integrand of (3-1), we get

$$G_{q,r}(f)(z) \le M_{q,r}(f)(z) < \infty.$$

Then for $j = 1, 2, \ldots$, we can pick $h_j \in H(B(z,r))$ such that

$$M_{q,r}(f-h_j)(z) \to G_{q,r}(f)(z) \tag{3-4}$$

as $j \to \infty$. Hence, for j sufficiently large,

$$M_{q,r}(h_j)(z) \le C\{M_{q,r}(f-h_j)(z) + M_{q,r}(f)(z)\} \le C M_{q,r}(f)(z). \tag{3-5}$$

This shows that $\{h_j\}_{j=1}^\infty$ is a normal family. Thus, we can find a subsequence $\{h_{j_k}\}_{k=1}^\infty$ and a function $h \in H(B(z,r))$ so that $\lim_{k\to\infty} h_{j_k}(w) \to h(w)$ for $w \in B(z,r)$. By (3-4), applying Fatou's lemma, we have

$$G_{q,r}(f)(z) \le M_{q,r}(f-h)(z) \le \liminf_{k\to\infty} M_{q,r}(f-h_{j_k})(z) = G_{q,r}(f)(z),$$

which proves (3-2). It remains to note that, with the plurisubharmonicity of $|h|^q$, for $w \in B(z,r/2)$, we have

$$|h(w)| \le M_{q,r/2}(h)(w) \le C M_{q,r}(h)(z) \le C M_{q,r}(f)(z),$$

which completes the proof. □

Corollary 3.4. *For $0 < s < r$, there is a constant $C > 0$ such that for $f \in L^q_{\mathrm{loc}}$ and $w \in B(z, r-s)$, it holds that*

$$G_{q,s}(f)(w) \le M_{q,s}(f-h)(w) \le C G_{q,r}(f)(z), \tag{3-6}$$

where h is as in Lemma 3.3.

Proof. For $0 < s < r$ and $w \in B(z, r - s)$, we have $B(w, s) \subset B(z, r)$. Then, the first estimate in (3-6) comes from the definition of $G_{q,s}(f)$, while (3-2) yields

$$M_{q,s}(f - h)(w) \le C M_{q,r}(f - h)(z) = C G_{q,r}(f)(z),$$

which completes the proof. □

For $z \in \mathbb{C}^n$ and $r > 0$, let

$$A^q(B(z, r), dv) = L^q(B(z, r), dv) \cap H(B(z, r))$$

be the q-th Bergman space over $B(z, r)$. Denote by $P_{z,r}$ the corresponding Bergman projection induced by the Bergman kernel for $A^2(B(z, r), dv)$. It is well known that $P_{z,r}(f)$ is well-defined for $f \in L^1(B(z, r), dv)$.

Lemma 3.5. *Suppose* $1 \le q < \infty$ *and* $0 < s < r$. *There is a constant* $C > 0$ *such that, for* $f \in L^q_{\text{loc}}$ *and* $w \in B(z, r - s/(2))$,

$$G_{q,s}(f)(w) \le M_{q,s}(f - P_{z,r}(f))(w) \le C G_{q,r}(f)(z) \quad \textit{for } z \in \mathbb{C}^n. \tag{3-7}$$

Proof. We only need to prove the second inequality. Suppose $1 < q < \infty$. Notice that $P_{0,1}$ is the standard Bergman projection on the unit ball of $\mathbb{C}^n$. Theorem 2.11 of [Zhu 2005] implies that

$$\|P_{0,1}\|_{L^q(B(0,1),dv) \to A^q(B(0,1),dv)} < \infty.$$

Now for $r > 0$ fixed and $f \in L^q((B(0, r), dv)$, set $f_r(w) = f(rw)$. Then

$$\|f_r\|_{L^q(B(0,1),dv)} = r^{-2n/q} \|f\|_{L^q(B(0,1),dv)}.$$

Furthermore, it is easy to verify that the operator $f \mapsto P_{0,1}(f_r)(\cdot/r)$ is self-adjoint and idempotent, and it maps $L^2((B(0, r), dv)$ onto $A^2((B(0, r), dv)$. Therefore,

$$P_{0,r}(f)(z) = P_{0,1}(f_r)\left(\frac{z}{r}\right) \quad \text{for } f \in L^q(B(0, r), dv),$$

and hence

$$\|P_{0,r}\|_{L^q(B(0,r),dv) \to A^q(B(0,r),dv)} = \|P_{0,1}\|_{L^q(B(0,1),dv) \to A^q(B(0,1),dv)}.$$

Now for $z \in \mathbb{C}^n$ and $r > 0$, using a suitable dilation, it follows that

$$\|P_{z,r}\|_{L^q(B(z,r),dv) \to A^q(B(z,r),dv)} = \|P_{0,1}\|_{L^q(B(0,1),dv) \to A^q(B(0,1),dv)} < \infty. \tag{3-8}$$

Unfortunately, $P_{z,r}$ is not bounded on $L^1(B(z, r), dv)$, but with the same approach as above, by Fubini's theorem and Theorem 1.12 of [Zhu 2005], we have

$$\|P_{z,r}\|_{L^1(B(z,r),dv) \to A^1(B(z,r),(r^2 - |\cdot - z|^2)dv)} \le C \tag{3-9}$$

for $z \in \mathbb{C}^n$ and $r > 0$.

Choose h as in Lemma 3.3. Then $h \in A^q(B(z,r), dv)$ because $f \in L^q_{\mathrm{loc}}$. Thus, $P_{z,r}(h) = h$. Now for $w \in B(z, (r-s)/2)$ and $1 \le q < \infty$,

$$\begin{aligned}
\Big\{\int_{B(w,s)} &|f - P_{z,r}(f)|^q \, dv\Big\}^{1/q} \\
&\le C\Big\{\int_{B(z,(r+s)/2)} |f - P_{z,r}(f)|^q \, dv\Big\}^{1/q} \\
&\le C\Big\{\int_{B(z,r)} |f(\xi) - P_{z,r}(f)(\xi)|^q (r^2 - |\xi - z|^2) \, dv(\xi)\Big\}^{1/q} \\
&\le C\Big\{\Big[\int_{B(z,r)} |f - h|^q \, dv\Big]^{1/q} + \Big[\int_{B(z,r)} |P_{z,r}(f-h)(\xi)|^q (r^2 - |\xi - z|^2) \, dv(\xi)\Big]^{1/q}\Big\} \\
&\le C\Big\{\int_{B(z,r)} |f - h|^q \, dv\Big\}^{1/q}. \qquad (3\text{-}10)
\end{aligned}$$

From this and Lemma 3.3, (3-7) follows. □

Given $t > 0$, let $\{a_j\}_{j=1}^\infty$ be a $(t/2)$-lattice, set $J_z = \{j : z \in B(a_j, t)\}$ and denote by $|J_z|$ the cardinal number of J_z. By (2-9), $|J_z| = \sum_{j=1}^\infty \chi_{B(a_j,t)}(z) \le N$. Choose a partition of unity $\{\psi_j\}_{j=1}^\infty$, $\psi_j \in C^\infty(\mathbb{C}^n)$, subordinate to $\{B(a_j, t/2)\}$ such that

$$\begin{gathered}
\operatorname{supp} \psi_j \subset B(a_j, t/2), \quad \psi_j(z) \ge 0, \quad \sum_{j=1}^\infty \psi_j(z) = 1, \\
|\bar\partial \psi_j(z)| \le C t^{-1}, \quad \sum_{j=1}^\infty \bar\partial \psi_j(z) = 0.
\end{gathered} \qquad (3\text{-}11)$$

Given $f \in L^q_{\mathrm{loc}}$, for $j = 1, 2, \dots$, pick $h_j \in H(B(a_j, t))$ as in Lemma 3.3 so that

$$M_{q,t}(f - h_j)(a_j) = G_{q,t}(f)(a_j).$$

Define

$$f_1 = \sum_{j=1}^\infty h_j \psi_j \quad \text{and} \quad f_2 = f - f_1. \qquad (3\text{-}12)$$

Notice that $f_1(z)$ is a finite sum for every $z \in \mathbb{C}^n$ and hence well-defined because we have $\operatorname{supp} \psi_j \subset B(a_j, t/2) \subset B(a_j, t)$.

Inspired by a similar treatment on pages 254–255 of [Luecking 1992], using the partition of unity, we can prove the following estimate.

Lemma 3.6. *Suppose* $0 < q < \infty$. *For* $f \in L^q_{\mathrm{loc}}$ *and* $t > 0$, *decomposing* $f = f_1 + f_2$ *as in* (3-12), *we have* $f_1 \in C^2(\mathbb{C}^n)$ *and*

$$|\bar\partial f_1(z)| + M_{q,t/2}(\bar\partial f_1)(z) + M_{q,t/2}(f_2)(z) \le C G_{q,2t}(f)(z) \qquad (3\text{-}13)$$

for $z \in \mathbb{C}^n$, *where the constant C is independent of f.*

Proof. Observe first that $f_1 \in C^2(\mathbb{C}^n)$ follows directly from the properties of the functions h_j and ψ_j. For $z \in \mathbb{C}^n$, we may assume $z \in B(a_1, t/2)$ without loss of generality. Then for those j that satisfy $\bar\partial \psi_j(z) \ne 0$,

$|h_j - h_1|^q$ is plurisubharmonic on $B(z, t/2) \subset B(a_j, t)$. Hence, by Corollary 3.4,

$$\begin{aligned}
|\bar\partial f_1(z)| = \left|\sum_{j=1}^{\infty}(h_j(z) - h_1(z))\bar\partial\psi_j(z)\right| &\le \sum_{j=1}^{\infty} |h_j(z) - h_1(w)|\,|\bar\partial\psi_j(z)| \\
&\le C \sum_{\{j:\,|a_j - z| < t/2\}} M_{q,t/4}(h_j - h_1)(z) \\
&\le C \sum_{\{j:\,|a_j - z| < t/2\}} [M_{q,t/4}(f - h_j)(z) + M_{q,t/4}(f - h_1)(w)] \\
&\le C \sum_{\{j:\,|a_j - z| < t/2\}} G_{q,t}(f)(a_j).
\end{aligned}$$

Thus, using Corollary 3.4 again, we get

$$|\bar\partial f_1(z)| \le C G_{q,3t/2}(f)(z) \quad \text{for } z \in \mathbb{C}^n,$$

and so,

$$M_{q,t/2}(\bar\partial f_1)(z)^q \le C \frac{1}{|B(z, t/2)|} \int_{B(z,t/2)} G_{q,3t/2}(f)(w)^q\, dw \le C G_{q,2t}(f)(z)^q.$$

Similarly, we have $|f_2(\xi)|^q \le C \sum_{j=1}^{\infty} |f(\xi) - h_j(\xi)|^q \psi_j(\xi)^q$, and so

$$M_{q,t/2}(f_2)(z)^q \le C \sum_{j=1}^{\infty} \frac{1}{|B(z, t/2)|} \int_{B(z,t/2)} |f - h_j|^q \psi_j^q\, dv \le C \sum_{\{j:\,|a_j - z| < t/2\}} G_{q,t}(f)(a_j)^q.$$

Therefore,

$$M_{q,t/2}(f_2)(z) \le C G_{q,3t/2}(f)(z).$$

Combining this and the other two estimates above gives (3-13). □

Given $\{\psi_j\}$ as in (3-11), we have another decomposition $f = \mathfrak{F}_1 + \mathfrak{F}_2$, where

$$\mathfrak{F}_1 = \sum_{j=1}^{\infty} P_{a_j,t}(f)\psi_j \quad \text{and} \quad \mathfrak{F}_2 = f - \mathfrak{F}_1. \tag{3-14}$$

When $q = 2$, the two decompositions coincide.

Corollary 3.7. *Suppose* $1 \le q < \infty$. *For* $f \in L^q_{\text{loc}}$ *and* $t > 0$, *we have* $\mathfrak{F}_1 \in C^2(\mathbb{C}^n)$ *and*

$$|\bar\partial\mathfrak{F}_1(z)| + M_{q,t/2}(\bar\partial\mathfrak{F}_1)(z) + M_{q,t/2}(\mathfrak{F}_2)(z) \le C G_{q,2t}(f)(z) \tag{3-15}$$

for $z \in \mathbb{C}^n$, *where the constant* C *is independent of* f.

Proof. The proof can be carried out as that of Lemma 3.6 using (3-7) instead of (3-6). We omit the details. □

3B. *The decomposition.* In our analysis, we will appeal to $\bar\partial$-techniques several times. Let $\Omega \subset \mathbb{C}^n$ be strongly pseudoconvex with C^4 boundary, and let S be a $\bar\partial$-closed (0, 1) form on Ω with L^p coefficients,

$1 \le p \le \infty$. As in [Krantz 1992], we denote by $\mathrm{H}_\Omega(S)$ the Henkin solution of $\bar\partial$-equation $\bar\partial u = S$ on Ω. We observe that Theorem 10.3.9 of that work implies that, for $1 \le q < \infty$,

$$\|\mathrm{H}_\Omega(S)\|_{L^q(\Omega,dv)} \le C\|S\|_{L^q(\Omega,dv)}, \tag{3-16}$$

where the constant C is independent of S and of "small" perturbations of the boundary. (We note that the second item in Theorem 10.3.9 of [Krantz 1992] is stated incorrectly and should read $\|u\|_{L^q} \le C_p\|f\|_p$ instead.) Indeed, to deduce (3-16), we consider three cases. First, for $1 \le q < (2n+2)/(2n+1)$,

$$\|\mathrm{H}_\Omega(S)\|_{L^q(\Omega,dv)} \le C\|S\|_{L^1(\Omega,dv)} \le C\|S\|_{L^q(\Omega,dv)}.$$

For $q = (2n+2)/(2n+1)$, take $1 < p = q < 2n+2$ and $q_1 = (2n+2)/(2n) > q$. Then $1/q_1 = 1/p - 1/(2n+2)$, and by the second item in Theorem 10.3.9 of [Krantz 1992], we have

$$\|\mathrm{H}_\Omega(S)\|_{L^q(\Omega,dv)} \le C\|\mathrm{H}_\Omega(S)\|_{L^{q_1}(\Omega,dv)} \le C\|S\|_{L^p(\Omega,dv)}.$$

Finally, for $q > (2n+2)/(2n+1)$, choose p so that $1/q = 1/p - 1/(2n+2)$. Then $1 < p < 2n+2$ and $p < q$. Now Theorem 10.3.9 of [Krantz 1992] implies

$$\|\mathrm{H}_\Omega(S)\|_{L^q(\Omega,dv)} \le C\|S\|_{L^p(\Omega,dv)} \le C\|S\|_{L^q(\Omega,dv)}.$$

Theorem 3.8. *Suppose* $1 \le q < \infty$, $0 < s < \infty$, *and* $f \in L^q_{\mathrm{loc}}$. *Then* $f \in \mathrm{IDA}^{s,q}$ *if and only if* f *admits a decomposition* $f = f_1 + f_2$ *such that*

$$f_1 \in C^2(\mathbb{C}^n), \quad M_{q,r}(\bar\partial f_1) + M_{q,r}(f_2) \in L^s \tag{3-17}$$

for some (*or any*) $r > 0$. *Furthermore, for fixed* $\tau, r > 0$, *it holds that*

$$\|f\|_{\mathrm{IDA}^{s,q}} \simeq \|G_{q,\tau}(f)\|_{L^s} \simeq \inf\{\|M_{q,r}(\bar\partial f_1)\|_{L^s} + \|M_{q,r}(f_2)\|_{L^s}\}, \tag{3-18}$$

where the infimum is taken over all possible decompositions $f = f_1 + f_2$ *that satisfy* (3-17) *with a fixed* r.

Proof. First, given $0 < r < R < \infty$, we have some $\boldsymbol{a}_1, \boldsymbol{a}_1, \dots, \boldsymbol{a}_m \in \mathbb{C}^n$ so that $B(0,R) \subset \bigcup_{j=1}^m B(\boldsymbol{a}_j, r)$. Then, for $g \in L^q_{\mathrm{loc}}$,

$$M_{q,R}(g)(z)^s \le C\sum_{j=1}^m M_{q,r}(g)(z+\boldsymbol{a}_j)^s, \quad z \in \mathbb{C}^n,$$

and

$$\int_{\mathbb{C}^n} M_{q,R}(g)(z)^s\,dv(z) \le C\sum_{j=1}^m \int_{\mathbb{C}^n} M_{q,r}(g)(z+\boldsymbol{a}_j)^s\,dv(z) \le C\int_{\mathbb{C}^n} M_{q,r}(g)(z)^s\,dv(z). \tag{3-19}$$

This implies that (3-17) holds for some r if and only if it holds for any r.

Suppose that $f \in L^q_{\mathrm{loc}}$ with $\|G_{q,\tau}(f)\|_{L^s} < \infty$ for some $\tau > 0$ and decompose $f = f_1 + f_2$ as in Lemma 3.6 with $t = \tau/2$. Then $f_1 \in C^2(\mathbb{C}^n)$ and

$$|\bar\partial f_1(z)| + M_{q,\tau/4}(\bar\partial f_1)(z) + M_{q,\tau/4}(f_2)(z) \le CG_{q,\tau}(f)(z).$$

Now for any $r > 0$, we have

$$\|M_{q,r}(\bar{\partial} f_1)\|_{L^s} + \|M_{q,r}(f_2)\|_{L^s} \le C\|G_{q,\tau}(f)\|_{L^s}. \tag{3-20}$$

This implies that, $f = f_1 + f_2$ satisfies (3-17).

Conversely, suppose $f = f_1 + f_2$ with $f_1 \in C^2(\mathbb{C}^n)$ and $M_{q,r}(\bar{\partial} f_1) + M_{q,r}(f_2) \in L^s$ for some $r > 0$ as in Theorem 3.8. Then, for any $\tau > 0$,

$$\|G_{q,\tau}(f_2)\|_{L^s} \le C\|M_{q,\tau}(f_2)\|_{L^s} \le C\|M_{q,r}(f_2)\|_{L^s}. \tag{3-21}$$

So $f_2 \in \mathrm{IDA}^{s,q}$. To consider f_1, we write $u = \mathrm{H}_{B(z,2\tau)}(\bar{\partial} f_1)$ for the Henkin solution of the equation $\bar{\partial} u = \bar{\partial} f_1$ on $B(z, 2\tau)$. From (3-16) and (3-17), u satisfies

$$M_{q,2\tau}(u)(z) \le C M_{q,2\tau}(\bar{\partial} f_1)(z) \quad \text{for } z \in \mathbb{C}^n, \tag{3-22}$$

which implies that $u \in L^q(B(z, 2\tau), dv)$. Similarly to (3-10),

$$M_{q,\tau}(\mathrm{P}_{z,2\tau}(u))(z) \le C M_{q,2\tau}(u)(z).$$

Thus,

$$\begin{aligned} M_{q,\tau}(u - \mathrm{P}_{z,2\tau}(u))(z) &\le M_{q,\tau}(u)(z) + M_{q,\tau}(\mathrm{P}_{z,2\tau}(u))(z) \\ &\le C M_{q,2\tau}(u)(z). \end{aligned} \tag{3-23}$$

Since

$$f_1 - u \in L^q(B(z, 2\tau), dv) \quad \text{and} \quad \bar{\partial}(f_1 - u) = 0,$$

we have

$$f_1 - u \in A^q(B(z, 2\tau), dv).$$

Notice that $P_{z,2\tau}|_{A^q(B(z,2\tau),dv)} = \mathrm{I}$, and so

$$f_1(\xi) - \mathrm{P}_{z,2\tau}(f_1)(\xi) = u(\xi) - \mathrm{P}_{z,2\tau}(u)(\xi) \quad \text{for } \xi \in B(z, 2\tau). \tag{3-24}$$

Combining (3-22), (3-23) and (3-24), we get

$$\begin{aligned} M_{q,\tau}(f_1 - \mathrm{P}_{z,2\tau}(f_1))(z) &= M_{q,\tau}(u - \mathrm{P}_{z,2\tau}(u))(z) \\ &\le M_{q,2\tau}(u)(z) \le C M_{q,2\tau}(\bar{\partial} f_1)(z). \end{aligned}$$

Therefore, by (3-19),

$$\begin{aligned} \|G_{q,\tau}(f_1)\|_{L^s} &\le \|M_{q,r}(f_1 - \mathrm{P}_{z,2\tau}(f_1))\|_{L^s} \\ &\le C\|M_{q,2\tau}(\bar{\partial} f_1)\|_{L^s} \le C\|M_{q,r}(\bar{\partial} f_1)\|_{L^s}. \end{aligned}$$

This and (3-21) yield

$$\|G_{q,\tau}(f)\|_{L^s} \le C\{\|M_{q,r}(\bar{\partial} f_1)\|_{L^s} + \|M_{q,r}(f_2)\|_{L^s}\}. \tag{3-25}$$

Thus, $f = f_1 + f_2 \in \mathrm{IDA}^{s,q}$.

It remains to note that the norm equivalence (3-18) follows from (3-20) and (3-25). □

With a similar proof we have the following corollary.

Corollary 3.9. *Suppose* $1 \le q < \infty$, *and* $f \in L^q_{\mathrm{loc}}$. *Then* $f \in \mathrm{BDA}^q$ (*or* VDA^q) *if and only if* $f = f_1 + f_2$, *where*

$$f_1 \in C^2(\mathbb{C}^n), \quad \bar\partial f_1 \in L^\infty_{0,1} \quad (\textit{or } \lim_{z\to\infty} |\bar\partial f_1| = 0) \tag{3-26}$$

and

$$M_{q,r}(f_2) \in L^\infty \quad (\textit{or } \lim_{z\to\infty} M_{q,r}(f_2) = 0) \tag{3-27}$$

for some (*or any*) $r > 0$. *Furthermore*,

$$\|f\|_{\mathrm{BDA}^q} \simeq \inf\{\|\bar\partial f_1\|_{L^\infty_{0,1}} + \|M_{q,r}(f_2)\|_{L^\infty}\},$$

where the infimum is taken over all possible decompositions $f = f_1 + f_2$, *with* f_1 *and* f_2 *satisfying the conditions in* (3-26) *and* (3-27).

Corollary 3.10. *Suppose* $1 \le q < \infty$. *Different values of* r *give equivalent seminorms* $\|G_{q,r}(\,\cdot\,)\|_{L^s}$ *on* $\mathrm{IDA}^{s,q}$ *when* $0 < s < \infty$ *and on both* BDA^q *and* VDA^q *when* $s = \infty$.

Remark 3.11. Recall that each f in BMO^q can be decomposed as $f = f_1 + f_2$, where f_1 is of bounded oscillation BO and f_2 has a bounded average BA^q (see [Zhu 2012] for the one-dimensional case and [Lv 2019] for the general case). Furthermore, we may choose f_1 to be a Lipschitz function in $C^2(\mathbb{C}^n)$ (see Corollary 3.37 of [Zhu 2012]); that is, $f \in \mathrm{BMO}^q$ if and only if $f = f_1 + f_2$ with all $\partial f_1/\partial x_j \in L^\infty$ for $j = 1, 2, \dots, 2n$ and $f_2 \in \mathrm{BA}^q$, or in the language of complex analysis both $\bar\partial f_1$ and $\bar\partial \bar f_1$ are bounded. Therefore, $f \in \mathrm{BMO}^q$ if and only if $f, \bar f \in \mathrm{BDA}^q$. For a similar relationship between IMO^q and the IDA spaces, see Lemma 6.1 of [Hu and Virtanen 2022] and Theorem 7.1 below.

3C. *IDA as a Banach space.*

We next prove that $\mathrm{IDA}^{s,q}/H(\mathbb{C}^n)$ with $1 \le s, q < \infty$ is a Banach space when equipped with the induced norm

$$\|f + H(\mathbb{C}^n)\| = \|f\|_{\mathrm{IDA}^{s,q}} \tag{3-28}$$

for $f \in \mathrm{IDA}^{s,q}$.

Theorem 3.12. *For* $1 \le s, q < \infty$, *the quotient space* $\mathrm{IDA}^{s,q}/H(\mathbb{C}^n)$ *is a Banach space with the norm induced by* $\|\cdot\|_{\mathrm{IDA}^{s,q}}$.

Proof. Obviously $H(\mathbb{C}^n) \subset \mathrm{IDA}^{s,q}$. Now given $f \in \mathrm{IDA}^{s,q}$ and $h \in H(\mathbb{C}^n)$, we have $G_{q,r}(f) = G_{q,r}(f+h)$. This means that the norm in (3-28) is well-defined on $\mathrm{IDA}^{s,q}/H(\mathbb{C}^n)$. If $\|f\|_{\mathrm{IDA}^{s,q}} = 0$, then $G_{q,r}(f)(z) = 0$ in $\mathbb{C}^n$. By Lemma 3.3, $f \in H(B(z,r))$ and hence $f \in H(\mathbb{C}^n)$.

Let $f_1, f_2 \in \mathrm{IDA}^{s,q}$ and $z \in \mathbb{C}^n$. According to Lemma 3.3, there are functions h_j holomorphic in $B(z,r)$ such that

$$M_{q,r}(f_j - h_j)(z) = G_{q,r}(f_j)(z) \quad \text{for } j = 1, 2.$$

Then, since

$$M_{q,r}((f_1 + f_2) - (h_1 + h_2))(z) \le M_{q,r}(f_1 - h_1)(z) + M_{q,r}(f_2 - h_2)(z),$$

we have

$$G_{q,r}(f_1 + f_2)(z) \le G_{q,r}(f_1)(z) + G_{q,r}(f_2)(z) \quad \text{for } z \in \mathbb{C}^n.$$

Hence, $\|f_1+f_2\|_{\mathrm{IDA}^{s,q}} \le \|f_1\|_{\mathrm{IDA}^{s,q}} + \|f_2\|_{\mathrm{IDA}^{s,q}}$. In addition, $\|f\|_{\mathrm{IDA}^{s,q}} \ge 0$ and $\|af\|_{\mathrm{IDA}^{s,q}} = |a|\|f\|_{\mathrm{IDA}^{s,q}}$ for $a \in \mathbb{C}$. Therefore, $\|\cdot\|_{\mathrm{IDA}^{s,q}}$ induces a norm on $\mathrm{IDA}^{s,q}/H(\mathbb{C}^n)$.

It remains to prove that the norm is complete. Suppose that $\{f_m\}_{m=1}^\infty$ is a Cauchy sequence in

$$\|\cdot\|_{\mathrm{IDA}^{s,q}} = \|G_{q,1}(\cdot)\|_{L^s}.$$

According to Corollary 3.10, we may assume that $\{f_m\}_{m=1}^\infty$ is a Cauchy sequence in $\|G_{q,r}(\cdot)\|_{L^s}$ with $r>0$ fixed. We now embark on proving that, for some $f \in \mathrm{IDA}^{s,q}$, $\lim_{m\to\infty} \|G_{q,r/2}(f_m - f)\|_{L^s} = 0$, which implies $\{f_m\}_{m=1}^\infty$ converges to some $f \in \mathrm{IDA}^{s,q}$ in the $\|\cdot\|_{\mathrm{IDA}^{s,q}}$-topology. For this purpose, let $\{a_j\}_{j=1}^\infty$ be some $t=(r/4)$-lattice. We decompose each f_m similarly to (3-14) as

$$f_{m,1} = \sum_{j=1}^\infty P_{a_j,r}(f_m)\psi_j \quad \text{and} \quad f_{m,2} = f_m - f_{m,1},$$

where $\{\psi_j\}_{j=1}^\infty$ is the partition of unity subordinate to $\{B(a_j,r/4)\}_{j=1}^\infty$ as in (3-11). It follows from Corollary 3.7 that

$$\begin{aligned} M_{q,r/8}(f_{m,2}-f_{k,2})(z)^s &= M_{q,r/8}\biggl((f_m - f_k) - \sum_{j=1}^\infty P_{a_j,t}(f_m - f_k)\psi_j\biggr)(z)^s \\ &\le C G_{q,r/2}(f_m - f_k)(z)^s \\ &\le C\int_{B(z,r/2)} G_{q,r}(f_m - f_k)(\xi)^s\, dv(\xi). \end{aligned}$$

This implies that $\{f_{m,2}\}_{j=1}^\infty$ converges to some function f_2 in the L^q_{loc}-topology. In addition, by Lemma 3.5, we have

$$M_{q,r/2}(f_{m,2} - f_{k,2} - \mathrm{P}_{z,r}(f_{m,2} - f_{k,2}))(z) \le C G_{q,r}(f_{m,2} - f_{k,2})(z).$$

Letting $k \to \infty$ and applying Fatou's lemma, we get

$$\begin{aligned} G_{q,r/2}(f_{m,2} - f_2)(z)^s &\le M_{q,r/2}(f_{m,2} - f_2 - \mathrm{P}_{z,r}(f_{m,2} - f_2))(z)^s \\ &\le C \liminf_{k\to\infty} G_{q,r}(f_{m,2} - f_{k,2})(z)^s. \end{aligned}$$

Integrate both sides over $\mathbb{C}^n$ and apply Fatou's lemma again to obtain the estimate

$$\int_{\mathbb{C}^n} G_{q,r/2}(f_{m,2} - f_2)^s\, dv \le C \liminf_{k\to\infty} \|f_{m,2} - f_{k,2}\|_{\mathrm{IDA}^{s,q}}.$$

Therefore,

$$\lim_{m\to\infty} \|f_{m,2} - f_2\|_{\mathrm{IDA}^{s,q}} = 0. \tag{3-29}$$

Next we consider $\{f_{m,1}\}_{m=1}^\infty$. Applying the estimate (3-15) to $f_m - f_k$,

$$|\bar\partial(f_{m,1} - f_{k,1})(z)| \le C G_{q,r/2}(f_m - f_k)(z). \tag{3-30}$$

Hence, $\{\bar\partial f_{m,1}\}_{m=1}^\infty$ is a Cauchy sequence in $L^s_{0,1}$ (see (2-11)). We may assume $\bar\partial f_{m,1} \to S = \sum_{j=1}^n S_j\, d\bar z_j$ under the $L^s_{0,1}$-norm. Since $\bar\partial^2 = 0$, $\bar\partial f_{m,1}$ is trivially $\bar\partial$-closed, and so, as the $L^s_{0,1}$ limit of $\{\bar\partial f_{m,1}\}_{m=1}^\infty$, S is also $\bar\partial$-closed weakly. Let $\phi(z) = \frac{1}{2}|z|^2$ and $g = 1 \in \Gamma$, and define

$$f_1(z) = A_\phi(S) \quad \text{and} \quad f^*_{m,1} = A_\phi(\bar\partial f_{m,1}).$$

Then, by Lemma 2.4,

$$f_1,\ f_{m,1}^* \in L^s(\phi) \subset L^s_{\mathrm{loc}}, \quad \bar\partial f_{m,1}^* = \bar\partial f_{m,1},$$

and $\{f_{m,1}^*\}_{m=1}^\infty$ converges to f_1 in $L^s(\phi)$. Therefore, for $\psi \in C_c^\infty(\mathbb{C}^n)$ (the family of all C^∞ functions with compact support) and $j = 1, 2, \dots, n$, it holds that

$$-\left\langle f_1, \frac{\partial\psi}{\partial z_j}\right\rangle_{L^2} = -\lim_{m\to\infty}\left\langle f_{m,1}^*, \frac{\partial\psi}{\partial z_j}\right\rangle_{L^2} = \lim_{m\to\infty}\left\langle \frac{\partial f_{m,1}^*}{\partial\bar z_j}, \psi\right\rangle_{L^2} = \lim_{m\to\infty}\left\langle \frac{\partial f_{m,1}}{\partial\bar z_j}, \psi\right\rangle_{L^2} = \langle S_j, \psi\rangle_{L^2}.$$

Hence, $\bar\partial f_1 = S$ weakly. Then for $\mathrm{H}_{B(z,r)}(\bar\partial f_{m,1} - S)$, the Henkin solution to the equation $\bar\partial u = \bar\partial f_{m,1} - S$ on $B(z,r)$, (3-16) gives

$$\|\mathrm{H}_{B(z,r)}(\bar\partial f_{m,1} - S)\|_{L^q(B(z,r),dv)} \le C\|\bar\partial f_{m,1} - S\|_{L^q(B(z,r),dv)}. \tag{3-31}$$

In addition, according to (3-24), it holds that

$$(f_{m,1} - f_1) - \mathrm{P}_{z,r}(f_{m,1} - f_1) = \mathrm{H}_{B(z,r)}(\bar\partial f_{m,1} - S) - \mathrm{P}_{z,r}(\mathrm{H}_{B(z,r)}(\bar\partial f_{m,1} - S))$$

on $B(z,r)$. Therefore, by (3-8), (3-9), and (3-31) we have

$$\begin{aligned}
&\|(f_{m,1} - f_1) - \mathrm{P}_{z,r}(f_{m,1} - f_1)\|^q_{L^q(B(z,r/2),\,dv)} \\
&\qquad = \|\mathrm{H}_{B(z,r)}(\bar\partial f_{m,1} - S) - \mathrm{P}_{z,r}(\mathrm{H}_{B(z,r)}(\bar\partial f_{m,1} - S))\|^q_{L^q(B(z,r/2),\,dv)} \\
&\qquad \le C\|\mathrm{H}_{B(z,r)}(\bar\partial f_{m,1} - S)\|^q_{L^q(B(z,r),\,dv)} \\
&\qquad \le C\|\bar\partial f_{m,1} - S\|^q_{L^q(B(z,r),dv)}.
\end{aligned} \tag{3-32}$$

Since $S = \lim_{k\to\infty}\bar\partial f_{k,1}$ in $L^s_{0,1}$, by Fatou's lemma,

$$\begin{aligned}
\|\bar\partial f_{m,1} - S\|^q_{L^q(B(z,r),dv)} &\le C\liminf_{k\to\infty}\|\bar\partial(f_{m,1} - f_{k,1})\|^q_{L^q(B(z,r),dv)} \\
&\le C\liminf_{k\to\infty} G_{q,2r}(f_{m,1} - f_{k,1})(z)^q,
\end{aligned} \tag{3-33}$$

where the last inequality follows from (3-30). We combine (3-32) and (3-33) to get

$$\|(f_{m,1} - f_1) - \mathrm{P}_{z,r}(f_{m,1} - f_1)\|^s_{L^q(B(z,r/2),dv)} \le C\liminf_{k\to\infty} G_{q,2r}(f_{m,1} - f_{k,1})(z)^s.$$

Integrating both sides over $\mathbb{C}^n$ with respect to dv and applying Fatou's lemma once more gives the estimates

$$\begin{aligned}
\|f_{m,1} - f_1\|^s_{\mathrm{IDA}^{s,q}} &\le C\int_{\mathbb{C}^n}\|(f_{m,1} - f_1) - \mathrm{P}_{z,r}(f_{m,1} - f_1)\|^s_{L^q(B(z,r/2))}\,dv \\
&\le C\int_{\mathbb{C}^n}\liminf_{k\to\infty} G_{q,2r}(f_{m,1} - f_{k,1})^s\,dv \\
&\le C\liminf_{k\to\infty}\|f_{m,1} - f_{k,1}\|^s_{\mathrm{IDA}^{s,q}}.
\end{aligned} \tag{3-34}$$

Therefore, $\lim_{m\to\infty}\|f_{m,1} - f_1\|_{\mathrm{IDA}^{s,q}} = 0$. Set $f = f_1 + f_2 \in L^q_{\mathrm{loc}}$. From (3-29) and (3-34) it follows that

$$\lim_{m\to\infty}\|f_m - f\|_{\mathrm{IDA}^{s,q}} \le \lim_{m\to\infty}(\|f_{m,1} - f_1\|_{\mathrm{IDA}^{s,q}} + \|f_{m,2} - f_2\|_{\mathrm{IDA}^{s,q}}) = 0,$$

which completes the proof of the completeness and of the theorem. □

Corollary 3.13. *Let $1 \le q < \infty$. With the norm induced by $\|\cdot\|_{\mathrm{BDA}^q}$, the quotient space $\mathrm{BDA}^q/H(\mathbb{C}^n)$ is a Banach space and VDA^q is a closed subspace of BDA^q.*

Proof. The proof of Theorem 3.12 works for $s = \infty$, so $\mathrm{BDA}^q/H(\mathbb{C}^n)$ is a Banach space in $\|\cdot\|_{\mathrm{BDA}^q}$. That VDA^q is a closed subspace of BDA^q can be proved in a standard way. □

4. Proof of Theorem 1.1

Given two F-spaces X and Y, we write $B(\mathrm{X})$ for the unit ball of X. A linear operator T from X to Y is bounded (or compact) if $\mathrm{T}(B(\mathrm{X}))$ is bounded (or relatively compact) in Y. The collection of all bounded (and compact) operators from X to Y is denoted by $\mathcal{B}(\mathrm{X}, \mathrm{Y})$ (and by $\mathcal{K}(\mathrm{X}, \mathrm{Y})$ respectively). We use $\|\mathrm{T}\|_{\mathrm{X}\to\mathrm{Y}}$ to denote the corresponding operator norm. In particular, we recall that when $0 < p < 1$, the Fock space $F^p(\varphi)$ with the metric given by $d(f, g) = \|f - g\|_{p,\varphi}^p$ is an F-space.

To deal with the boundedness and compactness of Hankel operators, we need an additional result involving positive measures and their averages. More precisely, given a positive Borel measure μ on $\mathbb{C}^n$ and $r > 0$, we write $\hat{\mu}_r(z) = \mu(B(z, r))$. Notice, in particular, $\hat{\mu}_r$ is a constant multiple of the averaging function induced by the measure μ.

Lemma 4.1. *Suppose $0 < p \le 1$ and $r > 0$. There is a constant C such that, for μ a positive Borel measure on $\mathbb{C}^n$, Ω a domain in $\mathbb{C}^n$, and $g \in H(\mathbb{C}^n)$, it holds that*

$$\left(\int_\Omega |g(\xi)e^{-\varphi(\xi)}|\, d\mu(\xi)\right)^p \le C\int_{\Omega_r^+} |g(\xi)e^{-\varphi(\xi)}|^p \hat{\mu}_r(\xi)^p\, dv(\xi),$$

where $\Omega_r^+ = \bigcup_{\{z\in\Omega\}} B(z, r)$.

Proof. Let $\{a_j\}_{j=1}^\infty$ be an $(r/4)$-lattice. Notice that

$$\hat{\mu}_{r/4}(a_j) \le C \inf_{w\in B(a_j, r/2)} \hat{\mu}_r(w)$$

for all $j \in \mathbb{N}$ and $(a+b)^p \le a^p + b^p$ for $a, b \ge 0$. Then

$$\begin{aligned}
\left(\int_\Omega |g(\xi)e^{-\varphi(\xi)}|\, d\mu(\xi)\right)^p &\le \sum_{j=1}^\infty \left(\int_{B(a_j,r/4)\cap\Omega} |g(\xi)e^{-\varphi(\xi)}|\, d\mu(\xi)\right)^p \\
&\le C \sum_{\{j:\, B(a_j,r/4)\cap\Omega\neq\varnothing\}} \sup_{\xi\in B(a_j,r/4)\cap\Omega} |g(\xi)e^{-\varphi(\xi)}|^p \hat{\mu}_{r/4}(a_j)^p \\
&\le C \sum_{\{j:\, B(a_j,r/4)\cap\Omega\neq\varnothing\}} \hat{\mu}_{r/4}(a_j)^p \int_{B(a_j,r/2)} |g(\xi)e^{-\varphi(\xi)}|^p\, dv(\xi) \\
&\le C \sum_{\{j:\, B(a_j,r/4)\cap\Omega\neq\varnothing\}} \int_{B(a_j,r/2)} |g(\xi)e^{-\varphi(\xi)}|^p \hat{\mu}_r(\xi)^p\, dv(\xi) \\
&\le C\int_{\Omega_r^+} |g(\xi)e^{-\varphi(\xi)}|^p \hat{\mu}_r(\xi)^p\, dv(\xi),
\end{aligned}$$

which completes the proof. □

Remark 4.2. To prove compactness of Hankel operators on spaces that are not necessarily Banach spaces, we use the following result. For $0 < p, q < \infty$, $H_f : F^p(\varphi) \to L^q(\varphi)$ is compact if and only if

$$\lim_{m\to\infty} \|H_f(g_m)\|_{q,\varphi} = 0$$

for any sequences $\{g_m\}_{m=1}^{\infty}$ in $B(F^p(\varphi))$ satisfying

$$\lim_{m\to\infty} \sup_{w\in E} |g_m(w)| = 0$$

for compact subsets E in $\mathbb{C}^n$.

Necessity is trivial. To prove sufficiency, we notice that $B(F^p(\varphi))$ is a normal family, so for any sequence $\{g_m\}_{m=1}^{\infty} \subset B(F^p(\varphi))$, there exist a holomorphic function g_0 on $\mathbb{C}^n$ and a subsequence $\{g_{m_j}\}_{j=1}^{\infty}$ such that

$$\lim_{j\to\infty} \sup_{w\in E} |g_{m_j}(w) - g_0(w)| = 0.$$

This and Fatou's lemma imply that $g_0 \in B(F^p(\varphi))$, and hence by the hypothesis, we get

$$\lim_{j\to\infty} \|H_f(g_{m_j}) - H_f(g_0)\|_{q,\varphi} = \lim_{j\to\infty} \|H_f(g_{m_j} - g_0)\|_{q,\varphi} = 0.$$

Thus, $H_f(B(F^p(\varphi)))$ is sequentially compact in $L^q(\varphi)$, that is, the Hankel operator $H_f : F^p(\varphi) \to L^q(\varphi)$ is compact.

4A. *The case* $0 < p \le q < \infty$ *and* $q \ge 1$.

Proof of Theorem 1.1(a). By (2-3)–(2-5),

$$\|k_z\|_{p,\varphi} \le C, \qquad \sup_{\xi\in B(z,r_0)} |k_z(\xi)|e^{-\varphi(\xi)} \ge C \quad \text{and} \quad \lim_{z\to\infty} \sup_{w\in E} |k_z(w)| = 0 \tag{4-1}$$

for any compact subset $E \subset \mathbb{C}^n$. As in the proof of Theorem 4.2 of [Hu and Lu 2019], there is an r_0 such that, for all $z \in \mathbb{C}^n$, we have

$$\begin{aligned}\|H_f(k_z)\|_{q,\varphi}^q &\ge \int_{B(z,r_0)} |fk_z - P(fk_z)|^q e^{-q\varphi}\, dv \\ &\ge C\frac{1}{|B(z,r_0)|} \int_{B(z,r_0)} \left| f - \frac{1}{k_z} P(fk_z)\right|^q dv \ge C G_{q,r_0}^q(f)(z).\end{aligned} \tag{4-2}$$

If $H_f \in \mathcal{B}(F^p(\varphi), L^q(\varphi))$,

$$\|f\|_{\mathrm{BDA}^q} \le C\|H_f\|_{F^p(\varphi)\to L^q(\varphi)} < \infty; \tag{4-3}$$

if $H_f \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$, then $f \in \mathrm{VDA}^q$ because

$$\lim_{z\to\infty} G_{q,r_0}^q(f)(z) \le C \lim_{z\to\infty} \|H_f(k_z)\|_{q,\varphi} = 0. \tag{4-4}$$

Next we prove sufficiency. Suppose that $f \in \mathrm{BDA}^q$ and decompose $f = f_1 + f_2$ as in (3-12). Write $d\mu = |f_2|^q dv$ and $d\nu = |\bar\partial f_1|^q dv$. According to Theorem 2.6 of [Hu and Lv 2014] and Corollary 3.9, both $d\mu$ and $d\nu$ are (p,q)-Fock Carleson measures. We claim that both $f_1, f_2 \in \mathcal{S}$. Indeed, since $q \geq 1$, we can use Lemma 4.1 with $\Omega = \mathbb{C}^n$ and the measure $|f_2| d\nu$ to get

$$\begin{aligned}\int_{\mathbb{C}^n} |f_2(\xi)K(\xi,z)|e^{-\varphi(\xi)}\,dv(\xi) &\leq C\int_{\mathbb{C}^n} M_{1,r}(f_2)(\zeta)|K(\zeta,z)|e^{-\varphi(\zeta)}\,dv(\zeta)\\ &\leq C\int_{\mathbb{C}^n} M_{q,r}(f_2)(\zeta)|K(\zeta,z)|e^{-\varphi(\zeta)}\,dv(\zeta). \qquad (4\text{-}5)\end{aligned}$$

Since $f \in \mathrm{BDA}^q$, Lemma 3.6 implies

$$\int_{\mathbb{C}^n} |f_2(\xi)K(\xi,z)|e^{-\varphi(\xi)}\,dv(\xi) \leq C\|f\|_{BDA^q}\int_{\mathbb{C}^n} |K(\xi,z)|e^{-\varphi(\xi)}\,dv(\xi) < \infty$$

for $z \in \mathbb{C}^n$. Hence, $f_2 \in \mathcal{S}$, and so also $f_1 = f - f_2 \in \mathcal{S}$ because $f \in \mathcal{S}$ by the hypothesis. Since the Bergman projection P is bounded on $L^q(\varphi)$ when $q \geq 1$, we have, for $g \in \Gamma$,

$$\begin{aligned}\|H_{f_2}(g)\|_{q,\varphi} &\leq (1 + \|P\|_{L^q(\varphi)\to F^q(\varphi)})\|f_2 g\|_{q,\varphi}\\ &\leq C\|M_{q,r}(f_2)\|_{L^\infty}\|g\|_{q,\varphi} \leq C\|M_{q,r}(f_2)\|_{L^\infty}\|g\|_{p,\varphi},\end{aligned}$$

where the second inequality follows from Lemma 4.1. For $H_{f_1}(g)$ with $g \in \Gamma$, Corollary 2.5 shows that $H_{f_1}(g) = A_\varphi(g\bar\partial f_1) - P(A_\varphi(g\bar\partial f_1))$. Lemma 2.4 implies

$$\|H_{f_1}(g)\|_{q,\varphi} \leq C\|g\,|\bar\partial f_1|\|_{q,\varphi} \leq C\|\bar\partial f_1\|_{L^\infty}\|g\|_{q,\varphi} \leq C\|\bar\partial f_1\|_{L^\infty}\|g\|_{p,\varphi}. \qquad (4\text{-}6)$$

From the above estimates and the fact that Γ is dense in $F^p(\varphi)$, it follows that, for $0 < p \leq q < \infty$, we have

$$\|H_f\|_{F^p(\varphi)\to L^q(\varphi)} \leq C\{\|\bar\partial f_1\|_{L^\infty} + \|M_{q,r}(f_2)\|_{L^\infty}\} \leq C\|f\|_{\mathrm{BDA}^q}, \qquad (4\text{-}7)$$

where the latter inequality follows from Lemma 3.6.

For compactness, suppose $f \in \mathrm{VDA}^q$ so that $f = f_1 + f_2$ is as (3-12). Notice that both $d\mu = |f_2|^q\,dv$ and $d\nu = |\bar\partial f_1|^q\,dv$ are vanishing (p,q)-Fock Carleson measures. Let $\{g_m\}$ be a bounded sequence in $F^p(\varphi)$ converging to zero uniformly on compact subsets of $\mathbb{C}^n$. Then

$$\|H_{f_2}(g_m)\|_{L^q(\varphi)} \leq \|g_m f_2\|_{q,\varphi} + \|P(g_m f_2)\|_{q,\varphi} \leq C\biggl(\int_{\mathbb{C}} |g_m e^{-\varphi}|^q\,d\mu\biggr)^{1/q} \to 0$$

as $m \to \infty$. To prove $\lim_{m\to\infty}\|H_{f_1}(g_m)\|_{L^q(\varphi)} = 0$, for each m we pick some $g_m^* \in \Gamma$ so that $\|g_m - g_m^*\|_{p,\varphi} < 1/m$. Clearly, $\{g_m^*\}_{m=1}^\infty$ is bounded in $F^p(\varphi)$, and $\lim_{z\to\infty}\sup_{w\in E}|g_m^*(w)| = 0$ for any compact subset E. Again by Corollary 2.5,

$$\|H_{f_1}(g_m^*)\|_{L^q(\varphi)} \leq C\|g_m^*\bar\partial f_1\|_{L^q(\varphi)} \leq C\|g_m^*\|_{L^q(\mathbb{C}^n,d\nu)} \to 0 \quad \text{as } m \to \infty.$$

Thus, since Lemma 3.6 guarantees $H_{f_1} \in \mathcal{B}(F^p(\varphi), L^q(\varphi))$, it follows that $\lim_{m\to\infty}\|H_{f_1}(g_m)\|_{L^q(\varphi)} = 0$, and so

$$H_f = H_{f_1} + H_{f_2} \in \mathcal{K}(F^p(\varphi), L^q(\varphi)).$$

Finally, it remains to notice that the norm equivalence (1-1) follows from (4-3) and (4-7). □

4B. ***The case $1 \le q < p < \infty$.*** We can now prove the case $q < p$ under the assumption that $q \ge 1$.

Proof of Theorem 1.1(b). Suppose that $H_f \in \mathcal{B}(F^p(\varphi), L^q(\varphi))$. Because the proof of sufficiency is similar to the implication $(A) \Rightarrow (C)$ of Theorem 4.4 in [Hu and Lu 2019], we only give the sketch here.

Indeed, take r_0 as in (4-1), and set $t = r_0/4$. Let $\{a_j\}_{j=1}^{\infty}$ be a $(t/2)$-lattice. By Lemma 2.4 of [Hu and Lv 2014], $\left\|\sum_{j=1}^{\infty} \lambda_j k_{a_j}\right\|_{p,\varphi} \le C\|\{\lambda_j\}\|_{l^p}$ for all $\{\lambda_j\}_{j=1}^{\infty} \in l^p$, where the constant C is independent of $\{\lambda_j\}_{j=1}^{\infty}$. Let $\{\phi_j\}_{j=1}^{\infty}$ be the sequence of Rademacher functions on the interval [0, 1]. Using the boundedness of H_f, we get

$$\left\|H_f\left(\sum_{j=1}^{\infty} \lambda_j \phi_j(s) k_{a_j}(\,\cdot\,)\right)\right\|_{q,\varphi} \le C\|H_f\|_{F^p(\varphi)\to L^q(\varphi)} \|\{|\lambda_j|^q\}\|_{l^{p/q}}^{1/q} \tag{4-8}$$

for $s \in [0, 1]$. On the other hand,

$$\int_{B(a_j,t)} |H_f(k_z)(\xi) e^{-\varphi(\xi)}|^q \, dv(\xi) \ge C G_{q,t}(f)(a_j)^q. \tag{4-9}$$

This and Khintchine's inequality yield

$$\int_0^1 \left\|H_f\left(\sum_{j=1}^{\infty} \lambda_j \phi_j(s) k_{a_j}(\,\cdot\,)\right)\right\|_{q,\varphi}^q dt \ge C \sum_{j=1}^{\infty} |\lambda_j|^q G_{q,t}(f)(a_j)^q.$$

Combining this with (4-8) gives

$$\sum_{j=1}^{\infty} |\lambda_j|^q G_{q,t}(f)(a_j)^q \le C\|H_f\|_{F^p(\varphi)\to L^q(\varphi)}^q \|\{|\lambda_j|^q\}\|_{l^{p/q}}$$

for all $\{|\lambda_j|^q\}_{j=1}^{\infty} \in l^{p/q}$. By duality with the exponentials p/q and its conjugate,

$$\sum_{j=1}^{\infty} G_{q,t}(f)(a_j)^{pq/(p-q)} \le C\|H_f\|_{F^p(\varphi)\to L^q(\varphi)}^{pq/(p-q)}.$$

Therefore, by (3-7),

$$\begin{aligned}\int_{\mathbb{C}^n} G_{q,t/2}(f)(z)^{pq/(p-q)} \, dv(z) &\le \sum_{j=1}^{\infty} \int_{B(a_j,t/2)} G_{q,t/2}(f)(z)^{pq/(p-q)} \, dv(z) \\ &\le C\|H_f\|_{F^p(\varphi)\to L^q(\varphi)}^{pq/(p-q)},\end{aligned} \tag{4-10}$$

which means that $f \in \mathrm{IDA}^{s,q}$ with the estimate $\|f\|_{\mathrm{IDA}^{s,q}} \le C\|H_f\|$.

It should be pointed out that the right-hand side of the estimate (4.24) (the analogue of (4-10) above) in [Hu and Lu 2019] should read $C\|H_f\|_{A^p_\omega\to L^q_\omega}^{pq/(p-q)}$, and not $C\|H_f\|_{A^p_\omega\to L^q_\omega}$ as stated there.

Conversely, suppose $f \in \mathrm{IDA}^{s,q}$. As before, decompose $f = f_1 + f_2$ as in (3-12). From Lemma 3.6 we know that $\|M_{q,r}(f_2)\|_{pq/(p-q)} \le C\|f\|_{\mathrm{IDA}^{s,q}}$. Applying Hölder's inequality to the right-hand side integral in (4-5) with exponent $pq/(p-q)$ and its conjugate exponent t, since we have $\|K(\,\cdot\,, z)\|_{t,\varphi} < \infty$, it follows that

$$\int_{\mathbb{C}^n} |f_2(\xi) K_z(\xi)| e^{-\varphi(\xi)} \, dv(\xi) \le C\|M_{q,r}(f_2)\|_{pq/(p-q)} \cdot \|K_z\|_{t,\varphi} < \infty.$$

This implies $f_2 \in \mathcal{S}$, and so also $f_1 \in \mathcal{S}$.

Now for $d\nu = |\bar\partial f_1|^q dv$, applying Hölder's inequality again with $p/(p-q)$ and its conjugate exponent p/q, we get

$$\begin{aligned}\|\hat\nu_r\|^{p/(p-q)}_{L^{p/(p-q)}} &= C\int_{\mathbb{C}^n}\left\{\int_{B(\xi,r)}|\bar\partial f_1(\zeta)|^q\,dv(\zeta)\right\}^{p/(p-q)} dv(\xi)\\ &\le C\int_{\mathbb{C}^n} dv(\xi)\int_{B(\xi,r)}|\bar\partial f_1(\zeta)|^{pq/(p-q)}\,dv(\zeta)\\ &\simeq C\int_{\mathbb{C}^n}|\bar\partial f_1(\zeta)|^{pq/(p-q)}\,dv(\zeta)<\infty. \end{aligned}\tag{4-11}$$

Theorem 2.8 of [Hu and Lv 2014] shows that ν is a vanishing (p,q)-Fock Carleson measure; that is, the multiplier $M_{f_1} : g \mapsto g|\bar\partial f_1|$ is compact from $F^p(\varphi)$ to $L^q(\varphi)$ (see Proposition 2.3). Therefore, by Lemma 2.4(A), $A_\varphi(\,\cdot\,\bar\partial f_1)$ is compact from $F^p(\varphi)$ to $L^q(\varphi)$. Moreover, Γ is dense in $F^p(\varphi)$ and, by Corollary 2.5, $H_{f_1}(g) = A_\varphi(g\,\bar\partial f_1) - P\circ A_\varphi(g\,\bar\partial f_1)$ for $g\in\Gamma$. Hence, $H_{f_1} : F^p(\varphi)\to L^q(\varphi)$ is compact and we obtain the norm estimate

$$\|H_{f_1}\|_{F^p(\varphi)\to L^q(\varphi)} \le C \sup_{\{g\in F^p(\varphi):\,\|g\|_{p,\varphi}\le 1\}} \|A_\varphi(g\bar\partial f_1)\|_{q,\varphi} \le C\|\bar\partial f_1\|_{pq/(p-q)}.\tag{4-12}$$

Similarly to (4-11), using Lemma 3.6, for $d\mu = |f_2|^q dv$, we get

$$\begin{aligned}\|\hat\mu_r\|^{p/(p-q)}_{L^{p/(p-q)}} &= C\int_{\mathbb{C}^n}\left\{\int_{B(\xi,r)}|f_2(\zeta)|^q\,dv(\zeta)\right\}^{p/(p-q)} dv(\xi)\\ &= C\|M_{q,r}(f_2)\|^{pq/(p-q)}_{pq/(p-q)} \le C\|f\|^s_{\mathrm{IDA}^{s,q}}<\infty.\end{aligned}$$

Hence, $d\mu = |f_2|^q\,dv$ is a vanishing (p,q)-Fock Carleson measure. It follows from Proposition 2.3 that the identity operator

$$\mathrm{I} : F^p(\varphi)\to L^q(\mathbb{C}^n, e^{-q\varphi}\,d\mu)$$

is compact. Using the inequality

$$\|H_{f_2}(g)\|_{q,\varphi} \le C\|f_2 g\|_{q,\varphi} = C\|\mathrm{I}(g)\|_{L^q(\mathbb{C},\,e^{-q\varphi}d\mu)},\tag{4-13}$$

we see that H_{f_2} is compact from $F^p(\varphi)$ to $L^q(\varphi)$.

It remains to notice that the norm equivalence in (1-2) follows from combining the estimates in (4-10), (4-12), and (4-13). □

Remark 4.3. In [Stroethoff 1992], it was proved that for bounded symbols f, the Hankel operator $H_f : F^2\to L^2$ is compact if and only if

$$\|(I-P)(f\circ\phi_\lambda)\|\to 0\tag{4-14}$$

as $|\lambda|\to\infty$, where $\phi_\lambda(z) = z+\lambda$. This characterization was recently generalized to F^p_α with $1<p<\infty$ in [Hagger and Virtanen 2021]. Here we note that, using a generalization of Lemma 8.2 of [Zhu 2012] to the setting of $\mathbb{C}^n$, one can prove that Stroethoff's result remains true for Hankel operators acting from F^p_α to L^q_α whenever $1\le p,q<\infty$ even for unbounded symbols.

4C. *The case $0 < p \le q \le 1$ with bounded symbols.* We start with the following preliminary lemma whose proof can be completed with a standard ε argument.

Lemma 4.4. *Suppose that $0 < p < \infty$, $h \in L^\infty$ and $\lim_{z\to\infty} h(z) = 0$. Then for any bounded sequence $\{g_j\}_{j=1}^\infty$ in L^p_φ satisfying $\lim_{j\to\infty} g_j(z) = 0$ uniformly on compact subsets of $\mathbb{C}^n$, it holds that $\lim_{j\to\infty} \|g_j h\|_{p,\varphi} = 0$.*

Proof. If R is sufficiently large, there is a $C > 0$ such that

$$\begin{aligned}\|g_j h\|^p_{p,\varphi} &= \left(\int_{B(0,R)} + \int_{\mathbb{C}^n\setminus B(0,R)}\right)|g_j(\xi)h(\xi)e^{-\varphi(\xi)}|^p\,dv(\xi)\\ &\le \|h\|^p_{L^\infty}\sup_{|\xi|\le R}|g_j(\xi)e^{-\varphi(\xi)}|^p + C\|g_j\|^p_{p,\varphi} \to 0\end{aligned}$$

as $j\to\infty$. □

Proof of Theorem 1.1(c). Suppose that $f \in \mathcal{S}$. Then $f \in L^q_{\mathrm{loc}}$ for $0 < q \le 1$, and we may decompose $f = f_1 + f_2$ as in (3-12) with $t = r/2$. We claim that, for $g \in \Gamma$,

$$\|H_{f_1}(g)\|^q_{q,\varphi} \le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q\,\|\bar\partial f_1\|^q_{L^\infty(B(\xi,r),dv)}\,dv(\xi), \tag{4-15}$$

$$\|H_{f_2}(g)\|^q_{q,\varphi} \le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q M_{1,r}(f_2)(\xi)^q\,dv(\xi). \tag{4-16}$$

To estimate $\|H_{f_1}(g)\|_{q,\varphi}$, we use the representation

$$H_{f_1}(g) = A_\varphi(g\bar\partial f_1) - P(A_\varphi(g\bar\partial f_1))$$

(see (2-14)), which suggests that we define a measure $d\mu_z$ as

$$d\mu_z(\xi) = |\bar\partial f_1(\xi)|\left\{\frac{1}{|\xi-z|} + \frac{1}{|\xi-z|^{2n-1}}\right\}e^{-m|\xi-z|}\,dv(\xi).$$

Then there is a constant C such that, for $w\in\mathbb{C}^n$,

$$\int_{B(w,r)}|\bar\partial f_1(\xi)|\left\{\frac{1}{|\xi-z|} + \frac{1}{|\xi-z|^{2n-1}}\right\}e^{-m|\xi-z|^2}\,dv(\xi) \le C\int_{B(w,r)} d\mu_z(\xi).$$

Also, it is easy to verify that

$$\widehat{(\mu_z)}_r(w) \le C\sup_{\eta\in B(w,r)}|\bar\partial f_1(\eta)|e^{-m|w-z|},$$

where the constant C is independent of $z, w \in \mathbb{C}^n$. Recall that

$$A_\varphi(g\bar\partial f_1)(z) = \int_{\mathbb{C}^n} e^{\langle 2\partial\varphi, z-\xi\rangle}\sum_{j<n} g(\xi)\bar\partial f_1(\xi)\wedge\frac{\partial|\xi-z|^2\wedge(2\bar\partial\partial\varphi(\xi))^j\wedge(\bar\partial\partial|\xi-z|^2)^{n-1-j}}{j!\,|\xi-z|^{2n-2j}}.$$

Therefore, using (2-13) and Lemma 4.1, we get

$$\begin{aligned}|A_\varphi(g\bar\partial f_1)(z)e^{-\varphi(z)}|^q &\le C\left(\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|\,d\mu_z(\xi)\right)^q\\ &\le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q\,\|\bar\partial f_1\|^q_{L^\infty(B(\xi,r),dv)}e^{-qm|\xi-z|}\,dv(\xi).\end{aligned} \tag{4-17}$$

Fubini's theorem yields

$$\begin{aligned}\|A_\varphi(g\bar\partial f_1)\|_{q,\varphi}^q &\le C\int_{\mathbb{C}^n} dv(z)\int_{\mathbb{C}^n} |g(\xi)e^{-\varphi(\xi)}|^q\,\|\bar\partial f_1\|^q_{L^\infty(B(\xi,r),dv)}e^{-qm|\xi-z|}\,dv(\xi)\\ &\le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q\,\|\bar\partial f_1\|^q_{L^\infty(B(\xi,r),dv)}\,dv(\xi). \end{aligned}\tag{4-18}$$

To deal with $P(A_\varphi(g\bar\partial f_1))$, we use Lemma 2.2 to obtain positive constants θ and C so that, for $z\in\mathbb{C}^n$, we have

$$\begin{aligned}\int_{\mathbb{C}^n}|K(w,z)|e^{-m|\xi-z|}e^{-\varphi(z)}\,dv(z) &\le Ce^{\varphi(w)}\int_{\mathbb{C}^n}e^{-m|\xi-z|}e^{-\theta|w-z|}\,dv(z)\\ &= Ce^{\varphi(w)}\biggl(\int_{\{z:|z-\xi|\ge|z-w|\}}+\int_{\{z:|z-\xi|<|z-w|\}}\biggr)e^{-m|w-z|}e^{-\theta|\xi-z|}\,dv(z)\\ &\le Ce^{\varphi(w)}e^{-\tau|\xi-w|},\end{aligned}$$

where $\tau=\min\{\theta,m\}$. Therefore, (4-17) and Fubini's theorem yield

$$\begin{aligned}|P(A_\varphi(g\bar\partial f_1))(w)| &\le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|\,\|\bar\partial f_1\|_{L^\infty(B(\xi,r/2),dv)}\,dv(\xi)\int_{\mathbb{C}^n}|K(w,z)|e^{-\theta|\xi-z|}e^{-\varphi(z)}\,dv(z)\\ &\le Ce^{\varphi(w)}\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|\,\|\bar\partial f_1\|_{L^\infty(B(\xi,r/2),dv)}e^{-\tau|\xi-w|}\,dv(\xi).\end{aligned}$$

Lemma 4.1 again gives

$$\|P(A_\varphi(g\bar\partial f_1))(w)\|_{q,\varphi}^q\le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q\,\|\bar\partial f_1\|^q_{L^\infty(B(\xi,r),dv)}\,dv(\xi).$$

Combining this and (4-18), we get (4-15).

For (4-16), notice first that

$$\|f_2g\|_{q,\varphi}^q\le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q M_{q,r}^q(f_2)(\xi)\,dv(\xi),\tag{4-19}$$

and, by Lemma 4.1 with the measure $M_{1,r/2}(f_2)\,dv$, we have

$$\begin{aligned}|P(f_2g)(z)|^q &\le C\biggl(\int_{\mathbb{C}^n}|g(\xi)K(z,\xi)e^{-2\varphi(\xi)}|M_{1,r/2}(f_2)(\xi)\,dv(\xi)\biggr)^q\\ &\le C\int_{\mathbb{C}^n}|g(\xi)K(z,\xi)e^{-2\varphi(\xi)}|^q M_{1,r}(f_2)(\xi)^q\,dv(\xi).\end{aligned}\tag{4-20}$$

Integrating both sides of (4-20) against $e^{-q\varphi}\,dv$ over $\mathbb{C}^n$ and using (2-5), we get

$$\|P(f_2g)\|_{q,\varphi}^q\le C\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q M_{1,r}(f_2)(\xi)^q\,dv(\xi).\tag{4-21}$$

This and (4-19) imply (4-16).

Now we suppose that $f\in L^\infty$ and $0<p\le q<1$. For $g\in H(\mathbb{C}^n)$, similarly to the proof of (4-16), we have

$$\|H_f(g)\|_{q,\varphi}\le C\biggl(\int_{\mathbb{C}^n}|g(\xi)e^{-\varphi(\xi)}|^q M_{1,r}(f)(\xi)^q\,dv(\xi)\biggr)^{1/q}\le C\|f\|_{L^\infty}\|g\|_{p,\varphi}.$$

This implies boundedness of H_f with the norm estimate (1-3).

For the second assertion, suppose first that $\lim_{|z|\to\infty} G_{q,r}(f)(z) = 0$ for some $r > 0$ and write $f = f_1 + f_2$ as above. Since the unit ball $B(F^p(\varphi))$ of $F^p(\varphi)$ is a normal family, to show that H_f is compact from $F^p(\varphi)$ to $L^q(\varphi)$, it suffices to prove that, for $k = 1, 2$,

$$\lim_{j\to\infty} \|H_{f_k}(g_j)\|_{q,\varphi} = \lim_{j\to\infty} \|f_k g_j - P(f_k g_j)\|_{q,\varphi} = 0$$

for any bounded sequence $\{g_j\}_{j=1}^{\infty}$ in $F^p(\varphi)$ with the property that

$$\lim_{j\to\infty} \sup_{w\in E} |g_j(w)| = 0$$

for E compact in $\mathbb{C}^n$. From the assumption that $\lim_{z\to\infty} M_{q,r}(f_2)(z) = 0$, it follows that $d\mu = |f_2|^q\, dv$ is a vanishing (p, q)-Fock Carleson measure (see Theorem 2.7 of [Hu and Lv 2014] and Proposition 2.3). Therefore, we get

$$\|f_2 g_j\|_{q,\varphi} = \|g_j\|_{L^q(\mathbb{C}^n, |f_2|^q dv)} \to 0 \quad \text{as } j \to \infty.$$

Notice also that $\|g\|_{q,\varphi} \le C\|g\|_{p,\varphi}$ for $g \in F^q(\varphi)$ and $p \le q$. Further, by (4-16), we obtain

$$M_{1,r}(f_2)(\xi) \le \|f_2\|_{L^\infty}^{1-q} M_{q,r}(f_2)(\xi)^q,$$

and applying Lemma 4.4 to $h = M_{q,r}(f_2)^{q^2}$, we get

$$\begin{aligned}\|H_{f_2} g_j\|_{q,\varphi}^q &\le C \int_{\mathbb{C}^n} |g_j(\xi) e^{-\varphi(\xi)}|^q M_{1,r}(f_2)(\xi)^q\, dv(\xi)\\ &\le C\|f_2\|_{L^\infty}^{(1-q)q} \int_{\mathbb{C}^n} |g_j(\xi) e^{-\varphi(\xi)}|^q M_{q,r}(f_2)(\xi)^{q^2}\, dv(\xi) \to 0\end{aligned}$$

as $j \to \infty$. So $H_{f_2} \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$. As for H_{f_1}, it follows from Lemma 3.6 that

$$\|\bar\partial f_1\|_{L^\infty(B(\xi,r),dv)} \le C G_{q,r}(f)(\xi) \to 0 \quad \text{when } \xi \to \infty.$$

Therefore, by (4-15),

$$\|H_{f_1}(g_j)\|_{q,\varphi}^q \le C \int_{\mathbb{C}^n} |g_j(\xi) e^{-\varphi(\xi)}|^q\, \|\bar\partial f_1\|_{L^\infty(B(\xi,r),dv)}^q\, dv(\xi) \to 0$$

as $j \to \infty$, and hence we have $H_{f_1} \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$.

Conversely, suppose that H_f is compact from $F^p(\varphi)$ to $L^q(\varphi)$. Then, as in (4-4), we have

$$\lim_{z\to\infty} G_{q,r}(f)(z) \le C \lim_{z\to\infty} \|H_f(k_z)\|_{q,\varphi} = 0 \tag{4-22}$$

for $r \in (0, r_0]$ fixed. We claim that (4-22) is valid for any $r > 0$. To see this, we consider the Hankel operator H_f on the Fock space F^p_α. From (4-22), using the sufficiency part, it follows that H_f is compact from F^p_α to $L^q(\mathbb{C}^n, e^{-(q\alpha/2)|z|^2}\, dv)$. Notice that the equality (1-5) yields

$$\inf_{w\in B(z,r)} |K(w, z)| \ge C > 0$$

for any $r > 0$ fixed, where the constant C is independent of $z \in \mathbb{C}^n$. As in (4-2), we have

$$\lim_{z\to\infty} G_{q,r}(f)(z) \le C \lim_{z\to\infty} \|H_f(k_z)\|_{L^q(\mathbb{C}^n, e^{-(q\alpha/2)|z|^2} dv)} = 0.$$

Thus, $f \in \mathrm{VDA}^q$. □

The following Corollary 4.5 is a direct consequence of the proof of Theorem 1.1(c) which we use to complement and extend the classical result of Berger and Coburn in the next section.

Corollary 4.5. *Suppose that* $0 < q < 1$ *and* $f \in L^\infty$. *Then the limit* $\lim_{z\to\infty} G_{q,r}(f)(z) = 0$ *is independent of* $r > 0$.

5. Proof of Theorem 1.2

Proof of the case $0 < p \le q < \infty$. For $R > 0$, let $\{a_k\}_{k=1}^\infty$ be the $(R/2)$-lattice

$$\left\{ \frac{R}{2\sqrt{n}}(m_1 + k_1\mathrm{i}, m_2 + k_2\mathrm{i}, \dots, m_n + k_n\mathrm{i}) \in \mathbb{C}^n : m_j, k_j \in \mathbb{Z},\ j = 1, 2, \dots, n \right\}.$$

Choose $\rho \in C^\infty(\mathbb{C}^n)$ such that

$$0 \le \rho \le 1, \quad \rho|_{B(0,1/2)} \equiv 1, \quad \operatorname{supp} \rho \subseteq B\big(0, \tfrac{3}{4}\big).$$

Then $\|\nabla\rho\|_{L^\infty} < \infty$ and

$$0 < \sum_{k=1}^\infty \rho((z - a_k)/R) \le C$$

for $z \in \mathbb{C}^n$. Define $\psi_{j,R} \in C^\infty(\mathbb{C}^n)$ by

$$\psi_{j,R}(z) = \frac{\rho((z - a_j)/R)}{\sum_{k=1}^\infty \rho((z - a_k)/R)}.$$

Then $\{\psi_{j,R}\}_{j=1}^\infty$ is a partition of unity subordinate to $\{B(a_j, R)\}_{j=1}^\infty$ and

$$R\|\nabla\psi_{j,R}(\,\cdot\,)\|_{L^\infty} \le C, \tag{5-1}$$

where the constant C is independent of j and R.

Now we suppose that $f \in L^\infty$ and $H_f \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$. Theorem 1.1 and Corollary 4.5 imply that

$$\lim_{z\to\infty} G_{q,2R}(f)(z) = 0 \tag{5-2}$$

for $R > 0$ fixed. As in (3-2), pick $h_{j,R} \in H(B(a_j, 2R))$ so that

$$\frac{1}{|B(a_j, 2R)|} \int_{B(a_j,2R)} |f - h_{j,R}|^q \, dv = G_{q,2R}(f)(a_j)^q. \tag{5-3}$$

By (3-3),

$$\sup_{z\in B(a_j,R)} |h_{j,R}(z)| \le C\|f\|_{L^\infty}.$$

Set

$$f_{1,R} = \sum_{j=1}^\infty \psi_{j,R}\, h_{j,R} \quad \text{and} \quad f_{2,R} = f - f_{1,R}.$$

From estimates (2-9) and (3-3), it follows that there is a positive constant C such that

$$\|f_{1,R}\|_{L^\infty} + \|f_{2,R}\|_{L^\infty} \le C\|f\|_{L^\infty} \tag{5-4}$$

for $R > 0$. Lemma 3.6 and (5-2) imply that

$$\lim_{z\to\infty} M_{q,R}(\bar{f}_{2,R})(z) = \lim_{z\to\infty} M_{q,R}(f_{2,R})(z) = 0,$$

and so

$$H_{\bar{f}_{2,R}} \in \mathcal{K}(F^p(\varphi), L^q(\varphi)). \tag{5-5}$$

Recall that $P_{z,R}$ is the standard Bergman projection from $L^2(B(z,R), dv)$ to $A^2(B(z,R), dv)$. Since $h_{j,R}$ is bounded on $B(a_j, R)$, we have $h_{j,R} = P_{a_j,R}(h_{j,R})$, that is,

$$\overline{h_{j,R}(z)} = \frac{1}{\pi}\int_{B(a_j,R)} \frac{R^2\overline{h_{j,R}(\xi)}\,dv(\xi)}{(R^2-(\xi-a_j)\cdot\overline{(z-a_j)})^{n+1}}, \quad z \in B(a_j, R).$$

Hence,

$$|\bar\partial\overline{h_{j,R}(z)}| \le C\frac{\|h_{j,R}\|_{L^\infty(B(z,R),dv)}}{R} \quad \text{for } z \in \overline{B(a_j, 3R/4)}. \tag{5-6}$$

Notice that $\operatorname{supp}\psi_{j,R}\, h_{j,R} \subseteq \overline{B(a_j, 3R/4)}$, and the estimates (5-1) and (5-6) imply that

$$|\bar\partial\, \bar{f}_{1,R}| \le \sum_{j=1}^\infty |(\bar\partial\psi_{j,R})\bar{h}_{j,R}| + \sum_{j=1}^\infty \psi_{j,R}|\bar\partial(\bar{h}_{j,R})| \le C\frac{\|f\|_{L^\infty}}{R}.$$

Therefore, using (4-6) (when $q \ge 1$) and (4-15) (when $q < 1$), we have

$$\|H_{\bar{f}_{1,R}}\|^p_{F^p(\varphi)\to L^q(\varphi)} \le C\|\bar\partial\, \bar{f}_{1,R}\|_{L^\infty} \le C\frac{\|f\|_{L^\infty}}{R}.$$

The constants C above are all independent of f and R. Therefore,

$$\|H_{\bar f} - H_{\bar{f}_{2,R}}\|_{F^p(\varphi)\to L^q(\varphi)} = \|H_{\bar{f}_{1,R}}\|_{F^p(\varphi)\to L^q(\varphi)} \le C\frac{\|f\|_{L^\infty}}{R} \to 0$$

as $R \to \infty$. Finally, using (5-5) and the fact that $\mathcal{K}(F^p(\varphi), L^q(\varphi))$ is closed under the operator norm, we see that $H_{\bar f} \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$, which completes the proof. □

To deal with the case $1 \le q < p < \infty$, we use the Ahlfors–Beurling operator, which is a very well-known Calderón–Zygmund operator on $L^p(\mathbb{C})$, $1 < p < \infty$, defined as

$$\mathfrak{T}(f)(z) = \text{p.v.} -\frac{1}{\pi}\int_{\mathbb{C}} \frac{f(\xi)}{(\xi-z)^2}\,dv(\xi),$$

where p.v. means the Cauchy principal value. The Ahlfors–Beurling operator connects harmonic analysis and complex analysis, and it is of fundamental importance in several areas of mathematics including PDE and quasiconformal mappings. See [Ahlfors 2006; Astala et al. 2009] for further details and examples.

Lemma 5.1. *Suppose $1 < s < \infty$. Then there is some constant C, depending only on s, such that, for $f \in C^2(\mathbb{C}^n) \cap L^\infty$ and $j = 1, 2, \ldots, n$,*

$$\left\|\frac{\partial f}{\partial z_j}\right\|_{L^s} \le C\left\|\frac{\partial f}{\partial \bar z_j}\right\|_{L^s}. \tag{5-7}$$

Proof. We consider the case $n = 1$ first. Let $f \in C^2(\mathbb{C}) \cap L^\infty$. If $\|\partial f/\partial \bar{z}\|_{L^s} = 0$, then $f \in H(\mathbb{C}) \cap L^\infty$, which implies that the function f is constant and the estimate (5-7) follows. Next we suppose that $\|\partial f/\partial \bar{z}\|_{L^s} > 0$. Take $\psi(r) \in C^\infty(\mathbb{R})$ to be decreasing such that $\psi(x) = 1$ for $x \le 0$, $\psi(x) = 0$ for $x \ge 1$, and $0 \le -\psi'(x) \le 2$ for $x \in \mathbb{R}$. For $R > 0$ fixed, we set $\psi_R(x) = \psi(x - R)$ for $x \in \mathbb{R}$ and define $f_R(z) = f(z)\psi_R(|z|)$ for $z \in \mathbb{C}$. Since $f \in C^2(\mathbb{C}) \cap L^\infty$, it is obvious that $f_R(z) \in C_c^2(\mathbb{C})$, the set of C^2 functions on $\mathbb{R}^2$ with compact support. From Theorem 2.1.1 of [Chen and Shaw 2001], it follows that

$$f_R(z) = \frac{1}{2\pi \mathrm{i}} \int_{\mathbb{C}} \frac{\partial f_R/\partial \bar{z}}{\xi - z}\, d\xi \wedge d\bar{\xi}.$$

Notice that $\partial f_R/\partial \bar{z} = \psi_R(\partial f/\partial \bar{z}) + f(\partial \psi_R/\partial \bar{z})$. By Lemma 2 on page 52 of [Ahlfors 2006], we get

$$\frac{\partial f_R}{\partial z}(z) = \mathfrak{T}\Big(\frac{\partial f_R}{\partial \bar{z}}\Big)(z) = \mathfrak{T}\Big(\psi_R \frac{\partial f}{\partial \bar{z}}\Big)(z) + \mathfrak{T}\Big(f \frac{\partial \psi_R}{\partial \bar{z}}\Big)(z). \tag{5-8}$$

Now for $r > 0$ and $|z| < r$, when R is sufficiently large, it holds that

$$\Big|\mathfrak{T}\Big(f \frac{\partial \psi_R}{\partial \bar{z}}\Big)\Big|(z) \le \frac{\|f\|_{L^\infty}}{\pi(R-r)^2} \int_{R \le |\xi| \le R+1} dv(\xi) \le \frac{3R\|f\|_{L^\infty}}{(R-r)^2},$$

and hence

$$\Big\|\mathfrak{T}\Big(f \frac{\partial \psi_R}{\partial \bar{z}}\Big)\Big\|_{L^s(D(0,r),dv)} \le \Big\|\frac{\partial f}{\partial \bar{z}}\Big\|_{L^s}, \tag{5-9}$$

where $D(0, r) = \{z \in \mathbb{C} : |z| < r\}$. In addition, by the boundedness of $\mathfrak{T}$ on L^s (see, for example, the estimate (11) on page 53 in [Ahlfors 2006]), we get

$$\Big\|\mathfrak{T}\Big(\psi_R \frac{\partial f}{\partial \bar{z}}\Big)\Big\|_{L^s} \le C\Big\|\psi_R \frac{\partial f}{\partial \bar{z}}\Big\|_{L^s} \le C\Big\|\frac{\partial f}{\partial \bar{z}}\Big\|_{L^s}. \tag{5-10}$$

For R sufficiently large, from (5-8), (5-9) and (5-10) it follows that

$$\Big\|\frac{\partial f}{\partial z}\Big\|_{L^s(D(0,r),dv)} = \Big\|\frac{\partial f_R}{\partial z}\Big\|_{L^s(D(0,r),dv)} \le C\Big\|\frac{\partial f}{\partial \bar{z}}\Big\|_{L^s}.$$

Therefore,

$$\Big\|\frac{\partial f}{\partial z}\Big\|_{L^s} \le C\Big\|\frac{\partial f}{\partial \bar{z}}\Big\|_{L^s}. \tag{5-11}$$

Now for $n \ge 2$ and $f \in L^\infty \cap C^2(\mathbb{C}^n)$, by (5-11), we have

$$\begin{aligned}\int_{\mathbb{C}^n} \Big|\frac{\partial f}{\partial z_1}(\xi)\Big|^s dv(\xi) &= \int_{\mathbb{C}^{n-1}} dv(\xi') \int_{\mathbb{C}} \Big|\frac{\partial f}{\partial z_1}(\xi_1, \xi')\Big|^s dv(\xi_1)\\ &\le C\int_{\mathbb{C}^{n-1}} dv(\xi') \int_{\mathbb{C}} \Big|\frac{\partial f}{\partial \bar{z}_1}(\xi_1, \xi')\Big|^s dv(\xi_1).\end{aligned}$$

This implies (5-7) for $j = 1$. Similarly, (5-7) holds for $j = 2, \dots, n$, and the proof is complete. □

Proof of the case $1 \le q < p < \infty$. Notice first that if $H_f \in \mathcal{K}(F^p(\varphi), L^q(\varphi)$, then by Theorem 1.1, we have $f \in \mathrm{IDA}^{s,q}$ with $s = pq/(p - q) > 1$. We use a decomposition $f = f_1 + f_2$ as in (3-17) with $r = 1$.

Furthermore, by (5-4), we may assume that $\|f_1\|_{L^\infty} \le C\|f\|_{L^\infty}$. Then, from Lemma 5.1 it follows that

$$\|\bar\partial\, \bar f_1\|_{L^s} \le C \sum_{j=1}^n \left\|\frac{\partial \bar f}{\partial \bar z_j}\right\|_{L^s} = C \sum_{j=1}^n \left\|\frac{\partial f}{\partial z_j}\right\|_{L^s} \le C \sum_{j=1}^n \left\|\frac{\partial f}{\partial \bar z_j}\right\|_{L^s} \le C\|\bar\partial f_1\|_{L^s}.$$

We also observe that $\|M_{q,r}(\bar f_2)\|_{L^s} = \|M_{q,r}(f_2)\|_{L^s} < \infty$. Now Theorem 3.8 implies that $\bar f = \bar f_1 + \bar f_2 \in \mathrm{IDA}^{s,q}$, and hence, by Theorem 1.1, we get $H_{\bar f} \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$. □

Remark 5.2. Notice that it follows from the preceding proof that

$$\|H_{\bar f}\|_{F^p(\varphi)\to L^q(\varphi)} \le C\|H_f\|_{F^p(\varphi)\to L^q(\varphi)}.$$

6. Application to Berezin–Toeplitz quantization

As an application and further generalization of our results, we consider deformation quantization in the sense of [Rieffel 1989; 1990] and focus on one of its essential ingredients in the noncompact setting of $\mathbb{C}^n$ that involves the limit condition

$$\lim_{t\to 0} \|T_f^{(t)} T_g^{(t)} - T_{fg}^{(t)}\|_{F_t^2(\varphi)\to F_t^2(\varphi)} = 0.$$

Recently this and related questions were studied in [Bauer and Coburn 2016; Bauer et al. 2018; Fulsche 2020], which also provide further physical background and references for this type of quantization.

Recall that $\varphi \in C^2(\mathbb{C}^n)$ is real-valued and $\mathrm{Hess}_{\mathbb{R}}\,\varphi \simeq \mathrm{E}$, where E is the $2n \times 2n$-unit matrix. For $t > 0$, we set

$$d\mu_t(z) = \frac{1}{t^n} \exp\left\{-2\varphi\left(\frac{z}{\sqrt{t}}\right)\right\} dv(z)$$

and denote by $L_t^2(\varphi)$ the space of all Lebesgue measurable functions f in $\mathbb{C}^n$ such that

$$\|f\|_t = \left\{\int_{\mathbb{C}^n} |f|^2\, d\mu_t(z)\right\}^{1/2}.$$

Further, we let $F_t^2(\varphi) = L_t^2(\varphi) \cap H(\mathbb{C}^n)$. Then clearly $F_1^2(\varphi) = F^2(\varphi)$ and $L_1^2(\varphi) = L^2(\varphi)$ in terms of the spaces that were considered in the previous sections. Given $f \in L^\infty$, we use the orthogonal projection $P^{(t)}$ from $L_t^2(\varphi)$ onto $F_t^2(\varphi)$ to define the Toeplitz operator $T_f^{(t)}$ and the Hankel operator $H_f^{(t)}$, respectively, by

$$T_f^{(t)} = P^{(t)} M_f \quad \text{and} \quad H_f^{(t)} = (\mathrm{I} - P^{(t)}) M_f.$$

Let U_t be the dilation acting on measurable functions in $\mathbb{C}^n$ as

$$U_t : f \mapsto f(\cdot\sqrt{t}).$$

It is easy to verify that U_t is a unitary operator from $L_t^2(\varphi)$ to $L^2(\varphi)$ (as well as a unitary operator from $F_t^2(\varphi)$ to $F^2(\varphi)$). Further, we have $U_t P^{(t)} U_t^{-1} = P^{(1)}$, which implies that

$$U_t T_f^{(t)} U_t^{-1} = T_{f(\cdot\sqrt{t})}, \quad U_t H_f^{(t)} U_t^{-1} = H_{f(\cdot\sqrt{t})}. \tag{6-1}$$

Therefore,

$$\|T_f^{(t)}\|_{F_t^2(\varphi)\to F_t^2(\varphi)} = \|T_{f(\cdot\sqrt{t})}\|_{F^2(\varphi)\to F^2(\varphi)} \tag{6-2}$$

and

$$\|H_f^{(t)}\|_{F_t^2(\varphi)\to L_t^2(\varphi)} = \|H_{f(\,\cdot\,\sqrt{t})}\|_{F^2(\varphi)\to L^2(\varphi)}. \tag{6-3}$$

Given $f \in L^2_{\mathrm{loc}}$, for $z \in \mathbb{C}^n$ and $r > 0$ set

$$MO_{2,r}(f)(z) = \left\{\frac{1}{|B(z,r)|}\int_{B(z,r)} |f - f_{B(z,r)}|^2\,dv\right\}^{1/2}$$

where $f_S = (1/|S|)\int_S f\,dv$ for $S \subset \mathbb{C}^n$ measurable.

The following definitions of BMO and VMO are analogous to the classical definition introduced in [John and Nirenberg 1961], but they differ from those widely used in the study of Bergman and Fock spaces.

Definition 6.1. We denote by BMO the set of all $f \in L^2_{\mathrm{loc}}$ such that

$$\|f\|_* = \sup_{z\in\mathbb{C}^n,\, r>0} MO_{2,r}(f)(z) < \infty$$

and by VMO the set of all $f \in \mathrm{BMO}$ such that

$$\lim_{r\to 0}\sup_{z\in\mathbb{C}^n} MO_{2,r}(f)(z) = 0.$$

Definition 6.2. We define BDA_* to be the family of all $f \in L^2_{\mathrm{loc}}$ such that

$$\|f\|_{\mathrm{BDA}_*} = \sup_{z\in\mathbb{C}^n, r>0} G_{2,r}(f)(z) < \infty$$

and VDA_* to be the subspace of all $f \in \mathrm{BDA}_*$ such that

$$\lim_{r\to 0}\sup_{z\in\mathbb{C}^n} G_{2,r}(f)(z) = 0.$$

Given a family X of functions on $\mathbb{C}^n$, we set $\overline{X} = \{\bar{f} : f \in X\}$.

Proposition 6.3. *It holds that*

$$\mathrm{BMO} = \mathrm{BDA}_* \cap \overline{\mathrm{BDA}_*} \quad \textit{and} \quad \mathrm{VMO} = \mathrm{VDA}_* \cap \overline{\mathrm{VDA}_*}.$$

Furthermore, we have

$$\|f\|_{\mathrm{BMO}_*} \simeq \|f\|_{\mathrm{BDA}_*} + \|\bar{f}\|_{\mathrm{BDA}_*} \tag{6-4}$$

for $f \in L^2_{\mathrm{loc}}$.

Proof. From a careful inspection of the proof of Proposition 2.5 in [Hu and Wang 2018], it follows that there is a constant $C > 0$ such that, for $f \in L^2_{\mathrm{loc}}$ and $z \in \mathbb{C}^n$, $r > 0$, there is a constant $c(z)$ for which

$$\left\{\frac{1}{|B(z,r)|}\int_{B(z,r)} |f - c(z)|^2\,dv\right\}^{1/2} \le C\{G_{2,r}(f)(z) + G_{2,r}(\bar{f})(z)\}.$$

It is easy to verify that

$$MO_{2,r}(f)(z) \le \left\{\frac{1}{|B(z,r)|}\int_{B(z,r)} |f - c(z)|^2\,dv\right\}^{1/2},$$

and hence

$$MO_{2,r}(f)(z) \le C\{G_{2,r}(f)(z) + G_{2,r}(\bar{f})(z)\}.$$

On the other hand, by definition, we have

$$G_{2,r}(f)(z) \le MO_{2,r}(f)(z).$$

Thus, we have C_1 and C_2, independent of f, r and z, such that

$$\begin{aligned} C_1\{G_{2,r}(f)(z) + G_{2,r}(\bar{f})(z)\} &\le MO_{2,r}(f)(z) \\ &\le C_2\{G_{2,r}(f)(z) + G_{2,r}(\bar{f})(z)\}. \end{aligned} \tag{6-5}$$

Therefore, $f \in \mathrm{BMO}$ (or $f \in \mathrm{VMO}$) if and only if $f \in \mathrm{BDA}_* \cap \overline{\mathrm{BDA}_*}$ (or $f \in \mathrm{VDA}_* \cap \overline{\mathrm{VDA}_*}$). The estimate in (6-4) follows from (6-5). □

Theorem 6.4. *Suppose $f \in L^\infty$. Then for all $g \in L^\infty$, it holds that*

$$\lim_{t\to 0} \|T_f^{(t)} T_g^{(t)} - T_{fg}^{(t)}\|_{F_t^2(\varphi)\to F_t^2(\varphi)} = 0 \tag{6-6}$$

if and only if $f \in \overline{\mathrm{VDA}_*}$.

Proof. Given $f \in L^\infty$, it follows from (6-3) that

$$\|(H_{\bar{f}}^{(t)})^*\|_{L_t^2(\varphi)\to F_t^2(\varphi)} = \|H_{\bar{f}}^{(t)}\|_{F_t^2(\varphi)\to L_t^2(\varphi)} = \|H_{f(\cdot\sqrt{t})}\|_{F^2(\varphi)\to L^2(\varphi)}.$$

This and Theorem 1.1 imply

$$\frac{1}{C}\|G_{2,1}(f(\cdot\sqrt{t}))\|_{L^\infty} \le \|(H_{\bar{f}}^{(t)})^*\|_{L_t^2(\varphi)\to F_t^2(\varphi)} \le C\|G_{2,1}(f(\cdot\sqrt{t}))\|_{L^\infty}, \tag{6-7}$$

where the constant C is independent of f and t.

Suppose $f \in \overline{\mathrm{VDA}_*}$. Then, by definition, we have

$$\lim_{r\to 0} \sup_{z\in\mathbb{C}^n} G_{2,r}(\bar{f})(z) = 0.$$

It is easy to verify that

$$G_{2,1}(f(\cdot\sqrt{t}))(z) = G_{2,\sqrt{t}}(f)(z\sqrt{t}).$$

Now by (6-7), we get

$$\lim_{t\to 0} \|(H_{\bar{f}}^{(t)})^*\|_{L_t^2(\varphi)\to F_t^2(\varphi)} \le C \lim_{t\to 0} \|G_{2,\sqrt{t}}(\bar{f})\|_{L^\infty} = 0. \tag{6-8}$$

In addition, for $f, g \in L^\infty$, it is easy to verify that

$$T_f^{(t)} T_g^{(t)} - T_{fg}^{(t)} = -(H_{\bar{f}}^{(t)})^* H_g^{(t)}. \tag{6-9}$$

Therefore, for all $g \in L^\infty$,

$$\lim_{t\to 0} \|T_f^{(t)} T_g^{(t)} - T_{fg}^{(t)}\|_{F_t^2(\varphi)\to F_t^2(\varphi)} \le \|g\|_{L^\infty} \lim_{t\to 0} \|(H_{\bar{f}}^{(t)})^*\|_{L_t^2(\varphi)\to F_t^2(\varphi)} = 0,$$

which gives (6-6).

Conversely, suppose that (6-6) holds for every $g \in L^\infty$. Let $g = \bar{f} \in L^\infty$. Then it follows from (6-9) that

$$\begin{aligned}\lim_{t\to 0} \|H_{\bar{f}}^{(t)}\|^2_{F_t^2(\varphi)\to L_t^2(\varphi)} &= \lim_{t\to 0} \|(H_{\bar{f}}^{(t)})^* H_{\bar{f}}^{(t)}\|_{F_t^2(\varphi)\to F_t^2(\varphi)} \\ &= \lim_{t\to 0} \|T_f^{(t)} T_{\bar{f}}^{(t)} - T_{|f|^2}^{(t)}\|_{F_t^2(\varphi)\to F_t^2(\varphi)} = 0.\end{aligned}$$

This and (6-7) imply that $f \in \overline{\mathrm{VDA}_*}$. □

Combining Proposition 6.3 with Theorem 6.4, we obtain the following corollary, which is the main result of [Bauer et al. 2018] when $\varphi(z) = \frac{1}{8}|z|^2$.

Corollary 6.5. *Suppose $f \in L^\infty$. Then for all $g \in L^\infty$, it holds that*

$$\lim_{t\to 0} \|T_f^{(t)} T_g^{(t)} - T_{fg}^{(t)}\| = 0 \quad \textit{and} \quad \lim_{t\to 0} \|T_g^{(t)} T_f^{(t)} - T_{fg}^{(t)}\| = 0 \tag{6-10}$$

if and only if $g \in \mathrm{VMO}$. Here $\|\cdot\| = \|\cdot\|_{F_t^2(\varphi)\to F_t^2(\varphi)}$.

7. Further remarks

For $1 \le p, q < \infty$, we have characterized those $f \in \mathcal{S}$ for which $H_f : F^p(\varphi) \to L^q(\varphi)$ is bounded (or compact). For small exponents $0 < p < q < 1$, we have proved that this characterization remains true for compactness when $f \in L^\infty$. We also note that when $p \le q$ and $q \ge 1$, boundedness and compactness of Hankel operators $H_f : F^p(\varphi) \to L^p(\varphi)$ depend on q (see Remark 3.2 and Theorem 1.1), while for $p > q$ we cannot say the same — we note that we have no statement analogous to Remark 3.2 for $\mathrm{IDA}^{s,q}$.

Moreover, for harmonic symbols $f \in \mathcal{S}$ and $0 < p, q < \infty$, using the Hardy–Littlewood theorem on the submean value (see Lemma 2.1 of [Hu et al. 2007], for example), we are able to characterize boundedness of $H_f : F^p(\varphi) \to L^q(\varphi)$ with the space $\mathrm{IDA}^{s,q}$. We will return to this topic in a future publication.

We also note that the space $F^\infty(\varphi)$ does not appear in our results because Γ is not dense in it. Instead, it may be possible to consider the space

$$f^\infty(\varphi) = \{f \in F^\infty(\varphi) : fe^{-\varphi} \in C_0(\mathbb{C}^n)\},$$

which can be viewed as the closure of Γ in $F^\infty(\varphi)$, and extend our results to this setting.

Regarding weights, the Fock spaces studied in this paper are defined with weights $\varphi \in C(\mathbb{C}^n)$ satisfying $\mathrm{Hess}_{\mathbb{R}}\, \varphi \simeq \mathrm{E}$. As stated in Section 2A, these weights are contained in the class considered in [Schuster and Varolin 2012]. Now, we note that for the weights φ in that work, $i\partial\bar{\partial}\varphi \simeq \omega_0$, and from Hörmander's theorem on the canonical solution to the $\bar{\partial}$-equation it follows that

$$\|H_f g\|^2_{2,\varphi} \le \int_{\mathbb{C}^n} |g\bar{\partial} f|^2_{i\partial\bar{\partial}} e^{-2\varphi}\, dv \le C\|g|\bar{\partial} f|\|^2_{2,\varphi},$$

and hence we know that the conclusions of Theorem 1.1 remain true when $q = 2$ (see Theorem 4.3 of [Hu and Virtanen 2022]). Upon these observations, we raise the following conjecture.

Conjecture 1. Suppose $\varphi \in C^2(\mathbb{C}^n)$ satisfying $i\partial\bar{\partial}\varphi \simeq \omega_0$. Then for $f \in \mathcal{S}$ and $0 < p, q < \infty$, $H_f \in \mathcal{B}(F^p(\varphi), L^q(\varphi))$ if and only if $f \in \mathrm{IDA}^{s,q}$, where $s = pq/(p-q)$ if $p > q$ and $s = \infty$ if $p \le q$.

In the literature, there are a number of interesting results on the simultaneous boundedness (and compactness) of Hankel operators H_f and $H_{\bar{f}}$. These types of characterizations often involve the function spaces BMO^q and $\mathrm{IMO}^{s,q}$ in their conditions; see, e.g., [Hu and Wang 2018; Zhu 2012]. For $1 \le q < \infty$ and $1 \le s \le \infty$, set $\overline{\mathrm{IDA}^{s,q}} = \{\bar{f} : f \in \mathrm{IDA}^{s,q}\}$. Then Proposition 2.5 of [Hu and Wang 2018] shows that $\mathrm{IDA}^{s,q} \cap \overline{\mathrm{IDA}^{s,q}} = \mathrm{IMO}^{s,q}$ and the results of Section 4 provide a description of the simultaneous boundedness (or compactness) of H_f and $H_{\bar{f}}$ as seen in the following theorem, where as before, we set $s = pq/(p-q)$ if $p > q$ and $s = \infty$ if $p \le q$.

Theorem 7.1. *Let $\varphi \in C^2(\mathbb{C}^n)$ be real-valued,* $\mathrm{Hess}_{\mathbb{R}}\, \varphi \simeq \mathrm{E}$, *and let $f \in \mathcal{S}$. For $1 \le p, q < \infty$, Hankel operators H_f and $H_{\bar{f}}$ are simultaneously bounded from $F^p(\varphi)$ to $L^q(\varphi))$ if and only if $f \in \mathrm{IMO}^{s,q}$.*

We state one more conjecture related to Theorem 1.2, in which we proved that for $f \in L^\infty$ and $0 < p < \infty$, H_f is compact on $F^p(\varphi)$ if and only if $H_{\bar{f}}$ in compact on $F^p(\varphi)$. Recall that this phenomenon does not occur for Hankel operators on the Bergman space or on the Hardy space. As predicted in [Zhu 2012], and verified for Hankel operators on the weighted Fock spaces $F^p(\alpha)$ with $1 < p < \infty$ in [Hagger and Virtanen 2021], a partial explanation for this difference is the lack of bounded holomorphic or harmonic functions on the entire complex plane. From this point of view it is natural to suggest that a similar result should remain true for Hankel operators mapping from $F^p(\varphi)$ to $L^q(\varphi)$.

Conjecture 2. Suppose that $\varphi \in C^2(\mathbb{C}^n)$ satisfies $i\partial\bar{\partial}\varphi \simeq \omega_0$ and $0 < p, q < \infty$. Then for $f \in L^\infty$, $H_f \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$ if and only if $H_{\bar{f}} \in \mathcal{K}(F^p(\varphi), L^q(\varphi))$.

Notice that $\mathrm{IDA}^{s,q} \cap L^\infty$ is a Banach algebra under the norm $\|\cdot\|_{\mathrm{IDA}^{s,q}} + \|\cdot\|_\infty$. We can also express Conjecture 2 in algebraic terms; that is, we conjecture that $\mathrm{IDA}^{s,q} \cap L^\infty$ on $\mathbb{C}^n$ is closed under the conjugate operation $f \mapsto \bar{f}$, where $1 < s \le \infty$ and $0 < q < \infty$.

Related to our work on quantization and Theorem 6.4 in particular, we conclude this section with the following problem: characterize those $f \in L^\infty$ for which it holds that

$$\lim_{t \to 0} \|T_f^{(t)} T_g^{(t)} - T_{fg}^{(t)}\|_{S_2} = 0$$

for all $g \in L^\infty$, where $\|\cdot\|_{S_2}$ stands for the Hilbert–Schmidt norm. It would also be important to consider this question for other Schatten classes S_p.

Acknowledgments

The authors thank the referee for valuable comments and suggestions.

Z. Hu was partially supported by the National Natural Science Foundation of China (12071130, 12171150). J. Virtanen was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) grant EP/T008636/1.

References

[Ahlfors 2006] L. V. Ahlfors, *Lectures on quasiconformal mappings*, 2nd ed., Univ. Lect. Series **38**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[Astala et al. 2009] K. Astala, T. Iwaniec, and G. Martin, *Elliptic partial differential equations and quasiconformal mappings in the plane*, Princeton Math. Series **48**, Princeton Univ. Press, 2009. MR Zbl

[Axler 1986] S. Axler, "The Bergman space, the Bloch space, and commutators of multiplication operators", *Duke Math. J.* **53**:2 (1986), 315–332. MR Zbl

[Bauer 2005] W. Bauer, "Mean oscillation and Hankel operators on the Segal–Bargmann space", *Integral Equations Operator Theory* **52**:1 (2005), 1–15. MR Zbl

[Bauer and Coburn 2016] W. Bauer and L. A. Coburn, "Uniformly continuous functions and quantization on the Fock space", *Bol. Soc. Mat. Mex.* (3) **22**:2 (2016), 669–677. MR Zbl

[Bauer et al. 2018] W. Bauer, L. A. Coburn, and R. Hagger, "Toeplitz quantization on Fock space", *J. Funct. Anal.* **274**:12 (2018), 3531–3551. MR Zbl

[Berger and Coburn 1986] C. A. Berger and L. A. Coburn, "Toeplitz operators and quantum mechanics", *J. Funct. Anal.* **68**:3 (1986), 273–299. MR Zbl

[Berger and Coburn 1987] C. A. Berger and L. A. Coburn, "Toeplitz operators on the Segal–Bargmann space", *Trans. Amer. Math. Soc.* **301**:2 (1987), 813–829. MR Zbl

[Berger and Coburn 1994] C. A. Berger and L. A. Coburn, "Heat flow and Berezin–Toeplitz estimates", *Amer. J. Math.* **116**:3 (1994), 563–590. MR Zbl

[Berndtsson and Andersson 1982] B. Berndtsson and M. Andersson, "Henkin–Ramirez formulas with weight factors", *Ann. Inst. Fourier* (*Grenoble*) **32**:3 (1982), 91–110. MR Zbl

[Berndtsson and Ortega Cerdà 1995] B. Berndtsson and J. Ortega Cerdà, "On interpolation and sampling in Hilbert spaces of analytic functions", *J. Reine Angew. Math.* **464** (1995), 109–128. MR Zbl

[Chen and Shaw 2001] S.-C. Chen and M.-C. Shaw, *Partial differential equations in several complex variables*, AMS/IP Stud. Adv. Math. **19**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Cho and Zhu 2012] H. R. Cho and K. Zhu, "Fock–Sobolev spaces and their Carleson measures", *J. Funct. Anal.* **263**:8 (2012), 2483–2506. MR Zbl

[Christ 1991] M. Christ, "On the $\overline{\partial}$ equation in weighted L^2 norms in $\mathbb{C}^1$", *J. Geom. Anal.* **1**:3 (1991), 193–230. MR Zbl

[Coburn et al. 2021] L. Coburn, M. Hitrik, J. Sjöstrand, and F. White, "Weyl symbols and boundedness of Toeplitz operators", *Math. Res. Lett.* **28**:3 (2021), 681–696. MR Zbl

[Delin 1998] H. Delin, "Pointwise estimates for the weighted Bergman projection kernel in $\mathbb{C}^n$, using a weighted L^2 estimate for the $\overline{\partial}$ equation", *Ann. Inst. Fourier* (*Grenoble*) **48**:4 (1998), 967–997. MR Zbl

[Fulsche 2020] R. Fulsche, "Correspondence theory on p-Fock spaces with applications to Toeplitz algebras", *J. Funct. Anal.* **279**:7 (2020), art. id. 108661. MR Zbl

[Hagger and Virtanen 2021] R. Hagger and J. A. Virtanen, "Compact Hankel operators with bounded symbols", *J. Operator Theory* **86**:2 (2021), 317–329. MR Zbl

[Hu and Lu 2019] Z. Hu and J. Lu, "Hankel operators on Bergman spaces with regular weights", *J. Geom. Anal.* **29**:4 (2019), 3494–3519. MR Zbl

[Hu and Lv 2014] Z. Hu and X. Lv, "Toeplitz operators on Fock spaces $F^p(\varphi)$", *Integral Equations Operator Theory* **80**:1 (2014), 33–59. MR Zbl

[Hu and Virtanen 2020] Z. Hu and J. A. Virtanen, "Fredholm Toeplitz operators with VMO symbols and the duality of generalized Fock spaces with small exponents", *Proc. Roy. Soc. Edinburgh Sect. A* **150**:6 (2020), 3163–3186. MR Zbl

[Hu and Virtanen 2022] Z. Hu and J. A. Virtanen, "Schatten class Hankel operators on the Segal–Bargmann space and the Berger–Coburn phenomenon", *Trans. Amer. Math. Soc.* **375**:5 (2022), 3733–3753. MR Zbl

[Hu and Wang 2018] Z. Hu and E. Wang, "Hankel operators between Fock spaces", *Integral Equations Operator Theory* **90**:3 (2018), art. id. 37. MR Zbl

[Hu et al. 2007] Z. Hu, M. Pavlović, and X. Zhang, "The mixed norm spaces of polyharmonic functions", *Potential Anal.* **27**:2 (2007), 167–182. MR Zbl

[Janson et al. 1987] S. Janson, J. Peetre, and R. Rochberg, "Hankel forms and the Fock space", *Rev. Mat. Iberoam.* **3**:1 (1987), 61–138. MR Zbl

[John and Nirenberg 1961] F. John and L. Nirenberg, "On functions of bounded mean oscillation", *Comm. Pure Appl. Math.* **14** (1961), 415–426. MR Zbl

[Krantz 1992] S. G. Krantz, *Function theory of several complex variables*, 2nd ed., Wadsworth & Brooks/Cole, Pacific Grove, CA, 1992. MR Zbl

[Li 1994] H. Li, "Hankel operators on the Bergman spaces of strongly pseudoconvex domains", *Integral Equations Operator Theory* **19**:4 (1994), 458–476. MR Zbl

[Lindholm 2001] N. Lindholm, "Sampling in weighted L^p spaces of entire functions in $\mathbb{C}^n$ and estimates of the Bergman kernel", *J. Funct. Anal.* **182**:2 (2001), 390–426. MR Zbl

[Luecking 1992] D. H. Luecking, "Characterizations of certain classes of Hankel operators on the Bergman spaces of the unit disk", *J. Funct. Anal.* **110**:2 (1992), 247–271. MR Zbl

[Lv 2019] X. Lv, "Hankel operators on Fock spaces $F^p(\varphi)$", *Complex Var. Elliptic Equ.* **64**:9 (2019), 1522–1533. MR Zbl

[Pau et al. 2016] J. Pau, R. Zhao, and K. Zhu, "Weighted BMO and Hankel operators between Bergman spaces", *Indiana Univ. Math. J.* **65**:5 (2016), 1639–1673. MR Zbl

[Perälä et al. 2014] A. Perälä, A. Schuster, and J. A. Virtanen, "Hankel operators on Fock spaces", pp. 377–390 in *Concrete operators, spectral theory, operators in harmonic analysis and approximation* (Seville, Spain, 2011), edited by M. Cepedello Boiso et al., Oper. Theory Adv. Appl. **236**, Birkhäuser, Basel, 2014. MR Zbl

[Rieffel 1989] M. A. Rieffel, "Deformation quantization of Heisenberg manifolds", *Comm. Math. Phys.* **122**:4 (1989), 531–562. MR Zbl

[Rieffel 1990] M. A. Rieffel, "Deformation quantization and operator algebras", pp. 411–423 in *Operator theory: operator algebras and applications, I* (Durham, NH, 1988), edited by W. B. Arveson and R. G. Douglas, Proc. Sympos. Pure Math. **51.1**, Amer. Math. Soc., Providence, RI, 1990. MR Zbl

[Rudin 1973] W. Rudin, *Functional analysis*, McGraw-Hill, New York, 1973. MR Zbl

[Schuster and Varolin 2012] A. P. Schuster and D. Varolin, "Toeplitz operators and Carleson measures on generalized Bargmann–Fock spaces", *Integral Equations Operator Theory* **72**:3 (2012), 363–392. MR Zbl

[Stroethoff 1992] K. Stroethoff, "Hankel and Toeplitz operators on the Fock space", *Michigan Math. J.* **39**:1 (1992), 3–16. MR Zbl

[Zhu 2005] K. Zhu, *Spaces of holomorphic functions in the unit ball*, Grad. Texts in Math. **226**, Springer, 2005. MR Zbl

[Zhu 2012] K. Zhu, *Analysis on Fock spaces*, Grad. Texts in Math. **263**, Springer, 2012. MR Zbl

Received 10 Nov 2020. Revised 24 Mar 2022. Accepted 29 Apr 2022.

ZHANGJIAN HU: huzj@zjhu.edu.cn
Huzhou University, Huzhou, China

JANI A. VIRTANEN: j.a.virtanen@reading.ac.uk
jani.virtanen@helsinki.fi
University of Reading, Reading, United Kingdom

and

University of Helsinki, Helsinki, Finland

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2079–2107
DOI: 10.2140/apde.2023.16.2079

GLOBAL STABILITY OF SPACETIMES WITH SUPERSYMMETRIC COMPACTIFICATIONS

LARS ANDERSSON, PIETER BLUE, ZOE WYATT AND SHING-TUNG YAU

This paper proves the stability, with respect to the evolution determined by the vacuum Einstein equations, of the Cartesian product of higher-dimensional Minkowski space with a compact, Ricci-flat Riemannian manifold that admits a spin structure and a nonzero parallel spinor. Such a product includes the example of Calabi–Yau and other special holonomy compactifications, which play a central role in supergravity and string theory. The stability result proved in this paper shows that Penrose's instability argument [2003] does not apply to localised perturbations.

1. Introduction

Let $(\mathbb{R}^{1+n}, \eta_{\mathbb{R}^{1+n}})$ be the $(1+n)$-dimensional Minkowski spacetime, and let (K, k) be a compact, Ricci-flat Riemannian manifold that has a cover that admits a spin structure and a nonzero parallel spinor. The spacetime $\mathcal{M} = \mathbb{R}^{1+n} \times K$ with metric

$$\hat{g} = \eta_{\mathbb{R}^{1+n}} + k \tag{1}$$

is globally hyperbolic and Ricci flat, i.e, it is a solution to the $(1+n+d)$-dimensional vacuum Einstein equations. Such spacetimes play an essential role in supergravity and string theory [Candelas et al. 1985]. In this paper we refer to $(\mathcal{M}, \hat{g})$ as a spacetime with a supersymmetric (SUSY) compactification and (K, k) as the internal manifold.

The simplest spacetime with a supersymmetric compactification, which has been studied since the 1920s, is the Kaluza–Klein spacetime $(\mathbb{R}^{1+3} \times \mathbb{S}^1_\theta, \eta_{\mathbb{R}^{1+3}} + \mathrm{d}\theta^2)$ [Kaluza 1921; Klein 1926]. As shown by Witten in an influential paper [1982], this spacetime is unstable at the semiclassical level. Nonetheless in the same work Witten argued that the spacetime should be classically linearly stable.

By contrast, Penrose has sketched an argument intended to show that spacetimes with supersymmetric compactifications are generically classically unstable, for every dimension n and all internal manifolds, except possibly when the internal manifold is a flat d-dimensional torus [Penrose 2003; 2005]. There are theorems motivated by these considerations that generalise the classical singularity theorems to trapped surfaces of arbitrary codimension [Cipriani and Senovilla 2019; Galloway and Senovilla 2010]. However, the results of the present paper show that for spacetimes with supersymmetric compactifications the instability argued by Penrose does not hold for $n \geq 9$, and we conjecture here that in fact stability holds for $n \geq 3$. The nonnegativity of the spectrum of the Lichnerowicz Laplacian on symmetric 2-tensors, which holds for the internal spaces by the result of Dai, Wang, and Wei [Dai et al. 2005], plays a crucial role

MSC2020: 35L15, 35Q76, 53C25, 83E30.
Keywords: Einstein equation, stability, initial-value problem, SUSY.

in our stability proof. In fact, this nonnegativity, which is conjectured to hold for all compact Ricci-flat manifolds, is sufficient for our result. See Section 2A for details.

In order to state our main theorem, we need to introduce some notation. For the product spacetime $\mathbb{R}^{1+n} \times K$ we denote spacetime indices by $\alpha, \mu, \nu, \ldots$, Minkowski indices by $i, j, k, \ldots$, and internal indices by $A, B, C, \ldots$. For a general pseudo-Riemannian metric g, let $\nabla[g]$ denote its Levi-Civita connection, $\mathrm{Riem}[g]$ its Riemann curvature tensor, $\mathrm{Ric}[g]$ its Ricci curvature, and $\mathrm{d}\mu_g$ its volume form. Define the contraction

$$(R[g] \circ u)_{\mu\nu} = R_{\mu\rho\nu\lambda}[g] u^{\rho\lambda}, \tag{2}$$

which acts on symmetric (0, 2)-tensors $u_{\mu\nu}$. Given the supersymmetric spacetime metric $\hat{g}$ on $\mathbb{R}^{1+n} \times K$, let

$$(g_E)_{\mu\nu} = \hat{g}_{\mu\nu} + 2(\mathrm{d}t)_\mu (\mathrm{d}t)_\nu, \tag{3}$$

where $\mathrm{d}t$ is with respect to the standard Cartesian coordinates on $\mathbb{R}^{1+n}$. On K and $\mathbb{R}^{1+n} \times K$, define the inner products on (0, 2) tensors, respectively, as

$$\langle u, v\rangle_k = k^{AC} k^{BD} u_{AB} v_{CD} \quad \text{and} \quad \langle u, v\rangle_E = g_E^{\mu\nu} g_E^{\rho\sigma} u_{\mu\rho} v_{\nu\sigma}. \tag{4}$$

Define $|u|_k = (\langle u, u\rangle_k)^{1/2}$, and similarly for $|u|_E$.

The following is our main result. The details of some of the concepts appearing in the statement of the theorem appear in Definitions 2.10, 2.11, 2.12, 2.14 and Theorem 2.15.

Theorem 1.1. *Let $n, d \in \mathbb{Z}^+$ be such that $n \geq 9$, and let $N \in \mathbb{Z}^+$ be sufficiently large. Consider a spacetime $(\mathbb{R}^{1+n} \times K, \hat{g} = \eta_{\mathbb{R}^{1+n}} + k)$ with a supersymmetric compactification. Let g_S denote the Schwarzschild metric in the $\eta_{\mathbb{R}^{1+n}}$-wave gauge with mass parameter $C_S \geq 0$.*

There is an $\epsilon > 0$ such that if $(\mathbb{R}^n \times K, \gamma, \kappa)$ is an initial data set satisfying that outside the unit ball the initial data coincides with the product of Schwarzschild initial data with the unperturbed internal metric (i.e., $\gamma = g_S + k$ and $\kappa = 0$ where $|x| \geq 1$) and satisfying

$$\sum_{|I|\leq N} \|\nabla[\gamma]^I (\gamma - \hat{g}|_{t=0})\|^2_{L^2(\mathbb{R}^n\times K)} + \sum_{|I|\leq N-1} \|\nabla[\gamma]^I \kappa\|^2_{L^2(\mathbb{R}^n\times K)} + C_S^2 \leq \epsilon, \tag{5}$$

then there is a solution g of the vacuum Einstein equations on $\mathbb{R}^{1+n} \times K$ with initial data $(\mathbb{R}^n \times K, \gamma, \kappa)$ and satisfying the $\hat{g}$-wave gauge. There is the bound

$$\sup_{(t,x^i,\omega)\in\Sigma_s\times K} t^{2\delta(n)} |g(t, x^i, \omega) - \hat{g}(t, x^i, \omega)|^2_E \lesssim \epsilon, \tag{6}$$

where the decay rate is given by

$$\delta(n) = \tfrac{1}{4}(n-2). \tag{7}$$

Finally $(\mathbb{R}^{1+n} \times K, g)$ is globally hyperbolic and causally geodesically complete.

The stability result obtained in Theorem 1.1 covers a large class of product spacetimes, including many special holonomy compactifications relevant in supergravity and string theory. Although this paper succeeds in its goal of providing a counterexample to the dimension-independent argument in

[Penrose 2003], from a PDE perspective, Theorem 1.1 should be seen as a preliminary result, and we expect that the assumptions that $n \geq 9$ and that the Cauchy data is Schwarzschild near infinity can be relaxed. In fact we make the following conjecture.

Conjecture 1.2. *Spacetimes with a supersymmetric compactification*[1] *and $n = 3$ are nonlinearly stable.*

As explained below, this paper uses a relatively simple vector field argument, while, for example, the proof of global stability for the coupled Einstein–Klein–Gordon system in (1+3)-dimensions [LeFloch and Ma 2016] has required combining vector field arguments with estimates arising from control on the fundamental solution for the wave equation. Such detailed analysis is beyond the scope of this paper, but we intend to explore this in future work. Note that our current method can be easily used to show linear stability as far as $n = 3$.

The decay rate of $|h| \lesssim t^{-\delta(n)}$ arises essentially as a linear estimate. The linearisation of the Einstein equation is

$$(\Box_\eta + \Delta_k + 2R[\hat{g}] \circ) h_{\mu\nu} = 0. \tag{8}$$

To study conservation properties of the linear equations we introduce a novel stress-energy tensor

$$T[h]^\mu{}_\nu = \hat{g}^{\mu\alpha} \langle \nabla[\hat{g}]_\alpha h, \nabla[\hat{g}]_\nu h \rangle_E - \tfrac{1}{2} \hat{g}^{\alpha\beta} \langle \nabla[\hat{g}]_\beta h, \nabla[\hat{g}]_\alpha h \rangle_E \delta^\mu_\nu + \langle R[\hat{g}] \circ h, h \rangle_E \delta^\mu_\nu, \tag{9}$$

which is specifically adapted to the tensorial operator appearing in (8). The conditions on (K, k) imply (see Section 2A) that the energy integral derived from (9) is nonnegative.

The conditions on (K, k) imply that the operator $-(\Delta_k + 2R\circ)$ has a nonnegative discrete spectrum, so a spectral decomposition can be applied to solutions h of the linearised Einstein equation (8). The spectral component corresponding to the zero eigenvalue satisfies an effective wave equation $\Box_\eta (h^0)_{\mu\nu} = 0$, and the components corresponding to positive eigenvalues λ satisfy effective Klein–Gordon equations $(\Box_\eta - \lambda)(h^\lambda)_{\mu\nu} = 0$. A decomposition of this type has been used in the analysis of wave guides, where K is replaced by a compact subset of $\mathbb{R}^d$ with Neumann boundary conditions; see e.g., [Metcalfe and Stewart 2008; Metcalfe et al. 2005]. When applying the vector field method to the wave and Klein–Gordon equations, there is a unified approach using a basic energy of the form $\int \sum_{i=0}^n |\partial_i h|^2 + \lambda |h|^2 \, \mathrm{d}\mu$ that can be strengthened by commuting the equation with Γ, the set of generators of translations, rotations, and boosts. The use of this set of vector fields in the vector field method, with particular application to Klein–Gordon equations, goes back to [Klainerman 1985].

This unified approach then bifurcates: the Klein–Gordon equation does not admit any further commuting first-order operators but the energy has a nonvanishing lower-order term $\lambda|h|^2$; in contrast, the wave equation allows for commutation with the generator of dilations, $S = t\partial_t + r\partial_r$, but the lower-order term in the energy vanishes. For the quasilinear Einstein equation, we refrain from performing a spectral decomposition into wave and Klein–Gordon components. Thus, we use only the unified part of the approach (following especially the treatment of quasilinear Klein–Gordon equations in [Hörmander 1997]),

[1] Recall that the definition of a spacetime with a supersymmetric compactification, as introduced in the opening paragraph of this paper, includes the assumption that the spacetime is a fibre bundle with base space $(\mathbb{R}^{1+n}, \eta_{\mathbb{R}^{1+n}})$. Stability for a certain class of cosmological spacetimes as base spaces is proved in [Branding et al. 2019].

leaving us with a decay rate that is far from the sharp decay rates of the wave and Klein–Gordon equations. In particular, the vector field method can be used to prove decay rates, for the wave and Klein–Gordon equations, of $t^{-(n-1)/2}$ and $t^{-n/2}$, respectively.

In light of this, it seems likely that some novel refinement should allow for a significantly better decay rate than $t^{-\delta(n)}$ with $\delta(n) = \frac{1}{4}(n-2)$. This paper contains two types of refinement. First, the decay rate is shown to be $s^{-2\delta(n)}$, where $s^2 = t^2 - x^2$ inside light cones. The exponent $2\delta(n) = \frac{1}{2}(n-2)$ is much closer to the decay rate for the wave and Klein–Gordon equations. Second, the same decay rates are proved for $\Gamma^I h$ as for h, but, since the Γ contain t- and x-dependent weights, with respect to a translation invariant basis in Minkowski space, derivatives decay faster than the field h itself.

Having obtained a linear estimate that improves with increasing n, we take n large enough that $2\delta(n) - 2 > 1$, so that the nonlinear terms decay sufficiently fast for the linear estimates to remain valid. In particular, we take n large enough that we can ignore all nonlinear structure in the Einstein equation. It is well known that global existence results for semilinear equations in (1+3)-dimensions depend delicately on the nonlinearities, for example the null condition [Klainerman 1986]. Christodoulou and Klainerman [1993] used the vector field method to prove the stability of Minkowski spacetime. One of the major advances in the simplified vector field argument in [Lindblad and Rodnianski 2003; 2005; 2010] was the introduction of the weak null condition and the observation that the Einstein equations in the harmonic gauge satisfy this condition. LeFloch and Ma [2016] identified the relevant nonlinear structures for Klein–Gordon equations coupled to the (1+3)-dimensional Einstein equation.

The dimension of the compact manifold only appears in the required regularity of the initial data, which is given explicitly in Theorem 5.1. The restriction to initial data which is exactly Schwarzschild outside of a compact set mirrors the proof of Minkowski stability in (1+3)-dimensions by Lindblad and Rodnianski [2005].

Background and previous work. Theories of higher-dimensional gravity are of great interest in supergravity and string theory as possible models of quantum gravity. Many of these theories are built around the spacetimes with supersymmetric compactifications discussed above.

The background spacetimes considered in this paper are of the form $\mathbb{R}^{n+1} \times K$, with K compact and Ricci flat, and are hence anisotropic. Among the first stability results for anisotropic spacetimes of a related form was the proof of future stability of flat cosmological spacetimes of the form $M^3 \times S^1$, where M^3 is a flat (2+1)-dimensional Milne spacetime with metric $-dt^2 + t^2 H^2$ and H^2 is a hyperbolic surface, was considered by Choquet-Bruhat and Moncrief [2001]. See also [Andersson 2014; Reiris 2010].

Until now, the only nonlinear stability results for spacetimes with supersymmetric compactification have concerned the simplest Kaluza–Klein case when the internal space is the circle $\mathbb{S}^1$, or in slightly more generality, the flat d-dimensional torus. It was shown by one of the authors [Wyatt 2018] that this spacetime is classically stable to toroidal-independent perturbations. A model problem to remove this restriction with toroidal internal space has recently appeared [Huneau and Stingo 2021]. We remark that in the physics literature, these are known as zero-mode perturbations. An analogous result for cosmological Kaluza–Klein spacetimes, where the Minkowski spacetime is replaced by the four-dimensional Milne spacetime, has also recently been shown [Branding et al. 2019].

The spacetimes of importance in supergravity and string theory involve a nontrivial (i.e., nontoroidal) internal manifold with parallel spinors, such as a Calabi–Yau, G_2 or Spin(7) manifold. Note that a solution of the 10- or 11-dimensional *vacuum* Einstein equations can be considered as a particular solution of the supergravity equations. Local-in-time existence results are known for both the vacuum Einstein equations [Choquet-Bruhat 1952; Choquet-Bruhat and Geroch 1969] and for the supergravity equations [Choquet-Bruhat 1985]. Furthermore, global-in-time existence and decay results for a nonlinear wave equation for 3-form fields, on a fixed background spacetime with compact internal dimensions have been shown in [Ettinger 2015]. The field equation studied in that paper is modelled on the supergravity equations with the gravitational interaction turned off. In our present work, we consider the stability of spacetimes with supersymmetric compactifications as solutions to the vacuum Einstein equations. In future work we intend to study their stability under the supergravity equations.

In addition to determining the dynamics, the Einstein equations also imply that any initial data set must satisfy the constraint equations, which are themselves an important topic of study and have important consequences. A positive mass theorem holds for initial data (Σ, γ, κ) provided that $\Sigma \backslash \Sigma_0$ for some compact Σ_0 is topologically $(\mathbb{R}^n \backslash B) \times K$ for some ball B, that the dimension of the base space is at least $n \geq 3$, that the initial data (Σ, γ, κ) is asymptotically flat in the sense that the metric (including its derivatives) converges to $\delta + k$ sufficiently fast and that κ converges to zero sufficiently rapidly, that the background internal space (K, k) is a simply connected Calabi–Yau manifold, and that the scalar curvature is nonnegative [Dai 2004]. Recent work has shown the existence of such solutions in the case $(K, k) = (\mathbb{T}^d, \delta)$ [Huneau and Vâlcu 2021].

L^2 stability and L^∞ instability. Several people have suggested that the instability argument of Penrose [2003; 2005] should be interpreted as a statement with respect to perturbations that are not localised.[2] This unlocalised interpretation could be stated as saying that SUSY compactifications are unstable against perturbations of the initial data that depend only upon the position in the internal space K but are independent of $x \in \mathbb{R}^n$. Considering the behaviour of the initial data in $x \in \mathbb{R}^n$, this distinction can be interpreted as a being between unlocalised perturbations that merely have a small supremum (for the metric and a suitable number of derivatives) and localised perturbations that have finite and small norms based on the square integral of the perturbation (again including a suitable number of derivatives), such as we use in (5) of Theorem 1.1. We view this as a distinction between, on the one hand, instability in L^∞-based Sobolev spaces and, on the other, stability in L^2-based Sobolev spaces.

Although it is true that SUSY compactifications are unstable against perturbations in L^∞-based Sobolev spaces, this instability does not arise from the presence of the internal space but is already present in Minkowski space for $n \geq 3$. In particular, there is the explicit Kasner solution

$$g = -\,\mathrm{d}t^2 + (1+\epsilon t)^{4/3}\,\mathrm{d}(x^1)^2 + (1+\epsilon t)^{4/3}\,\mathrm{d}(x^2)^2 + (1+\epsilon t)^{-2/3}\,\mathrm{d}(x^3)^2.$$

This is typically considered with (x^1, x^2, x^3) being taken as coordinates on the torus $\mathbb{T}^3$, but it applies equally well on $\mathbb{R}^3$. By taking a tensor product with $(\mathbb{R}^{n-3}, \delta_{\mathbb{R}^{n-3}})$ or $(\mathbb{R}^{n-3} \times K, \delta_{\mathbb{R}^{n-3}} + k)$ one can

[2]We thank the first reviewer for emphasising this perspective.

extend this example to show L^∞ instability also for higher-dimensional Minkowski space and for SUSY compactifications.

The L^∞ instability of Minkowski space and SUSY compactifications can be viewed as part of a broader set of instability phenomena. The L^∞ instability of Minkowski space can be viewed as essentially equivalent to the instability of $(\mathbb{R} \times \mathbb{T}^n, -\mathrm{d}t^2 + \delta_{\mathbb{T}^n})$. Bartnik [1988] has conjectured that a globally hyperbolic spacetime with compact Cauchy surface and satisfying the strong energy condition is either causally incomplete or split as a metric product (and hence flat in the (3+1)-dimensional case). See also [Galloway 2019]. One heuristic justification for this conjecture follows a contradiction argument, which begins by considering what would happen if there were not some major divergence from the original solution. In this case, the metric perturbations would satisfy something close to energy conservation, would exhibit something close to Poincaré recurrence, and would eventually be found in any configuration compatible with the bound on the initial energy. However, just as it is possible to imagine black holes of arbitrarily small mass, it is possible to form trapped surfaces with arbitrarily small energy. Thus, the Poincaré recurrence would imply the eventual formation of trapped surfaces and hence of singularities. This would imply instability, which concludes the contradiction argument. There is a further extension of this belief that if a spacetime with a compact hypersurface does not expand sufficiently rapidly, then metric perturbations will not decay sufficiently rapidly and singularities will form. It is essential to make the distinction between L^∞ and L^2 perturbations when making PDE estimates.

Outline of paper. In Section 2 we introduce: the Lichnerowicz Laplacian, the foliation by hyperboloids, the gauge condition, and the higher-dimensional Schwarzschild-product spacetime. In Section 3 we prove a Sobolev estimate on hyperboloids with respect to wave-like energies. In Section 4 we define an energy functional adapted to the internal manifold and to hyperboloids. Finally in Section 5 we prove the main theorem.

There are four key elements that we add to the standard energy-estimates framework to prove the stability of SUSY compactifications. First, we observe that we can obtain arbitrarily rapid decay by going to sufficiently high dimension and that this decay allows us to control nonlinear terms. Second, the new Sobolev estimates in Section 3 give decay estimates that do not require decomposing metric perturbations into massive and massless parts. Following an argument of Hörmander, the Sobolev estimate in Lemma 3.2 holds on hyperboloids to exploit the fact that the initial data is essentially trivial outside the unit ball. Third, it is possible to introduce an energy that simultaneously enjoys several desirable properties. Namely, the energy introduced in Definition 4.1 is not merely the energy constructed from the energy-momentum tensor (9) for the linearised Einstein equation (8), but we show it is positive using known results on Ricci-flat compact manifolds with special holonomy which we review in Section 2A, and it is the basis for the Sobolev norms in Section 3. Fourth, in defining pointwise norms of derivatives (e.g., Definition 2.4), we commute the equation with the second-order Δ_k rather than just first-order vector fields, which are sufficient in Minkowski space. The higher-order Sobolev estimate in Corollary 4.7 has to use separate indices to count the Minkowski and internal derivatives, because our L^∞-norms use only an even number of derivatives in internal directions, while our L^2-norms use integer number of derivatives. Once we have used these four elements, it is possible to control the nonlinear (including quasilinear) terms in the Einstein equation using standard energy-estimate techniques.

2. Preliminaries

2A. *Parallel spinors and the Lichnerowicz Laplacian.* Our main theorem has been stated for an internal manifold that has a cover that admits a spin structure and a nonzero parallel spinor. In this subsection we detail how this condition relates to a linear stability condition involving the eigenvalues of an operator closely related to the Lichnerowicz Laplacian.

Definition 2.1 (Riemannian linear stability). Define $\Delta_k = k^{AB}\nabla[k]_A\nabla[k]_B$ to be the standard Laplacian on (K, k). Let u_{AB} be a symmetric $(0, 2)$ tensor defined on K. Define $\mathcal{L}$ to act on such tensors by

$$(\mathcal{L}u)_{AB} = -\Delta_k u_{AB} - 2(R[k] \circ u)_{AB}. \tag{10}$$

We define a Ricci-flat manifold (K, k) to be Riemannian linearly stable if and only if

$$\int_K \langle \mathcal{L}u, u\rangle_k \, \mathrm{d}\mu_k \geq 0, \tag{11}$$

for all symmetric $(0, 2)$-tensors u_{AB}.

The operator $\mathcal{L}$ is closely related to the Lichnerowicz Laplacian Δ_L, which acts on symmetric tensors by

$$(\Delta_L u)_{AB} = (\mathcal{L}u)_{AB} + \mathrm{Ric}[k]_{AC}u^C{}_B + \mathrm{Ric}[k]^C{}_B u_{AC}. \tag{12}$$

Clearly on a Ricci-flat space these operators are equivalent. The operator $\mathcal{L}$ is self-adjoint and elliptic, and consequently by the compactness of K and spectral theory, it has a discrete set of eigenvalues of finite multiplicity. Hence definition (11) amounts to a condition $\lambda_{\min} \geq 0$ on the lowest eigenvalue $\lambda_{\min}$ of $\mathcal{L}$. For further details see, e.g., [Besse 1987].

Our main Theorem 1.1 in fact applies more generally to internal manifolds which are Riemannian linearly stable. For the purposes of this paper, the crucial relation between spacetimes with a supersymmetric compactification and with an internal space that is Riemannian linearly stable is the following.

Theorem 2.2 [Dai et al. 2005, Theorem 1.1]. *If a compact, Ricci-flat Riemannian manifold (K, k) has a cover which is spin and admits a nonzero parallel spinor then it is Riemannian linearly stable.*

Note that some of the ideas established in [Dai et al. 2005] date back to work of Wang [1991] on the deformation theory of parallel and Killing spinors. A spin manifold (K, k) with a nonzero parallel spinor is Ricci flat and has special holonomy; see [Wang 1989] for a classification. It is not known if any hypotheses on the internal space beyond Ricci flatness are necessary for stability to hold, as all known examples of compact Ricci-flat manifolds admit a spin cover with nonzero parallel spinors. The problem of constructing Ricci-flat manifolds including ones with nonspecial holonomy has been widely studied. A few relevant references on the topic are [Biquard 2013; Brendle and Kapouleas 2017; Tian and Yau 1990; 1991].

The spatial equivalent of the $\hat{g}$-wave gauge was used in the proof of Milne stability [Andersson and Moncrief 2011]. This led to terms involving $\mathcal{L}$ appearing in their PDEs, which were treated using Riemannian linear stability properties specific to the Milne spacetime. Further results on Riemannian linear stability for Einstein manifolds can be found in [Kröncke 2015].

2B. *Cartesian, hyperbolic, and hyperbolic polar coordinates.*

Definition 2.3 (Minkowski space). Let $n \geq 1$ be an integer, let $(x^0, x^1, \dots, x^n) = (t, x^1, \dots, x^n) = (t, \vec{x})$ be Cartesian coordinates parametrising $\mathbb{R}^{1+n}$, and define

$$\eta_{\mathbb{R}^{1+n}} = -\,\mathrm{d}t^2 + \sum_{i=1}^{n} (\mathrm{d}x^i)^2. \tag{13}$$

Define, for $i \in \{1, \dots, n\}$, the translation vector fields T and X_i so that, in the Cartesian coordinates, they are given by

$$X_i = \partial_{x^i}, \quad T = X_0 = \partial_t. \tag{14}$$

Define, for $i, j \in \{0, \dots, n\}$, the vector fields Z_{ij} so that, in the Cartesian coordinates, they are given by

$$Z_{ij} = (\eta_{\mathbb{R}^{1+n}})_{jk} x^k \partial_i - (\eta_{\mathbb{R}^{1+n}})_{ik} x^k \partial_j. \tag{15}$$

Define the collection of Lorentz generators by

$$Z = \{Z_{ij}, T, X_i\}. \tag{16}$$

Define $|x|^2 = \sum_{i=1}^{n} (x^i)^2$ and define, in the region $t \geq |x|$, the hyperboloidal coordinates to be

$$s = (t^2 - |x|^2)^{1/2}, \quad y = x. \tag{17}$$

Define, for $i \in \{1, \dots, n\}$, the vector fields Y_i so that, in the hyperboloidal coordinates, they are given by

$$Y_i = \partial_{y^i}. \tag{18}$$

For $s_0 \geq 0$, define the spacelike hyperboloidal hypersurface

$$\Sigma_{s_0} = \{(t, x) \in \mathbb{R}^{1+n} : t > 0,\ s = s_0\}. \tag{19}$$

Note that, because $(\eta_{\mathbb{R}^{1+n}})_{00} = -1$, we have $Z_{0i} = t\partial_{x^i} + x_i \partial_t$. Furthermore the collection Z is closed under commutation and forms a basis for the Poincaré Lie algebra.

Definition 2.4 (pointwise derivative norms based on commuting operators). On $\mathbb{R}^{1+n} \times K$, define, for $i \in \{0, \dots, n\}$, X_i, Y_i, and Z_{ij} to be as in $\mathbb{R}^{1+n}$. Let primed roman letters denote spatial indices $i', j' \in \{1, \dots, n+d+1\}$. Define the following collection of vector fields

$$\Gamma = Z \cup \{\Delta_k\}. \tag{20}$$

Note that $[Z, \Delta_k] = 0$. Define $\mathbb{N} = \{0, 1, 2, \dots\}$. We will now define $\{Z_i\}_{i=1}^{(n+1)(n+2)/2}$ to be a reindexing of $\{X_i\}_{i=0}^{n} \cup \{Z_{ij}\}_{0 \leq i < j \leq n}$, define a multi-index to be an ordered list of arbitrary length of elements from $\left\{1, \dots, \frac{1}{2}(n+1)(n+2)\right\}$, and for a multi-index $I = (i_1, \dots, i_k)$ define the length $|I| = k$ and the differential operator $Z^I = Z_{i_k} \circ \cdots \circ Z_{i_1}$. For $I \in \mathbb{N}$ and $u_{\mu\nu}$ a tensor defined on $\mathbb{R}^{1+n} \times K$, define the generalised multi-index notation

$$|\Gamma^I u|_E^2 = \sum_{I_1 : |I_1| + 2j = |I|} |Z^{I_1} \Delta_k^j u|_E^2, \tag{21}$$

where the sum is taken over all multi-indices I_1 of length $|I_1| = k$ and integers j such that $k + 2j = |I|$.

Definition 2.5 (Sobolev norms). Let $u_{\mu\nu}$ be a tensor defined on $\mathbb{R}^{1+n} \times K$ and let $j \in \mathbb{N}$. Define

$$|\nabla[k]^j u|_E^2 = k^{A_1 B_1} \cdots k^{A_j B_j} g_E^{\mu\nu} g_E^{\rho\sigma} (\nabla[k]_{A_j} \cdots \nabla[k]_{A_1} u_{\mu\rho})(\nabla[k]_{B_j} \cdots \nabla[k]_{B_1} u_{\nu\sigma}). \tag{22}$$

For $\ell \in \mathbb{N}$, define the norms

$$\|u(\,\cdot\,,\,\cdot\,,\omega)\|_{H^\ell(K)} = \left(\int_K \sum_{0 \le j \le \ell} |\nabla[k]^j u(\,\cdot\,,\,\cdot\,,\omega)|_E^2 \,\mathrm{d}\mu_k \right)^{1/2}, \tag{23}$$

$$\|u(t,x,\omega)\|_{L^2(\Sigma_s \times K)} = \left(\int_{\Sigma_s \times K} |u(t,x,\omega)|_E^2 \,\mathrm{d}x \,\mathrm{d}\mu_k \right)^{1/2}, \tag{24}$$

where $\mathrm{d}x = \mathrm{d}x^1 \cdots \mathrm{d}x^n$ is defined to be the flat Euclidean volume form.

Lemma 2.6. $\quad Y_i = X_i + (x_i/t)T, \quad Z_{0i} = tY_i, \quad Z_{ij} = y_i Y_j - y_j Y_i.$

Proof. Since $t = \sqrt{s^2 + y^2}$, by the chain rule, for $j \in \{1, \dots, n\}$,

$$\frac{\partial}{\partial y^j} = \frac{\partial x^i}{\partial y^j} \frac{\partial}{\partial x^i} = \frac{\partial}{\partial x^j} + \frac{\partial t}{\partial y^j} \frac{\partial}{\partial t} = \frac{\partial}{\partial x^j} + \frac{y_j}{t} \frac{\partial}{\partial t},$$

which gives the first result. The second follows from multiplying both sides of the first by t. The third follows from

$$Z_{ij} = x_i X_j - x_j X_i = x_i(X_j + x_j t^{-1} T) - x_j(X_i + x_i t^{-1} T). \qquad \square$$

The following two lemmas relate the t coordinate to the s coordinate.

Lemma 2.7. *Let $s \ge 1$. Suppose $(t_0, x_0) \in \Sigma_s$ and $(t, x) \in \Sigma_s$ with $|x - x_0| \le \frac{1}{2} t_0$. In this case, $\frac{1}{2} t_0 \le t \le 2t_0$.*

Proof. For the graph $t = \sqrt{s^2 + |x|^2}$, the gradient

$$\left| \frac{\partial t}{\partial x} \right| = \left| \frac{x}{\sqrt{s^2 + |x|^2}} \right| \le 1, \tag{25}$$

so the change from t to t_0 is less than the change from $|x|$ to $|x_0|$. $\square$

Lemma 2.8. *There is a constant $C > 0$ such that for all $s > 1$, in the portion of Σ_s where $|x| \le t - 1$, one has $2t - 1 \le s^2 \le t^2$.*

Proof. Observe that $t^2 = s^2 + |x|^2 \ge s^2$. Since $|x|^2 \le t^2 - 2t + 1$, one has $s^2 = t^2 - |x|^2 \ge 2t - 1$. $\square$

The following are standard elliptic estimates; see for example [Besse 1987, Appendix H].

Lemma 2.9 (elliptic estimates on (K, k)). *For $\ell \in \mathbb{N}$ and $u_{\mu\nu}$ a sufficiently regular tensor defined on $\mathbb{R}^{1+n} \times K$, there exist constants $c_1, c_2, c_3 > 0$ such that*

$$\|u\|_{H^{2\ell}(K)} \le c_1 \|(\Delta_k)^\ell u\|_{L^2(K)} + c_2 \|u\|_{L^2(K)} \le c_3 \|u\|_{H^{2\ell}(K)}. \tag{26}$$

In Lemma 2.9, if u is orthogonal to the kernel of Δ_k, then there is a c_1 such that the first estimate holds with $c_2 = 0$.

2C. ***The Einstein equations.*** The theory of the Einstein equations is well known. In this section, we review this theory, for the sake of providing a self-contained presentation in this paper, and in particular to provide a self-contained statement of our main Theorem 1.1.

Definition 2.10 (geometric initial data set). Let $m \in \mathbb{N}^+$. An m-dimensional initial data set is defined to be a triple (Σ, γ, κ) such that Σ is an m-dimensional manifold, $\gamma_{i'j'}$ is a Riemannian metric on Σ, $\kappa_{i'j'}$ is a symmetric 2-tensor on Σ, and the following equations (the constraint equations) are satisfied:

$$R[\gamma] - |\kappa|^2 + (\operatorname{tr}(\kappa))^2 = 0, \quad \nabla[\gamma]_{i'} \operatorname{tr}(\kappa) - \nabla[\gamma]^{j'}(\kappa)_{i'j'} = 0, \tag{27}$$

where $\operatorname{tr}(\kappa) = \gamma^{i'j'}\kappa_{i'j'}$.

Definition 2.11 (solution of the Einstein equations with specified initial data). Let $\mathcal{M}$ be a manifold. A Lorentzian metric g on $\mathcal{M}$ is defined to be a solution of the vacuum Einstein equations if and only if its Ricci curvature vanishes,

$$\operatorname{Ric}[g]_{\mu\nu} = 0. \tag{28}$$

Let (Σ, γ, κ) be a geometric initial data set. A solution to the (geometric) Einstein equations with initial data (Σ, γ, κ) is defined to be a Lorentzian metric g on $I \times \Sigma$ for some interval I where one has: $0 \in I$, g is a solution of the Einstein equations (28), $\{0\} \times \Sigma$ and g restricted to vectors in $T(\{0\} \times \Sigma)$ are isometric in the category of Riemannian manifolds to (Σ, γ), and, with the identification given by this isometry, the second fundamental form of the embedding of $\{0\} \times \Sigma$ into $I \times \Sigma$ is κ.

As is well known, Definition 2.11 is stated in a more restrictive form than necessary. In Definition 2.11, for convenience, we have required that the initial data be specified at $t = 0$. This may initially appear more restrictive than definitions that are stated in other sources. By a translation in the t variable, Definition 2.11 could be stated on any level set of t. Furthermore, because of the freedom to introduce new coordinate systems on the manifold $I \times \Sigma$, Definition 2.11 is actually equivalent to definitions that allow initial data to specified on more general spacelike hypersurfaces.

2D. ***The reduced Einstein equations.*** To obtain a well-posed evolution problem for the Einstein equations we choose a gauge with respect to a fixed Lorentzian metric $\hat{e}_{\mu\nu}$ defined on $\mathcal{M}$.

Definition 2.12 ($\hat{e}$-wave gauge). For Lorentzian metrics g and $\hat{e}$ defined on some manifold $\mathcal{M}$, let $\nabla[g]$ and $\nabla[\hat{e}]$ be the Levi-Civita connections with corresponding Christoffel symbols $\Gamma[g]$ and $\Gamma[\hat{e}]$ in local coordinates. Define the vector field V^γ in local coordinates by

$$V^\gamma = g^{\alpha\beta}(\Gamma^\gamma_{\alpha\beta}[g] - \Gamma^\gamma_{\alpha\beta}[\hat{e}]). \tag{29}$$

Define also $V_\lambda = g_{\lambda\beta}V^\beta$. The $\hat{e}$-wave gauge condition is given by

$$V^\gamma = 0. \tag{30}$$

Recall that the difference of two Christoffel symbols is a tensor, and so V^γ is in fact a well-defined vector field on $\mathcal{M}$.

Definition 2.13 (reduced Einstein equations). Let $\mathcal{M}$ be a manifold with Lorentzian metric $\hat{e}$. A Lorentzian metric g on $\mathcal{M}$ is defined to be a solution of the reduced Einstein equations if and only if

$$g^{\alpha\beta}\nabla[\hat{e}]_\alpha\nabla[\hat{e}]_\beta g_{\mu\nu} - g^{\gamma\delta}(g_{\mu\lambda}\hat{e}^{\lambda\rho}\operatorname{Riem}[\hat{e}]_{\rho\gamma\nu\delta} + g_{\nu\lambda}\hat{e}^{\lambda\rho}\operatorname{Riem}[\hat{e}]_{\rho\gamma\mu\delta}) = Q_{\mu\nu}[g](\nabla[\hat{e}]g, \nabla[\hat{e}]g), \quad (31a)$$

where we have defined

$$Q_{\mu\nu}[g](\nabla[\hat{e}]g, \nabla[\hat{e}]g) = g^{\gamma\delta}g^{\alpha\beta}\Big(\nabla[\hat{e}]_\nu g_{\delta\beta}\nabla[\hat{e}]_\alpha g_{\mu\gamma} + \nabla[\hat{e}]_\mu g_{\gamma\alpha}\nabla[\hat{e}]_\beta g_{\nu\delta} - \tfrac{1}{2}\nabla[\hat{e}]_\nu g_{\delta\beta}\nabla[\hat{e}]_\mu g_{\gamma\alpha} + \nabla[\hat{e}]_\gamma g_{\mu\alpha}\nabla[\hat{e}]_\delta g_{\nu\beta} - \nabla[\hat{e}]_\gamma g_{\mu\alpha}\nabla[\hat{e}]_\beta g_{\nu\delta}\Big). \quad (31b)$$

2E. ***The higher-dimensional Schwarzschild spacetime.*** In this subsection, the higher-dimensional Schwarzschild solution is considered and its relationship to the initial data for the Einstein equations (28) and the reduced Einstein equations (31) is discussed. The form of the metric follows.

Definition 2.14. Let $n \in \mathbb{Z}$ be such that $n \geq 5$, and let $C_S \in [0, \infty)$. In Schwarzschild coordinates, the Schwarzschild metric is defined, for $(t, \bar{r}, \omega) \in \mathbb{R} \times (C_S^{1/(n-2)}, \infty) \times S^{n-1}$, to be

$$g_S = -\left(1 - \frac{C_S}{\bar{r}^{n-2}}\right)\mathrm{d}t^2 + \left(1 - \frac{C_S}{\bar{r}^{n-2}}\right)^{-1}\mathrm{d}\bar{r}^2 + \bar{r}^2\sigma_{S^{n-1}}. \quad (32)$$

The above metric can also be written in the wave gauge. For $n = 3$, it is sufficient to replace

$$(t, \bar{r}, \omega) \in \mathbb{R} \times (C_S^{1/(n-2)}, \infty) \times S^{n-1}$$

by $(t, x) = (t, r\omega)$ with $r = \bar{r} - M$; the resulting explicit metric can be found in [LeFloch and Ma 2016; Lindblad and Rodnianski 2005]. Although the case $n = 4$ leads to complicated terms involving logarithms, for $n \geq 5$, there is the following theorem.

Theorem 2.15 [Choquet-Bruhat et al. 2006, Section 5.2]. *Let $n \in \mathbb{Z}$ be such that $n \geq 5$, and let $C_S \in [0, \infty)$. There are coordinates (t, x) related to those in Definition 2.14 by $(x^i)_{i=0}^n = (t, r(\bar{r})\omega)$ with*

$$r(\bar{r}) = \bar{r} - \frac{C_S}{2\bar{r}^{n-3}} + O(\bar{r}^{5-2n}),$$

such that the $(x^i)_{i=0}^n$ satisfy the harmonic gauge, that is, the $\eta_{\mathbb{R}^{1+n}}$-wave gauge. Furthermore, there exist functions $h_{00}(R)$, $h(R)$, and $\hat{h}(R)$, defined on an interval around $R = 0$, that are analytic and bounded by a multiple of C_S near $R = 0$, and such that

$$g_S = -\left(1 - \frac{h_{00}(r^{-1})}{r^{n-2}}\right)(\mathrm{d}x^0)^2 + \sum_{i,j=1}^n \left[\left(1 + \frac{h(r^{-1})}{r^{n-2}}\right)\delta^{ij} + \frac{\hat{h}(r^{-1})}{r^{n-2}}\frac{x^i x^j}{r^2}\right]\mathrm{d}x^i\,\mathrm{d}x^j. \quad (33)$$

In particular, the difference between the components of g_S with respect to the harmonic coordinates and the corresponding components of the Minkowski metric are such that any ∂^I derivative decays at least as fast as $C_S r^{-(n-2)-|I|}$.

Note a result in [Dai 2004] ensures that $C_S \geq 0$ for the spacetimes of interest in our main Theorem 1.1.

3. Sobolev estimates on hyperboloids

We begin in Lemma 3.1 by recalling Hörmander's proof of a Sobolev estimate on hyperboloids. This allows us to introduce some of the key ideas that appear in our proof of the main result of this section, Lemma 3.2. The use of the vector field method to prove Sobolev estimates on hyperboloids originates in [Klainerman 1985].

Lemma 3.1 (Sobolev estimate for compactly supported functions on hyperboloids in Minkowski space [Hörmander 1997, Lemma 7.6.1]). *Let ν be the smallest integer greater than $\frac{1}{2}n$, and let $v \in C^{\nu}(\mathbb{R}^{1+n})$ have support in $|x| < t-1$. There is a constant C such that*

$$\sup_{\Sigma_s} t^n |v(t,x)|^2 \leq C \sum_{|I|\leq\nu} \int_{\Sigma_s} |Z^I v|^2 \,\mathrm{d}x. \tag{34}$$

Proof. Consider a point $(t_0, x_0) \in \Sigma_s$ with $|x_0|^2 \leq t_0^2 - 1$. Set $r_0 = \frac{1}{2}t_0$ and $y_0 = x_0$. Set Σ to be the portion of Σ_s on which $|x - x_0| \leq r_0$. Let $(t,x) \in \Sigma$. This implies $|t - t_0| \leq r_0$, which implies $\frac{1}{2}t \leq t_0 \leq 2t$. Thus,

$$\sum_{|I|\leq\nu} \int_{\Sigma_s} |Z^I v(t,x)|^2 \,\mathrm{d}x \geq C \sum_{|I|\leq\nu} \int_{\Sigma_s} |t_0^{|I|} Y^I v(t,y)|^2 \,\mathrm{d}y.$$

The right side can be rewritten, by introducing rescaled coordinates

$$\tilde{y} = 2t_0^{-1}(y - y_0) \quad \text{and} \quad \tilde{v}(\tilde{y}) = v(t,y).$$

One can now decompose the portion of Σ_s where $|x| \leq t-1$ into many subregions where t does not vary by more than a factor of 2. Let $\chi(\tilde{y})$ be a smooth cut-off such that χ is 1 on a neighbourhood of 0 and is 0 for $|\tilde{y}| \geq \frac{1}{2}$, it can further be bounded from below. A Sobolev estimate can then be applied to give a further lower bound on v. Combining these yields

$$\begin{aligned}
\sum_{|I|\leq\nu} \int_{\Sigma_s} |t_0^{|I|} Y^I v(t,x)|^2 \,\mathrm{d}y &= \sum_{|I|\leq\nu} \int_{|\tilde{y}|\leq 1} |\partial_{\tilde{y}}^I \tilde{v}(\tilde{y})|^2 t_0^n \,\mathrm{d}\tilde{y} \\
&\geq C t_0^n \sum_{|I|\leq\nu} \int_{|\tilde{y}|\leq 1} |\partial_{\tilde{y}}^I ((\chi\tilde{v})(\tilde{y}))|^2 \,\mathrm{d}\tilde{y} \\
&\geq C t_0^n |\tilde{v}(0)|^2 \\
&= C t_0^n |v(t_0,x_0)|^2,
\end{aligned}$$

which completes the proof. □

In the following lemma we obtain a Sobolev estimate for functions supported on product spacetimes with specified properties outside a compact set. In particular we obtain a pointwise estimate (36) in terms of the hyperboloidal time s, as well as a t-weighted pointwise estimate on a fixed hyperboloid (37).

Lemma 3.2 (Sobolev estimate for eventually prescribed functions on hyperboloids foliating product spacetimes). *Let $n \geq 4$, let $\tilde{d}$ be the smallest even integer larger than $\frac{1}{2}d$, and let $\tilde{\nu}$ be the smallest integer greater than $\frac{1}{2}n + \tilde{d}$. Let $u_{\mu\nu}$ and $f_{\mu\nu}$ be tensors on $\mathbb{R}^{1+n} \times K$ with f depending only on the Minkowski* *coordinates x^i. Let $u \in C^{\tilde{\nu}}(\mathbb{R}^{1+n} \times K)$ satisfy $u = f$ for $|x| \geq t-1$. Let $f \in C^{\infty}(\mathbb{R}^{1+n} \times K)$ be smooth and be such that, for all $I \in \mathbb{N}$, there is a C_I such that*[3]

$$|\nabla[\hat{g}]^I f|_E \leq C_{|I|}|x|^{-(n-1)/2-|I|}. \tag{35}$$

Let $\delta(n) = \frac{1}{4}(n-2)$. There is a constant C such that

$$\sup_{(t,x^i,\omega)\in\Sigma_s\times K} s^{4\delta(n)}|u(t,x^i,\omega)|_E^2 \leq C\sum_{|I|\leq\tilde{\nu}}\sum_{i=1}^n \int_{\substack{\Sigma_s\times K\\ |x|\leq t-1}} |Y_i Z^I u|_E^2 \,\mathrm{d}x\,\mathrm{d}\mu_k + C\sum_{|I|\leq\tilde{\nu}-1} C_I^2. \tag{36}$$

Furthermore there is a constant C such that

$$\sup_{(t,x^i,\omega)\in\Sigma_s\times K} t^{2\delta(n)}|u(t,x^i,\omega)|_E^2 \leq C\sum_{|I|\leq\tilde{\nu}}\sum_{i=1}^n \int_{\substack{\Sigma_s\times K\\ |x|\leq t-1}} |Y_i Z^I u|_E^2 \,\mathrm{d}x\,\mathrm{d}\mu_k + C\sum_{|I|\leq\tilde{\nu}-1} C_I^2. \tag{37}$$

Proof. Lemma 2.9 and the standard Sobolev estimate imply

$$\sup_{\omega\in K}|u(\cdot\,,\,\cdot\,,\omega)|_E \leq \|u\|_{H^{\tilde{d}}(K)} \leq \|(\Delta_k)^{\tilde{d}/2}u\|_{L^2(K)} + \|u\|_{L^2(K)},$$

for $\tilde{d}$ the smallest even integer greater than $\frac{1}{2}d$. This choice of $\tilde{d}$ being even is simply to make the elliptic estimate cleaner. Note the trivial estimate

$$\sum_{|I|\leq\tilde{\nu}-\tilde{d}} (|Y_i Z^I(\Delta_k)^{\tilde{d}/2}u|_E^2 + |Y_i Z^I u|_E^2) \leq \sum_{|I|+2j\leq\tilde{\nu}} |Y_i Z^I(\Delta_k)^j u|_E^2.$$

It is thus sufficient to prove in Minkowski space that

$$\sup_{\Sigma_s} s^{n-2}|u(t,x)|_E^2 \leq C\sum_{|I|\leq\tilde{\nu}-\tilde{d}}\sum_{i=1}^n \int_{\Sigma_s} |Y_i Z^I u|_E^2\,\mathrm{d}x + C\sum_{|I|\leq\tilde{\nu}-1} C_I^2, \tag{38}$$

since this would then imply

$$\begin{aligned}\sup_{\Sigma_s\times K} s^{n-2}|u(t,x^i,\omega)|_E^2 &\lesssim \sum_{|I|\leq\tilde{\nu}-\tilde{d}}\sum_{i=1}^n \|\sup_K(Y_i Z^I u)\|_{L_x^2}^2 + C\sum_{|I|\leq\tilde{\nu}-1} C_I^2\\ &\lesssim \sum_{|I|\leq\tilde{\nu}-\tilde{d}}\sum_{i=1}^n \|Y_i Z^I(\Delta_k)^{\tilde{d}/2}u\|_{L_x^2 L_K^2}^2 + C\sum_{|I|\leq\tilde{\nu}-1} C_I^2.\end{aligned}$$

For $|x| \geq t-1$ and $(t,x)\in\Sigma_s$, one has $t \sim |x|$, and so

$$s^{n-2}|u(t,x)|_E^2 \leq t^{n-2}|u(t,x)|_E^2 \leq C|x|^{n-2}|u(t,x)|_E^2 \leq C|x|^{n-2}|f(x)|_E^2 \leq CC_0^2.$$

Thus, it remains to prove (38) for $|x| \leq t-1$.

[3]The exponent on f is set to match that corresponding to the exponent arising from the pointwise estimate (36) on u in the region $|t-r| \leq C$. The limiting factor on the exponent in (36) arises from estimates on the hyperboloid, not from the decay of the prescribed function f. If a faster decay rate $t^{-\beta}$ could be proved (using similar methods) on hyperboloids for compact data, then a similar $t^{-\beta}$ decay could be proved for prescribed functions satisfying $f \leq r^{-\beta}$.

Consider the region $|x| \le t-1$. Set $t_{\max} = \frac{1}{2}(s^2+1)$, which is the value of t at which Σ_s intersects $|x| = t-1$ and which satisfies $t \le t_{\max} \le \frac{1}{2}(t^2+1)$ on the portion of Σ_s where $|x| \le t-1$ by Lemma 2.8. Let $\chi : \mathbb{R} \to [0,1]$ be a smooth cut-off function such that $\chi(\alpha) = 1$ for $\alpha < 1$ and $\chi(\alpha) = 0$ for $\alpha > 2$, and define the $(0,2)$ tensor $v_{\mu\nu}(t,x) = \chi(|x|/t_{\max}) u_{\mu\nu}(t,x)$. Observe that $u_{\mu\nu} = v_{\mu\nu}$ in the region $|x| \le t-1$.

Hormander's proof of Lemma 3.1 relies on a carefully chosen rescaling of a portion of the hyperboloid, and the rest of this proof follows the same idea, although the scaling is chosen differently. Recall both the Cartesian (t,x) and hyperboloidal (s,y) coordinates in Minkowski space, which are related via $(s,y) = (\sqrt{t^2-|x|^2}, x)$. Given a choice of s, define $\tilde{y} = s^{-1}y$ and set $\tilde{v}(\tilde{y})$ to be the value of v at hyperboloidal coordinates $(s, s\tilde{y})$. With this, $\mathrm{d}^n\tilde{y} = s^{-n}\,\mathrm{d}y$ and $\partial_{\tilde{y}^i} = s\partial_{y^i} = sY_i$. Recall that $Z_i = tY_i$. Thus, by a Sobolev estimate that exploits the fact that $1 < \frac{1}{2}n < \frac{1}{2}n+1$,

$$\sup_{\Sigma_s} |v(t,x)|_E^2 = \sup |\tilde{v}(\tilde{y})|_E^2 \lesssim \sum_{1 \le |J| \le \frac{n}{2}+1} \int |\partial_{\tilde{y}}^J \tilde{v}|_E^2 \,\mathrm{d}^n\tilde{y}.$$

From rescaling and the facts that $s \le t$ and that $Z_{0i} = tY_i$, it follows that

$$\begin{aligned}
\sup_{\Sigma_s} |v(t,x)|_E^2 &\lesssim s^{-n} \sum_{1 \le |J| \le \frac{n}{2}+1} \int |(sY)^J v|_E^2 \,\mathrm{d}^n y \lesssim s^{-n+2} \sum_{0 \le |J| \le \frac{n}{2}} \sum_i \int s^{2|J|} |Y^J Y_i v|_E^2 \,\mathrm{d}^n y \\
&\lesssim s^{-n+2} \sum_{0 \le |J| \le \frac{n}{2}} \sum_i \int t^{2|J|} |Y^J Y_i v|_E^2 \,\mathrm{d}^n y \lesssim s^{-n+2} \sum_{0 \le |J| \le \frac{n}{2}} \sum_i \int |Y_i Z^J v|_E^2 \,\mathrm{d}^n y.
\end{aligned}$$

The last integral can be decomposed into the regions where $|x| \le t-1$ and $|x| > t-1$. Where $|x| \le t-1$, the integral can be bounded by the integral term on the right-hand side of (38) since $\tilde{\nu} - \tilde{d} > \frac{1}{2}n$. Now consider the region $|x| > t-1$. Because of the support of χ, it is sufficient to consider the region $t_{\max} - 1 \le |x| \le 2(t_{\max}-1)$. In this region, $v = \chi f$. When a derivative is applied to v, it is applied to either χ or to f, in which case one obtains an additional factor of $t_{\max}^{-1}$ or $|x|^{-1}$, from the properties of χ and f, respectively. Since $|x|/t_{\max} \in [1,2]$ in the support of $\partial\chi$, effectively one obtains an extra factor of $|x|^{-1}$ in all cases, so $|Y_i Z^J v|_E \le C C_{|J|+1} |x|^{-(n-1)/2-1}$, and

$$\int_{|x| \ge t_{\max}-1} |Y_i Z^J u|_E^2 \,\mathrm{d}x \le C C_{|J|+1}^2 \int_{\mathbb{S}^{n-1}} \int_{t_{\max}-1}^{2(t_{\max}-1)} (|r|^{-(n-1)/2-1})^2 |r|^{n-1} \,\mathrm{d}r \,\mathrm{d}^{n-1}\omega_{\mathbb{S}^{n-1}} \le C C_{|J|+1}^2.$$

Observing that $s \ge C t^{1/2}$ in the region $|x| \le t-1$ allows us to obtain

$$\begin{aligned}
\sup_{\Sigma_s \times K} t^{2\delta(n)} |u|_E^2 &\le \sup_{\Sigma_s \times K \cap \{|x| \le t-1\}} t^{2\delta(n)} |u|_E^2 + \sup_{\Sigma_s \times K \cap \{|x| > t-1\}} t^{2\delta(n)} |u|_E^2 \\
&\lesssim \sup_{\Sigma_s \times K \cap \{|x| \le t-1\}} s^{4\delta(n)} |u|_E^2 + \sup_{\Sigma_s \times K \cap \{|x| > t-1\}} r^{2\delta(n)} |f|_E^2 \\
&\lesssim \sum_{|I| \le \tilde{\nu}} \sum_{i=1}^{n} \int_{\substack{\Sigma_s \times K \\ |x| \le t-1}} |Y_i Z^I u|_E^2 \,\mathrm{d}x \,\mathrm{d}\mu_k + \sum_{|I| \le \tilde{\nu}-1} C_I^2 + C_0 \sup_{\Sigma_s \times K \cap \{|x| > t-1\}} r^{(n-2)/2} r^{-(n-1)/2}.
\end{aligned}$$

In the final line we applied estimate (36) to the first term and assumption (35) to the second term. □

4. Energy integrals and inequalities

4A. ***Basic properties of the energy.*** The energy introduced in the following definition is related to the standard energy used to study quasilinear hyperbolic PDEs, albeit with additional terms included in order to be compatible with the linearised equations (8).

Definition 4.1 (Lichnerowicz-type energy on hyperboloids). Let $n \in \mathbb{Z}^+$ and let $U^{\mu\nu}$ and $u_{\mu\nu}$ be tensors defined on $\mathbb{R}^{1+n} \times K$. For $u, U \in C^1(\mathbb{R}^{1+n} \times K)$ and $s \geq 2$, define

$$\begin{aligned}\mathcal{E}[U; u; s] = \int_{\Sigma_s \times K} \biggl((s/t)^2 |\partial_t u|_E^2 + \sum_{i=1}^{n} |Y_i u|_E^2 + \langle \nabla[k]^A u, \nabla[k]_A u\rangle_E - 2\langle R[\hat{g}] \circ u, u\rangle_E \\ - 2U^{\alpha\beta} \langle \nabla[\hat{g}]_\beta u, \partial_t u\rangle_E n_\alpha + U^{\alpha\beta} \langle \nabla[\hat{g}]_\alpha u, \nabla[\hat{g}]_\beta u\rangle_E \biggr) \mathrm{d}x \, \mathrm{d}\mu_k, \end{aligned} \tag{39}$$

where $n_0 = 1$ and $n_i = -x_i/t$ for $i \in \{1, \dots, n\}$ and $n_A = 0$, and $\mathrm{d}x$ is the flat Euclidean volume form.

The final terms on the first line could equally well be written as $\langle \nabla[k]^A u, \nabla[k]_A u\rangle_E - 2\langle R[k] \circ u, u\rangle_E$, since the covariant derivative with respect to $\hat{g}$ in directions tangent to K are given by the covariant derivative with respect to k, and similarly for the curvature.

The terms on the second line of (39) are chosen so that, for solutions to the wave equation (42), the change in energy $\mathcal{E}[U; u; s_2] - \mathcal{E}[U; u; s_a]$ is given in (43) by an integral which has an integrand with no terms involving $(\nabla[\hat{g}]u)(\nabla[\hat{g}]\nabla[\hat{g}]u)$. The relevant cancellations to eliminate such terms follow from the properties of $T[U; u]^\mu{}_\nu$ introduced in the proof of Lemma 4.2.

Note that, following [Hörmander 1997; LeFloch and Ma 2016], we have defined $\mathcal{E}[U; u; s]$ so that it is not the naturally induced energy associated with the metric $\hat{g} + U$. This is because we have endowed Σ_s with the flat Euclidean volume form $\mathrm{d}x$, instead of the induced Riemannian volume form $(s/t)\,\mathrm{d}x$.

The following lemma provides us with an energy functional which allows us to measure the perturbation of the spacetime. Note that in (40) we require some weighted t-decay on hyperboloids which we recover from (37) in Lemma 3.2.

Lemma 4.2 (basic properties of the energy). *Take the conditions of Definition 4.1.*

(i) *There is an $\epsilon_n > 0$ such that if*

$$\sup_{\Sigma_s \times K} t|U|_E \leq C\epsilon_n, \tag{40}$$

then for $s \geq 2$,

$$\tfrac{1}{2}\mathcal{E}[U; u; s] \leq \mathcal{E}[0; u; s] \leq 2\mathcal{E}[U; u; s]. \tag{41}$$

(ii) *If $u_{\mu\nu}$ is a solution of*

$$(\hat{g} + U)^{\alpha\beta} \nabla[\hat{g}]_\alpha \nabla[\hat{g}]_\beta u_{\mu\nu} + 2(R[\hat{g}] \circ u)_{\mu\nu} = F_{\mu\nu}, \tag{42}$$

then

$$\begin{aligned}\mathcal{E}[U; u; s_1] = \mathcal{E}[U; u; s_2] + \int_{s_1}^{s_2} \int_{\Sigma_s \times K} \langle F, \partial_t u\rangle_E (s/t) \, \mathrm{d}y \, \mathrm{d}\mu_k \, \mathrm{d}s \\ + \int_{s_1}^{s_2} \int_{\Sigma_s \times K} (-2(\nabla[\hat{g}]_\alpha U^{\alpha\beta}) \langle \nabla[\hat{g}]_\beta u, \partial_t u\rangle_E + (\partial_t U^{\alpha\beta}) \langle \nabla[\hat{g}]_\alpha u, \nabla[\hat{g}]_\beta u\rangle_E)(s/t) \, \mathrm{d}y \, \mathrm{d}\mu_k \, \mathrm{d}s.\end{aligned} \tag{43}$$

Proof. We first derive the energy $\mathcal{E}[U;u;s]$ by considering the following nonlinear version of the stress energy tensor (9)

$$T[U;u]^{\mu}{}_{\nu}=(\hat{g}+U)^{\mu\alpha}\langle\nabla[\hat{g}]_{\alpha}u,\nabla[\hat{g}]_{\nu}u\rangle_E-\tfrac{1}{2}(\hat{g}+U)^{\alpha\beta}\langle\nabla[\hat{g}]_{\beta}u,\nabla[\hat{g}]_{\alpha}u\rangle_E\delta^{\mu}_{\nu}+\langle R[\hat{g}]\circ u,u\rangle_E\delta^{\mu}_{\nu}. \quad (44)$$

We calculate

$$\begin{aligned}\nabla[\hat{g}]_{\mu}T[U;u]^{\mu}{}_{\nu}&=\langle(\hat{g}+U)^{\alpha\beta}\nabla[\hat{g}]_{\alpha}\nabla[\hat{g}]_{\beta}u,\nabla[\hat{g}]_{\nu}u\rangle_E+(\hat{g}+U)^{\mu\alpha}\langle\nabla[\hat{g}]_{\alpha}u,\nabla[\hat{g}]_{\mu}\nabla[\hat{g}]_{\nu}u\rangle_E\\&\quad-(\hat{g}+U)^{\alpha\beta}\langle\nabla[\hat{g}]_{\nu}\nabla[\hat{g}]_{\beta}u,\nabla[\hat{g}]_{\alpha}u\rangle_E+\nabla[\hat{g}]_{\nu}\langle R[\hat{g}]\circ u,u\rangle_E\\&\quad+(\nabla[\hat{g}]_{\mu}U^{\mu\alpha})\langle\nabla[\hat{g}]_{\alpha}u,\nabla[\hat{g}]_{\nu}u\rangle_E-\tfrac{1}{2}(\nabla[\hat{g}]_{\nu}U^{\alpha\beta})\langle\nabla[\hat{g}]_{\alpha}u,\nabla[\hat{g}]_{\beta}u\rangle_E. \quad (45)\end{aligned}$$

Let X^{μ} be a vector field on $\mathbb{R}^{1+n}\times K$ tangent to $\mathbb{R}^{1+n}$. We have

$$\nabla[\hat{g}]_{\alpha}\nabla[\hat{g}]_{\beta}u_{\gamma\delta}=\nabla[\hat{g}]_{\beta}\nabla[\hat{g}]_{\alpha}u_{\gamma\delta}+\mathrm{Riem}[\hat{g}]_{\alpha\beta\gamma}{}^{\rho}u_{\rho\delta}+\mathrm{Riem}[\hat{g}]_{\alpha\beta\delta}{}^{\rho}u_{\rho\gamma}.$$

Since $(\mathbb{R}^{1+n},\eta_{\mathbb{R}^{1+n}})$ has zero Riemann curvature, and since the Riemann curvature for a product manifold is given by $\mathrm{Riem}[\hat{g}]=\mathrm{Riem}[\eta_{\mathbb{R}_{1+n}}]+\mathrm{Riem}[k]$, it follows that all components of the Riemann curvature $\mathrm{Riem}[\hat{g}]_{\alpha\beta\gamma}{}^{\delta}$ vanish unless all the indices $\alpha,\beta,\gamma,\delta$ correspond to internal directions tangent to K. Thus, the contraction with a vector tangent to $\mathbb{R}^{1+n}$ vanishes, and, in particular,

$$\mathrm{Riem}[\hat{g}]_{\alpha\beta\gamma\delta}X^{\delta}=0. \quad (46)$$

Consequently

$$\langle\nabla[\hat{g}]_{\alpha}\nabla[\hat{g}]_{\beta}u,\nabla[\hat{g}]_{\nu}u\rangle_E X^{\alpha}=\langle\nabla[\hat{g}]_{\beta}\nabla[\hat{g}]_{\alpha}u,\nabla[\hat{g}]_{\nu}u\rangle_E X^{\alpha}.$$

and also

$$\nabla[\hat{g}]_{\nu}\langle R[\hat{g}]\circ u,u\rangle_E X^{\nu}=2\langle R[\hat{g}]\circ u,X^{\nu}\nabla[\hat{g}]_{\nu}u\rangle_E.$$

This allows us to calculate

$$\begin{aligned}\nabla[\hat{g}]_{\mu}(T[U;u]^{\mu}{}_{\nu}X^{\nu})&=T^{\mu}{}_{\nu}[U]\nabla[\hat{g}]_{\mu}X^{\nu}+\langle F,X^{\nu}\nabla[\hat{g}]_{\nu}u\rangle_E+(\nabla[\hat{g}]_{\mu}U^{\mu\alpha})\langle\nabla[\hat{g}]_{\alpha}u,X^{\nu}\nabla[\hat{g}]_{\nu}u\rangle_E\\&\quad-\tfrac{1}{2}(X^{\nu}\nabla[\hat{g}]_{\nu}U^{\alpha\beta})\langle\nabla[\hat{g}]_{\alpha}u,\nabla[\hat{g}]_{\beta}u\rangle_E.\end{aligned}$$

Consider the hyperboloidal energy

$$\begin{aligned}\mathcal{E}[U;u;s]&=\int_{\Sigma_s\times K}-2T[U;u]^{\mu}{}_{\nu}(\partial_t)^{\nu}n_{\mu}\,\mathrm{d}x\,\mathrm{d}\mu_k\\&=\int_{\Sigma_s\times K}\Big(|\partial_t u|_E^2+\sum_{i=1}^{n}|\partial_i u|_E^2+\sum_{i=1}^{n}2(x^i/t)\langle\partial_t u,\partial_i u\rangle_E+k^{AB}\langle\nabla[\hat{g}]_A u,\nabla[\hat{g}]_B u\rangle_E\\&\qquad-2\langle R[\hat{g}]\circ u,u\rangle_E-2U^{\mu\rho}\langle\nabla[\hat{g}]_{\rho}u,\partial_t u\rangle_E n_{\mu}+U^{\rho\lambda}\langle\nabla[\hat{g}]_{\rho}u,\nabla[\hat{g}]_{\lambda}u\rangle_E\Big)\,\mathrm{d}x\,\mathrm{d}\mu_k,\end{aligned}$$

where $n_0=1$, $n_i=-\eta_{ij}x^j/t$ for $i\in\{1,\dots,n\}$ and $n_A=0$. Note that

$$\begin{aligned}\mathcal{E}[0;u;s]=\int_{\Sigma_s\times K}\Big(&|\partial_t u|_E^2+\sum_{i=1}^{n}|\partial_i u|_E^2+2(x^i/t)\langle\partial_t u,\partial_i u\rangle_E\\&+\langle\nabla[\hat{g}]^A u,\nabla[\hat{g}]_A u\rangle_E-2\langle R[\hat{g}]\circ u,u\rangle_E\Big)\,\mathrm{d}x\,\mathrm{d}\mu_k, \quad (47)\end{aligned}$$

which alternatively can be written in hyperboloidal coordinates as

$$\mathcal{E}[0; u; s] = \int_{\Sigma_s \times K} \left((s/t)^2 |\partial_t u|_E^2 + \sum_{i=1}^{n} |Y_i u|_E^2 + \langle \nabla[\hat{g}]^A u, \nabla[\hat{g}]_A u \rangle_E - 2\langle R[\hat{g}] \circ u, u \rangle_E \right) \mathrm{d}x \, \mathrm{d}\mu_k. \quad (48)$$

Since the contraction of $R[\hat{g}]$ with any direction tangent to $\mathbb{R}^{1+n}$ vanishes, and since $|w|_E \geq |w|_k$ for any tensor field w, it follows from the definition of $\mathcal{L}$ that

$$\begin{aligned}\int_K (\langle \nabla[\hat{g}]^A u, \nabla[\hat{g}]_A u \rangle_E - 2\langle R[\hat{g}] \circ u, u \rangle_E) \, \mathrm{d}\mu_k &\geq \int_K (\langle \nabla[\hat{g}]^A u, \nabla[\hat{g}]_A u \rangle_k - 2\langle R[\hat{g}] \circ u, u \rangle_k) \, \mathrm{d}\mu_k \\ &= \int_K \langle \mathcal{L}u, u \rangle_k \, \mathrm{d}\mu_k.\end{aligned}$$

Thus, from Theorem 2.2 and the condition of Riemannian linear stability (11), it follows that

$$\int_K (\langle \nabla[\hat{g}]^A u, \nabla[\hat{g}]_A u \rangle_E - 2\langle R[\hat{g}] \circ u, u \rangle_E) \, \mathrm{d}\mu_k \geq 0. \quad (49)$$

Thus, $\mathcal{E}[0, u, s] \geq 0$.

Using our previously calculated expression for the divergence of $T[U; u]^\mu{}_\nu X^\nu$, we obtain

$$\begin{aligned}\mathcal{E}[U; u; s_1] = \mathcal{E}[U; u; s_2] &+ \int_{s_1}^{s_2} \int_{\Sigma_s \times K} \langle -2F, \partial_t u \rangle_E (s/t) \, \mathrm{d}y \, \mathrm{d}\mu_k \, \mathrm{d}s \\ &+ \int_{s_1}^{s_2} \int_{\Sigma_s \times K} \big(-2(\nabla[\hat{g}]_\alpha U^{\alpha\beta}) \langle \nabla[\hat{g}]_\beta u, \partial_t u \rangle_E \\ &\qquad + (\partial_t U^{\alpha\beta}) \langle \nabla[\hat{g}]_\alpha u, \nabla[\hat{g}]_\beta u \rangle_E \big) (s/t) \, \mathrm{d}y \, \mathrm{d}\mu_k \, \mathrm{d}s\end{aligned}$$

via Stoke's theorem. This proves equality (43).

Condition (40) combined with $s \geq Ct^{1/2}$ implies $\sup_{\Sigma_s \times K} |U|_E (t/s)^2 \leq C\varepsilon_n$. For simplicity denote $k^{AB} \langle \nabla[\hat{g}]_A u, \nabla[\hat{g}]_B u \rangle_E$ by $|\partial_A u|_E^2$, then

$$\begin{aligned}\frac{s^2}{2t^2} \left(|\partial_t u|_E^2 \sum_i + |\partial_i u|_E^2 + |\nabla[k] u|_E^2 \right) &\leq \left(|\partial_t u|^2 + \sum_i |\partial_i u|^2 + |\nabla[k] u|_E^2 \right) (1 - |x|/t) \\ &\leq |\partial_t u|_E^2 + |\partial_i u|_E^2 + 2(x^i/t) \langle \partial_t u, \partial_i u \rangle_E + |\nabla[k] u|_E^2. \quad (50)\end{aligned}$$

Using this and Young's inequality we find

$$\begin{aligned}|\mathcal{E}[U; u; s] - \mathcal{E}[0; u; s]| &= \left| \int_{\Sigma_s \times K} (2U^{\alpha\beta} \langle \nabla[\hat{g}]_\alpha u, \partial_t u \rangle_E n_\beta - U^{\alpha\beta} \langle \nabla[\hat{g}]_\alpha u, \nabla[\hat{g}]_\beta u \rangle_E) \, \mathrm{d}x \, \mathrm{d}\mu_k \right| \\ &\leq C\varepsilon_n \mathcal{E}[0; u; s],\end{aligned}$$

and thus the energies are equivalent for sufficiently small ε_n. This proves estimate (41) and the lemma. □

Having defined the energy involving first-order derivatives, we now introduce higher-order energies.

Definition 4.3 (symmetry boosted energy). Let $(\mathbb{R}^{1+n} \times K, \hat{g})$ be a spacetime with a supersymmetric compactification and $N \in \mathbb{N}$. For $k \leq N$, define the energy of a symmetric tensor field g to be

$$\mathcal{E}_{k+1}(s) = \sum_{|I| \leq k} \mathcal{E}[g^{-1} - \hat{g}^{-1}; \Gamma^I g; s]. \quad (51)$$

We end this section with the following Hardy estimate on hyperboloids. The proof is standard; see for example [LeFloch and Ma 2016, Lemma 2.4].

Lemma 4.4 (Hardy estimate on hyperboloids). *Let $u_{\mu\nu}$ be a tensor defined on $\mathbb{R}^{1+n}$. Then one has*

$$\|r^{-1}u\|_{L^2(\Sigma_s)} \lesssim \sum_{i=1}^{n} \|Y_i u\|_{L^2(\Sigma_s)}. \tag{52}$$

4B. ***Preliminary L^2 and L^∞ estimates.*** In our nonlinear estimates we will estimate terms of the form

$$Z^I(\Delta_k)^j(uv) = \sum_{\substack{|I_1|+|I_2|=|I| \\ |J_1|+|J_2|=2j}} Z^{I_1}\nabla[k]^{J_1}u \cdot Z^{I_2}\nabla[k]^{J_2}v. \tag{53}$$

In the following lemma we estimate terms which appear as factors in the right-hand side of (53) in L^2 by using the elliptic estimates of Lemma 2.9 and the Hardy estimate of Lemma 4.4. Note the use of elliptic estimates allows us to avoid commuting derivatives, such as $[\nabla[k], \Delta_k]$, which shortens the argument.

Lemma 4.5 (L^2 estimate for distributed derivatives). *Let $u_{\mu\nu}$ be a tensor defined on $\mathbb{R}^{1+n}\times K$. Suppose N is even, $\ell \in \mathbb{N}$, and $\ell \le N+1$, then*

$$\sum_{|I|+|J|\le \ell} \|t^{-1}Z^I\nabla[k]^J u\|_{L^2(\Sigma_s\times K)} \lesssim \mathcal{E}_{N+1}(s)^{1/2}. \tag{54}$$

Proof. We prove the estimate by considering separately the cases of $|I|=0$ and $|I|\neq 0$. Firstly take $|I|\ge 1$, suppose $|J|=2m$ where $m\in\mathbb{N}$, and consider $|I|+|J|=\ell\le N+1$. Using the elliptic estimates of Lemma 2.9 we find

$$\begin{aligned}\|t^{-1}Z^I\nabla[k]^J u\|_{L^2(\Sigma_s\times K)} &\lesssim \big\|\|t^{-1}Z^I u\|_{H^{2m}(K)}\big\|_{L^2(\Sigma_s)} \lesssim \|t^{-1}Z^I(\Delta_k)^m u\|_{L^2(\Sigma_s\times K)} + \|t^{-1}Z^I u\|_{L^2(\Sigma_s\times K)}\\ &\lesssim \sum_{i=1}^n \|Y_i Z^{I-1}(\Delta_k)^m u\|_{L^2(\Sigma_s\times K)} + \sum_{i=1}^n \|Y_i Z^{I-1}u\|_{L^2(\Sigma_s\times K)}\\ &\lesssim \mathcal{E}[0;\, Z^{I-1}(\Delta_k)^m u;\, s]^{1/2} + \mathcal{E}[0;\, Z^{I-1}u;\, s]^{1/2} \lesssim \mathcal{E}_\ell(s)^{1/2}.\end{aligned}$$

Next take $|I|\ge 1$ and suppose $|J|=2m+1$ where $m\in\mathbb{N}$. For $|I|+|J|=\ell\le N+1$, again using Lemma 2.9, we have

$$\begin{aligned}&\|t^{-1}Z^I\nabla[k]^J u\|_{L^2(\Sigma_s\times K)}\\ &\qquad\lesssim \big\|\|t^{-1}Z^I u\|_{H^{2m+1}(K)}\big\|_{L^2(\Sigma_s)}\\ &\qquad\lesssim \sum_{i=1}^n \|Y_i Z^{I-1}u\|_{L^2(\Sigma_s\times K)} + \sum_{i=1}^n \|Y_i Z^{I-1}(\Delta_k)^m u\|_{L^2(\Sigma_s\times K)} + \|\nabla[k](Z^I(\Delta_k)^m u)\|_{L^2(\Sigma_s\times K)}\\ &\qquad\lesssim \mathcal{E}[0;\, Z^{I-1}u;\, s]^{1/2} + \mathcal{E}[0;\, Z^{I-1}(\Delta_k)^m u;\, s]^{1/2} + \mathcal{E}[0;\, Z^I(\Delta_k)^m u;\, s]^{1/2} \lesssim \mathcal{E}_\ell(s)^{1/2}.\end{aligned}$$

We now turn to the case $|I|=0$. Again we split into the cases of $|J|$ being even and odd. Start with $|J|=2m$ for $m\in\mathbb{N}$. Note that N is chosen to be even so that we have the strict inequality $2m < N+1$.

Applying the Hardy estimate from Lemma 4.4, and recalling that $t \geq r$ on the hyperboloid, yields

$$\begin{aligned}\|t^{-1}\nabla[k]^J u\|_{L^2(\Sigma_s\times K)} &\lesssim \big\| \|r^{-1}u\|_{H^{2m}(K)}\big\|_{L^2(\Sigma_s)} \lesssim \|r^{-1}(\Delta_k)^m u\|_{L^2(\Sigma_s\times K)} + \|r^{-1}u\|_{L^2(\Sigma_s\times K)}\\ &\lesssim \sum_{i=1}^n \|Y_i(\Delta_k)^m u\|_{L^2(\Sigma_s\times K)} + \sum_{i=1}^n \|Y_i u\|_{L^2(\Sigma_s\times K)}\\ &\lesssim \mathcal{E}[0,(\Delta_k)^m u; s]^{1/2} + \mathcal{E}[0,u;s]^{1/2} \lesssim \mathcal{E}_{N+1}(s)^{1/2}.\end{aligned}$$

Finally we have the case $|I| = 0$ and $|J| = 2m+1 \leq N+1$ for $m \in \mathbb{N}$. Again using Lemma 4.4 we obtain

$$\begin{aligned}\|t^{-1}\nabla[k]^J u\|_{L^2(\Sigma_s\times K)} &\lesssim \big\| \|r^{-1}u\|_{H^{2m+1}(K)}\big\|_{L^2(\Sigma_s)} \lesssim \|r^{-1}\nabla[k](\Delta_k)^m u\|_{L^2(\Sigma_s\times K)} + \|r^{-1}u\|_{L^2(\Sigma_s\times K)}\\ &\lesssim \|\nabla[k](\Delta_k)^m u\|_{L^2(\Sigma_s\times K)} + \sum_{i=1}^n \|Y_i u\|_{L^2(\Sigma_s\times K)}\\ &\lesssim \mathcal{E}[0,(\Delta_k)^m u; s]^{1/2} + \mathcal{E}[0,u;s]^{1/2} \lesssim \mathcal{E}_{|J|}(s)^{1/2}.\end{aligned}$$

Adding together the above estimates over all appropriate multi-indices gives the required result. □

Corollary 4.6 (L^2 estimate for eventually prescribed functions on hyperboloids foliating product spacetimes). *Let $n \geq 4$. Let $u_{\mu\nu}$ and $f_{\mu\nu}$ be tensors defined on $\mathbb{R}^{1+n}\times K$ with f depending only on the Minkowski coordinates. Suppose $u = f$ for $|x| \geq t-1$. Let $f \in C^\infty(\mathbb{R}^{1+n}\times K)$ be smooth and such that, for all $I \in \mathbb{N}$, there is a C_I such that*[4]

$$|\nabla[\hat{g}]^I f|_E \leq C_{|I|}|x|^{-(n+1)/2-|I|}. \tag{55}$$

Suppose N is even, $\ell \in \mathbb{N}$, and $\ell \leq N+1$, then

$$\sum_{|I|+|J|\leq \ell} \|(s/t)Z^I\nabla[k]^J u\|_{L^2(\Sigma_s\times K)} \lesssim s\mathcal{E}_{N+1}(s)^{1/2} + \sum_{|I|+|J|\leq\ell} C_{|I|,|J|}. \tag{56}$$

Proof. We will consider separately the regions $|x| \leq t-1$ and $|x| > t-1$. The estimate in the region $|x| \leq t-1$ follows by applying Lemma 4.5 with an additional factor of s. Next consider the region $|x| > t-1 \geq t_0 - 1$, where we let $t_0 = \frac{1}{2}(s^2+1)$ be the value of t at which Σ_s intersects $|x| = t-1$. Using assumption (55) we find

$$\begin{aligned}&\|(s/t)Z^I\nabla[k]^J u\|^2_{L^2(\Sigma_s\times K\cap\{|x|>t-1\})}\\ &\qquad\leq \int_{\Sigma_s\times K\cap\{|x|>t_0-1\}} |Z^I\nabla[k]^J u|^2_E \,\mathrm{d}x\,\mathrm{d}\mu_k \leq C\int_{\Sigma_s\cap\{|x|>t_0-1\}} |Z^I\nabla[k]^J f|^2_E\,\mathrm{d}x\\ &\qquad\leq CC^2_{|I|,|J|}\int_{\mathbb{S}^{n-1}}\int_{\Sigma_s\cap\{|x|\geq t_0-1\}} (|r|^{-(n+1)/2})^2|r|^{n-1}\,\mathrm{d}r\,\mathrm{d}\omega_{\mathbb{S}^{n-1}}\\ &\qquad\leq CC^2_{|I|,|J|}\int_{\mathbb{S}^{n-1}}\int_{\Sigma_s\cap\{|x|\geq t_0-1\}} r^{-2}\,\mathrm{d}r\,\mathrm{d}\omega_{\mathbb{S}^{n-1}} \leq CC^2_{|I|,|J|}.\end{aligned}$$

Adding together the above estimates over all appropriate multi-indices yields (56). □

[4]Note that the decay assumption on f is stronger here than the assumption (35) in Lemma 3.2.

We next use Lemma 3.2 to obtain L^∞ estimates for terms which appear as factors in the right-hand side of (53).

Corollary 4.7 (higher-order Sobolev estimates). *Let $n \geq 7$. Let $\tilde{d}$, $\tilde{\nu}$, $u_{\mu\nu}$, and $f_{\mu\nu}$ be as defined in Lemma 3.2. Then for $|I| + |J| = \ell \in \mathbb{N}$ there is a constant C such that*

$$\sup_{\Sigma_s \times K} (s^{4\delta(n)} |Z^I \nabla[k]^J u|_E^2 + s^{4\delta(n)-2} |(t/s) Z^I \nabla[k]^J u|_E^2) \\ \leq C \sum_{|I|+2j \leq \tilde{\nu}+\ell+1} \mathcal{E}[0;\, Z^I (\Delta_k)^j u;\, s] + C \sum_{|I| \leq \tilde{\nu}+\ell-1} C_{|I|}^2. \tag{57}$$

Proof. We consider the left-most term in (57) first. Let $\tilde{j}$ be the smallest even integer such that $\tilde{j} \geq |J|$. In particular this means

$$|I| + |J| \leq |I| + \tilde{j} \leq \ell + 1.$$

Recall that $\tilde{d}$ is the smallest even integer larger than $\frac{1}{2}d$ and $\tilde{\nu}$ is the smallest integer greater than $\frac{1}{2}n + \tilde{d}$. Applying Lemma 2.9 yields

$$\sup_K |\nabla[k]^J u|_E \leq \|u\|_{H^{\tilde{d}+\tilde{j}}(K)} \leq \|(\Delta_k)^{(\tilde{d}+\tilde{j})/2} u\|_{L^2(K)} + \|u\|_{L^2(K)}.$$

Thus, using in particular (38), we have

$$\begin{aligned} \sup_{(t,x,\omega) \in \Sigma_s \times K} & s^{4\delta(n)} |Z^I \nabla[k]^J u(t, x^i, \omega)|_E^2 \\ &\lesssim \sum_{|I_1| \leq \tilde{\nu}-\tilde{d}} \sum_{i=1}^n \|\sup_K (Y_i Z^{I_1} Z^I \nabla[k]^J u)\|_{L^2(\Sigma_s)}^2 + \sum_{|I_1| \leq \tilde{\nu}-1} C_{I_1}^2 \\ &\lesssim \sum_{|I_1| \leq \tilde{\nu}-\tilde{d}} \sum_{i=1}^n (\|Y_i Z^{I+I_1} u\|_{L^2(\Sigma_s \times K)}^2 + \|Y_i Z^{I+I_1} (\Delta_k)^{(\tilde{d}+\tilde{j})/2} u\|_{L^2(\Sigma_s \times K)}^2) + C \sum_{|I_1| \leq \tilde{\nu}-1} C_{I_1}^2 \\ &\lesssim \sum_{|I|+2j \leq \tilde{\nu}+\ell+1} \mathcal{E}[0;\, Z^I (\Delta_k)^j u;\, s] + C \sum_{|I| \leq \tilde{\nu}-1} C_I^2. \end{aligned}$$

To complete the proof for the second term of (57) we observe that $s \geq C t^{1/2}$ in the region $|x| \leq t-1$ while we only have $s \leq t \leq r$ in the region $|x| > t-1$. Since $n \geq 7$ we have $\delta(n) \geq 1$ and thus

$$\begin{aligned} \sup_{\Sigma_s \times K} & s^{4\delta(n)-2} |(t/s) Z^I \nabla[k]^J u|_E^2 \\ &\lesssim \sup_{\Sigma_s \times K \cap \{|x| \leq t-1\}} (t^2/s^4) s^{4\delta(n)} |Z^I \nabla[k]^J u|_E^2 + \sup_{\Sigma_s \times K \cap \{|x| > t-1\}} s^{4\delta(n)-4} r^2 |Z^I \nabla[k]^J f|_E^2 \\ &\lesssim \sum_{|I|+2j \leq \tilde{\nu}+\ell+1} \mathcal{E}[0;\, Z^I (\Delta_k)^j u;\, s] + \sum_{|I| \leq \tilde{\nu}+\ell-1} C_I^2 + C_I^2 \sup_{\Sigma_s \times K \cap \{|x| > t-1\}} r^{(n-2)-2} r^{-(n-1)}. \end{aligned}$$

Note in the final line we applied (35) and the first estimate of (57). $\square$

5. Proof of stability

5A. ***Stability for the reduced Einstein equations.*** We now restate our main Theorem 1.1 in terms of the reduced Einstein equations. For convenience we translate the initial data of Theorem 1.1 to $\{t=4\}$.

Theorem 5.1 (stability for the reduced Einstein equations). *Let $n, d \in \mathbb{Z}^+$ be such that $n \geq 9$, and let $N \in \mathbb{N}$ be an even integer strictly larger than $\frac{1}{2}(n+d+8)$. Let $(\mathbb{R}^{1+n} \times K, \hat{g} = \eta_{\mathbb{R}^{1+n}} + k)$ be a spacetime with a supersymmetric compactification.*

Let $(\{t=4\} \times \mathbb{R}^n \times K, g_0, g_1)$ be Cauchy data for the reduced Einstein equations (31). *Assume that, for $|x| \geq 1$ with respect to Minkowski coordinates on $\mathbb{R}^{1+n}$, $(g_0, g_1) = (g_S + k, 0)$ where g_S is the Schwarzschild metric in the $\eta_{\mathbb{R}^{1+n}}$-wave gauge with parameter $C_S \in [0, \infty)$.*

There is an $\epsilon > 0$ such that, if the initial data satisfies

$$\sum_{|I| \leq N} \|\nabla[g_0]^I (g_0 - \hat{g}|_{t=4})\|^2_{L^2(\mathbb{R}^n \times K)} + \sum_{|I| \leq N-1} \|\nabla[g_0]^I g_1\|^2_{L^2(\mathbb{R}^n \times K)} + C_S^2 \leq \epsilon, \tag{58}$$

then there is a future global solution $g_{\mu\nu}$ of the reduced Einstein equations (31) *with initial data $(h, \partial_t h)|_{t=4} = (g_0, g_1)$. Furthermore, there is the bound*

$$\sup_{(t,x,\omega) \in \Sigma_s \times K} s^{4\delta(n)} |g(t, x^i, \omega) - \hat{g}(t, x^i, \omega)|^2_E \lesssim \epsilon, \tag{59}$$

where $\delta(n)$ was defined in (7).

Proof. Let the perturbation and inverse perturbation be denoted, respectively, by

$$h_{\mu\nu} = g_{\mu\nu} - \hat{g}_{\mu\nu} \quad \text{and} \quad H^{\mu\nu} = g^{\mu\nu} - \hat{g}^{\mu\nu}.$$

Since g is a solution of the reduced Einstein equation (31), it follows that

$$(\hat{g}^{\alpha\beta} + H^{\alpha\beta}) \nabla[\hat{g}]_\alpha \nabla[\hat{g}]_\beta h_{\mu\nu} + 2(R[\hat{g}] \circ h)_{\mu\nu} = Q_{\mu\nu}[g](\nabla[\hat{g}]h, \nabla[\hat{g}]h) + F_{\mu\nu}(H, h), \tag{60}$$

where $Q_{\mu\nu}$ is defined in (31b) and $F_{\mu\nu}$ is defined by

$$F_{\mu\nu}(H, h) = H^{\alpha\beta}(h_{\alpha\delta} \operatorname{Riem}[\hat{g}]^\delta{}_{\mu\nu\beta} + h_{\alpha\delta} \operatorname{Riem}[\hat{g}]^\delta{}_{\nu\mu\beta}) + H^{\alpha\beta}(h_{\mu\delta} \operatorname{Riem}[\hat{g}]^\delta{}_{\alpha\nu\beta} + h_{\nu\delta} \operatorname{Riem}[\hat{g}]^\delta{}_{\alpha\mu\beta}).$$

By commuting the symmetries $Z^I (\Delta_k)^j$ through the system (60) we obtain

$$(\hat{g}^{\alpha\beta} + H^{\alpha\beta}) \nabla[\hat{g}]_\alpha \nabla[\hat{g}]_\beta (Z^I (\Delta_k)^j h_{\mu\nu}) - 2(R[\hat{g}] \circ Z^I (\Delta_k)^j h)_{\mu\nu} = \sum_{i=1}^{3} F^{i,I,j}_{\mu\nu}, \tag{61}$$

where

$$\begin{aligned} F^{1,I,j}_{\mu\nu} &= Z^I (\Delta_k)^j Q_{\mu\nu}[g](\nabla[\hat{g}]h, \nabla[\hat{g}]h), \\ F^{2,I,j}_{\mu\nu} &= Z^I (\Delta_k)^j F_{\mu\nu}(H, h), \\ F^{3,I,j}_{\mu\nu} &= [Z^I (\Delta_k)^j, H^{\alpha\beta} \nabla[\hat{g}]_\alpha \nabla[\hat{g}]_\beta] h_{\mu\nu}. \end{aligned} \tag{62}$$

The symmetry boosted energy is given by

$$\mathcal{E}_{k+1}(s) = \sum_{|I|+2j \leq k} \mathcal{E}[H; Z^I (\Delta_k)^j g; s]. \tag{63}$$

From Lemma 4.2 and the Cauchy–Schwarz inequality we obtain

$$\mathcal{E}_{N+1}(s')^{1/2} \leq \mathcal{E}_{N+1}(4)^{1/2} + \sum_{|I|+2j\leq N} \int_4^{s'} \left(\int_{\Sigma_s \times K} \left(\sum_{i=1}^{3} |F^{i,I,j}|_E^2 + |G^{I,j}|_E^2 \right) \mathrm{d}y\, \mathrm{d}\mu_k \right)^{1/2} \mathrm{d}s, \tag{64}$$

where the $G^{I,j}$ terms arise from applying $Z^I(\Delta_k)^j$ to the terms involving $\nabla[\hat{g}]\gamma$ or $\partial_t\gamma$ on the right side of the energy equality (43). In particular, these can be bounded by

$$|G^{I,j}|_E^2 \leq C|\nabla[\hat{g}]H|_E^2 |Z^I(\Delta_k)^j \nabla[\hat{g}]h|_E^2. \tag{65}$$

The reduced field equations (60) are a system of quasilinear, quasidiagonal wave equations for the perturbation $h_{\mu\nu}$ of the spacetime metric. The existence of unique local solutions emanating from Cauchy data is standard [Choquet-Bruhat 2009, Theorem 4.6 Appendix III].

The proof then follows a bootstrap argument (or continuous induction): we prove that there exist $C > 0$ and $\epsilon > 0$ such that, if $\mathcal{E}_{N+1}(4) + C_S < \epsilon$ and $\mathcal{E}_{N+1}(s) \leq C\epsilon$ for all s, then $\mathcal{E}_{N+1}(s) \leq \epsilon + C\epsilon^2$ for all s and hence $\mathcal{E}_{N+1}(s) \leq \frac{1}{2}C\epsilon$. We note that there is no loss of generality in placing our initial data at $t = 4$.

We consider the integral term on the right-hand side in (64) as the sum of integrals over $\Sigma_s \cap \{|x| \leq t-1\}$ and over $\Sigma_s \cap \{|x| > t-1\}$. Our approach is that, for sufficiently small C_S, in the latter exterior region the solution is identically the product of Schwarzschild with the internal manifold. Thus in the region $|x| \geq t-1$ the perturbation $h_{\mu\nu}$ is only nonzero on its Minkowski indices and on these indices it is identically Schwarzschild. We note that sufficiently small compactly supported initial data on $\{t = 4\} \cap \{|x| \leq 1\}$ can be extended to compactly supported initial data on Σ_4 [LeFloch and Ma 2014, Chapter 39].

Recall from Section 2E that the difference between components of the Minkowski metric and the Schwarzschild metric in wave coordinates decay as $C_S r^{-n+2}$ and the Christoffel symbols decay as $C_S r^{-n+1}$. Along a geodesic parametrised by λ, one has

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2} = \Gamma^i_{jk} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda}.$$

Since $C_S r^{-n+1}$ is integrable in r, there are geodesics along which t and r grow linearly and the $\mathrm{d}x^j/\mathrm{d}\lambda$ approach constant values, not all of which are vanishing. In particular, $\mathrm{d}r/\mathrm{d}t$ asymptotically approaches a constant, and this constant is 1 for null geodesics. The next-to-leading-order term in the geodesic equation arises from the metric, so it is of the form Cr^{-n+2}, which is again integrable. Furthermore, the smaller the mass C_S the sooner this asymptotic behaviour comes to dominate. In particular, if C_S is sufficiently small, then any causal curve launched from within $\Sigma_4 \cap \{|x| \leq t-2\}$ can never reach the region where $|x| \geq t-1$. Furthermore, by uniqueness of solutions to quasilinear wave equations, since the initial data on Σ_4 is identically Schwarzschild for $|x| > t-2$, the solution is identically Schwarzschild for $|x| > t-1$. In particular, when estimating the components of the solution to (61), we can use the Sobolev Lemma 3.2 and Corollary 4.6 on hyperboloids with eventually prescribed functions. (The conclusion of this paragraph is essentially Proposition 2.3 of [LeFloch and Ma 2016].)

The estimate (40) required by Lemma 4.2 is established by combining (37) with the bootstrap assumptions and noting that since $n \geq 9$ we certainly have $\delta(n) > 1$. Similarly since $n \geq 9$ the decay assumptions (55) in Corollary 4.6 and (35) in Lemma 3.2 are satisfied.

We are now in a position to apply the results from Section 4B to the nonlinearities in (64). In general we will distribute $(s/t)(t/s) = 1$ across the terms and estimate high-derivative terms with a factor of (s/t) using Corollary 4.6 and low-derivative terms with a factor of (t/s) using Corollary 4.7. We begin by estimating the term $G^{I,j}$. Using (65) we find

$$\begin{aligned}\sum_{|I|+2j\le N} \|G^{I,j}\|_{L^2(\Sigma_s\times K)} &\lesssim \sum_{|I|+|J|\le N} \biggl(\int_{\Sigma_s\times K} |(t/s)\nabla[\hat{g}]H|_E^2\,|(s/t)Z^I\nabla[k]^J\nabla[\hat{g}]h|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2}\\ &\le \sup_{\Sigma_s\times K}(|(t/s)\nabla[\hat{g}]h|_E)\biggl(\int_{\Sigma_s\times K}|(s/t)Z^I\nabla[k]^J\nabla[\hat{g}]h|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2}\\ &\lesssim \frac{1}{s^{2\delta(n)-1}}(\mathcal{E}_{\tilde{\nu}+3}(s)^{1/2}+C_S)(s\mathcal{E}_{N+1}(s)^{1/2}+C_S).\end{aligned}\tag{66}$$

The term $F^1_{\mu\nu}$ involves the standard quadratic derivative nonlinearities of the Einstein equations. Their weak null structure is of course not relevant here since the Minkowski dimension is taken so high. We first look at what type of terms are contained in $F^1_{\mu\nu}$:

$$\begin{aligned}&\sum_{|I|+2j\le N}\|F^{1,I,j}_{\mu\nu}\|_{L^2(\Sigma_s\times K)}\\ &\lesssim \sum_{|I|+|J|\le N}\biggl(\int_{\Sigma_s\times K}|(\hat{g}+H)^{-1}|_E^2\,|Z^I\nabla[k]^J(\nabla[\hat{g}]h\nabla[\hat{g}]h)|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2}\\ &\qquad+\sum_{\substack{|I_i|+|J_i|\le N\\ |I_1|+|J_1|\ge 1}}\biggl(\int_{\Sigma_s\times K}|Z^{I_1}\nabla[k]^{J_1}h|_E^2\,|Z^{I_2}\nabla[k]^{J_2}(\nabla[\hat{g}]h\nabla[\hat{g}]h)|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2}.\end{aligned}\tag{67}$$

We treat the first term on the right-hand side of (67) since the second term is higher-order and thus easier to estimate. Once again we estimate high-derivative terms with a factor of (s/t) using Corollary 4.6 and low-derivative terms with a factor of (t/s) using Corollary 4.7. This yields

$$\begin{aligned}&\sum_{|I|+|J|\le N}\biggl(\int_{\Sigma_s\times K}|(\hat{g}+H)^{-1}|_E^2\,|Z^I\nabla[k]^J(\nabla[\hat{g}]h\nabla[\hat{g}]h)|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2}\\ &\qquad\lesssim\sum_{\substack{|I_i|+|J_i|\le N\\ |I_2|+|J_2|\le N/2+1}}\biggl(\int_{\Sigma_s\times K}C|Z^{I_1}\nabla[k]^{J_1}\nabla[\hat{g}]h|\,|Z^{I_2}\nabla[k]^{J_2}\nabla[\hat{g}]h|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2},\end{aligned}\tag{68}$$

where by symmetry we can assume $|I_2|+|J_2|\le \frac{1}{2}N+1$. After using $(s/t)(t/s)=1$ we find

$$\begin{aligned}&\sum_{\substack{|I_i|+|J_i|\le N\\ |I_2|+|J_2|\le N/2+1}}\biggl(\int_{\Sigma_s\times K}C|(s/t)Z^{I_1}\nabla[k]^{J_1}\nabla[\hat{g}]h|\,|(t/s)Z^{I_2}\nabla[k]^{J_2}\nabla[\hat{g}]h|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2}\\ &\qquad\lesssim\sup_{\Sigma_s\times K}\biggl(\sum_{|I_2|+|J_2|\le N/2+1}|(t/s)Z^{I_2}\nabla[k]^{J_2}\nabla[\hat{g}]h|_E\biggr)\\ &\qquad\qquad\times\sum_{|I_1|+|J_1|\le N}\biggl(\int_{\Sigma_s\times K}|(s/t)Z^{I_1}\nabla[k]^{J_1}\nabla[\hat{g}]h|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\biggr)^{1/2}\end{aligned}$$

$$\lesssim \frac{1}{s^{2\delta(n)-1}}\bigg(\sum_{|I|+2j\le\tilde{\nu}+N/2+3}\mathcal{E}[0;\,Z^I(\Delta_k)^j u;\,s]^{1/2}+C_S\sum_{|I|\le\tilde{\nu}+N/2}C_I^2\bigg)(s\mathcal{E}_{N+1}(s)^{1/2}+C_S)$$

$$\lesssim \frac{1}{s^{2\delta(n)-2}}(\mathcal{E}_{\tilde{\nu}+N/2+4}(s)^{1/2}+C_S)(\mathcal{E}_{N+1}(s)^{1/2}+C_S). \tag{69}$$

The term $F^2_{\mu\nu}$ involves the new nonlinearities which are only nonzero when both $\mu,\nu\in\{A,\dots,B\}$. This means we can control $F^2_{\mu\nu}$ as follows:

$$\sum_{|I|+2j\le N}\|F^{2,I,j}_{\mu\nu}\|_{L^2(\Sigma_s\times K)}\lesssim \sup_{\Sigma_s\times K}\bigg(\sum_{|I_0|\le N}|\nabla[k]^{I_0}\,\mathrm{Riem}[k]|\bigg)$$
$$\times\sum_{|I_i|+|J_i|\le N}\bigg(\int_{\Sigma_s\times K}|Z^{I_1}\nabla[k]^{J_1}h|_E^2\,|Z^{I_2}\nabla[k]^{J_2}h|_E^2\,\mathrm{d}y\,\mathrm{d}\mu_k\bigg)^{1/2}. \tag{70}$$

The Riemann curvature components of k are bounded (since K is compact) which allows us to control the first factor in (70). To estimate the second factor in (70) we follow the same procedure as in $F^1_{\mu\nu}$, by controlling high-derivatives with a factor of (s/t) using Corollary 4.6 and low-derivatives with a compensating factor of (t/s) using Corollary 4.7. The result of this procedure leads to a term controlled by (69).

The final term $F^3_{\mu\nu}$ is a commutator involving the quasilinear perturbation of the principal part of the differential operator. Note first the identity

$$\sum_{|I|+2j\le N}|F^{3,I,j}_{\mu\nu}|_E\le C\sum_{\substack{|I_i|+|J_i|\le N\\|I_2|+|J_2|\le N-1}}|Z^{I_1}\nabla[k]^{J_1}H|_E\,|Z^{I_1}\nabla[k]^{J_1}\nabla[\hat{g}]\nabla[\hat{g}]h|_E. \tag{71}$$

Once again we distribute the product $(s/t)(t/s)=1$ across the two terms appearing here depending on where the derivatives land. The term with high-derivatives gains a factor of (s/t) and is controlled using Corollary 4.6 while the term with low-derivatives absorbs a compensating factor of (t/s) and is estimated using Corollary 4.7. Note that when the term $Z^{I_2}\nabla[k]^{J_2}(\nabla[\hat{g}]\nabla[\hat{g}]h)$ is estimated in L^∞, the Sobolev inequality will lead to a symmetry boosted energy at order $\tilde{\nu}+\frac{1}{2}N+5$. We eventually obtain

$$\sum_{|I|+2j\le N}\|F^{3,I,j}_{\mu\nu}\|_{L^2(\Sigma_s\times K)}\lesssim\frac{1}{s^{2\delta(n)-2}}(\mathcal{E}_{\tilde{\nu}+N/2+5}(s)^{1/2}+C_S)(\mathcal{E}_{N+1}(s)^{1/2}+C_S). \tag{72}$$

Putting these all together, inserting the bootstrap assumptions, and using also $C_S^2<\epsilon$, we find

$$\sum_{|I|+2j\le N}\int_4^{s'}\bigg(\int_{\Sigma_s\times K}\bigg(\sum_{i=1}^3|F^{i,I,j}|_E^2+|G^{I,j}|_E^2\bigg)\mathrm{d}y\,\mathrm{d}\mu_k\bigg)^{1/2}\mathrm{d}s\lesssim\epsilon\int_4^{s'}\frac{1}{s^{2\delta(n)-2}}\,\mathrm{d}s. \tag{73}$$

For integrability we require $2\delta(n)-2>1$, which is equivalent to each of the following:

$$\delta(n)>\tfrac{3}{2}\quad\text{and}\quad n>8. \tag{74}$$

This implies $n\ge 9$. For the Sobolev estimates we require

$$\tilde{\nu}+\tfrac{1}{2}N+4\le N. \tag{75}$$

Recalling the definition of $\tilde{\nu}$ given in Lemma 3.2, this holds provided $N>\frac{1}{2}(n+d+8)$ and N is even.

Consequently for sufficiently small ϵ and by Grönwall's inequality applied to the energy estimate (64) we find $\mathcal{E}_{\nu+1}(s) \le \frac{1}{2}C_1\epsilon$. We have thus obtained a future global solution $h_{\mu\nu} = g_{\mu\nu} - \hat{g}_{\mu\nu}$ to the reduced Einstein equations which clearly satisfies the decay bounds given in Theorem 5.1. □

Remark 5.2. The system (60) contains quadratic nonlinearities F_{AB} and F_{iA} that are new compared to the weak null terms identified in the proof of Minkowski stability in [Lindblad and Rodnianski 2003; 2010] and the proof of zero-mode Kaluza–Klein stability in [Wyatt 2018].

5B. ***Proof of Theorem 1.1.*** We are now in a position to use the results from Theorem 5.1 in order to prove our main result. Take an initial data set $(\mathbb{R}^n \times K, \gamma, \kappa)$ as specified in Theorem 1.1 with smallness conditions (5). We now transform this data into the form required by Theorem 5.1, which is a standard procedure; see for example [Lindblad and Rodnianski 2005]. We first set $((g_0)_{i'j'}, (g_1)_{i'j'}) = (\gamma_{i'j'}, \kappa_{i'j'})$. Diffeomorphism invariance allows us the freedom to choose the lapse and shift. We set the shift to be zero: $X_{i'} = 0$. We choose the lapse to be a smooth function satisfying

$$\begin{aligned} N(r) &= 1, & r &\le \tfrac{1}{2}, \\ |N-1| &\lesssim C_S, & \tfrac{1}{2} &\le r \le 1, \\ N(r) &= \left(1 - \frac{h_{00}(r^{-1})}{r^{n-2}}\right)^{1/2}, & r &\ge 1. \end{aligned}$$

We relate the lapse and shift with the Cauchy data for the reduced equations in Theorem 5.1 by setting $(g_0)_{00} = -N^2$ and $(g_0)_{0i'} = X_{i'}$. The initial data for $(\partial_t N, \partial_t X_{i'}) = ((g_1)_{00}, (g_1)_{0i'})$ is chosen by satisfying $V^\gamma = 0$. This amounts to solving the following equations on $\mathbb{R}^n \times K$:

$$\begin{aligned} N^{-3}((g_1)_{00} + N^2\gamma^{i'j'}\kappa_{i'j'}) &= {g_0}^{i'j'}\Gamma[\hat{e}]^0_{i'j'}, \\ -N^{-2}\gamma^{i'j'}(g_1)_{0j'} - N^{-1}\gamma^{i'j'}\partial_{j'}N + \gamma^{j'k'}\Gamma^{i'}_{j'k'}[\gamma] &= {g_0}^{j'k'}\Gamma[\hat{e}]^{i'}_{j'k'}. \end{aligned} \tag{76}$$

We have now brought the initial data of Theorem 1.1 into the form of Theorem 5.1. It remains to check that our assumptions on the lapse and shift are compatible with smallness conditions (58). To do this, recall the final sentence of Theorem 2.15. This implies that

$$\begin{aligned} \int_{\{r\ge1\}\cap\mathbb{R}^n} |\nabla[g_0]^I(-N^2-\eta_{00})|^2\,\mathrm{d}x &\le \int_{\{r\ge1\}\cap\mathbb{R}^n} C_S^2(r^{-(n-2)-|I|})^2 r^{n-1}\,\mathrm{d}r\,\mathrm{d}^{n-1}\omega_{\mathbb{S}^{n-1}} \\ &\le C_S^2\int_{\{r\ge1\}\cap\mathbb{R}^n} r^{-(n-3)-2|I|}\,\mathrm{d}r\,\mathrm{d}^{n-1}\omega_{\mathbb{S}^{n-1}} \le CC_S^2. \end{aligned}$$

By inverting the expressions (76) for $(\partial_t N, \partial_t X_{i'})$ it is clear that the smallness conditions (58) are satisfied. Furthermore it is a standard result, see for example [Choquet-Bruhat 2009, Theorem 8.3], that the future global solution constructed in Theorem 5.1 is in fact also a solution to the full Einstein equations.

Finally, note that the solution found in Theorem 5.1 is only defined to the future $t \ge 4$. Nonetheless, by time translation, we can treat the initial data as being on $\{t = 0\}$ instead of $\{t = 4\}$, so that Theorem 5.1 ensures the existence of a solution for $t \ge 0$. By time reversibility for the Einstein equation (and the reduced Einstein equation), we similarly obtain a solution for $t \le 0$. Thus, we can construct the global solution required in Theorem 1.1.

It now remains to prove the causal geodesic completeness of $(\mathbb{R}^{1+n} \times K, g)$.

Globally, the metrics g and $\hat{g}$ are very close, in the sense that, with respect to a basis constructed from the X_i and an orthonormal basis on K, their components vanish to order ϵ globally. Denote from now onwards $T = \mathrm{d}t$. This is a globally timelike one-form such that $|g(T,T) - 1| \lesssim \epsilon$. Thus, $g - 2TT$ defines a Riemannian metric. (Note that in the introduction, we used the slightly different Euclidean metric $\hat{g} - 2TT$.) Within this proof, we define, for a vector u, the Euclidean length to be

$$|u|^2 = u^\alpha u^\beta (g_{\alpha\beta} + 2T_\alpha T_\beta). \tag{77}$$

Note that the fact that g and $\hat{g}$ are very close implies the equivalence $|u|_E \sim |u|$.

Consider a causal geodesic γ that is affinely parametrised by λ. For the remainder of this paragraph, let $t = t(\lambda)$ denote the value of the Cartesian coordinate t at the point $\gamma(\lambda)$. By rescaling, we may assume that $\mathrm{d}t/\mathrm{d}\lambda = 1$ at $t = 0$. Let v be the (artificial, Euclidean) speed defined by $v \geq 0$ and

$$v^2 = \left| \frac{\mathrm{d}\gamma^\alpha}{\mathrm{d}\lambda} \right|^2. \tag{78}$$

Since g and $\hat{g}$ are very close, the rate of change in the t direction cannot be (much) greater than the Euclidean speed, i.e.,

$$\left| \frac{\mathrm{d}t}{\mathrm{d}\lambda} \right| = \left| \frac{\mathrm{d}\gamma^0}{\mathrm{d}\lambda} \right| \lesssim v.$$

On the other hand, since γ is causal, the component of $\mathrm{d}\gamma/\mathrm{d}\lambda$ in the T direction cannot vanish faster than the length of the component in the orthogonal spatial directions, and the square of the Euclidean velocity is the sum of the squares of the lengths of the T components and the orthogonal spatial component (up to order ϵ multiplicative errors); thus

$$\left| \frac{\mathrm{d}t}{\mathrm{d}\lambda} \right| = \left| \frac{\mathrm{d}\gamma^0}{\mathrm{d}\lambda} \right| \gtrsim v.$$

In particular, there is the equivalence $|\mathrm{d}t/\mathrm{d}\lambda| \sim v$.

Since $\nabla[g]g = 0$ and $\nabla[g]_{\mathrm{d}\gamma/\mathrm{d}\lambda} \mathrm{d}\gamma/\mathrm{d}\lambda = 0$, the rate of change of the velocity is given by

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} v^2 = 4 \left(\frac{\mathrm{d}\gamma^\alpha}{\mathrm{d}\lambda} T_\alpha \right) \left(\frac{\mathrm{d}\gamma^\beta}{\mathrm{d}\lambda} \nabla[g]_{\mathrm{d}\gamma/\mathrm{d}\lambda} T_\beta \right). \tag{79}$$

Since the absolute value of $(\mathrm{d}\gamma^\alpha/\mathrm{d}\lambda)T_\alpha = \mathrm{d}t/\mathrm{d}\lambda$ and the Euclidean length of $\mathrm{d}\gamma/\mathrm{d}\lambda$ are dominated by v,

$$\frac{\mathrm{d}v}{\mathrm{d}\lambda} \lesssim |\nabla[g]_{\mathrm{d}\gamma/\mathrm{d}\lambda} T| v. \tag{80}$$

The $\nabla[g]T$ can be expanded in terms of g and $\nabla[\hat{g}]g$. Both of these have norms that decay as $t^{-\delta(n)}$ due to (74). Thus,

$$\frac{\mathrm{d}v}{\mathrm{d}\lambda} \lesssim \epsilon t^{-\delta(n)} v^2. \tag{81}$$

Thus, for ϵ sufficiently small, a simple bootstrap argument shows that $v \sim 1$ along all of γ, and hence $\mathrm{d}t/\mathrm{d}\lambda \sim 1$. In particular, t is monotone along γ.

Let t_{sup} be the supremum of the t values that are achieved along γ. For contradiction, suppose $t_{\text{sup}} < \infty$. Since the length of the spatial component of $\mathrm{d}\gamma/\mathrm{d}\lambda$ is also uniformly equivalent to v, and hence to $\mathrm{d}t/\mathrm{d}\lambda$, it follows that, as $t \nearrow t_{\text{sup}}$, the curve γ has a limit in $\mathbb{R}^{1+n} \times K$. Because of the global bounds on g and its derivatives, by the standard Picard–Lindelöf theorem for ODEs, the curve γ must smoothly extend through this limiting point, contradicting the definition of t_{sup}. Thus, $t_{\text{sup}} = \infty$. The only other way in which γ can be future incomplete is if t diverges to ∞ in a finite λ interval, but this is also impossible, since $\mathrm{d}t/\mathrm{d}\lambda \sim 1$. By time symmetry, the same argument holds in the past. Thus, any causal geodesic is complete.

The previous construction shows that every causal geodesic goes through each level set of t. Thus, the level sets of t are Cauchy surfaces, and $(\mathbb{R}^{1+n} \times K, g)$ is globally hyperbolic.

Acknowledgements

Andersson thanks the Mathematics Department at Harvard University for hospitality and support during the spring term of 2011, when the initial investigation of this topic was carried out. Andersson is also grateful to the YMSC at Tsinghua University, Beijing, for support and hospitality during part of the work of this paper, and thanks Pin Yu and Po-Ning Chen for helpful discussions. The authors Andersson, Blue, and Wyatt thank the hospitality of the Institut Mittag-Leffler, Djursholm, Sweden during the winter semester of 2019. Wyatt is supported by MIGSAA, a Centre for Doctoral Training funded by EPSRC, the Scottish Funding Council, Heriot–Watt University, and the University of Edinburgh. Yau is supported by the Black Hole Initiative, award 61497, at Harvard University and by NSF grant DMS-1607871 “Analysis, Geometry and Mathematical Physics”. We thank the journal reviewers for their thoughtful comments, which helped us improve the paper.

References

[Andersson 2014] L. Andersson, “Cosmological models and stability”, pp. 277–303 in *General relativity, cosmology and astrophysics* (Prague, 2012), edited by J. Bičák and T. Ledvinka, Fundam. Theor. Phys. **177**, Springer, 2014. MR Zbl

[Andersson and Moncrief 2011] L. Andersson and V. Moncrief, “Einstein spaces as attractors for the Einstein flow”, *J. Differential Geom.* **89**:1 (2011), 1–47. MR Zbl

[Bartnik 1988] R. Bartnik, “Remarks on cosmological spacetimes and constant mean curvature surfaces”, *Comm. Math. Phys.* **117**:4 (1988), 615–624. MR Zbl

[Besse 1987] A. L. Besse, *Einstein manifolds*, Ergebnisse der Mathematik (3) **10**, Springer, 1987. MR Zbl

[Biquard 2013] O. Biquard, “Désingularisation de métriques d’Einstein, I”, *Invent. Math.* **192**:1 (2013), 197–252. MR Zbl

[Branding et al. 2019] V. Branding, D. Fajman, and K. Kröncke, “Stable cosmological Kaluza–Klein spacetimes”, *Comm. Math. Phys.* **368**:3 (2019), 1087–1120. MR Zbl

[Brendle and Kapouleas 2017] S. Brendle and N. Kapouleas, “Gluing Eguchi–Hanson metrics and a question of Page”, *Comm. Pure Appl. Math.* **70**:7 (2017), 1366–1401. MR Zbl

[Candelas et al. 1985] P. Candelas, G. T. Horowitz, A. Strominger, and E. Witten, “Vacuum configurations for superstrings”, *Nuclear Phys. B* **258**:1 (1985), 46–74. MR

[Choquet-Bruhat 1952] Y. Fourès-Bruhat, “Théorème d’existence pour certains systèmes d’équations aux dérivées partielles non linéaires”, *Acta Math.* **88** (1952), 141–225. MR Zbl

[Choquet-Bruhat 1985] Y. Choquet-Bruhat, “Causalité des théories de supergravité”, pp. 79–93 in *The mathematical heritage of Élie Cartan* (Lyon, 1984), Astérisque **S131**, Soc. Math. France, Paris, 1985. MR Zbl

[Choquet-Bruhat 2009] Y. Choquet-Bruhat, *General relativity and the Einstein equations*, Oxford Univ. Press, 2009. MR Zbl

[Choquet-Bruhat and Geroch 1969] Y. Choquet-Bruhat and R. Geroch, "Global aspects of the Cauchy problem in general relativity", *Comm. Math. Phys.* **14**:4 (1969), 329–335. MR Zbl

[Choquet-Bruhat and Moncrief 2001] Y. Choquet-Bruhat and V. Moncrief, "Future global in time Einsteinian spacetimes with U(1) isometry group", *Ann. Henri Poincaré* **2**:6 (2001), 1007–1064. MR Zbl

[Choquet-Bruhat et al. 2006] Y. Choquet-Bruhat, P. T. Chruściel, and J. Loizelet, "Global solutions of the Einstein–Maxwell equations in higher dimensions", *Classical Quantum Gravity* **23**:24 (2006), 7383–7394. MR Zbl

[Christodoulou and Klainerman 1993] D. Christodoulou and S. Klainerman, *The global nonlinear stability of the Minkowski space*, Princeton Math. Series **41**, Princeton Univ. Press, 1993. MR Zbl

[Cipriani and Senovilla 2019] N. Cipriani and J. M. M. Senovilla, "Singularity theorems for warped products and the stability of spatial extra dimensions", *J. High Energy Phys.* **2019**:4 (2019), art. id. 175. MR Zbl

[Dai 2004] X. Dai, "A positive mass theorem for spaces with asymptotic SUSY compactification", *Comm. Math. Phys.* **244**:2 (2004), 335–345. MR Zbl

[Dai et al. 2005] X. Dai, X. Wang, and G. Wei, "On the stability of Riemannian manifold with parallel spinors", *Invent. Math.* **161**:1 (2005), 151–176. MR Zbl

[Ettinger 2015] B. Ettinger, "Well-posedness of the equation for the three-form field in eleven-dimensional supergravity", *Trans. Amer. Math. Soc.* **367**:2 (2015), 887–910. MR Zbl

[Galloway 2019] G. J. Galloway, "Existence of CMC Cauchy surfaces and spacetime splitting", *Pure Appl. Math. Q.* **15**:2 (2019), 667–682. MR Zbl

[Galloway and Senovilla 2010] G. J. Galloway and J. M. M. Senovilla, "Singularity theorems based on trapped submanifolds of arbitrary co-dimension", *Classical Quantum Gravity* **27**:15 (2010), art. id. 152002. MR Zbl

[Hörmander 1997] L. Hörmander, *Lectures on nonlinear hyperbolic differential equations*, Math. Appl. **26**, Springer, 1997. MR Zbl

[Huneau and Stingo 2021] C. Huneau and A. Stingo, "Global well-posedness for a system of quasilinear wave equations on a product space", preprint, 2021. arXiv 2110.13982

[Huneau and Vâlcu 2021] C. Huneau and C. Vâlcu, "Einstein constraint equations for Kaluza–Klein spacetimes", preprint, 2021. arXiv 2111.14395

[Kaluza 1921] T. Kaluza, "Zum Unitätsproblem in der Physik", *S.-B. Preussische Akad. Wissenschaften* **1921** (1921), 966–972. Zbl

[Klainerman 1985] S. Klainerman, "Global existence of small amplitude solutions to nonlinear Klein–Gordon equations in four space-time dimensions", *Comm. Pure Appl. Math.* **38**:5 (1985), 631–641. MR Zbl

[Klainerman 1986] S. Klainerman, "The null condition and global existence to nonlinear wave equations", pp. 293–326 in *Nonlinear systems of partial differential equations in applied mathematics*, *I* (Santa Fe, NM, 1984), edited by B. Nicolaenko et al., Lectures Appl. Math. **23**, Amer. Math. Soc., Providence, RI, 1986. MR Zbl

[Klein 1926] O. Klein, "Quantentheorie und fünfdimensionale Relativitätstheorie", *Z. Physik* **37**:12 (1926), 895–906. Zbl

[Kröncke 2015] K. Kröncke, "On the stability of Einstein manifolds", *Ann. Global Anal. Geom.* **47**:1 (2015), 81–98. MR Zbl

[LeFloch and Ma 2014] P. G. LeFloch and Y. Ma, *The hyperboloidal foliation method*, Ser. Appl. Comput. Math. **2**, World Sci., Hackensack, NJ, 2014. MR Zbl

[LeFloch and Ma 2016] P. G. LeFloch and Y. Ma, "The global nonlinear stability of Minkowski space for self-gravitating massive fields", *Comm. Math. Phys.* **346**:2 (2016), 603–665. MR Zbl

[Lindblad and Rodnianski 2003] H. Lindblad and I. Rodnianski, "The weak null condition for Einstein's equations", *C. R. Math. Acad. Sci. Paris* **336**:11 (2003), 901–906. MR Zbl

[Lindblad and Rodnianski 2005] H. Lindblad and I. Rodnianski, "Global existence for the Einstein vacuum equations in wave coordinates", *Comm. Math. Phys.* **256**:1 (2005), 43–110. MR Zbl

[Lindblad and Rodnianski 2010] H. Lindblad and I. Rodnianski, "The global stability of Minkowski space-time in harmonic gauge", *Ann. of Math.* (2) **171**:3 (2010), 1401–1477. MR Zbl

[Metcalfe and Stewart 2008] J. Metcalfe and A. Stewart, "Almost global existence for quasilinear wave equations in waveguides with Neumann boundary conditions", *Trans. Amer. Math. Soc.* **360**:1 (2008), 171–188. MR Zbl

[Metcalfe et al. 2005] J. Metcalfe, C. D. Sogge, and A. Stewart, "Nonlinear hyperbolic equations in infinite homogeneous waveguides", *Comm. Partial Differential Equations* **30**:4-6 (2005), 643–661. MR Zbl

[Penrose 2003] R. Penrose, "On the instability of extra space dimensions", pp. 185–201 in *The future of the theoretical physics and cosmology* (Cambridge, 2002), edited by G. W. Gibbons et al., Cambridge Univ. Press, 2003. MR

[Penrose 2005] R. Penrose, *The road to reality*: *a complete guide to the laws of the universe*, Knopf, New York, 2005. MR Zbl

[Reiris 2010] M. Reiris, "The ground state and the long-time evolution in the CMC Einstein flow", *Ann. Henri Poincaré* **10**:8 (2010), 1559–1604. MR Zbl

[Tian and Yau 1990] G. Tian and S.-T. Yau, "Complete Kähler manifolds with zero Ricci curvature, I", *J. Amer. Math. Soc.* **3**:3 (1990), 579–609. MR Zbl

[Tian and Yau 1991] G. Tian and S.-T. Yau, "Complete Kähler manifolds with zero Ricci curvature, II", *Invent. Math.* **106**:1 (1991), 27–60. MR Zbl

[Wang 1989] M. Y. Wang, "Parallel spinors and parallel forms", *Ann. Global Anal. Geom.* **7**:1 (1989), 59–68. MR Zbl

[Wang 1991] M. Y. Wang, "Preserving parallel spinors under metric deformations", *Indiana Univ. Math. J.* **40**:3 (1991), 815–844. MR Zbl

[Witten 1982] E. Witten, "Instability of the Kaluza–Klein vacuum", *Nuclear Phys. B* **195**:3 (1982), 481–492. Zbl

[Wyatt 2018] Z. Wyatt, "The weak null condition and Kaluza–Klein spacetimes", *J. Hyperbolic Differ. Equ.* **15**:2 (2018), 219–258. MR Zbl

Received 30 Nov 2020. Revised 25 Feb 2022. Accepted 25 Mar 2022.

LARS ANDERSSON: lars.andersson@bimsa.cn
Albert Einstein Institute, Max–Planck Institute for Gravitational Physics, Potsdam, Germany

PIETER BLUE: p.blue@ed.ac.uk
School of Mathematics, Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh, United Kingdom

ZOE WYATT: zoe.wyatt@maths.cam.ac.uk
Department of Pure Mathematics & Mathematical Statistics, Faculty of Mathematics, University of Cambridge, Cambridge, United Kingdom

SHING-TUNG YAU: styau@tsinghua.edu.cn
Center of Mathematical Sciences and Applications, Harvard University, Cambridge, MA, United States
Current address: *Yau Mathematical Sciences Center, Tsinghua University, Beijing, China*

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2109–2145
DOI: 10.2140/apde.2023.16.2109

STABILITY OF TRAVELING WAVES FOR THE BURGERS–HILBERT EQUATION

ÁNGEL CASTRO, DIEGO CÓRDOBA AND FAN ZHENG

We consider smooth solutions of the Burgers–Hilbert equation that are a small perturbation δ from a global periodic traveling wave with small amplitude ϵ. We use a modified energy method to prove the existence time of smooth solutions on a time scale of $1/(\epsilon\delta)$, with $0 < \delta \ll \epsilon \ll 1$, and on a time scale of ϵ/δ^2, with $0 < \delta \ll \epsilon^2 \ll 1$. Moreover, we show that the traveling wave exists for an amplitude ϵ in the range $(0, \epsilon^*)$, with $\epsilon^* \sim 0.23$, and fails to exist for $\epsilon > 2/e$.

1. Introduction

1A. ***The Burgers–Hilbert equation*** **(*BH*).** We study the size and stability of traveling waves of the Burgers–Hilbert equation (BH),

$$f_t = Hf + ff_x \quad \text{for } (x,t) \in \Omega \times \mathbb{R}, \tag{1-1}$$

$$f(x,0) = f_0(x), \tag{1-2}$$

where Ω is the real line $\mathbb{R}$ or the torus $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ and Hf is the Hilbert transform which is defined for $f : \mathbb{R}$ (resp. $\mathbb{T}$) $\to \mathbb{R}$ by

$$Hf(x) = \frac{1}{\pi}\,\mathrm{P.V.}\!\int_{\mathbb{R}} \frac{f(y)}{x-y}\,dy \quad \text{resp. } Hf(x) = \frac{1}{2\pi}\,\mathrm{P.V.}\!\int_0^{2\pi} f(y)\cot\frac{x-y}{2}\,dy.$$

Its action in the frequency space is $\widehat{Hf}(k) = -i\,\mathrm{sgn}\,k\,\hat{f}(k)$ for $k \neq 0$, and $\widehat{Hf}(0) = 0$.

This equation arose in [Marsden and Weinstein 1983] as a quadratic approximation for the evolution of the boundary of a simply connected vorticity patch in two dimensions. Later, Biello and Hunter [2010] proposed the model as an approximation for describing the dynamics of small slope vorticity fronts in the 2-dimensional incompressible Euler equations. Recently, the validity of this approximation was proved in [Hunter et al. 2022].

MSC2020: 35F25, 76B47.
Keywords: Burgers–Hilbert, normal forms, traveling waves.

By standard energy estimates the initial value problem for (BH) is locally well-posed in H^s for $s > \frac{3}{2}$. Bressan and Nguyen [2014] established in global existence of weak solutions for initial data $f_0 \in L^2(\mathbb{R})$, with $f(x,t) \in L^\infty(\mathbb{R}) \cap L^2(\mathbb{R})$ for all $t > 0$. Bressan and Zhang [2017] constructed locally in time piecewise continuous solutions to the BH equation with a single discontinuity where the Hilbert transform generates a logarithmic singularity. Uniqueness for general global weak solutions of [Bressan and Nguyen 2014] is open. But piecewise continuous solutions are shown to be unique in [Krupa and Vasseur 2020].

The Burgers–Hilbert equation can indeed form shocks in finite time. Various numerical simulations have been performed in [Biello and Hunter 2010; Hunter 2018; Klein and Saut 2015]. Finite time singularities, in the $C^{1,\delta}$ norm, with $0 < \delta < 1$, were shown to exist in [Castro et al. 2010] for initial data f_0 in $L^2(\mathbb{R}) \cap C^{1,\delta}(\mathbb{R})$ that has a point $x_0 \in \mathbb{R}$ such that $H(f_0)(x_0) > 0$ and $f_0(x_0) \geq (32\pi \| f_0 \|_{L^2})^{1/3}$. Recently, with a different approach, Saut and Wang [2022] proved shock formation in finite time for (BH) and Yang [2021] constructed solutions that develop an asymptotic self-similar shock at one single point with an explicitly computable blowup profile for (BH).

In this paper we are concerned with the dynamics in the small amplitude regime where (BH) can be viewed as a perturbation of the linearized (BH) equation $f_t = H[f]$. Since the nonlinear term in (1-1) is quadratic and the Hilbert transform is orthogonal in L^2, standard energy estimates yield a time of existence of smooth solutions $T \sim 1/\| f_0 \|$. Thanks to the effect of the Hilbert transform and using the normal form method, Hunter, Ifrim, Tataru and Wong (see [Hunter and Ifrim 2012; Hunter et al. 2015]) were able to improve this time of existence. More precisely, if ϵ is the size of the initial data, they prove a lifespan $T \sim 1/\epsilon^2$ for small enough ϵ (see also [Ehrnström and Wang 2019] for a similar approach with a modified version of the (BH) equation). The proofs are based on the normal form method and on the modified energy method. Furthermore, Hunter [2018] showed for $0 < \epsilon \ll 1$ the existence of C^∞-traveling wave solutions of the form

$$f_\epsilon(x,t) = u_\epsilon(x + v_\epsilon t),$$

with

$$u_\epsilon(x) = \epsilon \cos(x) + O(\epsilon^2), \tag{1-3}$$

$$v_\epsilon = -1 + O(\epsilon^2). \tag{1-4}$$

Notice that, $(u_\epsilon(nx)/n, v_\epsilon/n)$ is also a C^∞-traveling wave solution.

Throughout the paper we will assume that the initial data f_0 has zero mean. Since (1-1) preserves the mean,

$$\int_0^{2\pi} f(x,t)\,dx = 0 \quad \text{for all } t.$$

Since in the construction above u_ϵ also has zero mean,

$$\int_0^{2\pi} f(x,t)\,dx = 0 \quad \text{for all } t.$$

1B. *The main theorem.* In the present work we extend the results in the small amplitude regime in the following way:

(1) Size of the traveling waves: We show that the traveling waves exist for an amplitude ϵ in the range $(0, \epsilon^*)$, with $\epsilon^* \sim 0.23$, and fail to exist for $\epsilon > 2/e$.

(2) Extended lifespan from a traveling wave: We prove that a δ-perturbation of u_ϵ lives, at least, for a time $T \sim 1/(\delta\epsilon)$ for $0 < \delta \ll \epsilon \ll 1$, and for a time $T \sim \epsilon/\delta^2$ for $0 < \delta \ll \epsilon^2 \ll 1$. This is an improvement compared with the time $T \sim 1/\epsilon^2$ provided by the results in [Hunter and Ifrim 2012; Hunter et al. 2015]. Indeed, our main theorem reads:

Theorem 1.1. *For* $0 < |\epsilon|,\ \delta \ll 1$ *let* $(u_\epsilon, v_\epsilon) \in C^\infty(\mathbb{T}) \times \mathbb{R}$ *be a traveling wave solution of* (1-1) *as in* (1-3) *and* (1-4) *and*

$$\| f_0 - u_\epsilon \|_{H^4(\mathbb{T})} < \delta.$$

Then there exist $0 < \epsilon_0 \ll 1,\ T(\epsilon, \delta) > 0$ *and a solution of* (1-1)

$$f(x,t) \in C([0, T(\epsilon,\delta)); H^4(\mathbb{T}))$$

such that

(1) *if* $\delta \ll |\epsilon|$ *and* $|\epsilon| \le \epsilon_0$, *then* $T(\epsilon,\delta) \sim 1/(\epsilon\delta)$,

(2) *if* $\delta \ll \epsilon^2$ *and* $|\epsilon| \le \epsilon_0$, *then* $T(\epsilon,\delta) \sim \epsilon/\delta^2$.

Moreover, there are two differentiable functions $\epsilon(t)$ *and* $a(t)$ *such that*

$$\| f(x,t) - u_{\epsilon(t)}(x + a(t)) \|_{H^4} \lesssim \delta.$$

1C. ***Sketch of the proof of Theorem 1.1.*** Now we briefly describe the proof of Theorem 1.1. Assume that the solution

$$f(x,t) = u_\epsilon(x + v_\epsilon t) + g(x + v_\epsilon t, t)$$

is a small perturbation around the traveling wave $u_\epsilon(x + v_\epsilon t)$. Then the linearization of the Burgers–Hilbert equation (1-1) is

$$L_\epsilon g := -v_\epsilon g_x + Hg + (u_\epsilon(x) g)_x = 0$$

so to the first order, the perturbation g solves the equation $g_t = L_\epsilon g$, with solution

$$g(x,t) = e^{tL_\epsilon} g(x,0).$$

Therefore the linear evolution of g is determined by the eigenvalues of L_ϵ.

The full nonlinear evolution of g is

$$g_t = L_\epsilon g + N(g,g),$$

where $N(g,g)$ is a nonlinearity that is (at least) quadratic in g. We plug in the linear solution to get

$$g_t = e^{tL_\epsilon} L_\epsilon g(x,0) + N(e^{tL_\epsilon} g(x,0), e^{tL_\epsilon} g(x,0))$$

to second order, which integrates to

$$g(x,t) = e^{tL_\epsilon} g(x,0) + e^{tL_\epsilon} \int_0^t e^{-sL_\epsilon} N(e^{sL_\epsilon} g(x,0), e^{sL_\epsilon} g(x,0))\, ds.$$

Expand (at least formally) the initial data and the nonlinearity in terms of the eigenvectors of L_ϵ as

$$g(x,0)=\sum_n c_n\varphi_n(x), \qquad N(\varphi_l,\varphi_m)=\sum_n c_{lmn}\varphi_n,$$

where the eigenvalue of φ_n is λ_n. Then

$$g(x,t)\approx\sum_n c_n e^{\lambda_n t}\varphi_n(x)+\sum_{l,m,n}\frac{e^{(\lambda_l+\lambda_m)t}-e^{\lambda_n t}}{\lambda_l+\lambda_m-\lambda_n}c_{lmn}\varphi_n(x) \tag{1-5}$$

to second order, provided that the denominator $\lambda_l+\lambda_m-\lambda_n$ is not equal to 0, i.e., that the eigenvalues are "nonresonant". Then we can integrate (1-1) up to a cubic error term, yielding the "cubic lifespan", i.e., initial data of size ϵ leads to a solution that exists for a time at least comparable to ϵ^{-2}. This is the "normal form transformation", first proposed by Poincaré in the setting of ordinary differential equations (see [Arnold 1983] for a book reference). Its application to partial differential equations was initiated by Shatah [1985] in the study of the nonlinear Klein–Gordon equation, and then extended to the water wave problem by Germain, Masmoudi and Shatah [Germain et al. 2012; 2015] and Ionescu and Pusateri [2015; 2018], the Burgers–Hilbert equation by Hunter, Ifrim, Tataru and Wang [Hunter et al. 2015], and more recently, the Einstein–Klein–Gordon equation by Ionescu and Pausader [2022].

Unfortunately, nonresonance fails for L_ϵ because 0 is an eigenvalue, and $0+\lambda_n-\lambda_n=0$. The eigenvalue 0 arises from the symmetry of (1-1). Indeed, the initial data $u_\epsilon(x+\delta)\approx u_\epsilon(x)+\delta u_\epsilon'(x)$ produces the solution

$$f(x,t)=u_\epsilon(x+v_\epsilon t+\delta)\approx u_\epsilon(x+v_\epsilon t)+\delta u_\epsilon'(x+v_\epsilon t).$$

In this case $g(x,t)=\delta u_\epsilon'(x)$, with $g_t=0$, so $u_\epsilon'\in\ker L_\epsilon$. Also, the initial data $u_{\epsilon+\delta}(x)\approx u_\epsilon(x)+\delta\partial_\epsilon u_\epsilon(x)$ produces the solution

$$f(x,t)=u_{\epsilon+\delta}(x+v_{\epsilon+\delta}t)\approx u_\epsilon(x+v_\epsilon t)+\delta\partial_\epsilon u_\epsilon(x+v_\epsilon t)+\delta v_\epsilon' t u_\epsilon'(x+v_\epsilon t).$$

In this case $g(x,t)=\delta\partial_\epsilon u_\epsilon(x)+\delta v_\epsilon' t u_\epsilon'(x)$, so

$$L_\epsilon g=\delta L_\epsilon\partial_\epsilon u_\epsilon=g_t=\delta v_\epsilon' u_\epsilon'\in\ker L_\epsilon,$$

and thus $\partial_\epsilon u_\epsilon$ is in the generalized eigenspace corresponding to the eigenvalue 0.

These perturbations generate translations and variations along the bifurcation curve. We treat them separately using a more sophisticated ansatz

$$f(x,t)=u_{\epsilon(t)}(x+a(t))+g(x+a(t),t).$$

We will show in Proposition 4.1 that if $|\epsilon_0|$ and $\|f-u_{\epsilon_0}\|_{H^2}/|\epsilon_0|$ are sufficiently small, then f can always be put in the form above, with $|\epsilon-\epsilon_0|/|\epsilon_0|$ also small and the expansion of g not involving any eigenvector with eigenvalue 0. This way we remove the resonance caused by the eigenvalue 0 from the evolution of g.

We also need to analyze the other eigenvalues of L_ϵ, a first-order differential operator with variable coefficients, and a quasilinear perturbation from $L_0=\partial_x+H$, whose eigenvectors are the Fourier modes e^{inx}. Just like the Schrödinger operator with potential $-\Delta+V$, with a basis of eigenvectors known as the "Jost functions", giving rise to the "distorted Fourier transform" (see [Agmon 1975]), L_ϵ can also be diagonalized using a combination of conjugation and perturbative analysis. More precisely, let $g = h_x$. Then

$$L_\epsilon g = ((u_\epsilon(x) - v_\epsilon)g)_x + Hg = ((u_\epsilon(x) - v_\epsilon)h_x + Hh)_x,$$

so L_ϵ is conjugate to the operator $h \mapsto (u_\epsilon(x) - v_\epsilon)h_x + Hh$. Let $h = \tilde{h} \circ \phi_\epsilon$, where $\phi_\epsilon'(x)$ is proportional to $(u_\epsilon(x) - v_\epsilon)^{-1}$. Then

$$L_\epsilon g = ((c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} \circ \phi_\epsilon)_x,$$

where $c_\epsilon \to 1$ as $\epsilon \to 0$, and R_ϵ is a small smoothing remainder (i.e., it gains derivatives of arbitrarily high orders). Thus L_ϵ is conjugate to $c_\epsilon \partial_x + H + R_\epsilon$, whose eigenvalues can be approximated by those of $c_\epsilon \partial_x + H$, which are $\pm(nc_\epsilon i - i)$, $n = 1, 2, \ldots$. The general theory of unbounded analytic operators developed in [Kato 1976] allows us to justify this approximation up to $O(\epsilon^6)$ (see Corollary 3.10), and to relate the eigenvectors of L_ϵ to the Fourier modes (see Lemma 3.7), in the sense that another linear map $\tilde{h} \mapsto \mathfrak{h}$ conjugates L_ϵ into a Fourier multiplier whose action on $e^{i(n+\operatorname{sgn} n)x}$ is multiplication by λ_n ($n \neq 0$).

At the end of the day we have the following estimate for small ϵ:

$$|\lambda_l + \lambda_m - \lambda_n| > \begin{cases} \frac{1}{2}, & l + m \neq n, \\ \frac{1}{5}\epsilon^2, & l + m = n; \end{cases}$$

see Proposition 3.11. Because this value appears in the denominator in (1-5), if g has size δ, a direct application of the normal form transformation yields a lifespan comparable to ϵ^2/δ^2. To improve on this, we will make use of the structure of the nonlinearity:

$$N(\mathfrak{h}, \mathfrak{h}) = \tfrac{1}{2}\mathfrak{h}_x^2 + O(|\epsilon|).$$

The first term is the usual product-style nonlinearity, which imposes the restriction $l + \operatorname{sgn} l + m + \operatorname{sgn} m = n + \operatorname{sgn} n$, and implies $l + m - n = \pm 1 \neq 0$, so the normal form transformation can be carried out as before. The second term is of size $|\epsilon|$ and gains a factor of $1/|\epsilon|$ in the lifespan. Thus the usual energy estimate can show a lifespan comparable to $1/|\epsilon\delta|$, and the normal form transformation can show a lifespan comparable to $|\epsilon|/\delta^2$. This decomposition of the nonlinearity into one part satisfying classical additive frequency restrictions and another part enjoying better estimates analytically was first used by Germain, Pusateri and Rousset [Germain et al. 2018] to show global well-posedness of the 1-dimensional Schrödinger equation with potential (see also [Chen and Pusateri 2022]). Our result shows that this approach can be adapted to quasilinear equations and to the case of discrete spectrum.

1D. ***Outline of the paper.*** In Section 2 we study the traveling waves solutions for (1-1). For sake of completeness we sketch the proof of existence which follows from bifurcation theory. In addition we analyze the size of the traveling waves. In Section 3 we study the linearization of (1-1) around the traveling waves. In Section 4, we introduce a new frame of reference which will help us to avoid the resonances found in Section 3. Finally, in Section 5 we prove Theorem 1.1.

2. Traveling waves

The existence of traveling waves for (1-1) was shown in [Hunter 2018]. Here we will study their size after we give some details about the existence proof. We look for solutions of (1-1) of the form

$$f_\epsilon(x,t) = u_\epsilon(x + v_\epsilon t);$$

thus we have to find $(u_\epsilon,\ v_\epsilon)$ solving

$$Hu_\epsilon - v_\epsilon u'_\epsilon + u_\epsilon u'_\epsilon = 0. \tag{2-1}$$

If (u_ϵ, v_ϵ) is a solution, so is $(u^n_\epsilon(x), v^n_\epsilon) = (u_\epsilon(nx)/n, v_\epsilon/n)$. Thus from one solution we can get n-fold symmetric solutions for all $n \geq 1$.

To solve (2-1) we can apply the Crandall–Rabinowitz theorem [1971] to

$$\begin{aligned} F : H^{k,+}_r(\mathbb{T}) \times \mathbb{C} &\to H^{k-1,-}_r(\mathbb{T}),\\ (u,\mu) &\mapsto Hu + uu' - (-1+\mu)u', \end{aligned}$$

where

$$H^{k,+}_r(\mathbb{T}) = \{2\pi\text{-periodic, mean zero, even functions analytic in the strip } \{|\mathrm{Im}(z)| < r\}\},$$

endowed with the norm

$$\|f\|_{H^{k,+}_r(\mathbb{T})} = \sum_{\pm} \|f(\cdot \pm ir)\|_{H^k(\mathbb{T})},$$

and

$$H^{k,-}_r(\mathbb{T}) = \{2\pi\text{-periodic, odd functions analytic in the strip } \{|\mathrm{Im}(z)| < r\}\},$$

endowed with the norm

$$\|f\|_{H^{k,-}_r(\mathbb{T})} = \sum_{\pm} \|f(\cdot \pm ir)\|_{H^k(\mathbb{T})}.$$

Here $\|\cdot\|_{H^k(\mathbb{T})}$ is the usual Sobolev norm, and it is enough to take $k \geq 1$ and $r = 1$.

We notice that $F(0,\mu) = 0$ and the derivative of F at $u = 0,\ \mu = 0,$

$$D_u F(0,0)h = Hh + h'$$

has a nontrivial element in its kernel belonging to $H^{k,+}_r(\mathbb{T})$, namely, $h = \cos(x)$.

Thus, the application of the Crandall–Rabinowitz theorem allows to show the existence of a branch of solutions $(u_\epsilon, v_\epsilon) \in (H^{1,+}_1, \mathbb{R})$, bifurcating from $(0,-1)$ for (2-1) with the leading-order term

$$u_\epsilon(x) = \epsilon\cos(x) + O(\epsilon^2), \quad v_\epsilon = -1 + O(\epsilon).$$

We remark that we obtain a bifurcation curve

$$\begin{aligned} \epsilon &\to (u_\epsilon, v_\epsilon),\\ B_\delta = \{z \in \mathbb{C} : |z| < \delta\} &\to (H^{k-1,-}_r, \mathbb{R}), \end{aligned} \tag{2-2}$$

which is differentiable and hence analytic on B_δ for δ small enough.

The rest of this section is devoted to proving further properties of these solutions.

Introducing the asymptotic expansion

$$u_\epsilon(x) = \sum_{n=1}^{\infty} u_n(x)\epsilon^n, \quad v_\epsilon = \sum_{n=0}^{\infty} v_n \epsilon^n, \tag{2-3}$$

taking $u_1 = \cos(x)$, $\lambda_0 = -1$ and comparing the coefficient in ϵ^n we obtain that

$$u_n' + Hu_n = -v_{n-1}\sin(x) + \sum_{m=1}^{n-2} v_m u_{n-m}' - \frac{1}{2}\partial_x \sum_{m=1}^{n-1} u_{n-m}u_m = -v_{n-1}\sin(x) + f_n$$

for $n = 2, 3, \dots$.

We notice that in order to solve the equation $Hu + u' = f$ we need $(f, \sin(x)) = 0$. Therefore we have to choose $v_{n-1} = \frac{1}{\pi}(\sin(x), f_n)$. This gives us a recurrence for (u_n, v_{n-1}), $n \geq 2$, in terms of $\{(u_m, v_{m-1})\}_{m=1}^{n-1}$. In order to study this recurrence we will introduce the ansatz

$$u_n = \sum_{k=2}^{n} u_{n,k}\cos(kx). \tag{2-4}$$

By induction, one can check that the rest of coefficients in the expansion on cosines of u_n must be zero. In addition, if $u_\epsilon(x)$ solves (2-1), $u_{-\epsilon}(x+\pi)$ is also a bifurcation curve in the direction of $\cos(x)$, and then by uniqueness, $u_\epsilon(x) = u_{-\epsilon}(x+\pi)$, which yields $u_{n,k} = 0$ if $n - k = 1 \pmod 2$.

Comparing the coefficient of $\sin(kx)$, with $k = n \pmod 2$, and $2 \leq k \leq n$, we have

$$(1-k)u_{n,k} + k\sum_{m=1}^{n-k} v_m u_{n-m,k} - \frac{k}{4}\sum_{m=1}^{n-1}\sum_{l=\max(1,k-n+m)}^{\min(m,k-1)} u_{m,l}u_{n-m,k-l} \\ - \frac{k}{2}\sum_{m=1}^{n-1}\sum_{l=1}^{\min(m,n-m-k)} u_{m,l}u_{n-m,k+l} = 0. \tag{2-5}$$

And comparing with $\sin(x)$ we have

$$v_{n-1} = \frac{1}{2}\sum_{m=1}^{n-1}\sum_{l=1}^{\min(m,n-m-1)} u_{m,l}u_{n-m,1+l}. \tag{2-6}$$

Up to order $O(\epsilon^4)$ we find

$$\begin{aligned} u_\epsilon(x) &= \epsilon\cos x - \tfrac{1}{2}\epsilon^2\cos 2x + \tfrac{3}{8}\epsilon^3\cos 3x + O(\epsilon^4), \\ v_\epsilon &= -1 - \tfrac{1}{4}\epsilon^2 + O(\epsilon^4). \end{aligned} \tag{2-7}$$

The recurrence (2-5)–(2-6) allows us to prove the following result.

Theorem 2.1. *The radius of convergence of the series* (2-3), *with the coefficients given by* (2-4)–(2-6), *is not bigger than* $2/e$.

Proof. From (2-5) and (2-6) we have

$$(1-n)u_{n,n} = \frac{1}{2}\sum_{k=1}^{n-1}(n-k)u_{k,k}u_{n-k,n-k}.$$

Let

$$y = y(x) = x + \sum_{n=2}^{\infty} u_{n,n}x^n.$$

Then $y - xy' = xyy'/2$, which, together with $y \sim x$ for small x, yields $y = 2W(x/2)$, where W is the Lambert W-function. Since the radius of convergence of W at 0 is $1/e$, the radius of convergence of y at 0 is $2/e$, so the radius of convergence of (2-5) and (2-6) is at most $2/e$. □

In addition we can get a bound for how large the traveling wave can be.

Theorem 2.2. *The series* (2-3), *with the coefficients given by* (2-4)–(2-6), *converges for any* $\epsilon < x^* \sim 0.23$.

Proof. This proof is based on the implicit function theorem.

First we introduce the spaces

$$L^{2,-} = \{\text{odd functions } f \in L^2(\mathbb{T})\},$$
$$H^{1,+} = \{\text{even functions } f \in H^1(\mathbb{T})\}.$$

The space X is the orthogonal complement of the span of $\cos(x)$ in $H^{1,+}$. We will equip $L^{2,-}$ with the norm

$$\|u\|^2_{L^{2,-}} = \frac{1}{\pi}\int_{-\pi}^{\pi} |u(x)|^2\, dx \tag{2-8}$$

in such a way that $\|\sin(nx)\|_{L^{2,-}} = 1$ for $n \geq 1$. We also define

$$\|u\|^2_X = \frac{1}{\pi}\int_{-\pi}^{\pi} (|u'(x)|^2 + |u(x)|^2 - 2u(x)\Lambda u(x))\, dx. \tag{2-9}$$

Thus $\|\cos(nx)\|_X = n - 1$ for $n \geq 2$. The reason why we take these norms is technical and it will arise below. Finally we define

$$\mathcal{X} = X \times \mathbb{R}$$

equipped with the norm

$$\|(\tilde{u}, \nu)\|_{\mathcal{X}} = \sqrt{\|\tilde{u}\|^2_X + |\nu|^2}.$$

Since $u_\epsilon = \epsilon \cos x - \frac{1}{2}\epsilon^2 \cos 2x + O(\epsilon^3)$ and $v_\epsilon = -1 + O(\epsilon^2)$, we can let

$$\begin{aligned}
&G(\epsilon, \tilde{u}, \mu)\\
&= \frac{1}{\epsilon^2} F\big(\epsilon \cos x - \tfrac{1}{2}\epsilon^2 \cos 2x + \epsilon^2 \tilde{u}, \epsilon\mu\big)\\
&= H\tilde{u} + \epsilon\big(\cos x(\sin 2x + \tilde{u}') + \big(\tfrac{1}{2}\cos 2x - \tilde{u}\big)(\sin x - \epsilon \sin 2x - \epsilon\tilde{u}')\big) + \tilde{u}' - \mu(\sin x - \epsilon \sin 2x - \epsilon\tilde{u}')
\end{aligned}$$

map $\mathbb{R} \times \mathcal{X}$ to $L^{2,-}$.

Because of the existence of traveling waves, we already know that there exists ϵ^* such that, for every $\epsilon \in [0, \epsilon^*)$, there exist $\tilde{u}_\epsilon$ and μ_ϵ satisfying

$$G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon) = 0.$$

In addition we have

$$\begin{aligned}
\left.\frac{dG(\epsilon, \tilde{u}_\epsilon + s\tilde{v}, \mu + s\nu)}{ds}\right|_{s=0} &\equiv dG_{\epsilon, \tilde{u}_\epsilon, \mu}(\tilde{v}, \nu)\\
&= H\tilde{v} + \epsilon\big(\tilde{v}' \cos x - \tilde{v}(\sin x - \epsilon \sin 2x - \epsilon\tilde{u}') - \epsilon\big(\tfrac{1}{2}\cos 2x - \tilde{u}\big)\tilde{v}'\big)\\
&\qquad + \tilde{v}' - \nu(\sin x - \epsilon \sin 2x - \epsilon\tilde{u}') + \epsilon\mu\tilde{v}'
\end{aligned}$$

maps $(\tilde{v}, \nu) \in \mathcal{X}$ linearly to $L^{2,-}$.

Thus as far as $dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}(\tilde{u},\mu)$ is invertible from $\mathcal{X}$ to $L^{2,-}$ for $\epsilon \in [0, x^*)$ we will be able to extend the solution $(u_\epsilon, \mu_\epsilon)$ from $[0, \epsilon^*)$ to $[0, x^*)$ by the implicit function theorem.

Note that

$$dG_{0,0,0}(\tilde{v}, \nu) = H\tilde{v} + \tilde{v}' - \nu \sin x$$

is an isometry from $\mathcal{X}$ to $L^{2,-}$ under the norms given by (2-8) and (2-9). Therefore one can compute

$$dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} = dG^{-1}_{0,0,0}(\mathbb{I} + dG^{-1}_{0,0,0}(dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} - dG_{0,0,0})).$$

By the Neumann series and the fact that $dG_{0,0,0}$ is an isometry, $dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}$ will be invertible, as long as $\|dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} - dG_{0,0,0}\|_{\mathcal{X}\to L^{2,-}} < 1$. In order to show this last inequality we will bound

$$A_\epsilon := \|dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon} - dG_{0,0,0}\|_{\mathcal{X}\to L^{2,-}}$$

in terms of $\|\tilde{u}_\epsilon\|_X$ and μ_ϵ. After that we will bound $\|\tilde{u}_\epsilon\|_X$ and μ_ϵ. To do it we will use the information we have about $\partial_\epsilon \tilde{u}_\epsilon$ and $\partial_\epsilon \mu_\epsilon$.

Along the bifurcation curve,

$$\begin{aligned} dG_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}&(\partial_\epsilon \tilde{u}_\epsilon, \mu'_\epsilon) \\ &= -\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon) \\ &= \cos x(\sin 2x + \tilde{u}'_\epsilon) + \tfrac{1}{2}\sin x(\cos 2x - 2\tilde{u}_\epsilon) - \epsilon(\cos 2x - 2\tilde{u}_\epsilon)(\sin 2x + \tilde{u}'_\epsilon) + \mu_\epsilon(\sin 2x + \tilde{u}'_\epsilon). \end{aligned} \tag{2-10}$$

Thus

$$(\partial_\epsilon \tilde{u}_\epsilon, \mu'_\epsilon) = dG^{-1}_{\epsilon,\tilde{u}_\epsilon,\mu_\epsilon}(-\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)).$$

Therefore

$$\sqrt{\|\partial_\epsilon \tilde{u}_\epsilon\|_X^2 + |\mu'_\epsilon|^2} \le \frac{1}{1 - A_\epsilon}\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}. \tag{2-11}$$

In addition we have, for $r_\epsilon = \sqrt{\|\tilde{u}_\epsilon\|_X^2 + |\mu_\epsilon|^2}$,

$$\partial_\epsilon r_\epsilon \le \sqrt{\|\partial_\epsilon \tilde{u}_\epsilon\|_X^2 + |\mu'_\epsilon|^2} \le \frac{1}{1 - A_\epsilon}\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}.$$

Thus, explicit estimates for A_ϵ and $\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}$ in terms of r_ϵ and ϵ give a differential inequality for r_ϵ which can be used to bound A_ϵ.

We will need the following lemmas to bound A_ϵ and the norm $\|\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)\|_{L^{2,-}}$, where $\partial_\epsilon G(\epsilon, \tilde{u}_\epsilon, \mu_\epsilon)$ is given by the right-hand side of (2-10).

Lemma 2.3. *If $f \in X$ then $\|f \sin x - f' \cos x\|_{L^2} \le \sqrt{3}\|f\|_X$.*

Lemma 2.4. *If $f \in X$ then $\|2f \sin 2x - f' \cos 2x\|_{L^2} \le \frac{1}{2}\sqrt{17}\|f\|_X$.*

Proof. We only show Lemma 2.3. The proof of Lemma 2.4 is similar.

Let $f = \sum_{n=2}^{\infty} f_n \cos nx$. Then

$$2(f \sin x - f' \cos x) = f_2 \sin x + 2f_3 \sin 2x + \sum_{n=3}^{\infty} n(f_{n-1} + f_{n+1}) \sin nx,$$

and

$$\begin{aligned}4\|f\sin(x)-f'\cos(x)\|_{L^2}^2 &= f_2^2+4f_3^2+\sum_{n=3}^{\infty}n^2(f_{n-1}+f_{n+1})^2\\ &\le 10f_2^2+20f_3^2+59f_4^2+88f_5^2+18f_2f_4+32f_3f_5+4\sum_{n=6}^{\infty}(n^2+1)f_n^2.\end{aligned}$$

The infinite sum is bounded by $1.48\sum_{n=6}^{\infty}(n-1)^2f_n^2$, and it remains a finite-dimensional problem to show that the remaining terms are bounded by $12\sum_{n=2}^{5}(n-1)^2f_n^2$. □

Lemma 2.5. *If $f, g \in X$ then $\|(fg)'\|_{L^2} \le B\|f\|_X\|g\|_X$, where*

$$B=\sqrt{\frac{\pi^2}{3}+\frac{869}{144}}\approx 3.05.$$

Proof. Let $f=\sum_{n=2}^{\infty}f_n\cos nx,\ g=\sum_{n=2}^{\infty}g_n\cos nx\in X$. Then

$$(fg)'=-\frac{1}{2}\sum_{n\ge 1}n\sum_{|m|\ge 2,|n-m|\ge 2}f_{|m|}g_{|n-m|}\sin nx,$$

so by Cauchy–Schwarz,

$$\|(fg)'\|_{L^2}^2=\frac{1}{8}\sum_{|n|\ge 1}n^2\Biggl(\sum_{|m|\ge 2,|n-m|\ge 2}f_{|m|}g_{|n-m|}\Biggr)^2\le C\|f\|_X^2\|g\|_X^2,$$

where

$$C=\frac{1}{2}\sup_{n=1}^{\infty}\sum_{|m|\ge 2,|n-m|\ge 2}\frac{n^2}{(|m|-1)^2(|n-m|-1)^2}=\frac{\pi^2}{3}+\frac{869}{144}.$$ □

Now, with Lemmas 2.3, 2.5 and 2.4 we are ready to bound the right-hand side of (2-10). Indeed,

$$\|\text{right-hand side of (2-10)}\|_{L^2}\le\frac{\sqrt{10+4\epsilon^2}}{4}+2r_\epsilon+\frac{\sqrt{17}}{2}\epsilon\|\tilde{u}_\epsilon\|_X+B\epsilon\|\tilde{u}_\epsilon\|_X^2+\|\tilde{u}_\epsilon\|_X^2+|\mu_\epsilon|^2.$$

Turning to the other side, we have

$$\begin{aligned}&(dG_{(\epsilon,\tilde{u}_\epsilon,\mu_\epsilon)}-dG_{(0,0,0)})(\tilde{v},\nu)\\ &\quad=\epsilon\bigl(\tilde{v}'\cos x-\tilde{v}(\sin x-\epsilon\sin 2x-\epsilon\tilde{u}_\epsilon')-\epsilon\bigl(\tfrac{1}{2}\cos 2x-\tilde{u}_\epsilon\bigr)\tilde{v}'\bigr)+\epsilon\nu(\sin 2x+\tilde{u}_\epsilon')+\epsilon\mu_\epsilon\tilde{v}'. \qquad (2\text{-}12)\end{aligned}$$

Again by Lemmas 2.3, 2.5 and 2.4 we find

$$\begin{aligned}\|\text{left-hand side of (2-12)}\|_{L^2}&\le\Bigl(\sqrt{3}\epsilon+\frac{\sqrt{17}}{4}\epsilon^2+B\epsilon^2\|\tilde{u}_\epsilon\|_X+2\epsilon|\mu_\epsilon|\Bigr)\|\tilde{v}\|_X+\epsilon(1+2\|\tilde{u}_\epsilon\|_X)|\nu|\\ &=\epsilon(\sqrt{3}+2|\mu_\epsilon|,1+2\|\tilde{u}_\epsilon\|_X)\cdot(\|\tilde{v}\|_X,|\nu|)+\Bigl(\frac{\sqrt{17}}{4}\epsilon^2+B\epsilon^2\|\tilde{u}_\epsilon\|_X\Bigr)\|\tilde{v}\|_X,\end{aligned}$$

so

$$A_\epsilon\le 2\epsilon+2\epsilon r_\epsilon+\frac{\sqrt{17}}{4}\epsilon^2+B\epsilon^2r_\epsilon.$$

Since $dG_{(0,0,0)}$ is an isometry, the Neumann series $(1-T)^{-1}=\sum_{n=0}^{\infty}T^n$ shows that if $A_\epsilon<1$, then $dG_{(\epsilon,\tilde{u}_\epsilon,\mu_\epsilon)}$ is invertible, and $\|dG^{-1}_{(\epsilon,\tilde{u}_\epsilon,\mu_\epsilon)}\|\le(1-A_\epsilon)^{-1}$, so

$$\sqrt{\|\partial_\epsilon\tilde{u}_\epsilon\|_X^2+|\mu_\epsilon'|^2}\le\frac{1}{1-A_\epsilon}\Bigl(\frac{\sqrt{10+4\epsilon^2}}{4}+2r_\epsilon+\frac{\sqrt{17}}{2}\epsilon\|\tilde{u}_\epsilon\|_X+B\epsilon r_\epsilon^2+r_\epsilon^2\Bigr).$$

Then $r_0 = 0$ and

$$r'_\epsilon \le \frac{\frac{1}{4}\sqrt{10+4\epsilon^2} + \big(2+\frac{\sqrt{17}}{2}\epsilon\big)r_\epsilon + B\epsilon r_\epsilon^2 + r_\epsilon^2}{1-2\epsilon-2\epsilon r_\epsilon - \frac{\sqrt{17}}{4}\epsilon^2 - B\epsilon^2 r_\epsilon}.$$

By the comparison principle, r_ϵ is bounded from above by the solution to

$$\frac{dy}{dx} = y' = \frac{\sqrt{10+4x^2} + (8+2\sqrt{17}x)y + 4Bxy^2 + 4y^2}{4-8x-8xy-\sqrt{17}x^2-4Bx^2y}, \tag{2-13}$$

with $y(0) = 0$, which is

$$(2Bx^2+4x)y^2 + (8x+\sqrt{17}x^2-4)y + x\sqrt{x^2+2.5} + 2.5\sinh^{-1}(\sqrt{0.4}x) = 0.$$

When $x > 0$, the quadratic coefficient and the constant are positive, so this equation has a nonnegative root if and only if

$$8x+\sqrt{17}x^2-4 \le -2\sqrt{(2Bx^2+4x)(x\sqrt{x^2+2.5}+2.5\sinh^{-1}(\sqrt{0.4}x))},$$

whose solution is $x \le x^* \approx 0.23$ numerically. Hence the solution can be extended to $\epsilon = x^* \approx 0.23$. In order to achieve this last conclusion we notice that the solution to (2-13), with $y(0) = 0$ can be extended only if $A_\epsilon < 1$, since $1 - A_\epsilon$ arises in the denominator.

The above argument shows that for $\epsilon \in (-x^*, x^*)$, the bifurcation curve produces a traveling wave $u_\epsilon = \epsilon\cos x - \frac{1}{2}\epsilon^2\cos 2x + \epsilon^2\tilde{u}_\epsilon$, which travels at speed $v_\epsilon = -1 - \epsilon\mu_\epsilon$. Since all the operators involved are analytic in all its arguments, the bifurcation curve is analytic in ϵ on $(-x^*, x^*)$. It may be the case, however, that the power series for u_ϵ and v_ϵ around $\epsilon = 0$ has a smaller radius of convergence than x^* (for example, the function $f(x) = (x^2+1)^{-1}$ is analytic on the whole real line, but the radius of convergence of its power series around 0 is only 1.) We now show that the radius of convergence of the power series for u_ϵ and v_ϵ are indeed at least x^*.

We note that the above argument also works if ϵ is replaced with ϵe^{ia} ($a \in \mathbb{R}$), so the bifurcation curve (u_ϵ, v_ϵ) is also analytic in a neighborhood of $\{\epsilon e^{ia} : \epsilon \in (-x^*, x^*)\}$. Hence the curve is analytic in the disk of radius x^* centered at 0, so the radius of convergence of its power series around 0 is at least x^*. □

3. Linearization around traveling waves

In this section we will analyze the spectrum of the operator

$$L_\epsilon g = -v_\epsilon g_x + Hg + (u_\epsilon(x)g)_x$$

corresponding to the linearization of (1-1) around the traveling wave (u_ϵ, v_ϵ) bifurcating from zero in the direction of the cosine studied in the previous section.

Actually, let

$$f(x,t) = f_\epsilon(x,t) + g(x+v_\epsilon t, t),$$

with $f_\epsilon(x,t) = u_\epsilon(x+v_\epsilon t)$. Then

$$f_t(x,t) = \partial_t f_\epsilon(x,t) + (v_\epsilon g_x + g_t)(x+v_\epsilon t, t)$$

and

$$(Hf + ff_x)(x,t) = (Hf_\epsilon + f_\epsilon \partial_x f_\epsilon)(x,t) + Hg(x + v_\epsilon t, t) \\ + \partial_x (f_\epsilon(x,t) g(x + v_\epsilon t, t)) + g(x + v_\epsilon t, t)\partial_x g(x + v_\epsilon t, t).$$

Putting these in (1-1), we get the equation for $g(x,t)$

$$\partial_t g(x,t) = -v_\epsilon g(x,t)_x + Hg(x,t) + (u_\epsilon(x) g(x,t))_x + g(x,t) g(x,t)_x.$$

The linearization around $g = 0$ is

$$\partial_t g = L_\epsilon g,$$

where

$$L_\epsilon g = -v_\epsilon g_x + Hg + (u_\epsilon g)_x = \underbrace{Hg + g_x}_{Lg} + \sum_{n=1}^{\infty} \epsilon^n \underbrace{((u^{(n)} - v^{(n)})g)_x}_{L^{(n)} g}. \tag{3-1}$$

3A. ***The eigenvalue 0.*** The action of L on the Fourier modes is

$$\mathcal{F}(Lg)(m) = i(m - \operatorname{sgn} m)\hat{g}(m),$$

with eigenvalues 0 (double), $\pm i$, $\pm 2i$, ... (on $L^2(\mathbb{T})$ with zero mean). We first study the perturbation of the eigenspace corresponding to the double eigenvalue of 0. By translational symmetry, for any $\delta \in \mathbb{R}$, $u_\epsilon(x + \delta)$ is also a solution to

$$Hu - v_\epsilon u + uu' = 0.$$

Differentiation with respect to δ then shows that

$$L_\epsilon u'_\epsilon = Hu'_\epsilon - v_\epsilon u'_\epsilon + (u_\epsilon u'_\epsilon)' = 0.$$

Also, since u_ϵ lies on a bifurcation curve, we can differentiate

$$Hu_\epsilon - v_\epsilon u'_\epsilon + u_\epsilon u'_\epsilon = 0,$$

with respect to ϵ to get

$$L_\epsilon \partial_\epsilon u_\epsilon = H\partial_\epsilon u_\epsilon - (\partial_\epsilon v_\epsilon) u'_\epsilon + u_\epsilon \partial_\epsilon u'_\epsilon + u'_\epsilon \partial_\epsilon u_\epsilon = (\partial_\epsilon v_\epsilon) u'_\epsilon,$$

so on the span V_ϵ of u'_ϵ and $\partial_\epsilon u_\epsilon$, L_ϵ acts nilpotently by the matrix

$$\begin{pmatrix} 0 & \partial_\epsilon v_\epsilon \\ 0 & 0 \end{pmatrix}.$$

3B. ***Simplifying the linearized operator.*** We want to solve the eigenvalue problem

$$L_\epsilon g = ((u_\epsilon - v_\epsilon) g)' + Hg = \lambda(\epsilon) g.$$

Let $g = h'$. Then the antiderivative of the above is

$$(u_\epsilon - v_\epsilon) h' + Hh = \lambda(\epsilon) h \pmod 1. \tag{3-2}$$

Let $h = \tilde{h} \circ \phi_\epsilon$, where ϕ_ϵ satisfies

$$\phi'_\epsilon = \frac{2\pi}{u_\epsilon - v_\epsilon} \left(\int_0^{2\pi} \frac{dy}{u_\epsilon(y) - v_\epsilon} \right)^{-1}. \tag{3-3}$$

Then

$$(u_\epsilon - v_\epsilon)\phi_\epsilon'(\tilde{h}' \circ \phi_\epsilon) + H(\tilde{h} \circ \phi_\epsilon) = \lambda(\epsilon)\tilde{h} \circ \phi_\epsilon \pmod 1.$$

When ϵ is small enough, ϕ_ϵ is a diffeomorphism of $\mathbb{R}/2\pi\mathbb{Z}$, so

$$2\pi\left(\int_0^{2\pi} \frac{dy}{u_\epsilon(y) - v_\epsilon}\right)^{-1} \tilde{h}' + H(\tilde{h} \circ \phi_\epsilon) \circ \phi_\epsilon^{-1} = \lambda(\epsilon)\tilde{h} \pmod 1.$$

By the change of variable $z = \phi_\epsilon(y)$,

$$\begin{aligned} H(\tilde{h} \circ \phi_\epsilon) \circ \phi_\epsilon^{-1}(x) &= \frac{1}{2\pi}\int_0^{2\pi} \tilde{h}(\phi_\epsilon(y)) \cot\left(\frac{\phi_\epsilon^{-1}(x) - y}{2}\right) dy \\ &= \frac{1}{2\pi}\int_0^{2\pi} \tilde{h}(z) \cot\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right)(\phi_\epsilon^{-1})'(z)\, dz. \end{aligned}$$

The convolution kernel of the operator

$$R_\epsilon \tilde{h} = H(\tilde{h} \circ \phi_\epsilon) \circ \phi_\epsilon^{-1} - H\tilde{h}$$

is

$$K_\epsilon(x, z) = \cot\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right)(\phi_\epsilon^{-1})'(z) - \cot\left(\frac{x - z}{2}\right) \tag{3-4}$$

and the ϵ-derivative of the kernel is

$$\begin{aligned} \partial_\epsilon K_\epsilon(x, z) = -\csc^2\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right)\frac{\partial_\epsilon\phi_\epsilon^{-1}(x) - \partial_\epsilon\phi_\epsilon^{-1}(z)}{2}(\phi_\epsilon^{-1})'(z) \\ + \cot\left(\frac{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)}{2}\right)\partial_\epsilon(\phi_\epsilon^{-1})'(z). \end{aligned}$$

Near $x = 0$, $\csc x - 1/x^2$ and $\cot x - 1/x$ are smooth, and $(\phi_\epsilon^{-1})'$ is smooth everywhere, so when $x - z$ is small enough, up to a smooth function in (x, z),

$$\begin{aligned} \frac{\partial_\epsilon K_\epsilon(x, z)}{2} &= -\frac{(\partial_\epsilon\phi_\epsilon^{-1}(x) - \partial_\epsilon\phi_\epsilon^{-1}(z))(\phi_\epsilon^{-1})'(z)}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2} + \frac{\partial_\epsilon(\phi_\epsilon^{-1})'(z)}{\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)} \\ &= \frac{\partial_\epsilon(\phi_\epsilon^{-1})'(z)(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z)) - (\partial_\epsilon\phi_\epsilon^{-1}(x) - \partial_\epsilon\phi_\epsilon^{-1}(z))(\phi_\epsilon^{-1})'(z)}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2} \\ &= \frac{\partial_\epsilon(\phi_\epsilon^{-1})'(z)(x - z)^2 \int_0^1 (1-t)(\phi_\epsilon^{-1})''((1-t)z + tx)\, dt}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2} \\ &\qquad - \frac{(\phi_\epsilon^{-1})'(z)(x - z)^2 \int_0^1 (1-t)\partial_\epsilon(\phi_\epsilon^{-1})''((1-t)z + tx)\, dt}{(\phi_\epsilon^{-1}(x) - \phi_\epsilon^{-1}(z))^2}, \end{aligned}$$

which is itself a smooth function of (x, z) when $x - z$ is small enough (because ϕ_ϵ^{-1} is smooth). Then

$$\|\partial_\epsilon R_\epsilon \tilde{h}^{(m)}\|_{\dot{H}^k} \lesssim_{k,m} \|\tilde{h}\|_{L^2/(1)}, \quad k, m = 0, 1, \dots,$$

where the constant does not depend on ϵ, for all $\tilde{h} \in H^m/(1)$, or, equivalently,

$$\|\partial_\epsilon R_\epsilon \tilde{h}\|_{\dot{H}^k} \lesssim_{k,m} \|\tilde{h}\|_{\dot{H}^{-m}}, \quad k, m = 0, 1, \dots, \tag{3-5}$$

where the dot over H means that the norm does not measure frequency zero.

Definition 3.1. We say an operator is of class $\mathcal{S}$ if it satisfies (3-5). We say a family of operators is of class $\mathcal{S}$ uniformly if for each k and m there is an implicit constant that makes (3-5) true for all operators in the family.

Thus $\partial_\epsilon R_\epsilon$ is of class $\mathcal{S}$ uniformly in ϵ. Since $R_0 = 0$, R_ϵ/ϵ is also of class $\mathcal{S}$ uniformly in ϵ.

Now the eigenvalue problem for $\tilde{h}$ is of the form

$$(c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} = \lambda(\epsilon)\tilde{h} \pmod 1$$

or, equivalently,

$$(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon)\tilde{h} = c_\epsilon^{-1}\lambda(\epsilon)\tilde{h} \pmod 1, \tag{3-6}$$

where

$$c_\epsilon = 2\pi \left(\int_0^{2\pi} \frac{dy}{u_\epsilon(y) - v_\epsilon} \right)^{-1} \tag{3-7}$$

and R_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ. Note that since u_ϵ and v_ϵ are analytic functions of ϵ on a neighborhood of 0, with $u_0 = 0$ and $v_0 = -1$, so are ϕ_ϵ, R_ϵ and c_ϵ with $\phi_0 = I$, $R_0 = 0$ and $c_0 = 1$.

3C. ***Spectral analysis of the linearization.*** The eigenvalue problem (3-6) is a perturbation of the eigenvalue problem

$$\tilde{h}' + H\tilde{h} = \lambda\tilde{h} \pmod 1,$$

with explicit eigenvalues

$$0 \text{ (double)}, \ ni, \quad n = \pm 1, \pm 2, \dots,$$

and eigenfunctions

$$e^{\pm ix}, \ e^{i(n + \operatorname{sgn} n)x}, \quad n = \pm 1, \pm 2, \dots.$$

They form an orthogonal basis of $H^k/(1)$ for any nonnegative integer k.

Definition 3.2. Let $T : \dot{H}^k(\mathbb{T}) \to \dot{H}^k(\mathbb{T})$ for $k \in \mathbb{N}$ be a linear operator. We will define

$$\|T\| := \|T\|_{\dot{H}^k(\mathbb{T}) \to \dot{H}^k(\mathbb{T})}.$$

The resolvent $(\partial_x + H - z)^{-1}$ is also a Fourier multiplier whose action on Fourier modes is

$$(\partial_x + H - z)^{-1} e^{\pm i(n+1)x} = (\pm ni - z)^{-1} e^{\pm i(n+1)x}, \quad n = 0, 1, \dots. \tag{3-8}$$

The circle

$$\Gamma_n = \left\{ z : |z - ni| = \tfrac{1}{2} \right\}, \quad n = \pm 1, \pm 2, \dots,$$

encloses a single eigenvalue $\pm ni$, and the circle

$$\Gamma_0 = \left\{ z : |z| = \tfrac{1}{2} \right\}$$

encloses the double eigenvalue 0. On Γ_n and Γ_0 we have

$$|z-mi| \geq \tfrac{1}{2}, \quad m \in \mathbb{Z}, \tag{3-9}$$

so by (3-8),

$$\|(\partial_x + H - z)^{-1}\| \leq 2, \quad z \in \Gamma_n,\ n \in \mathbb{Z}. \tag{3-10}$$

Moreover the projection

$$P_n = -\frac{1}{2\pi i}\int_{\Gamma_n} (\partial_x + H - z)^{-1}\, dz, \quad n = \pm 1, \pm 2, \dots,$$

is the projection on the span of $e^{i(n+\operatorname{sgn} n)x}$ and the projection

$$P_0 = -\frac{1}{2\pi i}\int_{\Gamma_0} (\partial_x + H - z)^{-1}\, dz$$

is the projection on the span of e^{ix} and e^{-ix}.

Now when ϵ is small enough and $z \in \Gamma_n$, we have

$$\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z = (\partial_x + H - z)(1 + (\partial_x + H - z)^{-1} R'_\epsilon),$$

where

$$R'_\epsilon = (\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon) - (\partial_x + H) = (c_\epsilon^{-1} - 1)H + c_\epsilon^{-1} R_\epsilon \tag{3-11}$$

is analytic in ϵ near 0, with $R'_0 = 0$, thanks to the analyticity of c_ϵ. Taking the inverse gives

$$(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1} = (1 + (\partial_x + H - z)^{-1} R'_\epsilon)^{-1} (\partial_x + H - z)^{-1}$$

and the Neumann series

$$(1 + (\partial_x + H - z)^{-1} R'_\epsilon)^{-1} = \sum_{n=0}^{\infty} ((\partial_x + H - z)^{-1} R'_\epsilon)^n \tag{3-12}$$

converges because

$$\|(\partial_x + H - z)^{-1} R'_\epsilon\| \leq 2\|R'_\epsilon\| \lesssim_k \epsilon < 1$$

when ϵ is small enough (depending on k). Moreover,

$$\|(1 + (\partial_x + H - z)^{-1} R'_\epsilon)^{-1} - I\| \lesssim_k \epsilon$$

and so

$$\|(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1} - (\partial_x + H - z)^{-1}\| \lesssim \epsilon$$

uniformly for $z \in \Gamma_n$. Hence the projections

$$Q_n(\epsilon) = -\frac{1}{2\pi i}\int_{\Gamma_n} (\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1}\, dz, \quad n \in \mathbb{Z}, \tag{3-13}$$

exist and satisfy

$$\|Q_n(\epsilon) - P_n\| \lesssim_k \epsilon, \quad n \in \mathbb{Z}, \tag{3-14}$$

uniformly in n. Then by [Kato 1976, Chapter I, Section 4.6], when ϵ is small enough, $Q_n(\epsilon)$ is conjugate to P_n. Thus $\dim \operatorname{ran} Q_n(\epsilon) = 1$ for $n \neq 0$ and $\dim \operatorname{ran} Q_0(\epsilon) = 2$. So $\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon$ has a single eigenvalue enclosed by Γ_n for $n \neq 0$. In Section 3A we showed that the action on the range of $Q_0(\epsilon)$ is given by a nonzero nilpotent 2-by-2 matrix. If z is outside all these circles, then (3-10) still holds and the Neumann series (3-12) still converges to show that $\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z$ is invertible, so it has no other eigenvalues.

3D. ***Analyticity of eigenvalues and eigenvectors.*** By (3-8) and (3-9), $(\partial_x + H - z)^{-1}$ is analytic in (z, ϵ) for z in a neighborhood U of $\bigcup_{n\in\mathbb{Z}} \Gamma_n$, and ϵ near 0. By (3-11), R'_ϵ is analytic in ϵ near 0, so the series (3-12) shows that $(\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon - z)^{-1}$ is analytic in (z, ϵ) for $z \in U$ and ϵ near 0, and the integral (3-13) shows that all the projections $Q_n(\epsilon)$ ($n \in \mathbb{Z}$) are analytic in a neighborhood of 0 independent of n.

Let $\psi_n(\epsilon)$ be the corresponding eigenvectors to $Q_n(\epsilon)$ for $n \neq 0$. Thanks to (3-14), a good choice is $\psi_n(\epsilon) = Q_n(\epsilon) e^{i(n+\operatorname{sgn} n)x}$, which is nonzero and analytic in a neighborhood of 0 independent of n. Then by (3-6),

$$Q_n(\epsilon)(\partial_x + c_\epsilon^{-1}(H + R_\epsilon)) e^{i(n+\operatorname{sgn} n)x} = (\partial_x + c_\epsilon^{-1}(H + R_\epsilon))\psi_n(\epsilon) = c_\epsilon^{-1}\lambda_n(\epsilon)\psi_n(\epsilon).$$

On the other hand, the left-hand side equals

$$(n + \operatorname{sgn} n) i\, Q_n(\epsilon) e^{i(n+\operatorname{sgn} n)x} + c_\epsilon^{-1} Q_n(\epsilon)(H + R_\epsilon) e^{\pm i(n+\operatorname{sgn} n)x},$$

which is another vector analytic in ϵ near 0. Then by the next lemma, all the eigenvalues $c_\epsilon^{-1}\lambda_n(\epsilon)$, and hence $\lambda_n(\epsilon)$, are analytic in a neighborhood of 0 independent of n.

Lemma 3.3. *Let $u(\epsilon)$ and $v(\epsilon)$ be two vectors analytic in $\epsilon \in U$ satisfying*

$$u(\epsilon) \neq 0 \quad \textit{and} \quad v(\epsilon) = \lambda(\epsilon) u(\epsilon), \quad \epsilon \in U.$$

Then $\lambda(\epsilon)$ is analytic in $\epsilon \in U$.

Proof. Without loss of generality assume that $0 \in U$. Since the result is local in ϵ, it suffices to show that $\lambda(\epsilon)$ is analytic in a smaller neighborhood of 0.

Since $u(0) \neq 0$, we can find a linear functional f such that $f(u(0)) \neq 0$. Then $f(u(\epsilon)) \neq 0$ in a neighborhood of 0, and so

$$\lambda(\epsilon) = \frac{f(v(\epsilon))}{f(u(\epsilon))}$$

is analytic in a neighborhood of 0. □

Regarding the double eigenvalue 0, in Section 3A we showed that u'_ϵ and $\partial_\epsilon u_\epsilon$ are two generalized eigenvectors of the operator L_ϵ. Using the relation given in Section 3B, they correspond to two generalized eigenvectors $\psi_0^-(\epsilon)$ and $\psi_0^+(\epsilon)$ of the operator $\partial_x + c_\epsilon^{-1} H + c_\epsilon^{-1} R_\epsilon$, via the relation $(\psi_0^-(\epsilon) \circ \phi_\epsilon)' = u'_\epsilon$ and $(\psi_0^+(\epsilon) \circ \phi_\epsilon)' = \partial_\epsilon u_\epsilon$. Then clearly $\psi_0^\pm(\epsilon)$ are both analytic in ϵ.

From the analyticity of the eigenvalues $c_\epsilon^{-1}\lambda_n(\epsilon)$, it is easy to derive bounds on their Taylor coefficients.

Proposition 3.4. *For $k \geq 1$ and $n \neq 0$, the coefficient of ϵ^k in $c_\epsilon^{-1}\lambda_n(\epsilon)$ is bounded in absolute value by C^k for a constant $C > 0$ independent of n,* *Proof.* At the end of Section 3C we showed that when ϵ is in a neighborhood of 0 independent of n, the eigenvalues $c_\epsilon^{-1}\lambda_n(\epsilon)$ are enclosed in the circle Γ_n. Then

$$|c_\epsilon^{-1}\lambda_n(\epsilon) - ni| < \tfrac{1}{2}, \quad n = \pm 1, \pm 2, \dots.$$

The result follows from Cauchy's integral formula for Taylor coefficients. □

Corollary 3.5. *For $k \geq 0$ and $n \neq 0$, the coefficient of ϵ^k in $\lambda_n(\epsilon)$ is bounded in absolute value by $|n|C^k$ for a constant $C > 0$ independent of n.*

Proof. Since c_ϵ is analytic in ϵ near 0 with $c_0 = 1$, and $\lambda_n(0) = ni$, the result follows from Leibniz's rule. □

3E. *Conjugation to a Fourier multiplier.* We have conjugated the eigenspaces of $T = \partial_x + c_\epsilon^{-1}H + c_\epsilon^{-1}R_\epsilon$ (and also of $c_\epsilon\partial_x + H + R_\epsilon$) to Fourier modes via the operator

$$1 + W_\epsilon = \sum_{n\in\mathbb{Z}} P_n Q_n(\epsilon),$$

where P_0 is the projection onto the span of $e^{\pm ix}$, $Q_0(\epsilon)$ is the projection onto the span of $\psi_0^{\pm}(\epsilon)$, P_n is the projection onto the span of $e^{i(n+\operatorname{sgn} n)x}$, and $Q_n(\epsilon)$ is the projection onto the span of $\psi_n(\epsilon)$, $n = \pm 1, \pm 2, \dots$.

We will view T as a perturbation of $\partial_x + c_\epsilon^{-1}H$ and follow the proof of [Kato 1976, Chapter V, Theorem 4.15a]. In the process we will extract more information from the fact that R_ϵ is of class $\mathcal{S}$. Since

$$P_n^2 = P_n, \quad \sum_{n\in\mathbb{Z}} P_n = 1, \tag{3-15}$$

we have

$$W_\epsilon = \sum_{n\in\mathbb{Z}} P_n(Q_n(\epsilon) - P_n) \tag{3-16}$$

and $W_0 = 0$.

Proposition 3.6. *W_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ.*

Proof. We bound each term on the right-hand side separately. By [Kato 1976, Chapter V, (4.38)],

$$Q_n(\epsilon) - P_n = -c_\epsilon^{-1} Q_n(\epsilon) R_\epsilon Z_n(\epsilon) - c_\epsilon^{-1} Z_n'(\epsilon) R_\epsilon P_n,$$

where

$$Z_n(\epsilon) = \frac{1}{2\pi i}\int_{\Gamma_n} (z - (n + (1 - c_\epsilon^{-1})\operatorname{sgn} n)i)^{-1} (\partial_x + c_\epsilon^{-1}H - z)^{-1}\, dz,$$
$$Z_n'(\epsilon) = \frac{1}{2\pi i}\int_{\Gamma_n} (z - c_\epsilon^{-1}\lambda_n(\epsilon))^{-1} (T - z)^{-1}\, dz.$$

We now bound the operator norms of the right-hand side, with uniformity in ϵ and decay in n, in order to show that the sum in n converges.

First note that it is clear from the frequency side that when ϵ is in a neighborhood of 0 independent of n and $z \in \bigcup_{n\in\mathbb{Z}} \Gamma_n$ for all $m \geq 0$, the operator $(\partial_x + c_\epsilon^{-1}H - z)^{-1}$ is bounded from H^m to H^m, uniformly in ϵ and z. Since R_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ (see (3-5) and notice that $R_0 = 0$), it follows from the Neumann series that $\|(T - z)^{-1}\|_{\dot{H}^m \to \dot{H}^m}$ is finite and only depends on m. Since $|z - (n + (1 - c_\epsilon^{-1})\operatorname{sgn} n)i|$ and $|z - c_\epsilon^{-1}\lambda_n(\epsilon)|$ are uniformly bounded from below, both $Z_n(\epsilon)$ and $Z_n'(\epsilon)$ are bounded from $\dot{H}^m$

to $\dot{H}^m$, uniformly in ϵ and n. Since $Q_n(\epsilon)$ is given by a similar integral (3-13), it also has this property, which is also trivially true for P_n. Now, for all $n, m, k \in \mathbb{Z}$, $m, k \geq 0$ and $\tilde{h} \in L^2$,

$$\begin{aligned}\|Z_n'(\epsilon)R_\epsilon P_n\tilde{h}\|_{\dot{H}^k} &\lesssim_k \|R_\epsilon P_n\tilde{h}\|_{\dot{H}^k} \lesssim_{m,k} |\epsilon|\,\|P_n\tilde{h}\|_{\dot{H}^{-m-2}} \\ &\lesssim_{m,k} |\epsilon|(1+|n|)^{-2}\|\tilde{h}\|_{\dot{H}^{-m}} \end{aligned} \tag{3-17}$$

because P_n is the projection onto very specific Fourier modes. For the first term we have

$$\|R_\epsilon Z_n(\epsilon)\tilde{h}\|_{\dot{H}^k} \lesssim_{m,k} |\epsilon|\,\|Z_n(\epsilon)\tilde{h}\|_{\dot{H}^{-m}} \lesssim_{m,k} |\epsilon|\,\|\tilde{h}\|_{\dot{H}^{-m}}.$$

To introduce the action of $Q_n(\epsilon)$, note that the image of $Q_n(\epsilon)$ lies in the eigenspace of the operator $c_\epsilon\partial_x + H + R_\epsilon$, with eigenvalue $\lambda_n(\epsilon)$, so for $n \neq 0$ and $u \in \operatorname{Im} Q_n(\epsilon)$ we have

$$u = \lambda_n(\epsilon)^{-1}(c_\epsilon u' + Hu + R_\epsilon u),$$

so $\|u\|_{\dot{H}^k} \lesssim_k |\lambda_n(\epsilon)|^{-1}\|u\|_{\dot{H}^{k+1}} \lesssim |n|^{-1}\|u\|_{\dot{H}^{k+1}}$. Hence

$$\|Q_n(\epsilon)R_\epsilon Z_n(\epsilon)\|_{\dot{H}^k} \lesssim_k n^{-2}\|R_\epsilon Z_n(\epsilon)\|_{\dot{H}^{k+2}} \lesssim_{m,k} |\epsilon|(1+|n|)^{-2}\|\tilde{h}\|_{\dot{H}^{-m}}. \tag{3-18}$$

This also holds for $n = 0$ because R_ϵ/ϵ is of class $\mathcal{S}$ uniformly, so W_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ thanks to the convergence of $\sum_{n\in\mathbb{Z}}(1+|n|)^{-2}$. □

Now for $k = 0, 1, \ldots,$ there is a neighborhood of 0 such that when ϵ is in this neighborhood, $\|W_\epsilon\|_{\dot{H}^k\to\dot{H}^k} < 1$, so $1 + W_\epsilon : \dot{H}^k \to \dot{H}^k$ is invertible. By (3-15) and (3-16) it follows easily that

$$(1+W_\epsilon)Q_n(\epsilon) = P_n(1+W_\epsilon), \tag{3-19}$$

so the eigenspace of T is conjugated to the (span of) Fourier modes, and hence T is conjugated to a Fourier multiplier.

We have proven the following lemma:

Lemma 3.7. *For ϵ small enough, there exists an operator W_ϵ such that W_ϵ/ϵ is of class $\mathcal{S}$, uniformly in ϵ. Moreover*:

(1) *$1 + W_\epsilon : \dot{H}^k \to \dot{H}^k$ is invertible.*

(2) *$(1+W_\epsilon)Q_n(\epsilon) = P_n(1+W_\epsilon)$, $n \in \mathbb{Z}$.*

(3) *If ψ is in the closed linear span of the eigenvectors $\psi_n(\epsilon)$ $(n \neq 0)$ of $c_\epsilon\partial_x + H + R_\epsilon$, then*

$$(1+W_\epsilon)(c_\epsilon\partial_x + H + R_\epsilon)\psi = \Lambda_\epsilon(1+W_\epsilon)\psi,$$

where Λ_ϵ is a multiplier such that

$$\Lambda_\epsilon e^{i(n+\operatorname{sgn} n)x} = \lambda_n(\epsilon)e^{i(n+\operatorname{sgn} n)x}, \quad n = \pm1, \pm2, \ldots.$$

3F. *Taylor expansion of eigenvalues.* Now we Taylor expand the eigenvalues $\lambda_n(\epsilon)$ for $n \neq 0$. To do so it is more convenient to study the eigenvalue problem (3-2) for h:

$$L_\epsilon g := ((u_\epsilon - v_\epsilon)g)' + Hg = \lambda(\epsilon)g.$$

Recall the operator $L = L_0 = \partial_x + H$ whose action on the Fourier modes is

$$\mathcal{F}(Lg)(m) = i(m - \operatorname{sgn} m)\hat{g}(m),$$

with eigenvalues 0 (double), $\pm i$, $\pm 2i$, ... (g mean zero).

Since (u_ϵ, v_ϵ) is analytic in ϵ on a neighborhood of 0, and

$$\|h'\|_{L^2} \le \|h' + Hh\|_{L^2} + \|Hh\|_{L^2} = \|Lh\|_{L^2} + \|h\|_{L^2},$$

by [Kato 1976, Chapter VII, Theorem 2.6], L_ϵ is a holomorphic family of operators of type (A), so by Chapter VII, Section 2.3, all the results in Chapter II, Sections 1 and 2 apply, and we can compute the Taylor coefficients of $\lambda(\epsilon)$ as if L_ϵ acted on a finite-dimensional vector space.

We start with computing the resolvent of L,

$$R(z) = (L - z)^{-1}$$

whose action on the Fourier modes is

$$\mathcal{F}(R(z)g)(m) = (i(m - \operatorname{sgn} m) - z)^{-1}\hat{g}(m).$$

Around the eigenvalue ni ($n = \pm 1, \pm 2, \dots$) we have the expansion

$$R(z) = (ni - z)^{-1}P_n + \sum_{k=0}^{\infty}(z - ni)^k S_n^{k+1},$$

where P_n is the projection on the span of $e^{i(n+\operatorname{sgn} n)x}$ and

$$\mathcal{F}(S_n g)(m) = \frac{\hat{g}(m)}{i(m - \operatorname{sgn} m - n)}, \quad m \ne n + \operatorname{sgn} n. \tag{3-20}$$

By [Kato 1976, (II.2.33)],

$$\lambda_n(\epsilon) = ni + \sum_{k=1}^{\infty} \epsilon^k \lambda_n^{(k)}, \quad n = \pm 1, \pm 2, \dots,$$

where

$$\lambda_n^{(k)} = \sum_{p=1}^{k} \frac{(-1)^p}{p} \sum_{\substack{v_1+\cdots+v_p=n,\ v_j \ge 1 \\ h_1+\cdots+h_p=p-1}} \operatorname{Tr} L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)} S_n^{(h_1)},$$

where $S_n^{(0)} = -P_n$ and, for $h \ge 1$, $S_n^{(h)} = S_n^h$, with S_n defined in (3-20), and $L^{(v)}$ is the coefficient of ϵ^v in the Taylor expansion of L_ϵ. Note that the constraints in the summation imply that there is some $j \in \{1, \dots, p\}$ such that $h_j = 0$ and so $S_n^{(h_j)} = -P_n$, so every summand is a finite-rank operator whose trace is thus well-defined.

Lemma 3.8. *If A is a finite-rank operator*, *then* $\operatorname{Tr} AB = \operatorname{Tr} BA$.

Proof. By linearity we can assume A has the form $A(\cdot) = f(\cdot)v$ for some (not necessarily continuous) linear functional f. Then $\operatorname{Tr} A = f(v)$. Since $AB(\cdot) = f(B\cdot)v$ and $BA(\cdot) = f(\cdot)Bv$, it follows that $\operatorname{Tr} AB = f(Bv) = \operatorname{Tr} BA$. □

Using the lemma above, we can simplify the sum in $\lambda_n^{(k)}$ a little. Indeed, there are p circular rotations of the tuple $(h_1, \dots, h_p)$. Since $\bigl(\sum_j h_j, p\bigr) = 1$, the p circular rotations are all distinct, so we can choose the lexicographically smallest one as a representative. For such a representative, $h_1 = \min_j h_j = 0$, so $S_n^{(h_1)} = -P_n$, and thus we only need to act $L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)}$ on $e^{i(n+\operatorname{sgn} n)x}$ and take the $(n+\operatorname{sgn} n)$-th mode to compute the trace. Thus

$$\lambda_n^{(k)} = \sum_{p=1}^{k} (-1)^{p-1} \sum_{\substack{v_1+\cdots+v_p=k,\ v_j\ge 1 \\ h_1+\cdots+h_p=p-1 \\ (h_1,\dots,h_p) \text{ is a representative}}} \mathcal{F}[L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)} e^{i(n+\operatorname{sgn} n)x}](n+\operatorname{sgn} n). \tag{3-21}$$

Let us compute some terms $\lambda_n^{(k)}$ by using the formula (3-21). We have

$$\lambda_n^{(1)} = \operatorname{Tr} L^{(1)} P_n = 0$$

because L_1 shifts the mode by 1, and

$$\lambda_n^{(2)} = \operatorname{Tr}(L^{(2)} P_n - L^{(1)} S_n L^{(1)} P_n).$$

Put $s = \operatorname{sgn} n$. We extract the $(n+s)$-th mode of each term:

$$\operatorname{Tr} L^{(2)} P_n = \mathcal{F}[L^{(2)} 2 e^{i(n+s)x}](n+s) = \frac{i(n+s)}{4},$$

$$\begin{aligned} L^{(1)} S_n L^{(1)} e^{i(n+s)x} &= \frac{i L^{(1)} S_n}{2}((n+s+1)e^{i(n+s+1)x} + (n+s-1)e^{i(n+s-1)x}) \\ &= \frac{L^{(1)}}{2}((n+s+1)e^{i(n+s+1)x} - (n+s-1)e^{i(n+s-1)x}), \\ \operatorname{Tr} L^{(1)} S_n L^{(1)} P_n &= \frac{i(n+s+1)(n+s) - i(n+s-1)(n+s)}{4} = \frac{i(n+s)}{2}, \end{aligned}$$

so

$$\lambda_n^{(2)} = \frac{i(n+s)}{4} - \frac{2i(n+s)}{4} = -\frac{i(n+s)}{4}.$$

We can further compute that

$$\lambda_n(\epsilon) = in - \frac{\epsilon^2 i(n+s)}{4} - \frac{11\epsilon^4 i(n+s)}{32} - \frac{527 i\epsilon^6 (n+s)}{768} + O_n(\epsilon^7)$$

for $n = \pm 1, \pm 2, \pm 3, \dots$.

Proposition 3.9. *For $n = \pm 1, \pm 2, \dots$,*

$$\lambda_n^{(k)} = \begin{cases} 0, & 2 \nmid k, \\ i c^{(k)}(n+\operatorname{sgn} n), & k \le 2|n|+2, \end{cases}$$

where $c^{(k)}$ is the k-th Taylor coefficient of c_ϵ as defined in (3-7).

When $k \ge 2|n|+4$, $\lambda_n^{(k)}$ is still purely imaginary but the formula $\lambda_n^{(k)} = i c^{(k)}(n+\operatorname{sgn} n)$ does not hold in general.

Proof. Firstly we notice that, for $n=\pm 1$, the coefficient of ϵ^6 in $\lambda_{\pm 1}(\epsilon)$ is

$$\lambda_1(\epsilon) = i - \frac{\epsilon^2 i}{2} - \frac{11\epsilon^4 i}{16} - \frac{529\epsilon^6 i}{384} + O(\epsilon^7),$$

which does not hold for $\lambda_{\pm 1}^{(6)} = \pm 2ic^{(6)}$.

Next, we prove the fist part of the lemma. In each summand of (3-21), all the coefficients are real, except that each operator L brings a factor of i to the Fourier coefficients (via the operator ∂_x), and each operator S_n removes a factor of i (see (3-20)). Hence each summand is purely imaginary, and so is $\lambda_n^{(k)}$.

In each summand of (3-21), the operator $S_n^{(h_j)}$ is a Fourier multiplier that does not shift the modes, while the operator $L^{(m)}g = ((u^{(m)} - v^{(m)})g)'$ shifts the modes by at most m because $u^{(m)}$ only contains modes up to $e^{\pm imx}$. Also the amount of shift is equal to $m \pmod 2$. Thus when acting the sequence $L^{(v_p)} S_n^{(h_p)} \cdots L^{(v_1)}$ on $e^{i(n+s)x}$, the mode is consecutively shifted by at most $v_1, v_2, \dots, v_p$, and the total amount of shift is equal to $\sum_j v_j = k \pmod 2$. Since in the end we are taking the $(n+s)$-th mode, the total amount of shift must be 0 in order to count, so when k is odd $\lambda_n^{(k)} = 0$. When k is even, the mode $e^{i(n+s)x}$ can only be shifted as far as $e^{i(n+s\pm k/2)x}$; otherwise it can never be shifted back. Hence when $k \le 2|n|+2 = 2|n+s|$, the frequency always has the same sign as n or becomes 0. In the former case we can take $\operatorname{sgn} m = \operatorname{sgn} n$ in (3-20), while in the latter case the derivative in L kills it, so it does not hurt if we still take $\operatorname{sgn} m = \operatorname{sgn} n$ in (3-20). Either way we can take $\operatorname{sgn} m = \operatorname{sgn} n$ in (3-20). Thus the action of S_n is the same as that of S_n', where

$$\mathcal{F}(S_n' g)(m) = \frac{\hat{g}(m)}{i(m-n-\operatorname{sgn} n)}, \quad m \ne n + \operatorname{sgn} n.$$

For $n>0$, the operator S_n' is the analog of S_n for L^+, with

$$\mathcal{F}(L^+ g)(m) = i(m-1)\hat{g}(m),$$

i.e., $L^+ g = g' - ig$. Hence $\lambda_n^{(k)}$ remains the same if we replace L with L^+. Now we have

$$L_\epsilon^+ g := L^+ g + \sum_{n=1}^{\infty} \epsilon^n L^{(n)} g = -v_\epsilon g' - ig + (u_\epsilon g)' = ((u_\epsilon - v_\epsilon)g)' - ig,$$

whose eigenvalue problem is

$$((u_\epsilon - v_\epsilon)g)' - ig = \lambda^+(\epsilon) g.$$

Using the same change of variable as in Section 3B, the problem above can be transformed to

$$\tilde{h}' - i c_\epsilon^{-1} \tilde{h} = c_\epsilon^{-1} \lambda^+(\epsilon) \tilde{h},$$

whose eigenvalues are

$$\lambda_{n'}^+(\epsilon) = n' c_\epsilon i - i.$$

Since when $\epsilon \to 0$, $\lambda_n(\epsilon) \to ni$ and $c_\epsilon \to 1$, we must have $n' = n+1$, and so

$$\lambda_n(\epsilon) = (n+1)c_\epsilon i - i + O_n(\epsilon^{2n+4}).$$

For $n<0$, note that since L preserves real-valued functions, its eigenvalues come in conjugate pairs, so $\lambda_n(\epsilon) = \overline{\lambda_{|n|}(\epsilon)} = -\lambda_{|n|}(\epsilon)$ has the same property. □

Corollary 3.10. *When $|\epsilon|$ is small enough,*

$$|\lambda_n(\epsilon)-(n+\operatorname{sgn} n)c_\epsilon i+i\operatorname{sgn} n|<|n|(C\epsilon)^{2|n|+4}<C'\epsilon^6,\quad n\in\mathbb{Z}\backslash\{0\},$$
$$|\lambda_n'(\epsilon)-(n+\operatorname{sgn} n)\partial_\epsilon c_\epsilon i|<|n|(C\epsilon)^{2|n|+3}<C'\epsilon^5,\quad n\in\mathbb{Z}\backslash\{0\},$$

for some constant $C, C'>0$ independent of n.

Proof. By Proposition 3.9. the Taylor expansions of $\lambda_n(\epsilon)$ and $(n+\operatorname{sgn} n)c_\epsilon i-i\operatorname{sgn} n$ differ only from the term $\epsilon^{2|n|+4}$. By Corollary 3.5, the error terms of the former sum up to $O\big(|n|\sum_{k=2|n|+4}^\infty(C\epsilon)^k\big)=O(|n|(C\epsilon)^{2|n|+4})$ if, say, $C|\epsilon|<\frac{1}{2}$. The error term of the latter clearly also satisfy this bound.

To extend the chain of inequalities it suffices to note that $|n|(C\epsilon)^{2|n|-2}$ is uniformly bounded for $n\neq 0$ if $|C\epsilon|<\frac{1}{2}$. □

3G. *Time resonance analysis.* For m, n and $l\in\mathbb{Z}$ we consider

$$\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)=(m+n+l)c_\epsilon i+(\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l)(c_\epsilon-1)i+O(\epsilon^6).$$

Proposition 3.11. *If $m,n,l\in\mathbb{Z}$ and $mnl\neq 0$, then when ϵ is small enough, $|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|>\frac{1}{5}\epsilon^2$.*

Proof. By (3-7) and (2-7),

$$\begin{aligned}\frac{c_\epsilon}{2\pi}&=\left(\int_0^{2\pi}\frac{dy}{u_\epsilon(y)-v_\epsilon}\right)^{-1}=\left(\int_0^{2\pi}\frac{dy}{1+\epsilon\cos y-\frac{1}{2}\epsilon^2\cos 2y+\frac{1}{4}\epsilon^2}\right)^{-1}+O(\epsilon^3)\\&=\left(\int_0^{2\pi}\big(1-\epsilon\cos y+\epsilon^2\cos^2 y+\tfrac{1}{2}\epsilon^2\cos 2y-\tfrac{1}{4}\epsilon^2\big)\,dy\right)^{-1}+O(\epsilon^3),\\c_\epsilon&=\big(1+\tfrac{1}{4}\epsilon^2\big)^{-1}+O(\epsilon^3)=1-\tfrac{1}{4}\epsilon^2+O(\epsilon^3).\end{aligned}$$

We distinguish three cases.

Case 1: $m+n+l\neq 0$. Then $|m+n+l|\geq 1$. Since $c_\epsilon-1\lesssim\epsilon^2$,

$$\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)=(m+n+l)c_\epsilon i+O(\epsilon^2).$$

Since $c_\epsilon\to 1$ as $\epsilon\to 0$, we have $|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|>\frac{1}{2}|m+n+l|$ for small ϵ.

Case 2: $m+n+l=0$ and $mnl\neq 0$. Then

$$\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)=-\tfrac{1}{4}(\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l)\epsilon^2 i+O(\epsilon^3).$$

Since $|\operatorname{sgn} m|=|\operatorname{sgn} n|=|\operatorname{sgn} l|=1$, we have $|\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l|\geq 1$, so $|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|>\frac{1}{5}\epsilon^2$ when ϵ is small enough. □

When $m+n+l=0$ and $mnl=0$, since $\lambda_n(\epsilon)$ is odd in n, it follows that $\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)=0$. We do have time resonance in this case. We will eliminate this case by choosing a new frame of reference.

4. A new frame of reference

Recall that the traveling wave solution

$$f_\epsilon(x,t)=u_\epsilon(x+v_\epsilon t)$$

satisfies

$$\partial_t f_\epsilon = H f_\epsilon + f_\epsilon \partial_x f_\epsilon,$$

i.e.,

$$v_\epsilon u'_\epsilon = H u_\epsilon + u_\epsilon u'_\epsilon.$$

Now we aim to find a new reference frame. Let $P_0^\pm(\epsilon)$ be the projection on the 1-dimensional space spanned by the eigenvector $\varphi_0^+(\epsilon) = \partial_\epsilon u_\epsilon$ and $\varphi_0^-(\epsilon) = -\epsilon^{-1} u'_\epsilon$, respectively. Then we aim to rewrite

$$f(x,t) = u_{\epsilon(t)}(x + a(t)) + g(x + a(t), t),$$

where $\epsilon, a \in \mathbb{R}$ and $P_0^\pm(\epsilon(t))g = 0$. We first show that it is always possible, provided that f is close to a traveling wave.

Proposition 4.1. *Let $k \geq 2$. Then there is $r = r(k) > 0$ such that if $|\epsilon_0| < r$ and $\|f - u_{\epsilon_0}\|_{H^k} < r|\epsilon_0|$, then there is $\epsilon \in \mathbb{R}$, $a \in \mathbb{R}/2\pi\mathbb{Z}$ and $g \in H^k$ such that*

$$f(x) = u_\epsilon(x + a) + g(x + a), \tag{4-1}$$

$$P_0^\pm(\epsilon) g = 0, \tag{4-2}$$

$$|\epsilon - \epsilon_0| + \|g\|_{H^k} \lesssim \|f - u_{\epsilon_0}\|_{H^k}. \tag{4-3}$$

Moreover, ϵ, a and g depend smoothly on f.

Proof. Define the map $F : (-r, r)^2 \to \mathbb{R}^2$, $(\epsilon, a) \mapsto (y^+, y^-)$, with

$$P_0^\pm(\epsilon)(f(x - a) - u_\epsilon(x)) = y^\pm \varphi_0^\pm(\epsilon). \tag{4-4}$$

We now find the solution to the equation $F(\epsilon, a) = 0$. Since $P_0^\pm(\epsilon)$ is uniformly bounded in L^2 and $\|\varphi_0^\pm(\epsilon)\|$ is uniformly bounded from below,

$$|F(\epsilon, a)| \lesssim \|f(x - a) - u_\epsilon\|_{L^2}. \tag{4-5}$$

Summing the two equations in (4-4) and taking the total derivative yields

$$-(P_0^+(\epsilon) + P_0^-(\epsilon))(f'(x-a))da - \varphi_0^+(\epsilon)d\epsilon + (\partial_\epsilon P_0^+(\epsilon) + \partial_\epsilon P_0^-(\epsilon))(f(x-a) - u_\epsilon(x))d\epsilon \tag{4-6}$$

$$= \varphi_0^+(\epsilon)dy^+ + \varphi_0^-(\epsilon)dy^- + y^+\partial_\epsilon\varphi_0^+(\epsilon)d\epsilon + y^-\partial_\epsilon\varphi_0^-(\epsilon)d\epsilon. \tag{4-7}$$

Since $\|f\|_{H^2} \leq \|u_{\epsilon_0}\|_{H^2} + r|\epsilon_0| \lesssim |\epsilon_0|$, we have

$$\begin{aligned}\|f(x-a) - u_\epsilon\|_{H^1} &\leq \|f(x-a) - f(x)\|_{H^1} + \|f - u_{\epsilon_0}\|_{H^1} + \|u_\epsilon - u_{\epsilon_0}\|_{H^1} \\ &\lesssim |a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|. \end{aligned} \tag{4-8}$$

Since both $P_0^\pm(\epsilon)$ and $\partial_\epsilon P_0^\pm(\epsilon)$ are uniformly bounded on L^2, and $u'_\epsilon = -\epsilon\varphi_0^-(\epsilon)$,

$$\|(4\text{-}6) - \epsilon\varphi_0^-(\epsilon)da + \varphi_0^+(\epsilon)d\epsilon\|_{L^2} \lesssim (|a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|)(|da| + |d\epsilon|).$$

By (4-5) and (4-8),

$$\|y^\pm \partial_\epsilon\varphi_0^\pm(\epsilon)\|_{L^2} \lesssim |F(\epsilon, a)| \lesssim |a\epsilon_0| + r|\epsilon_0| + |\epsilon - \epsilon_0|,$$

so

$$\|(4\text{-}7)-\varphi_0^+(\epsilon)dy^+-\varphi_0^-(\epsilon)dy^-\|_{L^2}\lesssim(|a\epsilon_0|+r|\epsilon_0|+|\epsilon-\epsilon_0|)|d\epsilon|.$$

Hence the equality between (4-6) and (4-7) gives an estimate of the differential

$$\left\|dF(\epsilon,a)-\begin{pmatrix}1&0\\0&-\epsilon\end{pmatrix}\right\|\lesssim|a\epsilon_0|+r|\epsilon_0|+|\epsilon-\epsilon_0|.$$

We assume that the solution (ϵ,a) satisfies $|\epsilon-\epsilon_0|+|a\epsilon_0|<r_0|\epsilon_0|$, where r_0 is small enough. This in particular implies $\frac{1}{2}|\epsilon_0|<|\epsilon|<2|\epsilon_0|$. Then

$$\left\|dF(\epsilon,a)-\begin{pmatrix}1&0\\0&-\epsilon_0\end{pmatrix}\right\|\lesssim(r_0+r)|\epsilon_0|$$

is also small enough. Let

$$G=\mathbb{I}+\left(dF(\epsilon,a)-\begin{pmatrix}1&0\\0&-\epsilon_0\end{pmatrix}\right)\begin{pmatrix}1&0\\0&-1/\epsilon_0\end{pmatrix}.$$

Then

$$dF=G\begin{pmatrix}1&0\\0&-\epsilon_0\end{pmatrix}$$

and

$$\|G-\mathbb{I}\|\lesssim r_0+r.$$

If r_0 and r are small enough, then $\|G\|$ and $\|G^{-1}\|<2$.

Let $(\epsilon_1,a_1)=(\epsilon_0,0)-dF(\epsilon_0,0)^{-1}F(\epsilon_0,0)$. Then (recalling (4-5))

$$|\epsilon_1-\epsilon_0|+|a_1\epsilon_0|\lesssim|G^{-1}F(\epsilon_0,0)|\lesssim|F(\epsilon_0,0)|\lesssim\|f-u_{\epsilon_0}\|_{L^2}\lesssim r|\epsilon_0|.$$

Since $|\partial_\epsilon^2F|$ and $|\partial_{a\epsilon}F|\lesssim1$, and $|\partial_a^2F|\lesssim\|f\|_{H^2}\lesssim|\epsilon_0|$, by Taylor's theorem,

$$|F(\epsilon_1,a_1)|\lesssim|\epsilon_1-\epsilon_0|^2+|\epsilon_1-\epsilon_0|\,|a_1|+|\epsilon_0|\,|a_1|^2\lesssim r|F(\epsilon_0,0)|.$$

Hence the iteration $(\epsilon_{n+1},a_{n+1})=(\epsilon_n,a_n)+dF(\epsilon_n,a_n)^{-1}F(\epsilon_n,a_n)$ converges when r is small enough. Moreover $|\epsilon_n-\epsilon_0|+|a_n\epsilon_0|\lesssim|F(\epsilon_0,0)|$. Then $(\epsilon,a):=\lim_{n\to\infty}(\epsilon_n,a_n)$ satisfies $F(\epsilon,a)=0$ and $|\epsilon-\epsilon_0|+|a\epsilon_0|\lesssim|F(\epsilon_0,0)|\lesssim r|\epsilon_0|<r_0|\epsilon_0|$ if r is small compared to r_0.

Let $g=f(x-a)-u_\epsilon$. Then (4-1) and (4-2) clearly hold. Moreover,

$$\begin{aligned}\|g\|_{H^k}=\|g(x+a)\|_{H^k}&=\|f(x)-u_\epsilon(x+a)\|_{H^k}\\&\le\|f-u_{\epsilon_0}\|_{H^k}+\|u_\epsilon(x+a)-u_{\epsilon_0}(x)\|_{H^k}\\&\lesssim\|f-u_{\epsilon_0}\|_{H^k}+|\epsilon-\epsilon_0|+|a\epsilon_0|\\&\lesssim\|f-u_{\epsilon_0}\|_{H^k}+|F(\epsilon_0,0)|\lesssim\|f-u_{\epsilon_0}\|_{H^k}\end{aligned}$$

showing (4-3). The smooth dependence of ϵ, a and g on f is also clear. □

By translation symmetry, if f is $r|\epsilon_0|$-close to $u_{\epsilon_0}(x+a)$ for some $a\in\mathbb{R}/2\pi\mathbb{Z}$, we can reach a similar conclusion. Then we can write

$$f(x,t)=u_{\epsilon(t)}(x+a(t))+g(x+a(t),t).$$

We will obtain an energy estimate for g. Combined with local well-posedness of (1-1) and Proposition 4.1, we can show that the solution extends as long as the energy estimate closes, see the end of Section 5B.

To get the energy estimate, we first need to derive an evolution equation for g. Since f is differentiable in t, so are $\epsilon(t)$, $a(t)$ and g, and we get

$$f_t(x,t) = a'(t)(u'_\epsilon + g_x)(x+a(t)) + \epsilon'(t)\partial_\epsilon u_\epsilon(x+a(t)) + g_t(x+a(t),t)$$

and

$$\begin{aligned}(Hf + ff_x)(x,t) = (Hu_\epsilon + u_\epsilon u'_\epsilon)(x+a(t)) + Hg(x+a(t),t)\\ + \partial_x(u_\epsilon(x+a(t))g(x+a(t),t)) + (gg_x)(x+a(t),t).\end{aligned}$$

The equation for g is then

$$\begin{aligned}g_t &= v_\epsilon u'_\epsilon - a'(t)(u'_\epsilon + g_x) - \epsilon'(t)\partial_\epsilon u_\epsilon + Hg + (u_\epsilon g)_x + gg_x\\ &= L_\epsilon g + (v_\epsilon - a'(t))(u'_\epsilon + g_x) - \epsilon'(t)\partial_\epsilon u_\epsilon + gg_x.\end{aligned}$$

Since $P_0^\pm(\epsilon)g(t) = 0$, we have $P_0^\pm(\epsilon)g_t = -\epsilon'(t)\partial_\epsilon P_0^\pm(\epsilon)g$, so the action of the projections $P_0^\pm(\epsilon)$ on the above equation is

$$\begin{aligned}(v_\epsilon - a'(t))P_0^+(\epsilon)g_x + \epsilon'(t)(\partial_\epsilon P_0^+(\epsilon)g - \partial_\epsilon u_\epsilon) + P_0^+(\epsilon)(gg_x) = 0,\\ (v_\epsilon - a'(t))(u'_\epsilon + P_0^-(\epsilon)g_x) + \epsilon'(t)\partial_\epsilon P_0^-(\epsilon)g + P_0^-(\epsilon)(gg_x) = 0.\end{aligned}$$

Since $P_0^\pm(\epsilon)$ are bounded on L^2, we have $\|P_0^\pm(\epsilon)g_x\|_{L^2} \lesssim \|g\|_{H^1}$. Since $P_0^\pm(\epsilon)$ are analytic in ϵ, we have $\|\partial_\epsilon P_0^\pm(\epsilon)g\|_{L^2} \lesssim \|g\|_{L^2}$. Since $P_0^\pm(\epsilon)$ is a projection, we have $P_0^\pm(\epsilon)^2 = P_0^\pm(\epsilon)$. Taking the derivative in ϵ and using the constraint $P_0^\pm(\epsilon)g = 0$, we have $P_0^\pm(\epsilon)\partial_\epsilon P_0^\pm(\epsilon)g = \partial_\epsilon P_0^\pm(\epsilon)g$, i.e., $\partial_\epsilon P_0^\pm(\epsilon)g$ is in the 1-dimensional space spanned by $\varphi_0^\pm(\epsilon)$. Hence

$$|P_0^\pm(\epsilon)g_x/\varphi_0^+(\epsilon)| \lesssim \|g\|_{H^1}, \quad |\partial_\epsilon P_0^\pm(\epsilon)g/\varphi_0^\pm(\epsilon)| \lesssim \|g\|_{L^2}.$$

Thus, dividing the two equations by $\varphi_0^\pm(\epsilon)$ we get

$$\left|\left(\begin{pmatrix}0 & 1\\ \epsilon & 0\end{pmatrix} + O(\|g\|_{H^1})\right)\begin{pmatrix}v_\epsilon - a'(t)\\ \epsilon'(t)\end{pmatrix}\right| = \left|\begin{pmatrix}P_0^+(\epsilon)(gg_x)/\varphi_0^+(\epsilon)\\ P_0^-(\epsilon)(gg_x)/\varphi_0^-(\epsilon)\end{pmatrix}\right| \lesssim \|g(t)\|_{H^1}^2.$$

Assuming $\|g(t)\|_{H^1}/|\epsilon|$ is small enough we have

$$\begin{pmatrix}v_\epsilon - a'(t)\\ \epsilon'(t)\end{pmatrix} = \begin{pmatrix}O(\|g(t)\|_{H^1}^2/|\epsilon|)\\ O(\|g(t)\|_{H^1}^2)\end{pmatrix}. \tag{4-9}$$

4A. ***Diagonalization.*** To find the evolution of other modes, we diagonalize the equation for g. Let $g = h_x$ and $h = \tilde h \circ \phi_\epsilon$, where ϕ_ϵ satisfies (3-3). Recall from (3-1) that $L_\epsilon g = -v_\epsilon g_x + Hg + (u_\epsilon g)_x$, so

$$h_t = -v_\epsilon h_x + Hh + u_\epsilon h_x - \epsilon'(t)\partial_\epsilon U_\epsilon + (v_\epsilon - a'(t))(u_\epsilon + h_x) + \tfrac{1}{2}h_x^2 \pmod 1,$$

where U_ϵ is a primitive of u_ϵ. Differentiating $h = \tilde h \circ \phi_\epsilon$ with respect to ϵ we get

$$h_t = \tilde h_t \circ \phi_\epsilon + \epsilon'(t)(\partial_\epsilon \phi_\epsilon)(\tilde h_x \circ \phi_\epsilon).$$

On the other hand,

$$(-v_\epsilon h_x + Hh + u_\epsilon h_x)_x = L_\epsilon g = (((c_\epsilon \partial_x + H + R_\epsilon)\tilde{h}) \circ \phi_\epsilon)_x,$$

so

$$\tilde{h}_t = (c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} - \epsilon'(t)(\partial_\epsilon \phi_\epsilon \circ \phi_\epsilon^{-1})\tilde{h}_x - \epsilon'(t)\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1} \\ + (v_\epsilon - a'(t))(u_\epsilon + h_x) \circ \phi_\epsilon^{-1} + \tfrac{1}{2} h_x^2 \circ \phi_\epsilon^{-1} \pmod 1.$$

By the chain rule, $h_x = \phi_\epsilon'(\tilde{h}_x \circ \phi_\epsilon)$, so $h_x \circ \phi_\epsilon^{-1} = (\phi_\epsilon' \circ \phi_\epsilon^{-1})\tilde{h}_x$, and

$$\tilde{h}_t = (c_\epsilon \partial_x + H + R_\epsilon)\tilde{h} + \Phi_\epsilon \tilde{h}_x - \epsilon'(t)\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1} + (v_\epsilon - a'(t))u_\epsilon \circ \phi_\epsilon^{-1} + \tfrac{1}{2}(\phi_\epsilon' \circ \phi_\epsilon^{-1})^2 \tilde{h}_x^2 \pmod 1,$$
$$\Phi_\epsilon = -\epsilon'(t)(\partial_\epsilon \phi_\epsilon \circ \phi_\epsilon^{-1}) + (v_\epsilon - a'(t))(\phi_\epsilon' \circ \phi_\epsilon^{-1}).$$

Using the operator W_ϵ from Lemma 3.7 we have

$$(1 + W_\epsilon)(c_\epsilon \partial_x + H + R_\epsilon) = \Lambda_\epsilon (1 + W_\epsilon),$$

where Λ_ϵ is a Fourier multiplier whose action on the Fourier mode $e^{i(n+\operatorname{sgn} n)x}$ is multiplication by $\lambda_n(\epsilon)$. Since W_ϵ/ϵ is of class $\mathcal{S}$, uniformly in ϵ, for any smooth function F, the operator

$$\tilde{h} \mapsto \mathcal{R}_\epsilon(F)\tilde{h} := (1 + W_\epsilon)(F\tilde{h}_x) - F((1 + W_\epsilon)\tilde{h})_x$$

is of class $\mathcal{S}$, with the implicit constants depending on the C^k norms of F.

Let $\mathfrak{h} = (1 + W_\epsilon)\tilde{h}$. Then

$$(1 + W_\epsilon)\tilde{h}_t = \Lambda_\epsilon \mathfrak{h} + \Phi_\epsilon \mathfrak{h}_x - \epsilon'(t)(1 + W_\epsilon)(\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1}) \\ + (v_\epsilon - a'(t))(1 + W_\epsilon)(u_\epsilon \circ \phi_\epsilon^{-1}) + N_\epsilon[\mathfrak{h}, \mathfrak{h}] + \mathcal{R}_\epsilon(\Phi_\epsilon)\tilde{h} \pmod 1,$$

where

$$N_\epsilon[\mathfrak{h}, \mathfrak{h}] = \tfrac{1}{2}(\phi_\epsilon' \circ \phi_\epsilon^{-1})^2((1 + W_\epsilon)^{-1}\mathfrak{h})_x^2. \tag{4-10}$$

Both $\mathcal{R}_\epsilon(\partial_\epsilon \phi_\epsilon \circ \phi_\epsilon^{-1})$ and $\mathcal{R}_\epsilon(\phi_\epsilon' \circ \phi_\epsilon^{-1} - 1)/\epsilon$ are of class $\mathcal{S}$, uniformly in ϵ when ϵ is small. Moreover, since W_ϵ is analytic in ϵ with $W_0 = 0$, so is $\mathcal{R}_\epsilon(1)$ with $\mathcal{R}_0(1) = 0$. Hence $R_\epsilon(1)/\epsilon$ is of class $\mathcal{S}$ uniformly in ϵ, and so is $\mathcal{R}_\epsilon(\phi_\epsilon' \circ \phi_\epsilon^{-1})/\epsilon$.

Since $\partial_\epsilon u_\epsilon$ and u_ϵ' are in the generalized eigenspace of L_ϵ associated with the eigenvalue 0, we have $\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1}$ and $u_\epsilon \circ \phi_\epsilon^{-1}$ are in the corresponding space of $c_\epsilon \partial_x + H + R_\epsilon$, so $(1 + W_\epsilon)(\partial_\epsilon U_\epsilon \circ \phi_\epsilon^{-1})$ and $(1 + W_\epsilon)(u_\epsilon \circ \phi_\epsilon^{-1})$ are in the space spanned by $\sin x$ and $\cos x$, according to Lemma 3.7.

Now we have

$$\begin{aligned} \mathfrak{h}_t &= (1 + W_\epsilon)\tilde{h}_t + \epsilon'(t)\partial_\epsilon W_\epsilon \tilde{h} \\ &= \Lambda_\epsilon \mathfrak{h} + \Phi_\epsilon \mathfrak{h}_x + N_\epsilon[\mathfrak{h}, \mathfrak{h}] + \text{Rest} \pmod{1, \sin x, \cos x}, \end{aligned} \tag{4-11}$$

where $N_\epsilon[\mathfrak{h}, \mathfrak{h}]$ is given by (4-10) and

$$\text{Rest} = \epsilon'(t)(\partial_\epsilon W_\epsilon)\tilde{h} + \mathcal{R}_\epsilon(\Phi_\epsilon)\tilde{h}$$

is also of class $\mathcal{S}$ uniformly in ϵ when ϵ is small.

Recall that $\epsilon'(t)$ and $a'(t)$ are chosen such that $P_0(\epsilon)g(t)=0$ for all t, where $P_0(\epsilon)$ is the projection onto the span of $\partial_\epsilon u_\epsilon$ and u'_ϵ. This implies $Q_0(\epsilon)\tilde{h}(t)=0$ for all t, where $Q_0(\epsilon)$ is the projection onto the span of $\partial_\epsilon U_\epsilon\circ\phi_\epsilon^{-1}$ and $u_\epsilon\circ\phi_\epsilon^{-1}$. Since $1+W_\epsilon$ maps the span of $\partial_\epsilon U_\epsilon\circ\phi_\epsilon^{-1}$ and $u_\epsilon\circ\phi_\epsilon^{-1}$ to the span of $\sin x$ and $\cos x$, we have $\hat{\mathfrak{h}}(1)=\hat{\mathfrak{h}}(-1)=0$ for all t.

5. Energy estimates

Since $\hat{\mathfrak{h}}(1)=\hat{\mathfrak{h}}(-1)=0$ for all t, for $k=0,1,\dots$ we define the energy

$$E_k=\tfrac{1}{2}\|\mathfrak{h}\|^2_{\dot{H}^k}=\tfrac{1}{2}\|\mathfrak{h}\|^2_{H^k/(1,\sin x,\cos x)}$$

and aim to control its growth.

Using the evolution equation (4-11) for $\mathfrak{h}$ and the anti-self-adjointness of Λ_ϵ we get

$$\begin{aligned}\frac{d}{dt}E_k(t)&=E_\Phi(t)+E_N(t)+E_{\mathrm{Rest}}(t),\\ E_\Phi(t)&=\langle\Phi_\epsilon\mathfrak{h}_x,\mathfrak{h}\rangle_{\dot{H}^k},\\ E_N(t)&=\langle N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\mathfrak{h}(t)\rangle_{\dot{H}^k},\\ E_{\mathrm{Rest}}(t)&=\langle\epsilon'(t)\partial_\epsilon W_\epsilon\tilde{h}(t)+\mathcal{R}_\epsilon(\Phi_\epsilon)\tilde{h}(t),\mathfrak{h}(t)\rangle_{\dot{H}^k}.\end{aligned}$$

Recall that $g=h_x$, $h=\tilde{h}\circ\phi_\epsilon$ and $\mathfrak{h}=(1+W_\epsilon)\tilde{h}$. When ϵ is small enough, the last two are bounded operators with bounded inverse between $\dot{H}^k$, $k=0,1,\dots$, so

$$\|g\|_{H^k}\approx_k\|h\|_{\dot{H}^{k+1}}\approx_k\|\tilde{h}\|_{\dot{H}^{k+1}}\approx_k\|\mathfrak{h}\|_{\dot{H}^{k+1}}. \tag{5-1}$$

Since $\mathcal{R}_\epsilon(\partial_\epsilon\phi_\epsilon\circ\phi_\epsilon^{-1})$, $\mathcal{R}_\epsilon(\phi'_\epsilon\circ\phi_\epsilon^{-1})/\epsilon$ and $\partial_\epsilon W_\epsilon$ are of class $\mathcal{S}$ uniformly in ϵ,

$$\begin{aligned}\|(\epsilon'(t)\partial_\epsilon W_\epsilon-\epsilon'(t)\mathcal{R}_\epsilon(\partial_\epsilon\phi_\epsilon\circ\phi_\epsilon^{-1}))\tilde{h}(t)\|_{\dot{H}^k}&\lesssim_k\|g(t)\|^2_{H^1}\|\tilde{h}(t)\|_{\dot{H}^1}\lesssim_k E_2(t)^{3/2},\\ \|(v_\epsilon-a'(t))\mathcal{R}_\epsilon(\phi'_\epsilon\circ\phi_\epsilon^{-1})\tilde{h}(t)\|_{\dot{H}^k}&\lesssim_k(\|g(t)\|^2_{H^1}/\epsilon)\epsilon\|\tilde{h}(t)\|_{\dot{H}^1}\lesssim_k E_2(t)^{3/2}\end{aligned}$$

so

$$|E_{\mathrm{Rest}}(t)|\lesssim_k E_2(t)^{3/2}E_k(t)^{1/2}. \tag{5-2}$$

To bound E_Φ we use (4-9) and (5-1) to get

$$\|\Phi'_\epsilon\|_{C^k}\lesssim_k\|g(t)\|^2_{H^1}+(\|g(t)\|^2_{H^1}/|\epsilon|)|\epsilon|\lesssim_k E_2(t).$$

Since E_Φ loses only one derivative in $\mathfrak{h}$, we have

$$|E_\Phi(t)-\langle\Phi_\epsilon\partial_x^{k+1}\mathfrak{h}(t),\partial_x^k\mathfrak{h}(t)\rangle_{L^2/(1)}|\lesssim_k E_2(t)E_k(t). \tag{5-3}$$

For the sake of bounding this term, since the inner product is taken in the space $L^2/(1)$, we can without loss of generality assume that $\hat{\mathfrak{h}}(0)=0$ (which is not true in general) and integrate by parts to get

$$2\langle\Phi_\epsilon\partial_x^{k+1}\mathfrak{h}(t),\partial_x^k\mathfrak{h}(t)\rangle_{L^2/(1)}=\int_0^{2\pi}\Phi_\epsilon\partial_x(\partial_x^k\mathfrak{h}(t))^2\,dx=-\int_0^{2\pi}\Phi'_\epsilon(\partial_x^k\mathfrak{h}(t))^2\,dx$$

so again by (4-9) and (5-1),

$$|E_\Phi(t)|\lesssim_k E_2(t)E_k(t). \tag{5-4}$$

Combining (5-2), (5-3) and (5-4) shows that

$$\left|\frac{d}{dt}E_k(t) - E_N(t)\right| \lesssim_k E_2(t)E_k(t). \tag{5-5}$$

5A. ***Normal form transformation.*** To bound E_N we recall the expression of N_ϵ from (4-10). Since N_ϵ does not depend on the constant mode of $\mathfrak{h}$, we can also assume without loss of generality that $\hat{\mathfrak{h}}(0) = 0$. We further have the decompositions

$$\begin{aligned} E_N(t) &= E_{N1}(t) + E_{N2}(t), \\ E_{N1}(t) &= \frac{1}{2}\int_0^{2\pi} \partial_x^k \mathfrak{h}(t)\partial_x^k(\partial_x \mathfrak{h}(t))^2\,dx = \sum_{j=2}^{[k/2]+1} c_{kj}\int_0^{2\pi} \partial_x^k \mathfrak{h}(t)\partial_x^{k+2-j}\mathfrak{h}(t)\partial_x^j \mathfrak{h}(t), \end{aligned} \tag{5-6}$$

where $c_{kj} \in \mathbb{R}$ are constants and we integrated by parts to get rid of the terms with $k+1$ derivatives falling on a single factor of $\mathfrak{h}$.

We use the normal form transformation to bound them. Define the trilinear map

$$\begin{aligned} D_\epsilon[f_1, f_2, f_3] = \sum_{mnl\neq 0} \frac{1}{\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} & \hat{f}_1(m+\operatorname{sgn} m)e^{i(m+\operatorname{sgn} m)x} \\ & \times \hat{f}_2(n+\operatorname{sgn} n)e^{i(n+\operatorname{sgn} n)x}\hat{f}_3(l+\operatorname{sgn} l)e^{i(l+\operatorname{sgn} l)x}\,dx \end{aligned}$$

and put

$$D_{1,k,j}(t) = D_{\epsilon(t)}[\partial_x^k \mathfrak{h}(t), \partial_x^{k+2-j}\mathfrak{h}(t), \partial_x^j \mathfrak{h}(t)].$$

Then

$$\begin{aligned} \frac{d}{dt}D_{1,k,j}(t) = \epsilon'(t)(\partial_\epsilon D_\epsilon)&[\partial_x^k \mathfrak{h}(t), \partial_x^{k+2-j}\mathfrak{h}(t), \partial_x^j \mathfrak{h}(t)] \\ &+ D_\epsilon[\partial_x^k \partial_t \mathfrak{h}(t), \partial_x^{k+2-j}\mathfrak{h}(t), \partial_x^j \mathfrak{h}(t)] \\ &+ D_\epsilon[\partial_x^k \mathfrak{h}(t), \partial_x^{k+2-j}\partial_t \mathfrak{h}(t), \partial_x^j \mathfrak{h}(t)] \\ &+ D_\epsilon[\partial_x^k \mathfrak{h}(t), \partial_x^{k+2-j}\mathfrak{h}(t), \partial_x^j \partial_t \mathfrak{h}(t)]. \end{aligned}$$

Note that $E_{N1}(t)$ is a linear combination of the last three lines on the right-hand side, with ∂_t replaced with Λ_ϵ, so $(d/dt)\sum_{j=2}^{[k/2]+1} c_{jk}D_{1,k,j}(t) - E_{N1}(t)$ is a linear combination of

$$\epsilon'(t)(\partial_\epsilon D_\epsilon)[\partial_x^k \mathfrak{h}(t), \partial_x^{k+2-j}\mathfrak{h}(t), \partial_x^j \mathfrak{h}(t)], \tag{5-7}$$

$$D_\epsilon[\partial_x^k(\partial_t - \Lambda_\epsilon)\mathfrak{h}(t), \partial_x^{k+2-j}\mathfrak{h}(t), \partial_x^j \mathfrak{h}(t)], \tag{5-8}$$

$$D_\epsilon[\partial_x^k \mathfrak{h}(t), \partial_x^{k+2-j}(\partial_t - \Lambda_\epsilon)\mathfrak{h}(t), \partial_x^j \mathfrak{h}(t)], \tag{5-9}$$

$$D_\epsilon[\partial_x^k \mathfrak{h}(t), \partial_x^{k+2-j}\mathfrak{h}(t), \partial_x^j(\partial_t - \Lambda_\epsilon)\mathfrak{h}(t)]. \tag{5-10}$$

We estimate these terms one by one.

By the definition of D_ϵ,

$$\begin{aligned} (5\text{-}7) = \epsilon'(t) &\sum_{mnl\neq 0} \frac{(\lambda'_m(\epsilon)+\lambda'_n(\epsilon)+\lambda'_l(\epsilon))}{2(\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon))^2} \\ &\times \int_0^{2\pi} \hat{\mathfrak{h}}(m+\operatorname{sgn} m, t)\partial_x^k e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n, t)\partial_x^{k+2-j}e^{i(n+\operatorname{sgn} n)x}\hat{\mathfrak{h}}(l+\operatorname{sgn} l, t)\partial_x^j e^{i(l+\operatorname{sgn} l)x}\,dx. \end{aligned}$$

We first bound the fraction. By Corollary 3.10, when ϵ is small enough,

$$\begin{aligned}\lambda_m'(\epsilon)+\lambda_n'(\epsilon)+\lambda_l'(\epsilon) &= (m+n+l+\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l)\partial_\epsilon c_\epsilon i+O(\epsilon^5)\\ &\lesssim (|m+n+l|+1)|\epsilon|.\end{aligned} \tag{5-11}$$

On the other hand, the integral vanishes unless

$$m+n+l+\operatorname{sgn} m+\operatorname{sgn} n+\operatorname{sgn} l=0, \tag{5-12}$$

in which case $m+n+l$ is an odd number, and so is nonzero. Then by Case 1 of Proposition 3.11,

$$|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|>\tfrac{1}{2}|m+n+l|, \tag{5-13}$$

so

$$\frac{|\lambda_m'(\epsilon)+\lambda_n'(\epsilon)+\lambda_l'(\epsilon)|}{|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|^2}\lesssim|\epsilon|. \tag{5-14}$$

Then for $k\ge 3$,

$$\begin{aligned}|(5\text{-}7)| &\lesssim |\epsilon(t)\epsilon'(t)|\sum_{\substack{mnl\neq 0\\ (5\text{-}12)}}|(m+\operatorname{sgn} m)^k\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)\\ &\qquad\qquad\times(n+\operatorname{sgn} n)^{k+2-j}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)(l+\operatorname{sgn} l)^j\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)|\\ &\approx |\epsilon(t)\epsilon'(t)|\left|\int_0^{2\pi}\partial_x^k H(x,t)\partial_x^{k+2-j}H(x,t)\partial_x^j H(x,t)\right|dx\\ &\lesssim_k |\epsilon(t)\epsilon'(t)|\,\|H(x,t)\|_{H_x^k}^2\|H(x,t)\|_{W_x^{[k/2]+1,\infty}}\lesssim_k|\epsilon(t)\epsilon'(t)|\,\|H(x,t)\|_{H_x^k}^3\end{aligned} \tag{5-15}$$

since $k\ge[k/2]+2$, where

$$H(x,t)=\sum_{m\neq 0}|\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)|e^{i(m+\operatorname{sgn} m)x}$$

satisfies

$$\|H(x,t)\|_{H_x^k}=\|\mathfrak{h}(t)\|_{\dot H^k}\lesssim E_k(t)^{1/2}$$

so by (4-9) and (5-1),

$$|(5\text{-}7)|\lesssim_k|\epsilon|E_2(t)E_k(t)^{3/2}. \tag{5-16}$$

To bound the other terms (5-8), (5-9) and (5-10), we use the evolution equation (4-11) of $\mathfrak{h}$, which loses one derivative in $\mathfrak{h}$, so

$$\|(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t)\|_{\dot H^{k-1}}\lesssim(\|g\|_{H^1}^2/|\epsilon|)\|\mathfrak{h}(t)\|_{\dot H^k}+\|\mathfrak{h}(t)\|_{\dot H^k}^2.$$

If $\|g(t)\|_{H^1}/|\epsilon|$ is small enough and $k\ge 2$, the first term is dominated by the second term thanks to (5-1). Since in the summation of D_ϵ it holds that $m+n+l\neq 0$, the denominator is uniformly bounded from below thanks to (5-13). Unless $j=2$ in (5-8) and (5-9), we can integrate by parts if necessary to ensure that at most $k-1$ derivatives in x hit each factor of j. Then similarly to (5-15) it follows that for $k\ge 5$,

$$|((5\text{-}8),\ j\ge 3)+((5\text{-}9),\ j\ge 3)+(5\text{-}10)|\lesssim_k E_k(t)^2. \tag{5-17}$$

For $j=2$, by symmetry of D_ϵ it is clear that

$$((5\text{-}9),\ j{=}2)=((5\text{-}8),\ j{=}2), \tag{5-18}$$

which according to (4-11) equals

$$D_\epsilon[\partial_x^k(\Phi_\epsilon \mathfrak{h}_x(t)+N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)]+R(t)),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)].$$

Similarly to (5-15),

$$|D_\epsilon[\partial_x^k R(t),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)]|\lesssim_k E_3(t)^2E_k(t)^{1/2}. \tag{5-19}$$

Similarly to (5-3),

$$|D_\epsilon[\partial_x^k(\Phi_\epsilon\mathfrak{h}_x(t))-\Phi_\epsilon\partial_x^{k+1}\mathfrak{h}(t),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)]|\lesssim_k E_3(t)^{3/2}E_k(t). \tag{5-20}$$

By the definition of D_ϵ,

$$\begin{aligned}&D_\epsilon[\Phi_\epsilon\partial_x^{k+1}\mathfrak{h}_x(t),\ \partial_x^k\mathfrak{h}(t),\ \partial_x^2\mathfrak{h}(t)]\\ &=\sum_{mm'nl\neq 0}\frac{1}{\lambda_{m'}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)}\int_0^{2\pi}\widehat{\Phi}_\epsilon(p)e^{ipx}\hat{\mathfrak{h}}(m+\operatorname{sgn}m,t)\partial_x^{k+1}e^{i(m+\operatorname{sgn}m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn}n,t)\\ &\qquad\qquad\times\partial_x^k e^{i(n+\operatorname{sgn}n)x}\hat{\mathfrak{h}}(l+\operatorname{sgn}l,t)\partial_x^2e^{i(l+\operatorname{sgn}l)x}\,dx,\end{aligned} \tag{5-21}$$

where $m'+\operatorname{sgn}m'=p+m+\operatorname{sgn}m\neq 0,\pm 1$. We break the summation into several parts.

Part 1: $|p|\geq\frac{1}{3}|m+\operatorname{sgn}m|$. Then we can transfer the extra derivative from $\mathfrak{h}$ to Φ_ϵ, and compute as in (5-3) to get

$$|\text{Part 1}|\lesssim_k E_3(t)^{3/2}E_k(t). \tag{5-22}$$

Part 2: $|p|<|m+\operatorname{sgn}m|/3$ but $|p|\geq|n+\operatorname{sgn}n|/3$. If $|n+\operatorname{sgn}n|\geq|m|/3$ then $|p|\geq|m|/9$, and we get the same bound as before. Otherwise, since the integral vanishes unless

$$p+m+n+l+\operatorname{sgn}m+\operatorname{sgn}n+\operatorname{sgn}l=0 \tag{5-23}$$

in which case we have $|l+\operatorname{sgn}l|>|n+\operatorname{sgn}n|/3$, we can transfer the extra derivative to the factor $\partial_x^2\mathfrak{h}$ to get (note that $\|\Phi_\epsilon\|_{C^k}\lesssim_k\|g\|_{H^1}^2/|\epsilon|$)

$$|\text{Part 2}|\lesssim_k(\|g(t)\|_{H^1}^2/|\epsilon|)E_4(t)^{1/2}E_k(t)\lesssim E_4(t)E_k(t) \tag{5-24}$$

provided that $\|g(t)\|_{H^1}/|\epsilon|$ is small enough.

Part 3: $|p|<\frac{1}{3}|m+\operatorname{sgn}m|$ and $|p|<\frac{1}{3}|n+\operatorname{sgn}n|$. Then $\operatorname{sgn}(m'+\operatorname{sgn}m')=\operatorname{sgn}(m+\operatorname{sgn}m)$, i.e., $\operatorname{sgn}m'=\operatorname{sgn}m$, so $m'=m+p$. By symmetry,

$$\begin{aligned}\text{Part 3}=&\sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn}m|/3\\ |p|<|n+\operatorname{sgn}n|/3}}\frac{\frac{1}{2}}{\lambda_{m+p}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)}\int_0^{2\pi}\hat{\mathfrak{h}}(l+\operatorname{sgn}l,t)\partial_x^2e^{i(l+\operatorname{sgn}l)x}\widehat{\Phi}_\epsilon(p)e^{ipx}\hat{\mathfrak{h}}(m+\operatorname{sgn}m,t)\\ &\qquad\qquad\times\partial_x^{k+1}e^{i(m+\operatorname{sgn}m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn}n,t)\partial_x^ke^{i(n+\operatorname{sgn}n)x}\,dx\\ &+\sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn}m|/3\\ |p|<|n+\operatorname{sgn}n|/3}}\frac{\frac{1}{2}}{\lambda_m(\epsilon)+\lambda_{n+p}(\epsilon)+\lambda_l(\epsilon)}\int_0^{2\pi}\hat{\mathfrak{h}}(l+\operatorname{sgn}l,t)\partial_x^2e^{i(l+\operatorname{sgn}l)x}\widehat{\Phi}_\epsilon(p)e^{ipx}\hat{\mathfrak{h}}(m+\operatorname{sgn}m,t)\\ &\qquad\qquad\times\partial_x^ke^{i(m+\operatorname{sgn}m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn}n,t)\partial_x^{k+1}e^{i(n+\operatorname{sgn}n)x}\,dx.\end{aligned}$$

Note that the two denominators are uniformly bounded from below. Also, $\operatorname{sgn}(m+p) = \operatorname{sgn} m$ and $|m+p| > \frac{1}{3}(2|m|-1)$, and similarly for l. Then by Corollary 3.10, the two denominators differ by $O(|m|\epsilon^{4|m|/3+3} + |n|\epsilon^{4|n|/3+3})$, so

$$\begin{aligned}
\text{Part 3} = &\sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn} m|/3\\ |p|<|n+\operatorname{sgn} n|/3}} \frac{\frac{1}{2}}{\lambda_{m+p}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} \hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\partial_x^2 e^{i(l+\operatorname{sgn} l)x}\widehat{\Phi}_\epsilon(p)e^{ipx}\partial_x(\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)\\
&\qquad\qquad \times \partial_x^k e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)\partial_x^k e^{i(n+\operatorname{sgn} n)x})\,dx\\
&+\sum_{(5\text{-}23)} O(|m|\epsilon^{4|m|/3+3}+|n|\epsilon^{4|n|/3+3})\int_0^{2\pi} |(l+\operatorname{sgn} l)^2\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\widehat{\Phi}_\epsilon(p)(m+\operatorname{sgn} m)^k\\
&\qquad\qquad \times\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)(n+\operatorname{sgn} n)^{k+1}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)|\\
= &\sum_{\substack{mnl\neq 0\\ |p|<|m+\operatorname{sgn} m|/3\\ |p|<|n+\operatorname{sgn} n|/3}} \frac{-\frac{1}{2}}{\lambda_{m+p}(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)} \int_0^{2\pi} \partial_x(\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\partial_x^2 e^{i(l+\operatorname{sgn} l)x}\widehat{\Phi}_\epsilon(p)e^{ipx})\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)\\
&\qquad\qquad \times \partial_x^k e^{i(m+\operatorname{sgn} m)x}\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)\partial_x^k e^{i(n+\operatorname{sgn} n)x}\,dx\\
&+\sum_{(5\text{-}23)} O(\epsilon^4)\int_0^{2\pi} |(l+\operatorname{sgn} l)^2\hat{\mathfrak{h}}(l+\operatorname{sgn} l,t)\\
&\qquad\qquad \times\widehat{\Phi}_\epsilon(p)(m+\operatorname{sgn} m)^k\hat{\mathfrak{h}}(m+\operatorname{sgn} m,t)(n+\operatorname{sgn} n)^k\hat{\mathfrak{h}}(n+\operatorname{sgn} n,t)|,
\end{aligned}$$

where we integrated by parts in the first integral and used the bounds $|m|\epsilon^{4|m|/3+3}$ and

$$|n(n+\operatorname{sgn} n)\epsilon^{4|n|/3+3}| \lesssim \epsilon^4$$

in the second. Then as in (5-15) it follows that

$$\begin{aligned}
|\text{Part 3}| &\lesssim_k (\|g(t)\|_{H^1}^2/|\epsilon|)E_4(t)^{1/2}E_k(t) + \epsilon^4(\|g(t)\|_{H^1}^2/|\epsilon|)E_3(t)^{1/2}E_k(t)\\
&\lesssim E_4(t)E_k(t)
\end{aligned} \tag{5-25}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

Combining (5-20), (5-22), (5-24) and (5-25) shows that

$$|D_\epsilon[\partial_x^k(\Phi_\epsilon\mathfrak{h}_x(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]| \lesssim_k E_4(t)(1+E_4(t)^{1/2})E_k(t) \tag{5-26}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

We now turn to $D_\epsilon[\partial_x^k N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]$. Similarly to (5-3),

$$\begin{aligned}
&\big|D_\epsilon[\partial_x^k N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]\\
&-D_\epsilon[(\phi_\epsilon'\circ\phi_\epsilon^{-1})^2\partial_x((1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}(1+W_\epsilon)^{-1}\mathfrak{h}(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]\big| \lesssim_k |\epsilon|E_3(t)^{1/2}E_k(t)^{3/2}.
\end{aligned} \tag{5-27}$$

Since W_ϵ/ϵ is of class $\mathcal{S}$ uniformly in ϵ, so is $((1+W_\epsilon)^{-1}-1)/\epsilon$, so

$$\begin{aligned}
&|D_\epsilon[(\phi_\epsilon'\circ\phi_\epsilon^{-1})^2(\partial_x(1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}((1+W_\epsilon)^{-1}-1)\mathfrak{h}(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]|\\
&\qquad\qquad \lesssim_k |\epsilon|E_3(t)^{3/2}E_k(t)^{1/2}.
\end{aligned} \tag{5-28}$$

Finally, $D_\epsilon[(\phi'_\epsilon \circ \phi_\epsilon^{-1})^2(\partial_x(1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}\mathfrak{h}(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]$ is of the same form as the left-hand side of (5-21), so we trace the same argument to get

$$\begin{aligned}
|\text{Part 1}| &\lesssim_k E_3(t)E_k(t),\\
|\text{Part 2}| &\lesssim_k E_4(t)E_k(t),\\
|\text{Part 3}| &\lesssim_k E_4(t)E_k(t)+\epsilon^4E_3(t)E_k(t) \lesssim E_4(t)E_k(t)
\end{aligned}$$

provided that ϵ is small enough. Hence

$$|D_\epsilon[(\phi'_\epsilon \circ \phi_\epsilon^{-1})^2(\partial_x(1+W_\epsilon)^{-1}\mathfrak{h}(t))(\partial_x^{k+1}\mathfrak{h}(t)),\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]| \lesssim_k E_4(t)E_k(t). \tag{5-29}$$

Combining (5-27), (5-28) and (5-29) shows that, for $k \ge 4$,

$$|D_\epsilon[\partial_x^k N_\epsilon[\mathfrak{h}(t),\mathfrak{h}(t)],\, \partial_x^k\mathfrak{h}(t),\, \partial_x^2\mathfrak{h}(t)]| \lesssim_k E_4(t)^{1/2}E_k(t)^{3/2} \tag{5-30}$$

provided that ϵ is small enough.

Combining (5-19), (5-26) and (5-30) shows that, for $k \ge 4$,

$$|((5\text{-}8),\, j=2)| \lesssim_k E_4(t)^{1/2}(1+E_4(t)^{1/2})E_k(t)^{3/2} \tag{5-31}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

Finally, combining (5-16), (5-17), (5-18) and (5-31) shows that, for $k \ge 5$,

$$\left|\frac{d}{dt}\sum_{j=2}^{[k/2]+1} c_{jk}D_{1,k,j}(t) - E_{N1}(t)\right| \lesssim_k (1+E_4(t)^{1/2})E_k(t)^2 \tag{5-32}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

5B. ***Lifespan when $\delta \ll \epsilon$.*** In this section we will obtain a preliminary bound for $E_{N2} = E_N - E_{N1}$ and show a lifespan of $1/(\epsilon\delta)$ when $\|g_0\|_{H^5(\mathbb{T})} = \delta \ll \epsilon$, i.e., $\delta \le c\epsilon$ for some $c > 0$ independent of ϵ.

Recall from (5-6) that

$$E_N(t) = \frac{1}{2}\int_0^{2\pi} \partial_x^k\mathfrak{h}(t)\partial_x^k((\phi'_\epsilon \circ \phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}(t))_x^2)\,dx.$$

Similarly to (5-3), for $k \ge 3$,

$$\left|E_N(t) - \int_0^{2\pi} (\phi'_\epsilon \circ \phi_\epsilon^{-1})^2\partial_x^k\mathfrak{h}(t)\partial_x^k(((1+W_\epsilon)^{-1}\mathfrak{h}_x(t))^2)\,dx\right| \lesssim_k |\epsilon|E_k(t)^{3/2}.$$

Since $((1+W_\epsilon)^{-1}-1)/\epsilon$ is of class $\mathcal{S}$ uniformly in ϵ,

$$\left|\int_0^{2\pi} (\phi'_\epsilon \circ \phi_\epsilon^{-1})^2\partial_x^k\mathfrak{h}(t)\partial_x^k(((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))^2)\,dx\right| \lesssim_k \epsilon^2E_k(t)^{3/2},$$

$$\begin{aligned}
2\bigg|\int_0^{2\pi} &(\phi'_\epsilon \circ \phi_\epsilon^{-1})^2\partial_x^k\mathfrak{h}(t)\partial_x^k(((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))\mathfrak{h}_x(t))\,dx\\
&-\int_0^{2\pi} (\phi'_\epsilon \circ \phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))\partial_x^k\mathfrak{h}(t)\partial_x^{k+1}\mathfrak{h}(t)\,dx\bigg| \lesssim_k |\epsilon|E_k(t)^{3/2}.
\end{aligned}$$

Finally, by integration by parts,

$$\left|2\int_0^{2\pi}(\phi_\epsilon'\circ\phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t))\partial_x^k\mathfrak{h}(t)\partial_x^{k+1}\mathfrak{h}(t)\,dx\right| \\ =\left|\int_0^{2\pi}\partial_x((\phi_\epsilon'\circ\phi_\epsilon^{-1})^2((1+W_\epsilon)^{-1}\mathfrak{h}_x(t)-\mathfrak{h}_x(t)))(\partial_x^k\mathfrak{h}(t))^2\,dx\right|\lesssim_k|\epsilon|E_k(t)^{3/2}.$$

Combining the bounds above shows that, for $k\geq 3$,

$$|E_{N_2}(t)|=|E_N(t)-E_{N_1}(t)|\lesssim_k|\epsilon|E_k(t)^{3/2} \tag{5-33}$$

provided that ϵ is small enough.

Now combining (5-5), (5-32) and (5-33) shows that, for $k\geq 5$,

$$\frac{d}{dt}\left|\sum_{j=2}^{[k/2]+1}c_{jk}D_{1,k,j}(t)-E_k(t)\right|\lesssim_k(1+E_4(t)^{1/2})E_k(t)^2+|\epsilon|E_k(t)^{3/2}. \tag{5-34}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough. Hence

$$E_k(t)-E_k(0)=\sum_{j=2}^{[k/2]+1}c_{jk}(D_{1,k,j}(t)-D_{1,k,j}(0))+O_k(\|(1+E_4^{1/2})E_k^2+|\epsilon|E_k^{3/2}\|_{L^1([0,t])}).$$

Similarly to (5-15),

$$|D_{1,k,j}(t)|=|D_{\epsilon(t)}[\partial_x^k\mathfrak{h}(t),\partial_x^{k+2-j}\mathfrak{h}(t),\partial_x^j\mathfrak{h}(t)]|\lesssim_k E_k(t)^{3/2}.$$

Now we are able to show a lifespan longer than what follows from local well-posedness. Assume that the initial data is

$$f(x,0)=u_\epsilon(x)+g(x),$$

where $|\epsilon|\leq\epsilon_0$ is small enough, the energy $E_k(0)$ computed from g is $E_k(0)=\delta^2$, and $|\delta/\epsilon|$ is also small enough. Let

$$T^*=\sup\bigl\{T:\text{there exists a solution } f(x,t)=u_{\epsilon(t)}(x+a(t))+g(x+a(t),t), \tag{5-35}$$

$$t\in[0,T]\text{ such that }\tfrac{1}{2}|\epsilon|\leq|\epsilon(t)|\leq 2|\epsilon|,\ E_k(t)\leq 4\delta^2\bigr\}. \tag{5-36}$$

Then the above conditions hold for all $t<T^*$. Moreover, the energy estimate implies

$$E_k(t)=\delta^2+O_k(\delta^3+t(\delta^4+|\epsilon|\delta^3))=\delta^2+O_k(\delta^3(1+t|\epsilon|)).$$

Then there is $c_k>0$ such that if $T^*\leq c_k/|\epsilon|\delta$, then $E_k(t)\leq 2\delta^2$. Also,

$$|\|f(x,t)\|_{L^2}-\|u_\epsilon\|_{L^2}|=|\|f(x,0)\|_{L^2}-\|u_\epsilon\|_{L^2}|\leq\|g\|_{L^2}\lesssim\delta$$

by conservation of the L^2 norm. Meanwhile $|\|f(x,t)\|_{L^2}-\|u_{\epsilon(t)}\|_{L^2}|\lesssim\delta$, so $|\|u_{\epsilon(t)}\|_{L^2}-\|u_\epsilon\|_{L^2}|\lesssim\delta$. When $|\epsilon|$ is small enough, $\|u_\epsilon\|_{L^2}$ is differentiable in ϵ with nonzero derivative at $\epsilon=0$. Since $|\delta/\epsilon|$ is small enough, $|\epsilon(t)-\epsilon|\lesssim\delta$.

By local well-posedness, the solution can be extended to a time $t^* > T^*$, with

$$\|f(x,t) - f(x,T^*)\|_{H^2} \lesssim (t^* - T^*)(\|f(x,t)\|_{H^3} + \|f(x,t)\|_{H^3}^2) \le (t^* - T^*)|\epsilon|$$

for $t \in [T^*, t^*]$. Then $\|f(x,t) - u_{\epsilon(T^*)}(x + a(T^*))\|_{H^2} \lesssim (t^* - T^*)|\epsilon| + \delta$. Take $t^* = T^* + \delta/|\epsilon|$. Then $f(x,t)$ satisfies the conditions in Proposition 4.1, so (5-35) holds up to time t^*. Since $f(x, T^*)$ is small in H^4, $f(x,t)$ is uniformly bounded in H^4 on $[T^*, t^*]$, so it stays within a compact set in H^2. Since ϵ is differentiable in $f \in H^2$, $|\epsilon(t) - \epsilon(T^*)| \lesssim (t^* - T^*)|\epsilon| \lesssim \delta$, so $|\epsilon(t) - \epsilon| \lesssim \delta$, so $|\epsilon|/2 \le |\epsilon(t)| \le 2|\epsilon|$ holds up to time t^*. The energy estimate then implies $E_k \le 3\delta^2$ also up to time t^*, so (5-36) holds up to time t^*, contradicting the definition of T^*. Hence the lifespan $T^* \gtrsim_k 1/|\epsilon|\delta$.

5C. ***Longer lifespan when $\delta \ll \epsilon^2$.*** When the perturbation g is very small compared to ϵ^2, that is, $\|g_0\|_{H^5(\mathbb{T})} = \delta \ll \epsilon^2$, we can obtain a longer lifespan by applying the normal form transformation to

$$E_{N2} = E_N - E_{N1} = E_{N21} + E_{N22} + E_{N23} + E_{N24},$$

where

$$E_{N21} = \sum_{j=1}^{[k/2]+1} c'_{kj} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t) \partial_x^{k+2-j}((\phi'_\epsilon \circ \phi_\epsilon^{-1})((1+W_\epsilon)^{-1} - 1)\mathfrak{h}(t)) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h}(t))\, dx,$$

$$E_{N22} = \sum_{j=1}^{[k/2]+1} \sum_{i=1}^{k+2-j} c_{kji} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t) \partial_x^i(\phi'_\epsilon \circ \phi_\epsilon^{-1}) \partial_x^{k+2-i-j} \mathfrak{h}(t) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h}(t))\, dx,$$

$$E_{N23} = \sum_{j=2}^{[k/2]+1} c_{kj} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t)(\phi'_\epsilon \circ \phi_\epsilon^{-1} - 1) \partial_x^{k+2-j} \mathfrak{h}(t) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h}(t))\, dx,$$

$$E_{N24} = \sum_{j=2}^{[k/2]+1} c_{kj} \int_0^{2\pi} \partial_x^k \mathfrak{h}(t) \partial_x^{k+2-j} \mathfrak{h}(t) \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\mathfrak{h} - \mathfrak{h})\, dx,$$

where c_{kj}, c'_{kj} and $c_{kji} \in \mathbb{R}$ are constants and we integrated by parts to get rid of the terms with $k+1$ derivatives falling on a single factor of $\mathfrak{h}$, except for the term with $j = 1$ in E_{N21}, in which the $k+1$ derivatives do not matter in view of the fact that the operator $(\phi'_\epsilon \circ \phi_\epsilon^{-1})((1+W_\epsilon)^{-1} - 1)$ is of class $\mathcal{S}$.

Now we define

$$\begin{aligned} D_{\epsilon,21}[f_1, f_2, f_3] &= \sum_{\substack{j=1 \\ mnl\neq 0}}^{[k/2]+1} \frac{c'_{kj}}{\lambda_m(\epsilon) + \lambda_n(\epsilon) + \lambda_l(\epsilon)} \int_0^{2\pi} \hat{f}_1(m + \operatorname{sgn} m) e^{i(m+\operatorname{sgn} m)x} \\ &\qquad \times \partial_x^{k+2-j}((\phi'_\epsilon \circ \phi_\epsilon^{-1})((1+W_\epsilon)^{-1} - 1)\hat{f}_2(n + \operatorname{sgn} n) e^{i(n+\operatorname{sgn} n)x}) \\ &\qquad \times \partial_x^j((\phi'_\epsilon \circ \phi_\epsilon^{-1})(1+W_\epsilon)^{-1}\hat{f}_3(l + \operatorname{sgn} l) e^{i(l+\operatorname{sgn} l)x}), \end{aligned}$$

and

$$D_{21}(t) = D_{\epsilon,21}[\mathfrak{h}(t), \mathfrak{h}(t), \mathfrak{h}(t)],$$

and similarly define D_{22}, D_{23} and D_{24}. Then

$$\frac{d}{dt}D_{21}(t)-E_{N21}(t)=\epsilon'(t)(\partial_\epsilon D_{\epsilon,21})[\mathfrak{h}(t),\mathfrak{h}(t),\mathfrak{h}(t)] \tag{5-37}$$

$$+D_{\epsilon,21}[(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t),\mathfrak{h}(t),\mathfrak{h}(t)] \tag{5-38}$$

$$+D_{\epsilon,21}[\mathfrak{h}(t),(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t),\mathfrak{h}(t)] \tag{5-39}$$

$$+D_{\epsilon,21}[\mathfrak{h}(t),\mathfrak{h}(t),(\partial_t-\Lambda_\epsilon)\mathfrak{h}(t)]. \tag{5-40}$$

We estimate these terms one by one.

For (5-37), (5-11) still holds, but there are nontrivial actions on $\mathfrak{h}$ in the slots, so no frequency restriction such as (5-12) exists. When $m+n+l\neq 0$, we are in Case 1 of Proposition 3.11, so (5-13), and hence (5-14), still hold. When $m+n+l=0$, by Case 2 of Proposition 3.11, when ϵ is small enough,

$$|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|>\tfrac{1}{5}\epsilon^2, \tag{5-41}$$

which, combined with (5-11), shows that the multiplier in $\partial_\epsilon D_\epsilon$ is bounded by

$$\frac{|\lambda_m'(\epsilon)+\lambda_n'(\epsilon)+\lambda_l'(\epsilon)|}{|\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)|^2}\lesssim|\epsilon|^{-3} \tag{5-42}$$

instead of (5-14). Since both $(\phi_\epsilon'\circ\phi_\epsilon^{-1})((1+W_\epsilon)^{-1}-1)/\epsilon$ and $\partial_\epsilon(\phi_\epsilon'\circ\phi_\epsilon^{-1})((1+W_\epsilon)^{-1}-1)$ are of class $\mathcal{S}$ uniformly in ϵ, it follows that, for $k\geq 3$,

$$|(5\text{-}37)|\lesssim_k|\epsilon'(t)|\epsilon^{-2}E_k(t)^{3/2}\lesssim\epsilon^{-2}E_2(t)E_k(t)^{3/2} \tag{5-43}$$

provided that ϵ is small enough.

The terms (5-38), (5-39) and (5-40) are like (5-8), (5-9) and (5-10) respectively, except that instead of the uniform lower bound of $\lambda_m(\epsilon)+\lambda_n(\epsilon)+\lambda_l(\epsilon)$ we now have (5-41), which loses two factors of ϵ, but we are helped by the ϵ-smallness of $(\phi_\epsilon'\circ\phi_\epsilon^{-1})((1+W_\epsilon)^{-1}-1)$, which wins back a factor of ϵ. All told we lose a factor of ϵ compared to (5-32), so, for $k\geq 5$,

$$|(5\text{-}38)+(5\text{-}39)+(5\text{-}40)|\lesssim_k|\epsilon|^{-1}(1+E_4(t)^{1/2})E_k(t)^2 \tag{5-44}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough.

Combining (5-43) and (5-44) shows that, for $k\geq 5$,

$$\left|\frac{d}{dt}D_{21}(t)-E_{N21}(t)\right|\lesssim_k|\epsilon|^{-1}(1+E_4(t)^{1/2})E_k(t)^2 \tag{5-45}$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough. We can also save a factor of ϵ in the other terms E_{N22}, E_{N23} and E_{N24} thanks to the ϵ-smallness of $(\phi_\epsilon'\circ\phi_\epsilon^{-1})'$ and $\phi_\epsilon'\circ\phi_\epsilon^{-1}-1$. Hence the bound (5-45) also holds for E_{N22}, E_{N23} and E_{N24}.

Combining (5-5), (5-32) and (5-45) shows that, for $k\geq 5$,

$$E_k(t)-E_k(0)=\sum_{j=2}^{[k/2]+1}c_{jk}(D_{1,k,j}(t)-D_{1,k,j}(0))+\sum_{j=1}^{4}(D_{2j}(t)-D_{2j}(0)) \\ +O_k(|\epsilon|^{-1}\|(1+E_4^{1/2})E_k^2\|_{L^1([0,t])})$$

provided that ϵ and $\|g(t)\|_{H^1}/|\epsilon|$ are small enough. Similarly to (5-33), for $k \geq 3$,

$$|D_{2,k,j}(t)| \lesssim_k \epsilon(t)^{-2}|\epsilon(t)|E_k(t)^{3/2} = E_k(t)^{3/2}/|\epsilon|.$$

Hence if $E_k(0) = \delta^2 \lesssim 1$ and $E_k \leq 2\delta^2$ on $[0, t]$ then

$$E_k(t) = \delta^2 + |\epsilon|^{-1}\delta^3 + O_k(t|\epsilon|^{-1}\delta^4).$$

Assume δ/ϵ^2 is small. Then the second term on the right-hand side is $\lesssim \delta^{5/2}$, so we close the estimate for a time $t \lesssim_k |\epsilon|/\delta^2$, which is also the lifespan in this case.

Acknowledgements

This work is supported in part by the Spanish Ministry of Science and Innovation, through the "Severo Ochoa Programme for Centres of Excellence in R&D" (CEX2019-000904-S) and MTM2017-89976-P. Castro was partially supported by the Europa Excelencia program ERC2018-092824. Córdoba and Zheng were partially supported by the ERC Advanced Grant 788250.

References

[Agmon 1975] S. Agmon, "Spectral properties of Schrödinger operators and scattering theory", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **2**:2 (1975), 151–218. MR Zbl

[Arnold 1983] V. I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, Grundlehren der Math. Wissenschaften **250**, Springer, 1983. MR Zbl

[Biello and Hunter 2010] J. Biello and J. K. Hunter, "Nonlinear Hamiltonian waves with constant frequency and surface waves on vorticity discontinuities", *Comm. Pure Appl. Math.* **63**:3 (2010), 303–336. MR Zbl

[Bressan and Nguyen 2014] A. Bressan and K. T. Nguyen, "Global existence of weak solutions for the Burgers–Hilbert equation", *SIAM J. Math. Anal.* **46**:4 (2014), 2884–2904. MR Zbl

[Bressan and Zhang 2017] A. Bressan and T. Zhang, "Piecewise smooth solutions to the Burgers–Hilbert equation", *Commun. Math. Sci.* **15**:1 (2017), 165–184. MR Zbl

[Castro et al. 2010] A. Castro, D. Córdoba, and F. Gancedo, "Singularity formations for a surface wave model", *Nonlinearity* **23**:11 (2010), 2835–2847. MR Zbl

[Chen and Pusateri 2022] G. Chen and F. Pusateri, "The 1-dimensional nonlinear Schrödinger equation with a weighted L^1 potential", *Anal. PDE* **15**:4 (2022), 937–982. MR Zbl

[Crandall and Rabinowitz 1971] M. G. Crandall and P. H. Rabinowitz, "Bifurcation from simple eigenvalues", *J. Funct. Anal.* **8** (1971), 321–340. MR Zbl

[Ehrnström and Wang 2019] M. Ehrnström and Y. Wang, "Enhanced existence time of solutions to the fractional Korteweg–de Vries equation", *SIAM J. Math. Anal.* **51**:4 (2019), 3298–3323. MR Zbl

[Germain et al. 2012] P. Germain, N. Masmoudi, and J. Shatah, "Global solutions for the gravity water waves equation in dimension 3", *Ann. of Math.* (2) **175**:2 (2012), 691–754. MR Zbl

[Germain et al. 2015] P. Germain, N. Masmoudi, and J. Shatah, "Global existence for capillary water waves", *Comm. Pure Appl. Math.* **68**:4 (2015), 625–687. MR Zbl

[Germain et al. 2018] P. Germain, F. Pusateri, and F. Rousset, "The nonlinear Schrödinger equation with a potential", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **35**:6 (2018), 1477–1530. MR Zbl

[Hunter 2018] J. K. Hunter, "The Burgers–Hilbert equation", pp. 41–57 in *Theory, numerics and applications of hyperbolic problems, II* (Aachen, Germany, 2016), edited by C. Klingenberg and M. Westdickenberg, Springer Proc. Math. Stat. **237**, Springer, 2018. MR Zbl

[Hunter and Ifrim 2012] J. K. Hunter and M. Ifrim, "Enhanced life span of smooth solutions of a Burgers–Hilbert equation", *SIAM J. Math. Anal.* **44**:3 (2012), 2039–2052. MR Zbl

[Hunter et al. 2015] J. K. Hunter, M. Ifrim, D. Tataru, and T. K. Wong, "Long time solutions for a Burgers–Hilbert equation via a modified energy method", *Proc. Amer. Math. Soc.* **143**:8 (2015), 3407–3412. MR Zbl

[Hunter et al. 2022] J. K. Hunter, R. C. Moreno-Vasquez, J. Shu, and Q. Zhang, "On the approximation of vorticity fronts by the Burgers–Hilbert equation", *Asymptot. Anal.* **129**:2 (2022), 141–177. MR Zbl

[Ionescu and Pausader 2022] A. D. Ionescu and B. Pausader, *The Einstein–Klein–Gordon coupled system: global stability of the Minkowski solution*, Ann. of Math. Stud. **213**, Princeton Univ. Press, 2022. MR Zbl

[Ionescu and Pusateri 2015] A. D. Ionescu and F. Pusateri, "Global solutions for the gravity water waves system in 2D", *Invent. Math.* **199**:3 (2015), 653–804. MR Zbl

[Ionescu and Pusateri 2018] A. D. Ionescu and F. Pusateri, *Global regularity for 2D water waves with surface tension*, Mem. Amer. Math. Soc. **1227**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Kato 1976] T. Kato, *Perturbation theory for linear operators*, 2nd ed., Grundlehren der Math. Wissenschaften **132**, Springer, 1976. MR Zbl

[Klein and Saut 2015] C. Klein and J.-C. Saut, "A numerical approach to blow-up issues for dispersive perturbations of Burgers' equation", *Phys. D* **295/296** (2015), 46–65. MR Zbl

[Krupa and Vasseur 2020] S. G. Krupa and A. F. Vasseur, "Stability and uniqueness for piecewise smooth solutions to a nonlocal scalar conservation law with applications to Burgers–Hilbert equation", *SIAM J. Math. Anal.* **52**:3 (2020), 2491–2530. MR Zbl

[Marsden and Weinstein 1983] J. Marsden and A. Weinstein, "Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids", *Phys. D* **7**:1-3 (1983), 305–323. MR Zbl

[Saut and Wang 2022] J.-C. Saut and Y. Wang, "The wave breaking for Whitham-type equations revisited", *SIAM J. Math. Anal.* **54**:2 (2022), 2295–2319. MR Zbl

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

[Yang 2021] R. Yang, "Shock formation of the Burgers–Hilbert equation", *SIAM J. Math. Anal.* **53**:5 (2021), 5756–5802. MR Zbl

Received 3 Mar 2021. Revised 22 Mar 2022. Accepted 29 Apr 2022.

ÁNGEL CASTRO: angel_castro@icmat.es
Instituto de Ciencias Matemáticas ICMAT-CSIC-UAM-UCM-UC3M, Madrid, Spain

DIEGO CÓRDOBA: dcg@icmat.es
Instituto de Ciencias Matemáticas ICMAT-CSIC-UAM-UCM-UC3M, Madrid, Spain

FAN ZHENG: fan.zheng@icmat.es
Instituto de Ciencias Matemáticas ICMAT-CSIC-UAM-UCM-UC3M, Madrid, Spain

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2147–2171
DOI: 10.2140/apde.2023.16.2147

DEFINING THE SPECTRAL POSITION OF A NEUMANN DOMAIN

RAM BAND, GRAHAM COX AND SEBASTIAN K. EGGER

A Laplacian eigenfunction on a two-dimensional Riemannian manifold provides a natural partition into Neumann domains, a.k.a. a Morse–Smale complex. This partition is generated by gradient flow lines of the eigenfunction, which bound the so-called Neumann domains. We prove that the Neumann Laplacian defined on a Neumann domain is self-adjoint and has a purely discrete spectrum. In addition, we prove that the restriction of an eigenfunction to any one of its Neumann domains is an eigenfunction of the Neumann Laplacian. By comparison, similar statements about the Dirichlet Laplacian on a nodal domain of an eigenfunction are basic and well-known. The difficulty here is that the boundary of a Neumann domain may have cusps and cracks, so standard results about Sobolev spaces are not available. Another very useful common fact is that the restricted eigenfunction on a nodal domain is the first eigenfunction of the Dirichlet Laplacian. This is no longer true for a Neumann domain. Our results enable the investigation of the resulting spectral position problem for Neumann domains, which is much more involved than its nodal analogue.

1. Introduction and statement of results

Let M be a closed, connected, orientable surface with a smooth Riemannian metric g. It is well known that the Laplace–Beltrami operator Δ is self-adjoint and has a purely discrete spectrum. We arrange the eigenvalues in increasing order

$$0 = \lambda_0 < \lambda_1 \le \lambda_2 \le \cdots \tag{1-1}$$

and let $\{f_n\}_{n=0}^{\infty}$ denote a corresponding complete system of orthonormal eigenfunctions, so that

$$\Delta f_n = \lambda_n f_n. \tag{1-2}$$

While we are motivated by the study of eigenfunctions, most of the results and constructions in this paper are valid for arbitrary Morse functions. It is well known that for a generic Riemannian metric all of the Laplace–Beltrami eigenfunctions are Morse [Uhlenbeck 1976].

The main objects of study in this paper are the Neumann domains of a Morse function, to be defined next. Given a smooth function f on M, we let $\varphi : \mathbb{R} \times M \to M$ denote the flow along the gradient vector field, i.e., the solution to

$$\partial_t \varphi(t,x) = -\operatorname{grad} f|_{\varphi(t,x)}, \quad \varphi(0,x) = x. \tag{1-3}$$

For a critical point $\boldsymbol{c}$ of f, we define its stable and unstable manifolds by

$$W^s(\boldsymbol{c}) := \bigl\{x \in M : \lim_{t\to\infty} \varphi(t,x) = \boldsymbol{c}\bigr\}, \quad W^u(\boldsymbol{c}) := \bigl\{x \in M : \lim_{t\to-\infty} \varphi(t,x) = \boldsymbol{c}\bigr\}. \tag{1-4}$$

We denote the sets of minima, maxima and saddles of f by $\mathrm{Min}(f)$, $\mathrm{Max}(f)$ and $\mathrm{Sad}(f)$, respectively.

MSC2020: primary 35P05, 58C40, 58J50; secondary 37C15, 57K20.
Keywords: Neumann domains, Neumann lines, nodal domains, Laplacian eigenfunctions, Morse–Smale complexes.

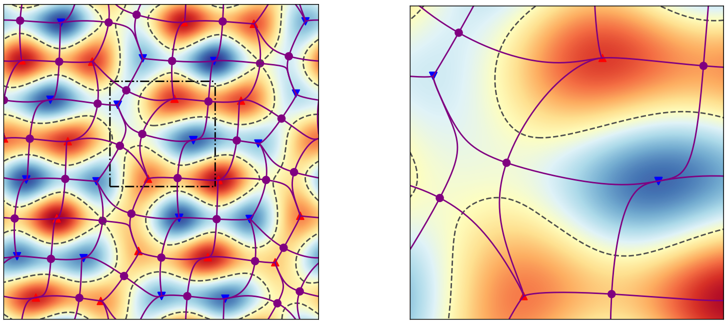

Figure 1. Left: An eigenfunction corresponding to the eigenvalue $\lambda = 17$ on the flat torus with fundamental domain $[0, 2\pi] \times [0, 2\pi]$. Circles mark saddle points and triangles mark extremal points (maxima by triangles pointing upwards and vice versa for minima). The nodal set is marked by dashed lines and the Neumann line set by solid lines. The Neumann domains are the domains bounded by the Neumann line set. Right: A magnification of the marked square from the left figure, showing Neumann domains with and without cusps. (This figure was produced using [Taylor 2018].)

Definition 1.1 [Band and Fajman 2016]. Let f be a Morse function on M.

(1) Let $\boldsymbol{p} \in \mathrm{Min}(f)$ and $\boldsymbol{q} \in \mathrm{Max}(f)$ such that $W^s(\boldsymbol{p}) \cap W^u(\boldsymbol{q}) \neq \varnothing$. Each of the connected components of $W^s(\boldsymbol{p}) \cap W^u(\boldsymbol{q})$ is called a *Neumann domain* of f.

(2) The *Neumann line set* of f is

$$\mathcal{N} := \overline{\bigcup_{\boldsymbol{r} \in \mathrm{Sad}(f)} W^s(\boldsymbol{r}) \cup W^u(\boldsymbol{r})}. \tag{1-5}$$

This defines a partition of the manifold M, which we call the Neumann partition. It is not hard to show that M equals the disjoint union of all Neumann domains and the Neumann line set, under the assumption that $\mathcal{N} \neq \varnothing$; see [Band and Fajman 2016, Proposition 1.3]. (Note that $\mathcal{N} = \varnothing$ means f has no saddle points; this is only possible when M is a sphere and f has exactly two critical points.) Figure 1 depicts the Neumann partition of a particular eigenfunction on the flat torus.

By construction we have that $\operatorname{grad} f$ is parallel to the boundary of any Neumann domain Ω, as the boundary is made up of gradient flow lines, so we conclude that the normal derivative vanishes, $\partial_\nu f|_{\partial\Omega} = 0$, assuming $\partial\Omega$ is sufficiently smooth. This formal observation motivates our study of the Neumann Laplacian on Ω, which we precisely define in Definition 4.1.

While the Dirichlet Laplacian on any bounded open set has a purely discrete spectrum, the same is not necessarily true of the Neumann Laplacian. Indeed, the essential spectrum may be nonempty, and in fact can be an arbitrary closed subset of $[0, \infty)$; see [Hempel et al. 1991]. Nevertheless, the Neumann Laplacian of a Neumann domain is well-behaved.

Theorem 1.2. *Let Ω be a Neumann domain of a Morse function f. Then the Neumann Laplacian Δ_Ω^N on Ω (see Definition 4.1) is a nonnegative, self-adjoint operator with purely discrete spectrum, i.e., consisting only of isolated eigenvalues of finite multiplicity.*

The main difficulty in proving this theorem is due to possible cusps on the boundary of the Neumann domain; see Proposition 2.5 and the discussion preceding it. Such cusps prevent the application of standard results on density and compact embeddings of Sobolev spaces. We overcome this difficulty in the proof of Theorem 1.2 by using some particular geometric properties that the Neumann domain boundary possesses.

It is well known that the restriction of f to any of its nodal domains is an eigenfunction of the Dirichlet Laplacian. Similarly, we have:

Theorem 1.3. *If Ω is a Neumann domain of a Morse function f, then $f|_\Omega \in \mathcal{D}(\Delta_\Omega^N)$. In particular, if f is an eigenfunction of Δ, then $f|_\Omega$ is an eigenfunction of Δ_Ω^N with the same eigenvalue.*

In fact, we prove much more: in Proposition 4.3 we completely characterize the domain of the Neumann Laplacian, and in Proposition 4.8 and Corollary 4.9 we give some easily verified sufficient conditions for a function to be in the domain of Δ_Ω^N. Given a Morse eigenfunction, by which we mean an eigenfunction of the Laplace–Beltrami operator that is also a Morse function, Theorem 1.2 allows us to define its spectral position as follows.

Definition 1.4. Let f be a Morse eigenfunction for an eigenvalue λ, and let Ω be a Neumann domain of f. We define the *spectral position* of Ω as the position of λ in the Neumann spectrum of Ω, i.e.,

$$N_\Omega(\lambda) := |\{\mu_n \in \sigma(\Delta_\Omega^N) : \mu_n < \lambda\}|, \tag{1-6}$$

where $\sigma(\Omega) := \{\mu_n\}_{n=0}^\infty$ is the Neumann spectrum of Ω (which is discrete by Theorem 1.2), containing multiple appearances of degenerate eigenvalues and including $\mu_0 = 0$.

From Theorem 1.3 we in fact have $\lambda = \mu_n$ for some n, and so we can equivalently write

$$N_\Omega(\lambda) = \min\{n : \mu_n = \lambda\}.$$

In particular, if $\lambda \in \sigma(\Delta_\Omega^N)$ is simple, then $\lambda = \mu_n$ for a unique n, and hence $N_\Omega(\lambda) = n$. This equality explains the terminology "spectral position" for $N_\Omega(\lambda)$.

The spectral position is a key notion for Neumann domains. Finding its value is a great challenge and is of major importance in studying Neumann domains and their properties [Band and Fajman 2016; Band et al. 2021; Alon et al. 2020]. The corresponding notion for a *nodal* domain is trivial: if D is a nodal domain of f, then $f|_D$ is always the first eigenfunction of the Dirichlet Laplacian on D. This is a basic observation which serves as an essential ingredient in many nodal domain proofs. No such result holds for Neumann domains, and in fact the spectral position of an eigenfunction restricted to a Neumann domain can be arbitrarily high, by [Band et al. 2021, Theorem 1.4].

Structure of the paper. In Section 2 we describe some essential geometric properties of Neumann domains, emphasizing the potentially singular nature of their boundary. In Section 3 we use this geometric structure to establish fundamental properties of Sobolev spaces on Neumann domains, including nonstandard density and compactness results. Finally, in Section 4 we use these properties to study the Neumann Laplacian, in particular proving Theorems 1.2 and 1.3.

2. Geometric properties of Neumann domains

As above, we take M to be a closed, connected, orientable surface with a smooth Riemannian metric g. Note that all of the statements in this section hold for arbitrary Morse functions, and not only for eigenfunctions. For convenience we recall the following definitions.

Definition 2.1. Let $f : M \to \mathbb{R}$ be a smooth function.

(1) f is said to be a *Morse function* if the Hessian, Hess $f(\boldsymbol{p})$, is nondegenerate at every critical point $\boldsymbol{p}$ of f.

(2) A Morse function f is said to be *Morse–Smale* if for all critical points $\boldsymbol{p}$ and $\boldsymbol{q}$, the stable and unstable manifolds $W^s(\boldsymbol{p})$ and $W^u(\boldsymbol{q})$ intersect transversely (see Lemma 2.4 for an equivalent definition in two dimensions).

We now recall some basic topological properties of Neumann domains.

Theorem 2.2 [Band and Fajman 2016, Theorem 1.4]. *Let f be a Morse function with a nonempty set of saddle points. Let $\boldsymbol{p} \in \mathrm{Min}(f)$, $\boldsymbol{q} \in \mathrm{Max}(f)$ with $W^s(\boldsymbol{p}) \cap W^u(\boldsymbol{q}) \neq \varnothing$, and let Ω be a connected component of $W^s(\boldsymbol{p}) \cap W^u(\boldsymbol{q})$, i.e., a Neumann domain. The following properties hold*:

(1) *The Neumann domain Ω is a simply connected open set.*

(2) *All critical points of f belong to the Neumann line set.*

(3) *The extremal points of f on $\overline{\Omega}$ are exactly $\boldsymbol{p}$ and $\boldsymbol{q}$.*

(4) *If f is a Morse–Smale function, then $\partial\Omega$ consists of Neumann lines connecting saddle points with $\boldsymbol{p}$ or $\boldsymbol{q}$. In particular, $\partial\Omega$ contains either one or two saddle points.*

(5) *If $c \in \mathbb{R}$ is such that $f(\boldsymbol{p}) < c < f(\boldsymbol{q})$, then $\overline{\Omega} \cap f^{-1}(c)$ is a smooth, non-self-intersecting one-dimensional curve in $\overline{\Omega}$, with its two boundary points lying on $\partial\Omega$.*

Parts (2) and (4) of this theorem motivate us to examine individual Neumann lines and their connectivity to the critical points of f.

Definition 2.3. (1) A *Neumann line* is the closure of a connected component of $W^s(\boldsymbol{r})\setminus\{\boldsymbol{r}\}$ or $W^u(\boldsymbol{r})\setminus\{\boldsymbol{r}\}$ for some $\boldsymbol{r} \in \mathrm{Sad}(f)$.

(2) For a critical point $\boldsymbol{c}$ of f, we define its degree, $\deg(\boldsymbol{c})$, to be the number of Neumann lines connected to $\boldsymbol{c}$.

Each Neumann line is thus the closure of a gradient flow line, connecting a saddle point to another critical point. Note that $\boldsymbol{r}$ is removed prior to taking the closure, as the closure of either $W^s(\boldsymbol{r})$ or $W^u(\boldsymbol{r})$ will consist of two Neumann lines meeting tangentially at $\boldsymbol{r}$. The connectivity of Neumann lines is directly related to the Morse–Smale property of f.

Lemma 2.4 [Alon et al. 2020]. *On a two-dimensional manifold a Morse function is Morse–Smale if and only if there is no Neumann line connecting two saddle points.*

The following properties of Neumann lines will be used throughout the rest of the paper.

Proposition 2.5. *Let f be a Morse function and Ω one of its Neumann domains.*

(1) *If $\boldsymbol{c}$ is a saddle point of f, then* $\deg(\boldsymbol{c}) = 4$ *and the angle between any two adjacent Neumann lines which meet at $\boldsymbol{c}$ is $\frac{\pi}{2}$.*

(2) *If $\boldsymbol{c}$ is an extremal point of f whose Hessian is not proportional to the metric g, then any two Neumann lines meet at $\boldsymbol{c}$ with angle* 0, $\frac{\pi}{2}$ *or* π.

(3) *Let $\boldsymbol{c}$ be an intersection point of a nodal line and a Neumann line. If $\boldsymbol{c}$ is a saddle point, then the angle between those lines is $\frac{\pi}{4}$. Otherwise, the angle is $\frac{\pi}{2}$.*

Remark 2.6. More generally, if $\boldsymbol{c}$ is a saddle point and there exist coordinates (x, y) near $\boldsymbol{c}$ in which f is given by the homogeneous harmonic polynomial $\mathrm{Re}(x+iy)^k$, then $\deg(\boldsymbol{c}) = 2k$. For a nondegenerate saddle the existence of such coordinates (with $k = 2$) is an immediate consequence of the Morse lemma, so we obtain Proposition 2.5(1) as a special case of this remark. Sufficient conditions for f to be written in this form are given in [Cheng 1976, Lemma 2.4].

The first and third parts of Proposition 2.5 were proved in [McDonald and Fulling 2014; Banyaga and Hurtubise 2004, Theorem 4.2; Alon et al. 2020, Proposition 4.1]. The second part of the proposition is proved below (see Remark 2.9 after the proof) using the following version of Hartman's theorem, which will also be used in the proofs of Lemma 3.2 and Proposition 3.3 to give a canonical description of the boundary of a Neumann domain near a cusp point.

Proposition 2.7 [Hartman 1960]. *Let E be an open neighbourhood of $\boldsymbol{p} \in \mathbb{R}^2$. Suppose $F \in C^2(E, \mathbb{R}^2)$ and let φ be the flow of the nonlinear system $\partial_t \varphi(t, x) = F(\varphi(t, x))$. Assume that $F(\boldsymbol{p}) = \boldsymbol{0}$ and the Jacobian* $\mathrm{DF}(\boldsymbol{p})$ *is diagonalizable and its eigenvalues have nonzero real part. Then, there exists a C^1-diffeomorphism $\Phi : U \to V$ of an open neighbourhood U of $\boldsymbol{p}$ onto an open neighbourhood V of the origin such that $D\Phi(\boldsymbol{p}) = \mathbb{I}$ and for each $x \in U$ the flow line through x is mapped by Φ to*

$$\Phi(\varphi(t, x)) = \mathrm{e}^{\mathrm{DF}(\boldsymbol{p})t}\,\Phi(x) \tag{2-1}$$

for small enough t values.

Remark 2.8. The textbook version of Hartman's theorem in n dimensions (see, for instance, [Perko 2001, p. 120]) only guarantees the existence of a homeomorphism Φ. For $n = 2$, the proposition above guarantees that Φ is a C^1-diffeomorphism, but for $n > 2$ further assumptions on the Jacobian are required to obtain this additional regularity. For instance, it suffices to assume that all of the eigenvalues of $\mathrm{DF}(\boldsymbol{p})$ are in the same (left or right) half-plane; see [Perko 2001, p. 127]. That version of the theorem would be sufficient for our purposes, since we only apply Proposition 2.7 at nondegenerate extrema, where all eigenvalues have the same sign. However, it is interesting to note that Proposition 2.7 also applies at saddle points in two dimensions.

Proof of Proposition 2.5(2). Let $\boldsymbol{c}$ be an extremal point of f whose Hessian is not proportional to g. Since $\mathrm{Hess}\, f(\boldsymbol{c})$ is nondegenerate, both eigenvalues of $\mathrm{Hess}\, f(\boldsymbol{c})$ are either strictly positive or strictly negative. We choose normal coordinates in an open neighbourhood $\widetilde{E}$ of $\boldsymbol{c}$, with respect to which $\widetilde{E}$ is represented by an open subset $E \subset \mathbb{R}^2$, $\boldsymbol{c}$ corresponds to the origin $\boldsymbol{0} \in \mathbb{R}^2$, and $g_{ij}(\boldsymbol{0}) = \delta_{ij}$.

We now apply Proposition 2.7 to $F = -\operatorname{grad} f$. Since $\mathrm{DF}(\mathbf{0}) = -\operatorname{Hess} f(\mathbf{0})$ is diagonalizable and has nonzero eigenvalues, there exist $U \subset E$ and $V \subset \mathbb{R}^2$, both containing the origin, and a C^1-diffeomorphism $\Phi : U \to V$ such that the gradient flow lines are mapped by Φ to the flow lines $\mathrm{e}^{-t \operatorname{Hess} f(\mathbf{0})}\Phi(x)$ of the linearized system. In [McDonald and Fulling 2014, Theorem 3.1; Alon et al. 2020, Proposition 4.1] it was shown that the angle between such flow lines at an extremal point is either 0, $\frac{\pi}{2}$ or π, under the assumption that $\operatorname{Hess} f(\mathbf{0})$ is not a scalar matrix. This assumption holds, as the Hessian is not proportional to the metric and we have chosen coordinates with respect to which $g(\mathbf{0})$ is the identity.

It is left to relate the meeting angle between the gradient flow lines in M and the corresponding flow lines $\mathrm{e}^{-t \operatorname{Hess} f(\mathbf{0})}\Phi(x)$ in V. Since the tangent map $D\Phi(\mathbf{0}) : T_{\mathbf{0}}U \to T_{\mathbf{0}}V$ is the identity, and $g_{ij}(\mathbf{0}) = \delta_{ij}$, the meeting angle of any two curves at $\mathbf{0}$ is preserved by Φ; hence this angle is either 0, $\frac{\pi}{2}$ or π. ☐

Remark 2.9. The argument for Proposition 2.5(2) given in [Alon et al. 2020, Proposition 4.1] is incomplete and hence we have supplied a complete proof here. In particular, the Taylor expansion argument used in the proofs of [McDonald and Fulling 2014, Theorem 3.1; Alon et al. 2020, Proposition 4.1] does not suffice. Substituting the Taylor expansion of F into $\partial_t \varphi(t,x) = F(\varphi(t,x))$ gives

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \mathrm{DF}(\boldsymbol{p}) \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} + \mathcal{O}(\|(x(t), y(t))\|^2_{\mathbb{R}^2}), \tag{2-2}$$

but this does not allow us to conclude that the flow may be approximated by $\mathrm{e}^{t\,\mathrm{DF}(\boldsymbol{p})}\binom{x_0}{y_0}$ due to the possible coupling of higher-order terms in (2-2). A simple example is $F(x,y) = (-\lambda_1 x, -\lambda_2 y + x^2)$, with $0 < 2\lambda_1 < \lambda_2$. For the resulting system $x' = -\lambda_1 x,\ y' = -\lambda_2 y + x^2$ we have $x(t) = x_0 e^{-\lambda_1 t}$, but

$$y(t) = \left[y_0 - \frac{x_0^2}{\lambda_2 - 2\lambda_1} \right] e^{-\lambda_2 t} + \frac{x_0^2}{\lambda_2 - 2\lambda_1} e^{-2\lambda_1 t} \tag{2-3}$$

is dominated by $e^{-2\lambda_1 t}$ for large t, and hence is not approximated by a solution to the linearized equation $y' = -\lambda_2 y$.

From Proposition 2.5(2) we see that the boundary of a Neumann domain may possess a cusp (when the meeting angle is 0) and so it can fail to be Lipschitz continuous. Furthermore, it may even fail to be of class C, where we recall that the boundary of a domain is of class C if it can be locally represented as the graph of a continuous function, alternatively, if the domain has the segment property (see [Edmunds and Evans 1987; Mazya and Poborchi 1997] for details). To demonstrate that this is a subtle property, we bring as an example the domains

$$\begin{aligned} \Omega_1 &= \left\{ (x,y) \in \mathbb{R}^2 : \tfrac{1}{2}x^2 < y < x^2,\ 0 < x < 1 \right\}, \\ \Omega_2 &= \{ (x,y) \in \mathbb{R}^2 : -x^2 < y < x^2,\ 0 < x < 1 \}, \end{aligned} \tag{2-4}$$

which are shown in Figure 2. The domain Ω_1 does not satisfy the segment property at the origin, and hence is not of class C, even though its boundary is the union of two smooth curves. On the other hand, Ω_2 (which contains Ω_1) *is* of class C. This example will be important later, in the proof of Proposition 3.3.

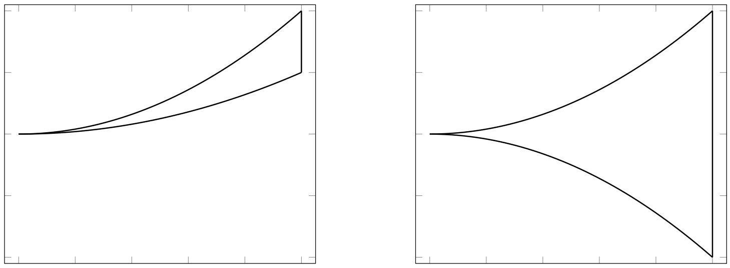

Figure 2. The regions Ω_1 (left) and Ω_2 (right) defined in (2-4) both have a cusp at the origin. However, Ω_1 is not of class C, whereas Ω_2 is.

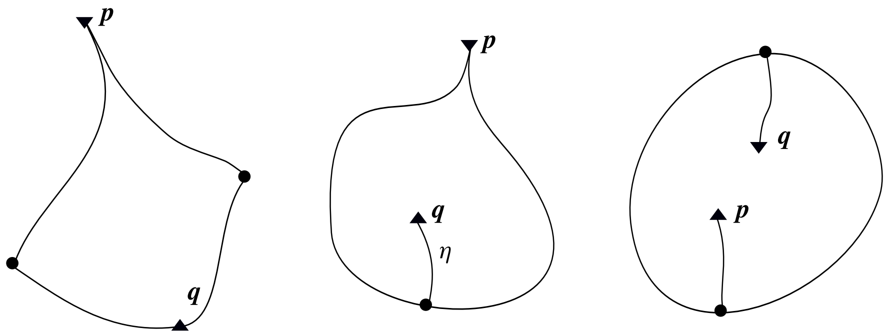

Figure 3. Possible types of Neumann domains for a Morse function: regular (left); cracked (centre); and doubly cracked (right). Saddle points are represented by balls, maxima by triangles pointing upwards and vice versa for minima. If f is Morse–Smale, its Neumann domains must look like one of the first two examples, with either one or two saddle points on the boundary. For the cracked domain shown in the centre, η is the only Neumann line connected to $\boldsymbol{q}$, hence $\deg(\boldsymbol{q}) = 1$. If f is not Morse–Smale, its Neumann domains can have additional saddle points on the boundary, and can have both extremal points of degree one, as shown on the right. (This last example has a Neumann line connecting two saddle points, which is not possible if f is Morse–Smale, by Lemma 2.4.)

We add that there is very little known in general regarding the asymptotic behaviour of Neumann lines near cusps. In particular, methods to treat cusps in a spectral-theoretic context, as in, e.g., [Jakšić et al. 1992; Flamencourt and Pankrashkin 2020; Band et al. 2021], have to be generalized for our purpose.

We end this section by examining Theorem 2.2 and its implications for the structure of Neumann domains. By the statement of the theorem, the boundary of a Neumann domain always contains a maximum and a minimum, but no other extrema. It follows that each Neumann domain must belong to one of the following two types (illustrated in Figure 3):

- a *regular Neumann domain* has on its boundary a maximum and a minimum, each of degree at least 2 (see Definition 2.3);
- a *cracked Neumann domain* has on its boundary an extremal point which is of degree 1.

Moreover, since the boundary is made up of Neumann lines, it must contain at least one saddle point. If f is Morse–Smale, the boundary contains at most two saddle points, by Theorem 2.2(4), but for a general Morse function it is possible to have more. The possible existence of additional saddle points is irrelevant for our analysis, however, since the boundary is Lipschitz near these points by Proposition 2.5(1).

Numerical observations suggest that generic Neumann domains are regular. However, it is not hard to construct Morse functions having cracked Neumann domains; see the Appendix. Theorems 1.2 and 1.3 apply to both types of domains, but in the proofs we need to pay careful attention to cracked domains. In particular, a cracked Neumann domain is not of class C, as the domain lies on both sides of its boundary.

Remark 2.10. In summary, a Neumann domain may fail to be of class C for two reasons: a cusp on the boundary or a crack in the domain, i.e., a Neumann line contained in the interior of $\overline{\Omega}$. These are the main technical obstacles to overcome in proving Theorems 1.2 and 1.3.

3. Sobolev spaces on Neumann domains

We now discuss properties of Sobolev spaces on Neumann domains. As described in the Introduction, and indicated in Proposition 2.5(2) (see also Remark 2.10), the difficulty is that the boundary of a Neumann domain need not be of class C, so standard density and compactness results do not apply.

In Section 3A we define Sobolev spaces on a Neumann domain and various subsets of its boundary. In Sections 3B and 3C we describe some technical constructions (dissection and truncation) that allow us to deal with cracks and cusps. Finally, in Section 3D we prove the main result of this section, Proposition 3.3, which establishes density and embedding properties for the space $W^{1,2}(\Omega)$ on a Neumann domain.

3A. ***Preliminaries.*** As above, we assume that (M, g) is a smooth, closed, connected, oriented Riemannian surface. For an open submanifold $N \subset M$, the Sobolev space $W^{k,2}(N)$ is defined to be the completion of $C^\infty(N)$ with respect to the norm

$$\|f\|^2_{W^{k,2}(N)} := \sum_{j=0}^{k} \int_N |\nabla^j f|^2, \tag{3-1}$$

where ∇ denotes the covariant derivative with respect to the metric g. The norm depends on g, but since M is compact, different metrics will produce equivalent norms. We will sometimes take advantage of this fact and compute the Sobolev norm using a metric $\tilde{g}$ defined in a local coordinate chart to have components $\tilde{g}_{ij} = \delta_{ij}$ (so that covariant derivatives become partial derivatives, the Riemannian volume form reduces to the Euclidean one, etc.). This allows us to apply standard methods in the theory of Sobolev spaces on Lipschitz domains in M.

Now suppose that $N \subset M$ is an open submanifold with Lipschitz boundary. We will later choose N to be a Neumann domain Ω, or a proper subset thereof (see Section 3C) if $\partial\Omega$ has a crack or a cusp. We define the boundary Sobolev spaces $H^s(\partial N)$ for $|s| \le 1$ via the Fourier transform and a suitable partition of unity, following [McLean 2000, p. 96], so that the dual space is given by $H^s(\partial N)^* = H^{-s}(\partial N)$. Moreover, for any open subset $\Gamma \subset \partial N$ we let

$$H^s(\Gamma) := \{f|_\Gamma : f \in H^s(\partial N)\}, \quad \widetilde{H}^s(\Gamma) := \text{closure of } C_0^\infty(\Gamma) \text{ in } H^s(\Gamma). \tag{3-2}$$

The space $\widetilde{H}^s(\Gamma)$ has an equivalent description that is often useful in practice:

$$\widetilde{H}^s(\Gamma) = \{f \in H^s(\partial N) : \operatorname{supp} f \subset \bar{\Gamma}\}. \tag{3-3}$$

This equivalence follows from [McLean 2000, Theorem 3.29]. Another convenient description, valid for $s \geq 0$, is

$$\widetilde{H}^s(\Gamma) = \{f \in L^2(\Gamma) : \tilde{f} \in H^s(\partial N)\}, \tag{3-4}$$

where $\tilde{f}$ denotes the extension of f by zero to $\partial N \setminus \Gamma$; this is [McLean 2000, Theorem 3.33].

It follows from the definitions that $\widetilde{H}^s(\Gamma) \subset H^s(\Gamma)$ for all $|s| \leq 1$, and it is well known that these spaces coincide for $|s| < \frac{1}{2}$. However, for $|s| \geq \frac{1}{2}$ we have $\widetilde{H}^s(\Gamma) \subsetneq H^s(\Gamma)$ whenever Γ is a proper subset of ∂N. To see this, consider the constant function $f \equiv 1$ on Γ, which is clearly in $H^s(\Gamma)$ for any s. It is easily verified that its extension $\tilde{f}$, which is just the indicator function χ_Γ, is not in $H^s(\partial N)$ for $s \geq \frac{1}{2}$, in which case we conclude from (3-4) that $f \notin \widetilde{H}^s(\Gamma)$. This distinction between the H^s and $\widetilde{H}^s$ spaces will be important when we consider the normal derivative of a function restricted to a subset of the boundary; see in particular Lemma 3.1 and its application in the proof of Proposition 4.3.

The $\widetilde{H}^s$ spaces arise naturally as duals to the H^s spaces. That is, for any $|s| \leq 1$ we have $H^s(\Gamma)^* = \widetilde{H}^{-s}(\Gamma)$, from [McLean 2000, Theorem 3.30]. In particular,

$$\widetilde{H}^{-1/2}(\Gamma) = H^{1/2}(\Gamma)^* \subsetneq \widetilde{H}^{1/2}(\Gamma)^* = H^{-1/2}(\Gamma). \tag{3-5}$$

Using (3-2) we obtain

$$\ell = 0 \quad \text{in } H^{-s}(\Gamma) \quad \iff \quad \ell(f) = 0 \quad \text{for all } f \in C_0^\infty(\Gamma). \tag{3-6}$$

We thus define for $0 \leq s \leq 1$ the mapping

$$\begin{gathered} \cdot^{\text{dual}} : L^2(\Gamma) \to H^s(\Gamma)^*, \\ g^{\text{dual}}(f) := \langle f, g \rangle_{L^2(\Gamma)}, \quad f \in H^s(\Gamma), \end{gathered} \tag{3-7}$$

observing that the L^2 inner product is well-defined because $H^s(\Gamma) \subset L^2(\Gamma)$ for $0 \leq s \leq 1$. As a result, we will often abuse notation and use integral notation to denote the action of $\ell \in H^s(\Gamma)^*$ on $f \in H^s(\Gamma)$, i.e., we will write

$$\ell(f) = \int_\Gamma \ell f$$

even when ℓ is not in the range of the map $\cdot^{\text{dual}}$; see in particular Green's identity (3-10) below.

Given a decomposition $\partial N = \bar{\Gamma}_1 \cup \bar{\Gamma}_2$, where Γ_1 and Γ_2 are disjoint, open subsets of ∂N, and a distribution $\ell \in H^{-s}(\partial N)$ for some $s \geq 0$, we have $\ell|_{\Gamma_i} \in H^{-s}(\Gamma_i)$ for $i = 1, 2$. For $0 \leq s < \frac{1}{2}$ we obtain the decomposition

$$\ell(\phi) = \ell|_{\Gamma_1}(\phi|_{\Gamma_1}) + \ell|_{\Gamma_2}(\phi|_{\Gamma_2})$$

for every $\phi \in H^s(\partial N)$. However, this is not true for $s \geq \frac{1}{2}$, and in fact the right-hand side is not even defined in this case, since $\phi|_{\Gamma_i} \in H^s(\Gamma_i)$, whereas $\ell|_{\Gamma_i} \in H^{-s}(\Gamma_i)$ might not be contained in $H^s(\Gamma_i)^*$, as indicated in (3-5). However, such a splitting does hold for ℓ if we assume that Γ_1 and Γ_2 are separated by a third subset Γ_0 on which ℓ vanishes.

Lemma 3.1. *Suppose $\partial N = \bar\Gamma_0 \cup \bar\Gamma_1 \cup \bar\Gamma_2$, where $\Gamma_0, \Gamma_1, \Gamma_2$ are disjoint, open subsets of ∂N with $\bar\Gamma_1 \cap \bar\Gamma_2 = \varnothing$. If $\ell \in H^{-1/2}(\partial N)$ vanishes on Γ_0, then $\ell|_{\Gamma_i} \in \widetilde{H}^{-1/2}(\Gamma_i)$ for $i = 1, 2$, and*

$$\ell(\phi) = \ell|_{\Gamma_1}(\phi|_{\Gamma_1}) + \ell|_{\Gamma_2}(\phi|_{\Gamma_2}) \tag{3-8}$$

for every $\phi \in H^{1/2}(\partial N)$.

Such a partition of the boundary is illustrated in Figure 6, where $N = \Omega_t \cap \Omega_{\mathrm{r}}$, $\bar\Gamma_0 = \gamma_{0,t}$, $\bar\Gamma_1 = \tilde\eta$ and $\bar\Gamma_2 = \gamma_{-,t}$.

Proof. We will use (3-3) to prove that $\ell|_{\Gamma_1} \in \widetilde{H}^{-1/2}(\Gamma_1)$. This does not follow immediately, however, since ℓ is not necessarily supported in $\bar\Gamma_1$. Therefore, we will create a modified distribution, ℓ_1, such that $\operatorname{supp} \ell_1 \subset \bar\Gamma_1$ and $\ell|_{\Gamma_1} = \ell_1|_{\Gamma_1}$.

Since $\bar\Gamma_1 \cap \bar\Gamma_2 = \varnothing$, we can find a smooth bump function χ_1 that equals 1 on Γ_1 and 0 on Γ_2. Consider the distribution $\ell_1(\phi) := \ell(\chi_1\phi)$, which is in $H^{-1/2}(\partial N)$. If $\operatorname{supp}\phi \subset \bar\Gamma_0 \cup \Gamma_2$, then $\operatorname{supp}(\chi_1\phi) \subset \Gamma_0$, and hence $\ell_1(\phi) = \ell(\chi_1\phi) = 0$, because ℓ vanishes on Γ_0. This shows that $\operatorname{supp}\ell_1 \subset \partial N \setminus (\bar\Gamma_0 \cup \Gamma_2) = \bar\Gamma_1$. On the other hand, if $\operatorname{supp}\phi \subset \Gamma_1$, then $\chi_1\phi = \phi$, and hence $\ell_1(\phi) = \ell(\phi)$. We have thus shown that $\ell|_{\Gamma_1} = \ell_1|_{\Gamma_1} \in \widetilde{H}^{-1/2}(\Gamma_1)$.

Similarly, we obtain $\ell_2 \in H^{-1/2}(\partial N)$, with $\operatorname{supp}\ell_2 \subset \bar\Gamma_2$ and $\ell|_{\Gamma_2} = \ell_2|_{\Gamma_2} \in \widetilde{H}^{-1/2}(\Gamma_2)$. It follows that the distribution

$$\hat\ell := \ell - \ell_1 - \ell_2 \in H^{-1/2}(\partial N)$$

has support in $\partial N \setminus (\Gamma_0 \cup \Gamma_1 \cup \Gamma_2) = (\bar\Gamma_0 \cap \bar\Gamma_1) \cup (\bar\Gamma_0 \cap \bar\Gamma_2)$, which is a finite set. However, a distribution in $H^{-1/2}$ cannot be supported on a finite set of points, by [McLean 2000, Lemma 3.39], so we conclude that $\hat\ell$ is identically zero, which completes the proof. □

Since N was assumed to have Lipschitz boundary, the trace map $\cdot|_{\partial N} : W^{1,2}(N) \to H^{1/2}(\partial N)$ is continuous. To define the normal derivative we first introduce the weak Laplace–Beltrami operator, $\Delta : W^{1,2}(\Omega) \to W_0^{1,2}(\Omega)^*$, where $\Omega \subset M$ is any open subset of M. By definition, $\Delta\psi = f$ means

$$\int_\Omega \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle = \int_\Omega f\phi \tag{3-9}$$

for all $\phi \in W_0^{1,2}(\Omega)$, where the integral on the right-hand side is shorthand for the action of $f \in W_0^{1,2}(\Omega)^*$ on $\phi \in W_0^{1,2}(\Omega)$. If $\Delta\psi = f \in L^2(\Omega)$, then this is a genuine L^2 inner product of f and ϕ.

The weak version of Green's identity then says that for any $\psi \in W^{1,2}(N)$ with $\Delta\psi \in L^2(N)$, there exists a unique $\partial_\nu\psi \in H^{-1/2}(\partial N)$ such that

$$\int_N \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle = \int_N (\Delta\psi)\phi + \int_{\partial N} (\partial_\nu\psi)\phi \tag{3-10}$$

for all $\phi \in W^{1,2}(N)$ [McLean 2000, Theorem 4.4]. The boundary term has to be understood as the action of $\partial_\nu\psi \in H^{-1/2}(\partial N)$ on $\phi|_{\partial N} \in H^{1/2}(\partial N)$, i.e., $(\partial_\nu\psi)(\phi|_{\partial N})$, but to simplify the presentation we use the integral notation of (3-10).

Finally, consider an open subset $\Gamma \subset \partial N$. For a function $\psi \in W^{1,2}(N)$ we define $\psi|_\Gamma$ to be the restriction of $\psi|_{\partial N} \in H^{1/2}(\partial N)$ to Γ, so that (3-2) implies $\psi|_\Gamma \in H^{1/2}(\Gamma)$. Similarly, if $\psi \in W^{1,2}(N)$

and $\Delta\psi \in L^2(N)$, we have $\partial_\nu\psi|_\Gamma \in H^{-1/2}(\Gamma)$. It is not necessarily true that $\partial_\nu\psi|_\Gamma$ is contained in the smaller space $\widetilde{H}^{-1/2}(\Gamma) = H^{1/2}(\Gamma)^*$; see (3-5). However, this will be the case if $\partial_\nu\psi$ vanishes on $\partial N \setminus \overline{\Gamma}$, by Lemma 3.1. This fact will play a crucial role in the proof of Proposition 4.3 below.

We conclude the section by explaining our decision to use $W^{k,2}$-Sobolev spaces on N but H^s-Sobolev spaces on ∂N. Recall that $H^1(\Omega) \subset W^{1,2}(\Omega)$ holds for any open set Ω, but the inclusion can be strict unless one has additional regularity of the boundary. In Definition 4.1 we construct the Neumann Laplacian as the self-adjoint operator corresponding to a nonnegative, symmetric bilinear form. For this we require the form to be closed, which is the case if the form domain is $W^{1,2}(\Omega)$, but need not be true if the form domain is $H^1(\Omega)$. On the other hand, the H^s-Sobolev spaces, defined via the Fourier transform, provide a more natural setting for the discussion of traces: If N is an open submanifold with Lipschitz boundary, there is a bounded, surjective trace map $\cdot|_{\partial N} : H^1(N) \to H^{1/2}(\partial N)$. For N Lipschitz we have the equality $H^1(N) = W^{1,2}(N)$, and hence a well-defined trace map $\cdot|_{\partial N} : W^{1,2}(N) \to H^{1/2}(\partial N)$.

3B. ***Dissections of Neumann domains.*** The boundary of a cracked Neumann domain cannot be of class C, whether or not there is a cusp on the boundary, due to the Neumann line η contained in the interior of $\overline{\Omega}$; see Figure 3. We deal with this by dissecting such a Neumann domain into two pieces, as shown in Figure 4, where one piece has Lipschitz boundary, and the other has boundary that is Lipschitz except possibly at a cusp point; i.e., it has the same regularity as a regular Neumann domain. For doubly cracked domains as in Figure 3, an analogous statement holds as the proof for that case is essentially the same. The dissection thus reduces many of the proofs for cracked domains to the corresponding results for regular domains.

This dissection is made possible by the following lemma.

Lemma 3.2. *Assume f is a Morse function and let γ be a Neumann line. Then γ has finite length $L(\gamma) < \infty$, and admits an arc-length parametrization with $\gamma \in C^1([0, L(\gamma)])$, i.e., boundary points are included.*

Proof. We decompose $\gamma = \gamma_0 \cup \gamma_1 \cup \gamma_2$, where γ_1 is defined in a small neighbourhood of the initial endpoint of γ and γ_2 is defined in a small neighbourhood of the terminal endpoint. Then it is enough to prove the corresponding statement for γ_0, γ_1 and γ_2.

The result for γ_0 follows by standard results for flows of smooth vector fields. Definition 2.3 implies that the endpoints of γ are critical points of f. If the initial endpoint (which we label $\boldsymbol{c}$) is a saddle, then the result for γ_1 follows, e.g., by [Banyaga and Hurtubise 2004, Theorem 4.2 (p. 94)]. On the other hand, if $\boldsymbol{c}$ is an extremum we use the map Φ from Proposition 2.7. Then $\Phi \circ \gamma_1$ is a flow line generated by $\mathrm{e}^{-\operatorname{Hess} f(\boldsymbol{c})t}$, and hence satisfies the properties of the claim, i.e., it is C^1 up to the endpoint and has finite length. As Φ^{-1} is a C^1 map and $\gamma_1 = \Phi^{-1} \circ (\Phi \circ \gamma_1)$ is a composition of C^1 functions, the claim for γ_1 follows. The proof for γ_2 is identical. □

Now suppose that Ω is a cracked Neumann domain. The doubly cracked case in Figure 3 can be treated analogously. Denote by $\boldsymbol{q}$ the extremum in the interior of $\overline{\Omega}$, and let η be the Neumann line attached to $\boldsymbol{q}$; see Figure 4. Choosing a Lipschitz curve $\tilde{\eta}$ in Ω that joins $\boldsymbol{q}$ with a noncusp point of $\partial\Omega$, we obtain a

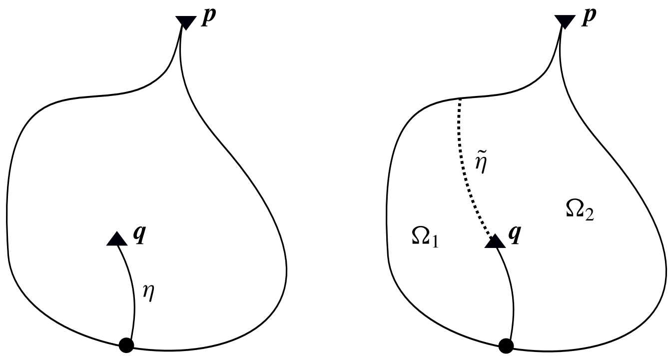

Figure 4. The dissection of a cracked Neumann domain, as given in (3-11). The Neumann line η is extended to a Lipschitz curve $\eta \cup \tilde{\eta}$, so that Ω_l is a Lipschitz domain and Ω_r possesses a cusp at $\boldsymbol{p}$.

dissection of Ω into disjoint parts Ω_l and Ω_r, i.e.,

$$\Omega \setminus \tilde{\eta} = \Omega_\mathrm{l} \cup \Omega_\mathrm{r}, \tag{3-11}$$

as shown in Figure 4. Lemma 3.2 guarantees that $\eta \cup \tilde{\eta}$ is a Lipschitz curve, so Ω_l has Lipschitz boundary, and Ω_r has Lipschitz boundary with the possible exception of a cusp at $\boldsymbol{p}$. This dissection induces an isometric embedding

$$\begin{aligned} W^{1,2}(\Omega) &\to W^{1,2}(\Omega_\mathrm{l}) \oplus W^{1,2}(\Omega_\mathrm{r}), \\ \phi &\mapsto (\phi|_{\Omega_\mathrm{l}}, \phi|_{\Omega_\mathrm{r}}). \end{aligned} \tag{3-12}$$

3C. ***Truncated Neumann domains.*** To deal with potential cusps at the maximum and minimum of f, we introduce truncated versions of Ω. Denoting by $\boldsymbol{p} \in \mathrm{Min}(f)$ and $\boldsymbol{q} \in \mathrm{Max}(f)$ the minimum and maximum of f in $\overline{\Omega}$, we observe that $f(\boldsymbol{q}) < f(\boldsymbol{p})$, since otherwise f would be constant on $\overline{\Omega}$, which is not possible as it is a Morse function. Adding a constant to f, which does not affect the gradient flow lines, we can thus assume that $f(\boldsymbol{q}) < 0 < f(\boldsymbol{p})$. (In the special case that f is an eigenfunction this condition holds automatically, so it is not necessary to shift f.)

We then define for each $0 < t < 1$ the truncated domains

$$\Omega_t := \begin{cases} \{x \in \Omega : f(x) < tf(\boldsymbol{q})\}, & \boldsymbol{q} \text{ is a cusp, } \boldsymbol{p} \text{ is not,} \\ \{x \in \Omega : tf(\boldsymbol{p}) < f(x)\}, & \boldsymbol{p} \text{ is a cusp, } \boldsymbol{q} \text{ is not,} \\ \{x \in \Omega : tf(\boldsymbol{p}) < f(x) < tf(\boldsymbol{q})\}, & \boldsymbol{q} \text{ and } \boldsymbol{p} \text{ are cusps,} \\ \Omega, & \text{otherwise.} \end{cases} \tag{3-13}$$

Some examples of this construction are shown in Figure 5.

The boundary of Ω_t can be decomposed as $\partial\Omega_t = \gamma_{\pm,t} \cup \gamma_{0,t}$, where $\gamma_{\pm,t}$ are level lines defined by

$$\gamma_{+,t} = \{x : f(x) = tf(\boldsymbol{q})\}, \quad \gamma_{-}, t = \{x : f(x) = tf(\boldsymbol{p})\}, \tag{3-14}$$

and $\gamma_{0,t} = \partial\Omega_t \cap \partial\Omega$ is the part of $\partial\Omega$ that remains after the truncation. Note that $\gamma_{0,t} \neq \varnothing$, and Proposition 2.5(3) implies that $\gamma_{\pm,t}$ meets $\partial\Omega$ perpendicularly, except for a finite number of exceptional times where $\gamma_{\pm,t}$ meets $\partial\Omega$ at a saddle point, in which case the meeting angle is $\frac{\pi}{4}$; see Figure 5.

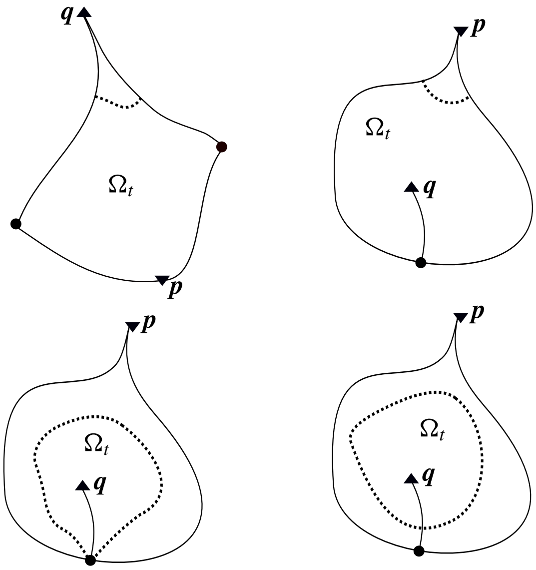

Figure 5. Neumann domains and their truncations, with the dotted line indicating the curve $\gamma_{\pm,t}$. The top two figures show regular and cracked domains for t close to 1. For the same cracked domain the bottom left figure shows an exceptional value of t, where $\gamma_{\pm,t}$ meets $\partial\Omega$ at angle $\frac{\pi}{4}$, and the bottom right figure shows a smaller value of t.

For a truncated Neumann domain Ω_t we denote its complement in Ω by $\Omega_t^c := \Omega \setminus \Omega_t$. For any $0 < t < 1$ and sufficiently small $\epsilon > 0$, we can find a smooth cutoff function χ on M such that

$$\chi(x) = \begin{cases} 0, & x \in \Omega_t, \\ 1, & x \in \Omega_{t+\epsilon}^c. \end{cases} \tag{3-15}$$

If desired, we can assume that χ is of the form $\alpha \circ f$ for some $\alpha \in C^\infty(\mathbb{R})$, in which case χ has the same level lines as f. For the arguments to follow, however, a generic smooth cutoff will suffice.

3D. ***Density and embedding results.*** We now state and prove the main result of this section.

Proposition 3.3. *Let (M, g) be a closed, connected, oriented Riemannian surface. If $\Omega \subset M$ is a Neumann domain of a Morse function f, the following hold*:

(1) *The embedding $W^{1,2}(\Omega) \to L^2(\Omega)$ is compact.*

(2) *If Ω is regular, then $C^1(\overline{\Omega})$ is dense in $W^{1,2}(\Omega)$.*

(3) *If Ω is cracked, then there exists $t \in (0, 1)$ such that the set of functions*

$$\{\phi \in W^{1,2}(\Omega) : \phi|_{\overline{\Omega}_t^c} \in C^1(\overline{\Omega}_t^c)\} \tag{3-16}$$

is dense in $W^{1,2}(\Omega)$.

The result is known if $\partial\Omega$ is of class C (see [Mazya and Poborchi 1997]) but, as noted above, the boundary of a Neumann domain does not need to have this property. If Ω is cracked and $\phi \in W^{1,2}(\Omega)$, its values on opposite sides of the crack η need not be close, so we cannot hope to approximate it by a function in $C^1(\overline{\Omega})$. However, by choosing t sufficiently large, we can ensure that $\overline{\Omega}_t^c$ is disjoint from η, and hence (3-16) holds.

The idea of the proof is to use Hartman's theorem (Proposition 2.7) to find a canonical description of the boundary near a cusp, and then apply the following lemma, which allows us to extend functions to a larger domain which still has a cusp but is of class C; see the domains Ω_1 and Ω_2 in Figure 2.

Lemma 3.4 [Mazya and Poborchi 1997, §5.4.1, Lemma 1, p. 285]. *Consider the domain*

$$\widetilde{\Omega} = \{(x, y) \in \mathbb{R}^2 : c_1\vartheta(x) < y < c_2\vartheta(x),\ 0 < x < 1\}$$

for some $c_1 < c_2$, where $\vartheta \in C^{0,1}([0, 1])$ is an increasing function with $\vartheta(0) = 0$ and $\vartheta'(t) \to 0$ as $t \to 0$, and define

$$G = \{(x, y) \in \mathbb{R}^2 : |y| < M\vartheta(x),\ 0 < x < 1\} \tag{3-17}$$

for $M \geq \max\{|c_1|, |c_2|\}$. Then there exists a continuous extension operator $\mathcal{E} : W^{1,2}(\widetilde{\Omega}) \to W^{1,2}(G)$.

We will apply this lemma with $\vartheta(x) = x^\alpha$ for some $\alpha > 1$.

Proof of Proposition 3.3. We first prove (1) and (2) for regular Neumann domains. Only the behaviour near the cusps has to be investigated, as they are the only possible non-Lipschitz points on $\partial\Omega$. A cusp is either a maximum or a minimum by Proposition 2.5. Without loss of generality, let $\boldsymbol{c} \in \operatorname{Max}(f)$ be the only cusp on $\partial\Omega$.

We localize at $\boldsymbol{c}$ by taking a smooth cutoff function χ, as in (3-15), that equals 1 in $\Omega_{t+\epsilon}^c$ and vanishes in Ω_t, and hence is supported in $\Omega_{\boldsymbol{c}} := \Omega_t^c$. Now take $\phi \in W^{1,2}(\Omega)$. We write $\phi = \chi\phi + (1-\chi)\phi$ and observe that $\chi\phi,\ (1-\chi)\phi \in W^{1,2}(\Omega)$. Thus, it is sufficient to prove the statements for both functions separately. For the latter function the observation that it is supported in a Lipschitz domain implies both (1) and (2) in Proposition 3.3.

For the former we choose t close to 1 and employ Proposition 2.7. Let Φ be the resulting C^1-diffeomorphism and define $\widetilde{\Omega}_{\boldsymbol{c}} = \Phi(\Omega_{\boldsymbol{c}})$. Owing to (2-1), the image in $\widetilde{\Omega}_{\boldsymbol{c}}$ of the two boundary curves meeting at $\boldsymbol{c}$ consists of flow lines obeying $\partial_t\gamma = -\operatorname{Hess} f(\boldsymbol{c})\gamma$. These are generated by $\mathrm{e}^{-t \operatorname{Hess} f(\boldsymbol{c})}x_0$, where x_0 is a suitable point on γ. An easy calculation as in [McDonald and Fulling 2014, Section 3; Alon et al. 2020, Proof of Proposition 4.1] shows that the flow lines near the origin may be parametrized in suitable coordinates by $\gamma(x) = (x, cx^\alpha)$, where $\alpha > 0$ depends only on the eigenvalues of $\operatorname{Hess} f(\boldsymbol{c})$ (in fact only on their ratio). This implies that near the origin the domain $\widetilde{\Omega}_{\boldsymbol{c}}$ is described by

$$(x, y) \in \widetilde{\Omega}_{\boldsymbol{c}} \quad \Longleftrightarrow \quad c_1x^\alpha < y < c_2x^\alpha \text{ and } x > 0. \tag{3-18}$$

We can assume that $\alpha > 1$. (If $\alpha = 1$, then $\widetilde{\Omega}_{\boldsymbol{c}}$ is in fact Lipschitz near $\boldsymbol{c}$, so there is nothing to prove; if $\alpha < 1$, we exchange x and y to obtain a similar description of the boundary with α replaced by $1/\alpha$.) Now Lemma 3.4 says that there exists a continuous extension operator $\mathcal{E} : W^{1,2}(\widetilde{\Omega}_{\boldsymbol{c}}) \to W^{1,2}(G)$, where $\widetilde{\Omega}_{\boldsymbol{c}} \subset G$ and near $\boldsymbol{0}$ the domain G is characterized by

$$(x, y) \in G \quad \Longleftrightarrow \quad |y| < Mx^\alpha \text{ and } x > 0, \tag{3-19}$$

with M large enough. Since the boundary ∂G is of class C, we can now infer by [Edmunds and Evans 1987, Theorem 4.17, p. 267; Mazya and Poborchi 1997, §1.4.2, Theorem 1, p. 28] that $W^{1,2}(G)$ satisfies statements (1) and (2) of the proposition. In particular, $W^{1,2}(G) \to L^2(G)$ is compact and $C^1(\overline{G})$ is dense in $W^{1,2}(G)$.

Using the fact that Φ is a C^1-diffeomorphism, it is easily shown that the pull-back map

$$\Phi^* : W^{1,2}(\widetilde{\Omega}_{\boldsymbol{c}}) \to W^{1,2}(\Omega_{\boldsymbol{c}}), \quad \phi \mapsto \phi \circ \Phi, \tag{3-20}$$

is well-defined and bijective, with

$$\frac{1}{C'}\|\phi\|^2_{W^{1,2}(\widetilde{\Omega}_{\boldsymbol{c}})} < \|\phi \circ \Phi\|^2_{W^{1,2}(\Omega_{\boldsymbol{c}})} < C'\|\phi\|^2_{W^{1,2}(\widetilde{\Omega}_{\boldsymbol{c}})}$$

for some $C' > 0$. Therefore, the inclusion $W^{1,2}(\Omega_{\boldsymbol{c}}) \to L^2(\Omega_{\boldsymbol{c}})$ can be written as the composition of a compact operator

$$W^{1,2}(\Omega_{\boldsymbol{c}}) \xrightarrow{(\Phi^{-1})^*} W^{1,2}(\widetilde{\Omega}_{\boldsymbol{c}}) \xrightarrow{\mathcal{E}} W^{1,2}(G) \longrightarrow L^2(G)$$

and a bounded operator

$$L^2(G) \longrightarrow L^2(\widetilde{\Omega}_{\boldsymbol{c}}) \xrightarrow{\Phi^*} L^2(\Omega_{\boldsymbol{c}})$$

(where the first map is restriction), and hence is compact. This completes the proof of (1) for regular Neumann domains.

To prove (2), let $\phi \in W^{1,2}(\Omega_{\boldsymbol{c}})$, so that $\phi \circ \Phi^{-1} \in W^{1,2}(\widetilde{\Omega}_{\boldsymbol{c}})$ and $\mathcal{E}(\phi \circ \Phi^{-1}) \in W^{1,2}(G)$. For any $\delta > 0$, there exists $\tilde{\phi} \in C^1(\overline{G})$ with $\|\tilde{\phi} - \mathcal{E}(\phi \circ \Phi^{-1})\|_{W^{1,2}(G)} < \delta$, and hence

$$\begin{aligned}\|\tilde{\phi}|_{\widetilde{\Omega}_{\boldsymbol{c}}} \circ \Phi - \phi\|_{W^{1,2}(\Omega_{\boldsymbol{c}})} &\le C'\|\tilde{\phi}|_{\widetilde{\Omega}_{\boldsymbol{c}}} - \phi \circ \Phi^{-1}\|_{W^{1,2}(\widetilde{\Omega}_{\boldsymbol{c}})} \\ &\le C'\|\tilde{\phi} - \mathcal{E}(\phi \circ \Phi^{-1})\|_{W^{1,2}(G)} < C'\delta.\end{aligned}$$

Since $\tilde{\phi}|_{\widetilde{\Omega}_{\boldsymbol{c}}} \circ \Phi \in C^1(\overline{\Omega}_{\boldsymbol{c}})$, this completes the proof of (2).

We next prove (1) for cracked Neumann domains, using the decomposition (3-12). More precisely, using Lemma 3.2 we may dissect Ω as in (3-11) and, without loss of generality, assume that the cusp is located on the boundary of Ω_{r}, as in Figure 3. Note that

$$W^{1,2}(\Omega) \to W^{1,2}(\Omega_{\mathrm{l}}) \oplus W^{1,2}(\Omega_{\mathrm{r}}) \to L^2(\Omega_{\mathrm{l}}) \oplus L^2(\Omega_{\mathrm{r}}) = L^2(\Omega)$$

and so it is enough to prove compactness of the embedding $W^{1,2}(\Omega_\bullet) \to L^2(\Omega_\bullet)$ for $\bullet = \mathrm{l}, \mathrm{r}$. For $\bullet = \mathrm{l}$ this follows from the Lipschitz property of $\partial\Omega_{\mathrm{l}}$. For $\bullet = \mathrm{r}$ we observe that $\partial\Omega_{\mathrm{r}}$ is Lipschitz except at the cusp, and so the proof given above for regular domains applies.

Finally, we prove (3). For $0 < t < 1$ sufficiently close to 1 we have $\Omega_t^c \subset \Omega_\bullet$ for either $\bullet = \mathrm{l}$ or r (the case $\bullet = \mathrm{r}$ is shown in Figure 4), so we choose t sufficiently close to 1 and $\epsilon > 0$ small enough that $\Omega_{\boldsymbol{c}} = \Omega_t^c \subset \Omega_\bullet$. Now let $\phi \in W^{1,2}(\Omega)$. Given $\delta > 0$, there exists by (2) a function $\phi_\delta \in C^1(\overline{O}_{\boldsymbol{c}})$ such that $\|\phi - \phi_\delta\|_{W^{1,2}(\Omega_{\boldsymbol{c}})} < \delta$. Choosing a smooth cutoff function χ that equals 1 in $\Omega_{t+\epsilon}^c$ and vanishes in Ω_t, we define $\tilde{\phi}_\delta = \chi\phi_\delta + (1-\chi)\phi \in W^{1,2}(\Omega)$ and compute

$$\begin{aligned}\|\phi - \tilde{\phi}_\delta\|_{W^{1,2}(\Omega)} &= \|\chi(\phi - \phi_\delta)\|_{W^{1,2}(\Omega_c)} \\ &\le K\|\phi - \phi_\delta\|_{W^{1,2}(\Omega_c)} < K\delta,\end{aligned} \tag{3-21}$$

where K is a constant depending only on χ. Finally, since $\operatorname{supp}\chi \subset \overline{\Omega}_t^c$, we have

$$\tilde{\phi}_\delta|_{\overline{\Omega}_t^c} = \chi\phi_\delta|_{\overline{\Omega}_t^c} \in C^1(\overline{\Omega}_t^c). \qquad \square$$

4. The Neumann Laplacian on a Neumann domain

In this section we define the Neumann Laplacian on a Neumann domain Ω, and establish some of its fundamental properties, in particular proving Theorems 1.2 and 1.3. This relies on the technical results of the previous section, namely Proposition 3.3.

4A. ***Definition and proof of Theorem 1.2.*** We define the Neumann Laplacian in the usual way, via a symmetric bilinear form.

Definition 4.1. The Neumann Laplacian on an open set $\Omega \subset M$, denoted by Δ_Ω^N, is the unique self-adjoint operator corresponding to the bilinear form

$$a(\psi, \phi) := \int_\Omega \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle, \quad \mathcal{D}(a) := W^{1,2}(\Omega). \tag{4-1}$$

More precisely, Δ_Ω^N is an unbounded operator on $L^2(\Omega)$, with domain

$$\mathcal{D}(\Delta_\Omega^N) = \{\psi \in W^{1,2}(\Omega) : \text{there exists } f_\psi \in L^2(\Omega) \text{ with } a(\psi,\phi) = \langle f_\psi, \phi\rangle_{L^2(\Omega)} \text{ for all } \phi \in W^{1,2}(\Omega)\}, \tag{4-2}$$

and for any $\psi \in \mathcal{D}(\Delta_\Omega^N)$ we have $\Delta_\Omega^N\psi = f_\psi$. The existence and uniqueness of such an operator follows immediately from the completeness of the form domain $\mathcal{D}(a) = W^{1,2}(\Omega)$ and standard theory of self-adjoint operators, for instance [Reed and Simon 1972, Theorem VIII.15]. If $\psi \in \mathcal{D}(\Delta_\Omega^N)$, then (4-2) implies

$$\int_\Omega \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle = \int_\Omega (\Delta_\Omega^N\psi)\phi$$

for all $\phi \in W_0^{1,2}(\Omega)$, and hence $\Delta\psi = \Delta_\Omega^N\psi \in L^2(\Omega)$. That is, Δ_Ω^N acts as the weak Laplace–Beltrami operator Δ defined in (3-9).

The next result is nontrivial, and relies on the special geometric structure of Neumann domains.

Proposition 4.2. *If $\Omega \subset M$ is a Neumann domain for a Morse function, then Δ_Ω^N has compact resolvent, and hence has purely discrete spectrum $\sigma(\Delta_\Omega^N) \subset [0,\infty)$.*

Proof. Proposition 3.3(1) says that the form domain $W^{1,2}(\Omega)$ is compactly embedded in $L^2(\Omega)$, so the result follows from [Reed and Simon 1978, Theorem XIII.64]. $\square$

4B. ***Domain of the Neumann Laplacian.*** We now describe the domain of the Neumann Laplacian, working towards the proof of Theorem 1.3. Recalling the truncated domain Ω_t introduced in Section 3C, and the decomposition $\partial\Omega_t = \gamma_{\pm,t} \cup \gamma_{0,t}$ of its boundary in (3-14), we have the following.

Proposition 4.3. *Let Ω be a Neumann domain of a Morse function f. The domain of the Neumann Laplacian is given by*

$$\mathcal{D}(\Delta_\Omega^N) = \Big\{\psi \in W^{1,2}(\Omega) : \Delta\psi \in L^2(\Omega),\ \partial_\nu(\psi|_{\Omega_t})|_{\gamma_{0,t}^{\mathrm{o}}} = 0 \textit{ for all } 0 < t < 1 \\ \textit{and } \lim_{t\to 1} \int_{\gamma_{\pm,t}} \partial_\nu(\psi|_{\Omega_t})\phi = 0 \textit{ for all } \phi \in W^{1,2}(\Omega)\Big\} \tag{4-3}$$

if Ω is regular, *and*

$$\mathcal{D}(\Delta_\Omega^N) = \Big\{\psi \in W^{1,2}(\Omega) : \Delta\psi \in L^2(\Omega),\ \partial_\nu(\psi|_{\Omega_t\cap\Omega_{\mathrm{l}}})|_{\gamma_{0,t}^{\mathrm{o}}} = \partial_\nu(\psi|_{\Omega_t\cap\Omega_{\mathrm{r}}})|_{\gamma_{0,t}^{\mathrm{o}}} = 0 \\ \textit{for all } 0 < t < 1,\ \textit{and } \lim_{t\to 1} \int_{\gamma_{\pm,t}} \partial_\nu(\psi|_{\Omega_t})\phi = 0 \textit{ for all } \phi \in W^{1,2}(\Omega)\Big\} \tag{4-4}$$

if Ω is cracked.

That is, to be in the domain of Δ_Ω^N, a function must satisfy Neumann boundary conditions on the Lipschitz part of the boundary, as well as a limiting boundary condition at each cusp. While this completely characterizes the domain of Δ_Ω^N, the limiting boundary conditions on $\gamma_{\pm,t}$ may be difficult to check in practice. Therefore, in the following section we will give simple criteria (Proposition 4.8 and Corollary 4.9) which guarantee these limiting conditions are satisfied.

Remark 4.4. Our techniques actually give a more general result, not just valid for Neumann domains. The key points are that $\partial\Omega$ is Lipschitz except for a finite number of cusps and cracks, and the cracks admit a Lipschitz continuation. A stronger result will be given below (in Remark 4.7) that relies on the detailed structure of the cusps, which for Neumann domains is a consequence of Hartman's theorem.

In proving the proposition, we must take into account the fact that Ω_{l} and Ω_{r} need not be Lipschitz; see Figure 4, where Ω_{r} has a cusp on its boundary. We therefore combine the dissection and truncation of Sections 3B and 3C, respectively. The resulting domains are shown in Figure 6. Note that the boundaries of $\Omega_t \cap \Omega_{\mathrm{l}}$ and $\Omega_t \cap \Omega_{\mathrm{r}}$ can be partitioned into three parts: $\gamma_{\pm,t}$ coming from the truncation; $\tilde{\eta}$ coming from the dissection; and $\gamma_{0,t}$, coming from the original domain Ω. We emphasize that the dissection (3-11) is an auxiliary construction, and our analysis does not depend on the specific choice of $\tilde{\eta}$.

Since $\eta \cup \tilde{\eta}$ has a Lipschitz neighbourhood in both $\Omega_t \cap \Omega_{\mathrm{l}}$ and $\Omega_t \cap \Omega_{\mathrm{r}}$, see Figure 6, we have

$$(\phi|_{\Omega_{\mathrm{l}}})|_{\tilde{\eta}^{\mathrm{o}}} = (\phi|_{\Omega_{\mathrm{r}}})|_{\tilde{\eta}^{\mathrm{o}}} \in H^{1/2}(\tilde{\eta}^{\mathrm{o}}) \quad \text{for } \phi \in W^{1,2}(\Omega), \tag{4-5}$$

with $\cdot^{\mathrm{o}}$ denoting the interior in $\gamma_{0,t} \cup \tilde{\eta}$. Therefore, the map

$$\begin{gathered} W^{1,2}(\Omega) \to H^{1/2}(\eta^{\mathrm{o}}) \oplus H^{1/2}(\eta^{\mathrm{o}}) \oplus H^{1/2}(\tilde{\eta}^{\mathrm{o}}), \\ \phi \mapsto ((\phi|_{\Omega_{\mathrm{l}}})|_{\eta^{\mathrm{o}}}, (\phi|_{\Omega_{\mathrm{r}}})|_{\eta^{\mathrm{o}}}, \phi|_{\tilde{\eta}^{\mathrm{o}}}), \end{gathered} \tag{4-6}$$

is well-defined, where $\phi|_{\tilde{\eta}^{\mathrm{o}}}$ denotes the common value in (4-5). We first analyze the normal derivatives on $\tilde{\eta}$.

Lemma 4.5. *Let Ω be a cracked Neumann domain. If $\psi \in W^{1,2}(\Omega)$ and $\Delta\psi \in L^2(\Omega)$, then*

$$\partial_\nu(\psi|_{\Omega_{\mathrm{l}}})|_{\tilde{\eta}^{\mathrm{o}}} + \partial_\nu(\psi|_{\Omega_{\mathrm{r}}})|_{\tilde{\eta}^{\mathrm{o}}} = 0 \in H^{-1/2}(\tilde{\eta}^{\mathrm{o}}). \tag{4-7}$$

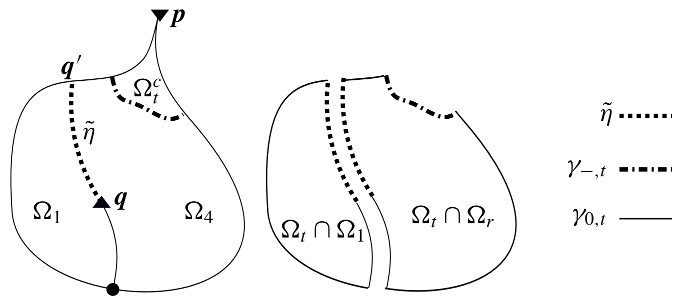

Figure 6. The dissected and truncated domains appearing in the proof of Proposition 4.3. Here $\gamma_{\pm,t}$ is a result of the truncation, $\tilde{\eta}$ is from the dissection, and $\gamma_{0,t}$ is the part of the original boundary, $\partial\Omega$, that remains after the truncation.

Proof. The hypothesis $\Delta\psi \in L^2(\Omega)$ means

$$\int_\Omega \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle = \langle \Delta\psi, \phi\rangle_{L^2(\Omega)} \tag{4-8}$$

for all $\phi \in C_0^\infty(\Omega)$. Together with Green's formula (3-10), this implies

$$\int_{\tilde{\eta}} (\partial_\nu(\psi|_{\Omega_{\mathrm{l}}}) + \partial_\nu(\psi|_{\Omega_{\mathrm{r}}}))\phi = 0 \tag{4-9}$$

for all $\phi \in C_0^\infty(\Omega)$. Since any function in $C_0^\infty(\tilde{\eta})$ can be realized as $\phi|_{\tilde{\eta}}$ for some $\phi \in C_0^\infty(\Omega)$, the result follows from (3-6). □

We next analyze the normal derivative on the Lipschitz part of the boundary, $\gamma_{0,t}$.

Lemma 4.6. *If* $\psi \in \mathcal{D}(\Delta_\Omega^N)$, *then*

$$\partial_\nu(\psi|_{\Omega_t\cap\Omega_\bullet})|_{\gamma_{0,t}^{\mathrm{o}}} = 0 \in H^{-1/2}((\partial(\Omega_t \cap \Omega_\bullet) \cap \gamma_{0,t})^{\mathrm{o}}) \tag{4-10}$$

for any $0 < t < 1$, *where* $\bullet = \mathrm{l}, \mathrm{r}$.

Proof. We prove the result for Ω_{l}, the argument for Ω_{r} is identical. For any test function $\phi \in W^{1,2}(\Omega_t \cap \Omega_{\mathrm{l}})$ with $\phi|_{\tilde{\eta}} = 0$ and $\phi|_{\gamma_{+,t}} = 0$, we get from Green's formula (3-10) that

$$\int_{\partial(\Omega_t\cap\Omega_{\mathrm{l}})\cap\gamma_{0,t}} \partial_\nu(\psi|_{\Omega_t\cap\Omega_{\mathrm{l}}})\phi = 0.$$

The image of the trace map restricted to

$$\{\phi \in W^{1,2}(\Omega_t \cap \Omega_{\mathrm{l}}) : \phi|_{\tilde{\eta}} = \phi|_{\gamma_{+,t}} = 0\}$$

is precisely $\widetilde{H}^{1/2}\big((\partial(\Omega_t\cap\Omega_{\mathrm{l}})\cap\gamma_{0,t})^{\mathrm{o}}\big)$, by (3-3) and [McLean 2000, Theorem 3.37], so the result follows. □

Now, equipped with our preliminary analysis of normal derivatives, we prove Proposition 4.3.

Proof of Proposition 4.3. We only prove (4-4); the proof of (4-3) for regular domains is similar but less involved, so we omit it. Let $\psi \in W^{1,2}(\Omega)$. From (4-2) we have that $\psi \in \mathcal{D}(\Delta_\Omega^N)$ if and only if

$\Delta\psi \in L^2(\Omega)$ and

$$\int_\Omega (\Delta\psi)\phi = \int_\Omega \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle \tag{4-11}$$

for all $\phi \in W^{1,2}(\Omega)$. Thus, we fix $\psi, \phi \in W^{1,2}(\Omega)$ with $\Delta\psi \in L^2(\Omega)$. Since $\langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle$ and $(\Delta\psi)\phi$ are in $L^1(\Omega)$, their integrals over Ω_t converge to their integrals over Ω as $t \to 1$ by the dominated convergence theorem; hence (4-11) is equivalent to

$$\lim_{t\to 1}\int_{\Omega_t} (\Delta\psi)\phi = \lim_{t\to 1}\int_{\Omega_t} \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle. \tag{4-12}$$

We now use the dissection (3-11), applying Green's formula on the truncated and dissected domains to obtain

$$\int_{\Omega_t} \langle \operatorname{grad}\psi, \operatorname{grad}\phi\rangle = \int_{\Omega_t} (\Delta\psi)\phi + \int_{\partial(\Omega_t\cap\Omega_\mathrm{l})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})\phi + \int_{\partial(\Omega_t\cap\Omega_\mathrm{r})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi.$$

Comparing with (4-12), we see that $\psi \in \mathcal{D}(\Delta_\Omega^N)$ if and only if

$$\lim_{t\to 1}\left\{\int_{\partial(\Omega_t\cap\Omega_\mathrm{l})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})\phi + \int_{\partial(\Omega_t\cap\Omega_\mathrm{r})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi\right\} = 0 \tag{4-13}$$

for each $\phi \in W^{1,2}(\Omega)$. Therefore, it suffices to show that (4-13) is equivalent to the conditions in (4-4).

We claim that if $\psi \in W^{1,2}(\Omega)$ satisfies $\Delta\psi \in L^2(\Omega)$ and $\partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})|_{\gamma_{0,t}^\circ} = \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})|_{\gamma_{0,t}^\circ} = 0$, then

$$\int_{\partial(\Omega_t\cap\Omega_\mathrm{l})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})\phi + \int_{\partial(\Omega_t\cap\Omega_\mathrm{r})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi = \int_{\gamma_{\pm,t}} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi \tag{4-14}$$

for any $0 < t < 1$ and $\phi \in W^{1,2}(\Omega)$. To prove this, we decompose the integrals over $\partial(\Omega_t \cap \Omega_\mathrm{l})$ and $\partial(\Omega_t \cap \Omega_\mathrm{r})$ into a sum of integrals over the different parts of the boundary. This is nontrivial, since this integral notation actually represents the action of the normal derivative distribution on a test function in $H^{1/2}$, and a distribution in $H^{-1/2}$ does not necessarily split into the sum of its restriction to different parts of the boundary, as discussed in Section 3A.

Here we make use of Lemma 3.1, as well as the assumption that $\partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})|_{\gamma_{0,t}^\circ} = \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})|_{\gamma_{0,t}^\circ} = 0$. Applying the lemma to $N = \Omega_t \cap \Omega_\mathrm{l}$, with the boundary decomposed into $\Gamma_1 = \tilde{\eta}$, $\Gamma_2 = \gamma_{+,t}$ and $\Gamma_0 = \partial(\Omega_t \cap \Omega_\mathrm{l}) \cap \gamma_{0,t}$, we obtain

$$\int_{\partial(\Omega_t\cap\Omega_\mathrm{l})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})\phi = \int_{\tilde{\eta}} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})\phi + \int_{\gamma_{+,t}} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})\phi.$$

Similarly, for $\Omega_t \cap \Omega_\mathrm{r}$ we get

$$\int_{\partial(\Omega_t\cap\Omega_\mathrm{r})} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi = \int_{\tilde{\eta}} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi + \int_{\gamma_{-,t}} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi.$$

Adding these together and using Lemma 4.5 to cancel the $\tilde{\eta}$ terms completes the proof of (4-14).

To finish the proof of the proposition, suppose that $\psi \in \mathcal{D}(\Delta_\Omega^N)$, so (4-13) holds. Lemma 4.6 implies that $\partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{l}})|_{\gamma_{0,t}^\circ} = \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})|_{\gamma_{0,t}^\circ} = 0$, so we can use (4-14) to conclude that

$$\lim_{t\to 1}\int_{\gamma_{-,t}} \partial_\nu(\psi|_{\Omega_t\cap\Omega_\mathrm{r}})\phi = 0$$

for any $\phi \in W^{1,2}(\Omega)$. Therefore, the boundary conditions given in (4-4) are satisfied. Conversely, if ψ satisfies the boundary conditions in (4-4), we take the limit of (4-14) to find that (4-13) holds and so $\psi \in \mathcal{D}(\Delta_\Omega^N)$. ☐

Remark 4.7. For the $\gamma_{\pm,t}$ boundary condition in (4-3) or (4-4), it is enough to check that

$$\lim_{t\to 1}\int_{\gamma_{\pm,t}} \partial_\nu(\psi|_{\Omega_t})\phi = 0 \tag{4-15}$$

for test functions $\phi \in W^{1,2}(\Omega)$ that are C^1 in a neighbourhood of the cusp. If ψ satisfies this, and the other conditions in (4-3) or (4-4), the proof of Proposition 4.3 shows that (4-11) holds for all such ϕ. It then follows from Proposition 3.3(3) that (4-11) in fact holds for all $\phi \in W^{1,2}(\Omega)$, and so $\psi \in \mathcal{D}(\Delta_\Omega^N)$.

4C. *Proof of Theorem 1.3.* If Ω has no cusps or cracks, then Proposition 4.3 says that $\mathcal{D}(\Delta_\Omega^N)$ simply consists of functions that are sufficiently regular and satisfy Neumann boundary conditions everywhere on $\partial\Omega$. On the other hand, when a cusp is present we must also impose the condition (4-15), which says the normal derivative of ψ on $\gamma_{\pm,t}$ does not blow up as the cusp is approached. We now give a simple condition that guarantees this is the case.

For simplicity we only state the result for a cusp at $\boldsymbol{q}$; the corresponding statement for a cusp at $\boldsymbol{p}$ is analogous. We define the "doubly truncated domain"

$$\Omega_t' = \{x \in \Omega : t_0 f(\boldsymbol{q}) < f(x) < t f(\boldsymbol{q})\} \tag{4-16}$$

for a fixed $0 < t_0 < 1$.

Proposition 4.8. *If $\psi \in W^{1,2}(\Omega)$ and there exists t_0 such that $\psi \in W^{2,2}(\Omega_t')$ for all $t_0 < t < 1$ and $(1-t)^{1/2}\|\psi\|^2_{W^{2,2}(\Omega_t')}$ is bounded near $t = 1$, then* (4-15) *holds.*

The proposition does not assume ψ is in $W^{2,2}(\Omega)$, but only that its $W^{2,2}(\Omega_t')$ norm does not blow up too quickly near the cusp. Of course this condition is automatically satisfied if $\psi \in W^{2,2}(\Omega)$.

Corollary 4.9. *If $\psi \in W^{2,2}(\Omega)$, then* (4-15) *holds.*

Since the Morse function f that generated the Neumann domain Ω was assumed to be smooth, $f|_\Omega$ satisfies the hypotheses of Corollary 4.9, and Theorem 1.3 follows immediately.

The main ingredient in the proof is a trace estimate for the doubly truncated domain Ω_t', with controlled dependence on t.

Lemma 4.10. *There exist constants $A, B > 0$ such that*

$$\int_{\gamma_{\pm,t}} u^2 \le \frac{A}{\sqrt{1-t}}\int_{\Omega_t'} u^2 + B\int_{\Omega_t'} |\nabla u|^2 \tag{4-17}$$

for all $u \in W^{1,2}(\Omega_t')$ and t sufficiently close to 1.

Proof. Increasing t_0 if necessary, we can assume that $\boldsymbol{q}$ is the only critical point of f in $\overline{\Omega}^c_{t_0}$. Consider the vector field

$$X := \chi u^2 \frac{\nabla f}{|\nabla f|},$$

where χ is a smooth cutoff function that vanishes on γ_{+,t_0} and equals 1 in a neighbourhood of $\boldsymbol{q}$. Since f is smooth and has no critical points in $\overline{\Omega}'_t$, we have $X \in W^{1,1}(\Omega'_t)$. Observe that $\nabla f/|\nabla f|$ is tangent to $\gamma_{0,t}$, whereas on $\gamma_{+,t}$ it coincides with the outward unit normal. This implies

$$\int_{\partial\Omega'_t} X \cdot \nu = \int_{\gamma_{+,t}} u^2$$

for any t large enough that $\chi|_{\gamma_{+,t}} \equiv 1$. On the other hand, the divergence theorem implies

$$\int_{\partial\Omega'_t} X \cdot \nu = \int_{\Omega'_t} \operatorname{div} X = \int_{\Omega'_t} \left(\nabla(\chi u^2) \cdot \frac{\nabla f}{|\nabla f|} + \chi u^2 \operatorname{div} \frac{\nabla f}{|\nabla f|} \right),$$

so we obtain

$$\int_{\gamma_{+,t}} u^2 \le B\|u\|^2_{H^1(\Omega'_t)} + \int_{\Omega'_t} u^2 \left| \operatorname{div} \frac{\nabla f}{|\nabla f|} \right| \tag{4-18}$$

for some constant B depending only on χ.

To estimate the integral on the right-hand side, we observe that the level sets of f have mean curvature $\operatorname{div}(\nabla f/|\nabla f|)$. Using the Morse lemma, we can find coordinates (x, y) in a neighbourhood of $\boldsymbol{q}$ such that $f(x, y) = f(\boldsymbol{q}) - x^2 - y^2$. A straightforward computation (see [Beck et al. 2021, Lemma 4.7]) gives

$$\left| \operatorname{div} \frac{\nabla f}{|\nabla f|}(x, y) \right| \le \frac{C}{\sqrt{x^2 + y^2}} = \frac{C}{\sqrt{f(\boldsymbol{q}) - f(x, y)}}$$

and so we have the uniform estimate

$$\left| \operatorname{div} \frac{\nabla f}{|\nabla f|} \right| \le \frac{C}{\sqrt{f(\boldsymbol{q})(t-1)}}$$

on $\overline{\Omega}'_t$. Substituting this into (4-18) completes the proof. ☐

The other ingredient in the proof of Proposition 4.8 is the following geometric estimate.

Lemma 4.11. *The length of $\gamma_{\pm,t}$ is $o((1-t)^{1/2})$ near $t = 1$.*

Proof. We prove the result for $\gamma_{+,t}$, assuming there is a cusp at the maximum $\boldsymbol{q}$; the proof for $\gamma_{-,t}$ is identical. Using the Morse lemma, we can find coordinates (x, y) near $\boldsymbol{q}$ such that $f(x, y) = f(\boldsymbol{q}) - x^2 - y^2$, and so $\gamma_{+,t}$ is contained in the circle of radius $\rho = \sqrt{(1-t) f(\boldsymbol{q})}$. More precisely, it is the arc bounded by the angles $\theta_1(t)$ and $\theta_2(t)$. Parametrizing this as $\gamma(\theta) = (\rho\cos\theta, \rho\sin\theta)$, we have $|\gamma'(\theta)|_g \le C\sqrt{1-t}$, where $|\cdot|_g$ denotes the length computed using the metric g and C is some constant depending on $f(\boldsymbol{q})$ and the components of g in this coordinate chart. This implies

$$L(\gamma_{+,t}) = \int_{\theta_1(t)}^{\theta_2(t)} |\gamma'(\theta)|_g \, d\theta \le C\sqrt{1-t}\,|\theta_2(t) - \theta_1(t)|.$$

Near $\boldsymbol{q}$, the boundary $\partial\Omega$ consists of two Neumann lines meeting tangentially at $\boldsymbol{q}$ (since there is a cusp). This implies $|\theta_2(t) - \theta_1(t)| \to 0$ as $t \to 1$ and completes the proof. ☐

We are now ready to prove Proposition 4.8.

Proof. Since $|\partial_\nu \psi| \le |\nabla\psi|$, it is enough to show that

$$\lim_{t\to 1} \int_{\gamma_{\pm,t}} |\nabla\psi|\phi = 0 \tag{4-19}$$

for all $\phi \in W^{1,2}(\Omega)$ that are C^1 in a neighbourhood of $\boldsymbol{q}$; see Remark 4.7. Fix such a ϕ and define $u = \chi|\nabla\psi|\phi$, where χ is a smooth cutoff function that equals 1 near $\boldsymbol{q}$ and is supported in the region where ϕ is smooth. The hypotheses on ψ imply $u \in L^2(\Omega)$ and $u \in W^{1,2}(\Omega_t')$ for all $t_0 < t < 1$, with

$$\|u\|_{L^2(\Omega)} \le C\|\psi\|_{W^{1,2}(\Omega)}, \quad \|u\|_{W^{1,2}(\Omega_t')} \le C\|\psi\|_{W^{2,2}(\Omega_t')}$$

for some constant C depending only on ϕ and χ.

Using Hölder's inequality and Lemma 4.10, we obtain

$$\begin{aligned}\left(\int_{\gamma_{\pm,t}} u\right)^2 &\le \left(\int_{\gamma_{\pm,t}} u^2\right) L(\gamma_{\pm,t})\\ &\le \left(\frac{A}{\sqrt{1-t}}\|u\|^2_{L^2(\Omega_t)} + B\|u\|^2_{W^{1,2}(\Omega_t')}\right) L(\gamma_{\pm,t})\\ &\le (A\|u\|^2_{L^2(\Omega)} + B\sqrt{1-t}\,\|u\|^2_{W^{1,2}(\Omega_t')})\frac{L(\gamma_{\pm,t})}{\sqrt{1-t}}.\end{aligned}$$

By Lemma 4.11 this tends to zero as $t \to 1$. □

Appendix: Morse–Smale functions with cracked Neumann domains

In this appendix we construct Morse–Smale functions having cracked Neumann domains. As in the rest of the paper, we assume M is a smooth, closed, connected orientable surface.

Theorem A.1. *Let f be a Morse–Smale function on M and Ω a Neumann domain of f. Then there exists a Morse–Smale function $\tilde f$ that has a cracked Neumann domain $\widetilde{\Omega} \subset \Omega$.*

We will see in the proof that $\tilde f$ can be chosen to agree with f outside an arbitrary open set $U \subset \Omega$. However, the difference $\tilde f - f$ may be large inside U. The existence of $\tilde f$ is given by the following general lemma.

Lemma A.2. *Let $U \subset M$ be an open subset and $f : U \to \mathbb{R}$ a smooth function having no critical points. There exists a smooth function $\tilde f : U \to \mathbb{R}$, with $\operatorname{supp}(\tilde f - f) \subset U$, whose only critical points are a nondegenerate maximum and a nondegenerate saddle.*

Proof. Since f has no critical points in U, we can invoke the canonical form theorem for smooth vector fields and find local coordinates (x, y) with respect to which $f(x, y) = Ax + B$ for $(x, y) \in (-1, 1) \times (-1, 1)$. Now choose a smooth function $\alpha(x)$ with $\operatorname{supp} \alpha \subset (-1, 1)$ and

$$\int_{-1}^{1} \alpha(x)\, dx = 0, \tag{A-1}$$

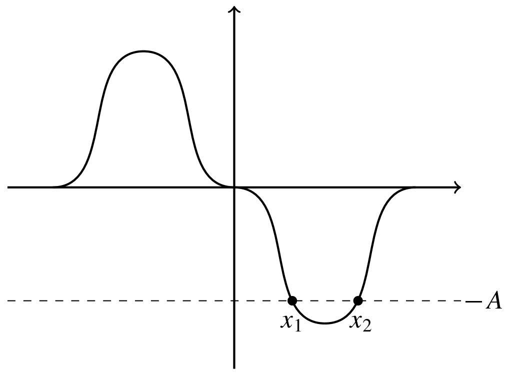

Figure 7. The function $\alpha(x)$ used in the proof of Lemma A.2.

so that there exist points $-1 < x_1 < x_2 < 1$ with

$$\begin{aligned}
&\alpha(x) > -A, && -1 < x < x_1,\\
&\alpha(x) = -A, && x = x_1,\\
&\alpha(x) < -A, && x_1 < x < x_2,\\
&\alpha(x) = -A, && x = x_2,\\
&\alpha(x) > -A, && x_2 < x < 1,
\end{aligned} \tag{A-2}$$

as shown in Figure 7.

We define

$$\tilde{f}(x, y) = f(x, y) + \beta(x)\gamma(y), \tag{A-3}$$

where $\beta(x) = \int_{-1}^{x} \alpha(t)\,dt$ and $\gamma(y) = \exp\{-1/(1-y^2)\}$. Note that γ is a nonnegative bump function supported in $(-1, 1)$ with $\gamma'(0) = 0$ and $\gamma''(0) < 0$. It follows that

$$\frac{\partial \tilde{f}}{\partial x} = A + \alpha(x)\gamma(y) \quad \text{and} \quad \frac{\partial \tilde{f}}{\partial y} = \beta(x)\gamma'(y),$$

and so the only critical points of $\tilde{f}$ in U are $(x_1, 0)$ and $(x_2, 0)$. We compute

$$\begin{aligned}
\frac{\partial^2 \tilde{f}}{\partial x^2}(x_1, 0) &= \alpha'(x_1)\gamma(0) < 0,\\
\frac{\partial^2 \tilde{f}}{\partial x^2}(x_2, 0) &= \alpha'(x_2)\gamma(0) > 0,\\
\frac{\partial^2 \tilde{f}}{\partial y^2}(x_i, 0) &= \beta(x_i)\gamma''(0) < 0,
\end{aligned}$$

and conclude $(x_1, 0)$ and $(x_2, 0)$ are a nondegenerate maximum and a nondegenerate saddle, respectively. □

Proof of Theorem A.1. If Ω is cracked we simply choose $\tilde{f} = f$ and there is nothing to prove. Therefore we assume that Ω is regular. Since f is Morse–Smale, Theorem 2.2 says the closure of Ω contains exactly four critical points, all of which are on the boundary: a maximum $\boldsymbol{q}$, a minimum $\boldsymbol{p}$, and saddle points $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$; see Figure 3.

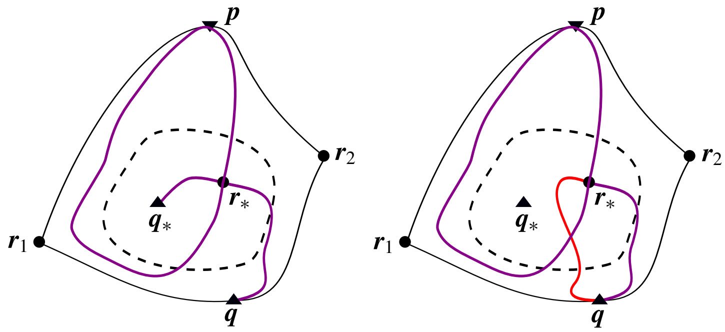

Figure 8. Left: The cracked Neumann domain constructed in Theorem A.1, with Neumann lines shown in purple. Right: If two Neumann lines connected $\boldsymbol{r}_*$ to $\boldsymbol{q}$, one of them (shown in red) would have to intersect another Neumann line, which is impossible. The dashed line represents the boundary of the set U containing $\operatorname{supp}(\tilde{f} - f)$.

Now choose $\tilde{f}$ according to Lemma A.2, for some open set $U \Subset \Omega$. By construction, $\tilde{f}$ has two critical points in Ω: a maximum $\boldsymbol{q}_*$ and a saddle point $\boldsymbol{r}_*$. Since $\tilde{f}$ is a Morse function, $\boldsymbol{r}_*$ has degree 4; i.e., there are four Neumann lines connected to $\boldsymbol{r}_*$. We obtain the result by studying the endpoints of these lines, as depicted in Figure 8. Since $\tilde{f}$ agrees with f in a neighbourhood of $\partial\Omega$, the invariant manifolds $W^s(\boldsymbol{r}_i)$ and $W^u(\boldsymbol{r}_i)$ are unchanged by the perturbation. As a result, it is not possible for any of the Neumann lines coming from $\boldsymbol{r}_*$ to end at $\boldsymbol{r}_1$ or $\boldsymbol{r}_2$. Therefore, the four Neumann lines from $\boldsymbol{r}_*$ can only end at $\boldsymbol{q}$, $\boldsymbol{p}$ or $\boldsymbol{q}_*$, so it follows from Lemma 2.4 that $\tilde{f}$ is Morse–Smale. The two lines along which $\tilde{f}$ is decreasing must end at $\boldsymbol{p}$, since it is the only minimum in $\overline{\Omega}$. This means the two lines along which f is increasing are connected to either $\boldsymbol{q}$ or $\boldsymbol{q}_*$. We claim that there is one Neumann line connected to each maximum.

Suppose instead that both ended at $\boldsymbol{q}$. Then the union of these Neumann lines forms a closed loop. Similarly, the union of the two lines ending at $\boldsymbol{p}$ is a closed loop. Both loops intersect at $\boldsymbol{r}_*$, where they are orthogonal by Proposition 2.5(1). Since Ω is simply connected, this can only happen if the loops also intersect at a point other than $\boldsymbol{r}_*$, but this is impossible since gradient flow lines cannot cross. The same argument shows that these lines cannot both be connected to $\boldsymbol{q}_*$; hence one must end at each maximum.

Since all of the Neumann lines in $\overline{\Omega}$ have been accounted for, this means $\boldsymbol{q}_*$ has degree 1; hence the Neumann domain with $\boldsymbol{q}_*$ on its boundary is cracked. □

Acknowledgments

Cox acknowledges the support of NSERC grant RGPIN-2017-04259. Band and Egger were supported by the ISF (grant no. 844/19).

References

[Alon et al. 2020] L. Alon, R. Band, M. Bersudsky, and S. Egger, "Neumann domains on graphs and manifolds", pp. 203–249 in *Analysis and geometry on graphs and manifolds* (Potsdam, Germany, 2017), edited by M. Keller et al., Lond. Math. Soc. Lect. Note Ser. **461**, Cambridge Univ. Press, 2020. MR Zbl

[Band and Fajman 2016] R. Band and D. Fajman, "Topological properties of Neumann domains", *Ann. Henri Poincaré* **17**:9 (2016), 2379–2407. MR Zbl

[Band et al. 2021] R. Band, S. K. Egger, and A. J. Taylor, "The spectral position of Neumann domains on the torus", *J. Geom. Anal.* **31**:5 (2021), 4561–4585. MR Zbl

[Banyaga and Hurtubise 2004] A. Banyaga and D. Hurtubise, *Lectures on Morse homology*, Kluwer Texts Math. Sci. **29**, Kluwer, Dordrecht, 2004. MR Zbl

[Beck et al. 2021] M. Beck, G. Cox, C. Jones, Y. Latushkin, and A. Sukhtayev, "A dynamical approach to semilinear elliptic equations", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **38**:2 (2021), 421–450. MR Zbl

[Cheng 1976] S. Y. Cheng, "Eigenfunctions and nodal sets", *Comment. Math. Helv.* **51**:1 (1976), 43–55. MR Zbl

[Edmunds and Evans 1987] D. E. Edmunds and W. D. Evans, *Spectral theory and differential operators*, Oxford Univ. Press, 1987. MR Zbl

[Flamencourt and Pankrashkin 2020] B. Flamencourt and K. Pankrashkin, "Strong coupling asymptotics for δ-interactions supported by curves with cusps", *J. Math. Anal. Appl.* **491**:1 (2020), art. id. 124287. MR Zbl

[Hartman 1960] P. Hartman, "On local homeomorphisms of Euclidean spaces", *Bol. Soc. Mat. Mexicana* (2) **5** (1960), 220–241. MR Zbl

[Hempel et al. 1991] R. Hempel, L. A. Seco, and B. Simon, "The essential spectrum of Neumann Laplacians on some bounded singular domains", *J. Funct. Anal.* **102**:2 (1991), 448–483. MR Zbl

[Jakšić et al. 1992] V. Jakšić, S. Molčanov, and B. Simon, "Eigenvalue asymptotics of the Neumann Laplacian of regions and manifolds with cusps", *J. Funct. Anal.* **106**:1 (1992), 59–79. MR Zbl

[Mazya and Poborchi 1997] V. G. Mazya and S. V. Poborchi, *Differentiable functions on bad domains*, World Sci., River Edge, NJ, 1997. MR Zbl

[McDonald and Fulling 2014] R. B. McDonald and S. A. Fulling, "Neumann nodal domains", *Philos. Trans. Roy. Soc. A* **372**:2007 (2014), art. id. 20120505. MR Zbl

[McLean 2000] W. McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge Univ. Press, 2000. MR Zbl

[Perko 2001] L. Perko, *Differential equations and dynamical systems*, 3rd ed., Texts in Appl. Math. **7**, Springer, 2001. MR Zbl

[Reed and Simon 1972] M. Reed and B. Simon, *Methods of modern mathematical physics*, *I*: *Functional analysis*, Academic Press, New York, 1972. MR Zbl

[Reed and Simon 1978] M. Reed and B. Simon, *Methods of modern mathematical physics*, *IV*: *Analysis of operators*, Academic Press, New York, 1978. MR Zbl

[Taylor 2018] A. J. Taylor, "pyneumann toolkit", Python code, 2018, available at https://github.com/inclement/neumann.

[Uhlenbeck 1976] K. Uhlenbeck, "Generic properties of eigenfunctions", *Amer. J. Math.* **98**:4 (1976), 1059–1078. MR Zbl

Received 23 Apr 2021. Accepted 28 Feb 2022.

RAM BAND: ramband@technion.ac.il
Department of Mathematics, Technion - Israel Institute of Technology, Haifa, Israel

and

Institute of Mathematics, University of Potsdam, Potsdam, Germany

GRAHAM COX: gcox@mun.ca
Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, Canada

SEBASTIAN K. EGGER: sebastian.k.egger@gmail.com
Department of Mathematics, Technion - Israel Institute of Technology, Haifa, Israel

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2173–2224
DOI: 10.2140/apde.2023.16.2173

A UNIQUENESS RESULT FOR THE TWO-VORTEX TRAVELING WAVE IN THE NONLINEAR SCHRÖDINGER EQUATION

DAVID CHIRON AND ELIOT PACHERIE

For the nonlinear Schrödinger equation in dimension 2, the existence of a global minimizer of the energy at fixed momentum has been established by Bethuel, Gravejat and Saut (2009) (see also work of Chiron and Mariş (2017)). This minimizer is a traveling wave for the nonlinear Schrödinger equation. For large momenta, the propagation speed is small and the minimizer behaves like two well-separated vortices. In that limit, we show the uniqueness of this minimizer, up to the invariances of the problem, hence proving the orbital stability of this traveling wave. This work is a follow up to two previous papers, where we constructed and studied a particular traveling wave of the equation. We show a uniqueness result on this traveling wave in a class of functions that contains in particular all possible minimizers of the energy.

1. Introduction and statement of the results

We consider the nonlinear Schrödinger equation

$$i\partial_t \Psi + \Delta\Psi - (|\Psi|^2 - 1)\Psi = 0 \tag{NLS}$$

in dimension 2 for $\Psi : \mathbb{R}_t \times \mathbb{R}^2_x \to \mathbb{C}$, also called the Gross–Pitaevskii equation without potential. The nonlinear Schrödinger equation is a physical model for Bose–Einstein condensation [1; 23; 37; 42], superfluidity [40] and nonlinear optics [30]. The condition at infinity for (NLS) will be

$$|\Psi| \to 1 \quad \text{as } |x| \to +\infty.$$

The (NLS) equation is associated with the Ginzburg–Landau energy

$$E(v) := \frac{1}{2}\int_{\mathbb{R}^2} |\nabla v|^2 + \frac{1}{4}\int_{\mathbb{R}^2} (1 - |v|^2)^2,$$

which is formally conserved by the (NLS) flow. We denote by $\mathcal{E}$ the set of functions with finite energy, that is,

$$\mathcal{E} := \{u \in H^1_{\mathrm{loc}}(\mathbb{R}^2, \mathbb{C}) : E(u) < +\infty\}.$$

Remark 1.1. The Cauchy problem for (NLS) is globally well-posed in the energy space; see [20; 21; 22].

Besides the energy, the momentum is another quantity formally conserved by the (NLS) flow and is associated with the invariance by translation of (NLS). Formally, the momentum of u is $\frac{1}{2}\int_{\mathbb{R}^2} \mathfrak{Re}(i\nabla u\bar{u}) \in \mathbb{R}^2$, but its precise definition requires some care in the energy space due to the condition at infinity (see [34]

MSC2020: 35A02, 35A15, 35B35, 35C07, 35Q56.
Keywords: Gross–Pitaevskii, uniqueness, traveling waves, vortices.

in dimension larger than 2 and [13] in dimension 2). If $u \in 1 + \mathcal{C}_c^\infty(\mathbb{R}^2)$ for instance, or if u is a traveling wave tending to 1 at infinity, then the expression of the momentum reduces to

$$\vec{P}(u) = (P_1(u), P_2(u)) = \frac{1}{2} \int_{\mathbb{R}^2} \mathfrak{Re}(i \nabla u (\bar{u} - 1)).$$

In addition to the translation invariance, the (NLS) equation is also phase-shift-invariant, that is, invariant by multiplication by a complex of modulus 1, and rotation-invariant.

1A. ***Traveling waves for* (NLS).** Following the works in the physical literature of Jones and Roberts [28; 29], there has been a large number of mathematical works on the question of existence and properties of traveling wave solutions in the (NLS) equation, which are solutions of

$$0 = (\mathrm{TW}_c)(u) := -ic\,\partial_{x_2} u - \Delta u - (1 - |u|^2)u$$

for some $c > 0$, corresponding to particular solutions of (NLS) of the form $\Psi(t, x) = u(x_1, x_2 + ct)$ (due to the rotational invariance, we may always assume that the traveling wave moves along the direction $-\vec{e}_2$). We refer to [9] for an overview on these problems in several dimensions. A natural approach is to look at the minimizing problem for $\mathfrak{p} > 0$

$$E_{\min}(\mathfrak{p}) := \inf_{u \in \mathcal{E}} \{E(u) : P_2(u) = \mathfrak{p}\}.$$

It was shown by Bethuel, Gravejat and Saut that there exists a minimizer to this problem.

Theorem 1.2 [10]. *For any $\mathfrak{p} > 0$, there exists a nonconstant function $u_\mathfrak{p} \in \mathcal{E}$ and $c(u_\mathfrak{p}) > 0$ such that $P_2(u_\mathfrak{p}) = \mathfrak{p}$, $u_\mathfrak{p}$ is a solution to $(\mathrm{TW}_{c(u_\mathfrak{p})})(u_\mathfrak{p}) = 0$ and*

$$E(u_\mathfrak{p}) = E_{\min}(\mathfrak{p}).$$

Furthermore, any minimizer for $E_{\min}(\mathfrak{p})$ is, up to a translation in x_1, even in x_1.

The strategy is to look at the corresponding minimization problem on larger and larger tori (this avoids the problems with the definition of the momentum), and then pass to the limit. For the minimizing problem $E_{\min}(\mathfrak{p})$, the compactness of minimizing sequences has been shown later on in [13] for the natural semidistance on $\mathcal{E}$

$$D_0(u, v) := \|\nabla u - \nabla v\|_{L^2(\mathbb{R}^2)} + \||u| - |v|\|_{L^2(\mathbb{R}^2)}.$$

Theorem 1.3 [13]. *For any $\mathfrak{p} > 0$ and any minimizing sequence $(u_n)_{n\in\mathbb{N}}$ for $E_{\min}(\mathfrak{p})$, there exists a subsequence $(u_{n_j})_{j\in\mathbb{N}}$, a sequence of translations $(y_j)_{j\in\mathbb{N}}$ and a nonconstant function $u_\mathfrak{p} \in \mathcal{E}$ such that $D_0(u_{n_j}, u_\mathfrak{p}) \to 0$, $P_2(u_{n_j}) \to P_2(u_\mathfrak{p}) = \mathfrak{p}$ and $E(u_{n_j}) \to E(u_\mathfrak{p}) = E_{\min}(\mathfrak{p})$ as $j \to +\infty$. In particular, there exists $c(u_\mathfrak{p}) > 0$ such that $P_2(u_\mathfrak{p}) = \mathfrak{p}$, $u_\mathfrak{p}$ is a solution to $(\mathrm{TW}_{c(u_\mathfrak{p})})(u_\mathfrak{p}) = 0$ and*

$$E(u_\mathfrak{p}) = E_{\min}(\mathfrak{p}).$$

Furthermore, the set $\mathcal{S}_\mathfrak{p} := \{v \in \mathcal{E} : P_2(v) = \mathfrak{p}$ and $E(v) = E_{\min}(\mathfrak{p})\}$ of minimizers for $E_{\min}(\mathfrak{p})$ is orbitally stable for the semidistance D_0.

An open and difficult question is to show, up to the invariances of the problem, the uniqueness of the energy minimizer at fixed momentum. In other words, the problem is to determine if $\mathcal{S}_{\mathfrak{p}}$ consists of a single orbit under phase shift and space translation; that is, do we have, for some minimizer $U_{\mathfrak{p}}$,

$$\mathcal{S}_{\mathfrak{p}} = \{U_{\mathfrak{p}}(\cdot - X)e^{i\gamma} : \gamma \in \mathbb{R},\ X \in \mathbb{R}^2\}?$$

The main consequence of our work is to solve this open problem of uniqueness for large momentum.

Theorem 1.4. *There exists $\mathfrak{p}_0 > 0$ such that, for any $\mathfrak{p} > \mathfrak{p}_0$, if $u, v \in \mathcal{E}$ with $P_2(u) = P_2(v) = \mathfrak{p}$ satisfy*

$$E(u) = E(v) = E_{\min}(\mathfrak{p}),$$

then, there exist $X \in \mathbb{R}^2$ and $\gamma \in \mathbb{R}$ such that

$$u = v(\cdot - X)e^{i\gamma}.$$

In fact, we will be able to show slightly stronger results than Theorem 1.4; see Theorem 1.11 below.

Even though we focus on the Ginzburg–Landau nonlinearity, it is plausible that our results hold true (still for large momentum) for more general nonlinearities, provided vortices exist. For the Ginzburg–Landau (cubic) nonlinearity, it is also possible that uniqueness of minimizers holds true for $E_{\min}(\mathfrak{p})$ for any $\mathfrak{p} > 0$. However, the numerical results given in [16] suggest that this may no longer be the case for more general nonlinearities.

In the analysis of the minimization problem in [10] (and also [13]), the following properties of $E_{\min}$ play a key role.

Proposition 1.5 [10]. *The function $E_{\min} : \mathbb{R}_+ \to \mathbb{R}$ is concave, nondecreasing and $\sqrt{2}$-Lipschitz continuous. In addition, there exists $K \geqslant 0$ such that, for any $\mathfrak{p} \geqslant 1$, we have*

$$E_{\min}(\mathfrak{p}) \leqslant 2\pi \ln \mathfrak{p} + K. \tag{1-1}$$

1B. ***A smooth branch of traveling waves for large momentum.*** There have been several ways of constructing traveling waves of the (NLS) equation, with different approaches. For instance, we may use variational methods, such as a mountain-pass argument in [3; 5], or by minimizing the energy at fixed kinetic energy [10; 13]. Also, we have constructed in [14] a traveling wave by perturbative methods, taking for ansatz a pair of vortices, by following the Lyapunov–Schmidt reduction method as initiated in [39]. Vortices are stationary solutions of (NLS) of degrees $n \in \mathbb{Z}^*$ (see [12; 23; 26; 37; 45]):

$$V_n(x) = \rho_n(r)e^{in\theta},$$

where $x = re^{i\theta}$, solving

$$\begin{cases} \Delta V_n - (|V_n|^2 - 1)V_n = 0, \\ |V_n| \to 1 \quad \text{as } |x| \to \infty. \end{cases}$$

In the previous paper [14], we constructed solutions of (TW_c) for small values of $c > 0$ as a perturbation of two well-separated vortices (the distance between their centers is large when c is small). We have shown the following result.

Theorem 1.6 [14, Theorem 1.1; 15, Proposition 1.2]. *There exists $c_0 > 0$ a small constant such that, for any $0 < c \leqslant c_0$, there exists a solution of* (TW_c) *of the form*

$$Q_c = V_1(\cdot - d_c\vec{e}_1)V_{-1}(\cdot + d_c\vec{e}_1) + \Gamma_c,$$

where $d_c = (1 + o_{c\to 0}(1))/c$ is a C^1 function of c. This solution has finite energy; that is, $Q_c \in \mathcal{E}$, and $Q_c \to 1$ at infinity.

Furthermore, for all $2 < p \leqslant +\infty$, there exists $c_0(p) > 0$ such that, if $0 < c \leqslant c_0(p)$, for the norm

$$\|h\|_p := \|h\|_{L^p(\mathbb{R}^2)} + \|\nabla h\|_{L^{p-1}(\mathbb{R}^2)}$$

and the space $X_p := \{f \in L^p(\mathbb{R}^2) : \nabla f \in L^{p-1}(\mathbb{R}^2)\}$, one has

$$\|\Gamma_c\|_p = o_{c\to 0}(1).$$

In addition,

$$c \mapsto Q_c - 1 \in C^1(]0, c_0(p)[, X_p),$$

with the estimate

$$\left\| \partial_c Q_c + \left(\frac{1 + o_{c\to 0}(1)}{c^2}\right)\partial_d(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1))_{|d=d_c} \right\|_p = o_{c\to 0}\left(\frac{1}{c^2}\right).$$

Finally, we have

$$\frac{d}{dc}(P_2(Q_c)) = \frac{-2\pi + o_{c\to 0}(1)}{c^2} < 0;$$

hence the C^1 mapping

$$\mathcal{P} :]0, c_0] \to \mathbb{R}, \quad c \mapsto P_2(Q_c),$$

is a strictly decreasing diffeomorphism from $]0, c_0]$ onto $[P_2(Q_{c_0}), +\infty[$.

Remark 1.7. With the same kind of approach, [33] also provides an existence result of traveling waves for (NLS), including some cases with more than two vortices. Our result has the advantage of showing the smoothness of the branch with respect to the speed. In particular, with the last part of Theorem 1.6, we see that we may also parametrize the branch $c \mapsto Q_c$ by its momentum $\mathcal{P}$.

It is conjectured that all these constructions yield the same branch of traveling waves (for large momentum) when they are all defined, and that they are the solutions numerically observed in [16; 28] for more general nonlinearities (see also [17]). We will show here that the construction of Theorem 1.6 yields the unique, up to the natural translation and phase invariances, constrained energy minimizers.

1C. ***A uniqueness result for symmetric functions.*** We have shown in [15] several coercivity results for the traveling waves constructed in Theorem 1.6. This will allow us to show the following uniqueness result for symmetric functions close to the branch constructed in Theorem 1.6. There, for $d \in \mathbb{R}$, we use the notation $\tilde{r}_d = \min(|\cdot - d\vec{e}_1|, |\cdot + d\vec{e}_1|)$.

Proposition 1.8. *There exists $\lambda_* > 1$ such that, for any $\lambda \geqslant \lambda_*$, there exists $\varepsilon(\lambda) > 0$ such that if a function $u \in \mathcal{E}$ satisfies*

(1) *for all $(x_1, x_2) \in \mathbb{R}^2$, $u(x_1, x_2) = u(-x_1, x_2)$,*

(2) *$u = V_1(x - d\vec{e}_1)V_{-1}(x + d\vec{e}_1) + \Gamma$, with $d > 1/\varepsilon(\lambda)$, $\|\Gamma\|_{L^\infty(\{\tilde{r}_d \leqslant 2\lambda\})} \leqslant \varepsilon(\lambda)$,*

(3) $\| |u| - 1 \|_{L^\infty(\{\tilde{r}_d \geqslant \lambda\})} \leqslant 1/\lambda_*$,

(4) $(\mathrm{TW}_c)(u) = 0$ *for some* $c > 0$ *such that* $|dc - 1| \leqslant \varepsilon(\lambda)$,

then, there exist $X \in \mathbb{R}$ *and* $\gamma \in \mathbb{R}$ *such that* $u = Q_c(\,\cdot\, - X\vec{e}_2)\mathrm{e}^{i\gamma}$, *where* Q_c *is defined in Theorem 1.6.*

Remark 1.9. In view of the symmetry assumption, we may replace the second hypothesis by

$$\|u - V_1(\,\cdot\, - d\vec{e}_1)\|_{L^\infty(B(d\vec{e}_1, 2\lambda))} \leqslant \varepsilon(\lambda).$$

We will discuss the main arguments of the proof of Proposition 1.8 in the next section. This result can be seen as a local uniqueness result, but the uniqueness turns out to be in a rather large class of functions. Indeed, two functions that satisfy the hypotheses of Proposition 1.8 can be very far from each other, for two main reasons. First, in condition (2), the vortices that compose one of them have no reason to be close to the ones composing the other function since d depends on u: their centers $\pm d\vec{e}_1$ only need to satisfy $|dc - 1| \leqslant \varepsilon(\lambda)$, but for instance both $d = 1/c$ and $d = 1/c + 1/\sqrt{c}$ satisfy these conditions for $c > 0$ small enough at fixed λ. Secondly, we only have that far from the vortices, the modulus is close to 1 from condition (3), but we have no information on the phase. The proof of Proposition 1.8 will rely on methods used in [15] in order to prove some coercivity, and we shall need to be very precise to take into account all these cases.

A way to see that Proposition 1.8 is a strong unicity result is that it implies the following local uniqueness result in L^∞ for even functions in x_1.

Corollary 1.10. *There exist* $c_0, \varepsilon > 0$ *such that, for* $0 < c < c_0$, *if a function* $u \in \mathcal{E}$ *satisfies*

(1) *for all* $(x_1, x_2) \in \mathbb{R}^2$, $u(x_1, x_2) = u(-x_1, x_2)$,

(2) $(\mathrm{TW}_c)(u) = 0$ *in the distributional sense*,

(3) $\|u - Q_c\|_{L^\infty(\mathbb{R}^2)} \leqslant \varepsilon$,

then, there exist $X \in \mathbb{R}$ *and* $\gamma \in \mathbb{R}$ *such that* $u = Q_c(\,\cdot\, - X\vec{e}_2)\mathrm{e}^{i\gamma}$.

We may now state our main result. It establishes that any traveling wave solution which is even in x_1 and within $\mathcal{O}(1)$ of the minimizing energy must be, for large momentum, the Q_c traveling wave constructed in Theorem 1.6, up to the natural translation and phase invariances.

Theorem 1.11. *For any* $\Lambda_0 > 0$ *there exists* $\mathfrak{p}_0(\Lambda_0) > 0$ *such that, if* $u \in \mathcal{E}$ *satisfies*

(1) *for all* $(x_1, x_2) \in \mathbb{R}^2$, $u(x_1, x_2) = u(-x_1, x_2)$,

(2) $(\mathrm{TW}_c)(u) = 0$ *for some* $c > 0$,

(3) $P_2(u) \geqslant \mathfrak{p}_0(\Lambda_0)$,

(4) $E(u) \leqslant 2\pi \ln P_2(u) + \Lambda_0$,

then, there exist $X \in \mathbb{R}$ *and* $\gamma \in \mathbb{R}$ *such that*

$$u = Q_c(\,\cdot\, - X\vec{e}_2)\mathrm{e}^{i\gamma},$$

where Q_c *is defined in Theorem 1.6. In particular,* $P_2(u) = \mathcal{P}(c)$ (*where* $\mathcal{P}$ *is defined in Theorem 1.6*).

Section 3 is devoted to the proof of this result. We show there that a function satisfying the hypotheses of Theorem 1.11 also satisfies the hypotheses of Proposition 1.8. Our result applies in particular to the constraint minimizers for the problem $E_{\min}(\mathfrak{p})$ for large $\mathfrak{p}$.

Corollary 1.12. *There exist $\mathfrak{p}_0 > 0$ such that, for any $\mathfrak{p} \geqslant \mathfrak{p}_0$ and any minimizer U for $E_{\min}(\mathfrak{p})$, there exist $\gamma \in \mathbb{R}$ and $X \in \mathbb{R}^2$ such that, with $c = \mathcal{P}^{-1}(\mathfrak{p})$,*

$$U = Q_c(\cdot - X)e^{i\gamma}.$$

Moreover, $(\mathrm{TW}_c)(U) = 0$.

Proof. By a first translation in x_1, we may assume, by Theorem 1.2, that this minimizer U is even in x_1. By Proposition 1.5, the last hypothesis (4) of Theorem 1.11 is satisfied; hence we may translate in x_2 and use phase shift and get that this minimizer U is Q_c. Necessarily, $P_2(U) = \mathfrak{p} = P_2(Q_c)$; thus $c = \mathcal{P}^{-1}(\mathfrak{p})$. □

Theorem 1.4 is a direct consequence of this corollary. This allows us to derive several interesting consequences on the function $E_{\min}$. This also shows that the branch of Theorem 1.6 coincides with the global energy minimizer at fixed momentum (up to translation and phase shift).

Theorem 1.13. *There exists $c_* > 0$ such that, for $0 < c \leqslant c_*$, Q_c is a minimizer for $E_{\min}(P_2(Q_c))$. Moreover, there exists $\mathfrak{p}_0 > 0$ such that the following statements hold*:

(1) *The function $E_{\min}$ is of class C^2 in $[\mathfrak{p}_0, +\infty[$ and*

$$0 > E''_{\min}(\mathfrak{p}) \sim -\frac{2\pi}{\mathfrak{p}^2}, \quad 0 < E'_{\min}(\mathfrak{p}) \sim \frac{2\pi}{\mathfrak{p}}, \quad E_{\min}(\mathfrak{p}) = 2\pi \ln \mathfrak{p} + \mathcal{O}(1).$$

(2) *For $\mathfrak{p} \geqslant \mathfrak{p}_0$, $\mathcal{S}_{\mathfrak{p}} = \{Q_{\mathcal{P}^{-1}(\mathfrak{p})}(\cdot - X)e^{i\gamma} : \gamma \in \mathbb{R},\ X \in \mathbb{R}^2\}$; hence, for any $\mathfrak{p} \geqslant \mathfrak{p}_0$, $E'_{\min}(\mathfrak{p})$ is the speed of any minimizer for $E_{\min}(\mathfrak{p})$.*

(3) *For any $\mathfrak{p} \geqslant \mathfrak{p}_0$, $Q_{\mathcal{P}^{-1}(\mathfrak{p})}$ is orbitally stable for the semidistance D_0 (or, equivalently, for $0 < c \leqslant c_*$, Q_c is orbitally stable for the semidistance D_0).*

(4) *For $\mathfrak{p} \geqslant \mathfrak{p}_0$ and any minimizer u for $E_{\min}(\mathfrak{p})$, up to a space translation and a phase shift, u enjoys the symmetry,*

$$\text{for all } (x_1, x_2) \in \mathbb{R}^2, \quad u(x_1, -x_2) = \bar{u}(x_1, x_2),$$

in addition to the symmetry in x_1.

(5) *For any $\Lambda > 0$, there exists $\mathfrak{p}_0(\Lambda) > 0$ such that, if $u \in \mathcal{E}$ satisfies $(\mathrm{TW}_c)(u) = 0$ for some $c > 0$, $P_2(u) \geqslant \mathfrak{p}_0(\Lambda)$ and u is even in x_1, then either $E(u) = E_{\min}(P_2(u))$ or $E(u) \geqslant E_{\min}(P_2(u)) + \Lambda$.*

Proof. By Theorems 1.2 and 1.3, we have the existence of at least one minimizer $U_{\mathfrak{p}}$ for $E_{\min}(\mathfrak{p})$, where $\mathfrak{p} > 0$. For large $\mathfrak{p}$, by applying Corollary 1.12, we have $U_{\mathfrak{p}} = Q_c(\cdot - X)e^{i\gamma}$ for some $X \in \mathbb{R}^2$ and $\gamma \in \mathbb{R}$, thus proving that Q_c is a minimizer for $E_{\min}(\mathfrak{p})$ and that $P_2(Q_c) = \mathcal{P}(c) = \mathfrak{p}$.

For (1), it suffices to notice that, in view of Corollary 1.12 applied to any minimizer (we have existence by Theorems 1.2 and 1.3) $E_{\min}(\mathfrak{p}) = E(Q_{\mathcal{P}^{-1}(\mathfrak{p})})$. We then conclude by using that $\mathcal{P}$ is a

C^1 diffeomorphism and that $c \mapsto E(Q_c)$ is also of class C^1 (see [15, Proposition 1.2]), that $E_{\min}$ is of class C^1 in $[\mathfrak{p}_0, +\infty[$ and that

$$E'_{\min}(\mathfrak{p}) = \frac{d}{dc}E(Q_c)_{|c=\mathcal{P}^{-1}(\mathfrak{p})} \times \frac{1}{\mathcal{P}'(\mathcal{P}^{-1}(\mathfrak{p}))} = \mathcal{P}^{-1}(\mathfrak{p}),$$

in view of the Hamilton-like relation (formally shown in [28] and rigorously proved for the branch constructed in Theorem 1.6 in [15])

$$\frac{d}{dc}E(Q_c) = c\frac{d}{dc}P_2(Q_c).$$

Since $\mathcal{P}$ is a C^1 diffeomorphism, we deduce that $E'_{\min}$ is of class C^1. The asymptotics for $E'_{\min}$ and $E''_{\min}$ then follow from Proposition 1.2 in [15]. Integration would yield $E_{\min}(\mathfrak{p}) \sim 2\pi \ln \mathfrak{p}$, but we may slightly improve this estimate. Indeed, Proposition 1.5 gives $E_{\min}(\mathfrak{p}) \leqslant 2\pi \ln \mathfrak{p} + \mathcal{O}(1)$, and the lower bound is a straightforward consequence of Theorem 3.4(i) and the study in Section 3B3.

Statement (2) is a rephrasing of Corollary 1.12, combined with the existence of at least one constrained minimizer. Statement (3) is then a direct consequence of Theorem 1.3. Statement (4) simply follows from the fact that Q_c enjoys by construction this symmetry (see [14]). Finally, statement (5) is also a rephrasing of Theorem 1.11. □

Remark 1.14. Concerning the stability given in statement (3) in the above theorem, we quote [32], where a linear "spectral" stability result is proved (through ad hoc hypotheses that were checked in [15]), namely that the linearized equation $i\partial_t v = L_{Q_c}(v)$ does not have exponentially growing solutions (in $\dot{H}_1(\mathbb{R}^2; \mathbb{C})$, say). Statement (3) in the above theorem does not rely on the result in [32], and is needed for the nonlinear (orbital) stability (following the Cazenave–Lions approach).

Let us conclude this section with several comments on our result. First, let us explain the relevance of the symmetry hypothesis, namely that we restrict to mappings that are even in x_1. This symmetry is used in the coercivity of the branch of Theorem 1.6, through the following arguments. The quadratic form around the traveling wave Q_c is decomposed in three areas, close to the two vortices, and far from them. In the latter region, the coercivity can be shown without any orthogonality condition. Close to the vortices, the quadratic form is close to the one of a single vortex, which was studied in [38]. Its coercivity requires three orthogonality conditions, two for the translation, and one for the phase. Therefore, we can show the coercivity of the full quadratic form with six orthogonality conditions, three for each vortex. However, the family of traveling waves of Theorem 1.6 has only five parameters (two for the speed, two for the translation, and one for the phase). The symmetry is then used to reduce the problem to three orthogonality conditions into a family with three parameters. With this symmetry, both orthogonality conditions on the phase for the two vortices become the same condition. It is possible to prove a coercivity result with only five orthogonality conditions without symmetry (see [15]), but then the coercivity constant goes to 0 when $c \to 0$. This would pose a problem for the uniqueness result. The last statement in Theorem 1.13 shows that, when restricting ourselves to symmetric traveling waves, there is an energy threshold under which there is no traveling wave except the Q_c branch.

Secondly, the proof of the fact that Q_c is a minimizer of the energy for fixed momentum relies on the existence of such minimizers. In particular, we have not been able to use our coercivity results in [15] in order to prove directly that Q_c is orbitally stable (for small c).

Thirdly, the symmetry in x_2 for the minimizers (statement (4)) is established as a consequence of the uniqueness result and not in itself. Notice that the numerical studies in [16; 17; 28] assume the two symmetries.

1D. ***The traveling wave Q_c and two other variational characterizations.*** Before providing other variational characterizations of Q_c, we have to define a distance on the energy space $\mathcal{E}$. One can use (see [22])

$$D_{\mathcal{E}}(\psi_1, \psi_2) := \|\psi_1 - \psi_2\|_{L^2(\mathbb{R}^2)+L^\infty(\mathbb{R}^2)} + \|\nabla\psi_1 - \nabla\psi_2\|_{L^2(\mathbb{R}^2)} + \||\psi_1| - |\psi_2|\|_{L^2(\mathbb{R}^2)},$$

which is adapted to the Cauchy problem. Actually, we may also use the pseudodistance[1]

$$D_0(\psi_1, \psi_2) := \|\nabla\psi_1 - \nabla\psi_2\|_{L^2(\mathbb{R}^2)} + \||\psi_1| - |\psi_2|\|_{L^2(\mathbb{R}^2)}.$$

Is it shown in [13, Corollary 4.13] that both the energy E and the momentum P_2 are continuous for the distance $D_{\mathcal{E}}$, and actually even for the pseudodistance D_0.

The traveling wave Q_c as a mountain-pass solution. Thanks to the results in Theorem 1.13, it is easy to show that we have locally, near Q_c, a mountain-pass geometry. Indeed, let $c_* > 0$ be small, and define

$$\Upsilon_{c_*} := \{\upsilon : [-1, +1] \to \mathcal{E} \text{ continuous} : \upsilon(-1) = Q_{3c_*/2},\ \upsilon(+1) = Q_{c_*/2}\},$$

the set of continuous paths from $Q_{3c_*/2}$ to $Q_{c_*/2}$ in $\mathcal{E}$. Then, we claim that

$$\inf_{\upsilon\in\Upsilon_{c_*}} \max_{t\in[-1,+1]} (E - c_* P_2)(\upsilon(t)) = (E - c_* P_2)(Q_{c_*}). \tag{1-2}$$

Indeed, let $\upsilon \in \Upsilon_{c_*}$. By the intermediate value theorem, there exists $t_* \in [-1, +1]$ such that $P_2(\upsilon(t)) = P_2(Q_{c_*})$ ($c \mapsto P_2(Q_c)$ is a C^1 function (see [15, Proposition 1.2]). Since Q_{c_*} is a minimizer for $E_{\min}(Q_{c_*})$, we infer

$$\max_{t\in[-1,+1]} (E - c_* P_2)(\upsilon(t)) \geqslant E(\upsilon(t_*)) - c_* P_2(Q_{c_*}) \geqslant E(Q_{c_*}) - c_* P_2(Q_{c_*}).$$

Moreover, considering the particular $\mathcal{C}^1$ path $\upsilon_* : [-1, +1] \to \mathcal{E}$ defined by $\upsilon(t) := Q_{c_*-tc_*/2}$, we see that

$$\frac{d}{dt}(E - c_* P_2)(\upsilon_*(t)) = -\frac{c_*}{2}\Big(\frac{d}{dc}E(Q_c) - c_*\frac{d}{dc}P_2(Q_c)\Big)_{|c=c_*-tc_*/2} = \frac{c_*^2 t}{4}\Big(\frac{d}{dc}P_2(Q_c)\Big)_{|c=c_*-tc_*/2}$$

in view of the Hamilton group relation $\frac{d}{dc}E(Q_c) = c\frac{d}{dc}P_2(Q_c)$ (see [15, Proposition 1.2]). Since $\frac{d}{dc}P_2(Q_c) < 0$, we deduce that $(E - c_* P_2)(\upsilon_*(t))$ increases in $[-1, 0]$ and decreases in $[0, +1]$, and hence has maximal value $E(Q_{c_*}) - c_* P_2(Q_{c_*})$, as wished.

[1] $D_0(\psi_1, \psi_2)$ is zero if and only if $\psi_2 - \psi_1$ is constant with $|\psi_1| - 1 = |\psi_2| - 1 \in L^2(\mathbb{R}^2)$.

Furthermore, by the asymptotics given in [15] and the above-mentioned Hamilton group relation $\frac{d}{dc}E(Q_c) = c\frac{d}{dc}P_2(Q_c)$, we have

$$(E - c_* P_2)(Q_{c_*}) - (E - c_* P_2)(Q_{c_*/2}) = \int_{c_*/2}^{c_*} (c - c_*)\frac{d}{dc}P_2(Q_c)\,dc > 0$$

since $c - c_* < 0$ and $\frac{d}{dc}P_2(Q_c) < 0$. Similarly, we prove that $(E - c_* P_2)(Q_{c_*}) - (E - c_* P_2)(Q_{3c_*/2}) < 0$.

We now claim that if $u \in \mathcal{E}$ is such that $(\mathrm{TW}_{c_*})(u) = 0$ and

$$(E - c_* P_2)(u) = \inf_{\upsilon \in \Upsilon_{c_*}} \max_{t \in [-1,+1]} (E - c_* P_2)(\upsilon(t)) = (E - c_* P_2)(Q_{c_*}), \tag{1-3}$$

by (1-2), that is, if u is a critical point of $E - c_* P_2$ at the good critical value, then we must have $P_2(u) = P_2(Q_{c_*})$. Indeed, by the Pohozaev identity (2-2), we have

$$c_* P_2(u) = \frac{1}{2}\int_{\mathbb{R}^2} (1 - |u|^2)^2\,dx \geqslant 0,$$

and hence $P_2(u) \geqslant 0$. Furthermore, we know that $E_{\min}$ is concave in $\mathbb{R}_+$ (Proposition 1.5), and that $E_{\min}$ is of class C^1 and strictly concave on $[\mathfrak{p}_0, +\infty[$ (by statement (1) of Theorem 1.13). Therefore, if $P_2(u) \neq P_2(Q_{c_*})$, then

$$\begin{aligned} E(u) \geqslant E_{\min}(P_2(u)) &> E_{\min}(P_2(Q_{c_*})) + E'_{\min}(P_2(Q_{c_*}))(P_2(u) - P_2(Q_{c_*})) \\ &= E(Q_{c_*}) + c_*(P_2(u) - P_2(Q_{c_*})), \end{aligned}$$

in contradiction with (1-3).

As a consequence, we have

$$E(u) = E(Q_{c_*}) = E_{\min}(P_2(u)) = E_{\min}(P_2(Q_{c_*})),$$

implying that u is a minimizer for $E_{\min}(P_2(Q_{c_*}))$; hence there exist $\gamma \in \mathbb{R}$ and $X \in \mathbb{R}^2$ such that $u = Q_{c_*}(\cdot - X)e^{i\gamma}$, proving a uniqueness result for mountain-pass-type traveling wave solutions. However, stating rigorously a useful uniqueness result for this kind of variational solution is not so easy: In [5], the mountain pass is implemented in the space $1 + H^1(\mathbb{R}^2)$, whereas we know (by the result in [25]) that the nontrivial traveling wave does not belong to this affine space; in [3], the solution is constructed by working first on $[-N, +N] \times \mathbb{R}$ and then passing to the limit, and it is then not immediate to compute the functional $E - cP$ on the solution; in addition, the method does not provide easily some explicit bounds on the energy or the momentum. We shall then not go further in this discussion even though the previous arguments indicate that mountain-pass solutions are (at least for small c) only the orbit of Q_c.

The traveling wave Q_c as a minimizer of $E - cP_2$ for fixed kinetic energy. In [13], for $\kappa \geqslant 0$, the following variational problem is investigated:

$$I_{\min}(\kappa) = \inf\left\{ \frac{1}{4}\int_{\mathbb{R}^2} (1 - |v|^2)^2\,dx - P_2(v),\ v \in \mathcal{E} : \frac{1}{2}\int_{\mathbb{R}^2} |\nabla v|^2\,dx = \kappa \right\}.$$

Any minimizer v for $I_{\min}(\kappa)$ is such that there exists $c > 0$ satisfying $(\mathrm{TW}_c)(v(\cdot/c)) = 0$. In two dimensions and for the Ginzburg–Landau nonlinearity, existence of minimizers for $\kappa > 0$ is established in

Theorem 1.2 there. Furthermore, it is shown in [13] (see Proposition 8.4 there) that if $\mathfrak{p} > 0$ and if U is a minimizer for $E_{\min}(\mathfrak{p})$ with speed c, then $U(c\,\cdot)$ is a minimizer for $I_{\min}(\kappa)$ with $\kappa = \frac{1}{2}\int_{\mathbb{R}^2} |\nabla U|^2\,dx$ (this last quantity is scale-invariant in two dimensions) and $I_{\min}$ is differentiable at this κ, with $I'_{\min}(\kappa) = -1/c^2$. Since Q_c is a minimizer for $E_{\min}(P_2(Q_c))$, if we prove that $c \mapsto \frac{1}{2}\int_{\mathbb{R}^2} |\nabla Q_c|^2\,dx$ is a decreasing C^1-diffeomorphism from $]0, c_0]$, for some small c_0, onto $[\kappa_0, +\infty[$, with $\kappa_0 := \frac{1}{2}\int_{\mathbb{R}^2} |\nabla Q_{c_0}|^2\,dx$, then we shall conclude that $I_{\min}$ is of class C^1 on $[\kappa_0, +\infty[$, and that (by the arguments in [13]) the only minimizer for $\kappa = \frac{1}{2}\int_{\mathbb{R}^2} |\nabla Q_c|^2\,dx$ (for some suitable $c \in]0, c_0]$) is $Q_c(c\,\cdot)$ up to the natural translation and phase invariances and, in addition, $I'_{\min}(\kappa) = -1/c^2$. In order to prove that statement, it suffices to use the Pohozaev identity (2-2) and deduce

$$\frac{1}{2}\int_{\mathbb{R}^2} |\nabla Q_c|^2\,dx = E(Q_c) - \frac{1}{4}\int_{\mathbb{R}^2} (1 - |Q_c|^2)^2\,dx = E(Q_c) - \frac{cP_2(Q_c)}{2}.$$

Therefore, by using the Hamilton-like relation $\frac{d}{dc}E(Q_c) = c\frac{d}{dc}P_2(Q_c)$ and then the asymptotics of $c \mapsto P_2(Q_c)$ obtained in [15], we arrive at

$$\frac{d}{2dc}\int_{\mathbb{R}^2} |\nabla Q_c|^2\,dx = \frac{d}{dc}(E(Q_c)) - \frac{c}{2}\frac{d}{dc}P_2(Q_c) - \frac{1}{2}P_2(Q_c) = \frac{c}{2}\frac{d}{dc}P_2(Q_c) - \frac{1}{2}P_2(Q_c) \sim -\frac{2\pi}{c} < 0.$$

The paper is organized as follows. In Section 2, we give the proof of the uniqueness result given in Proposition 1.8. Section 3 is devoted to the vortex analysis of traveling waves with energy $E_{\min}(\mathfrak{p}) + \mathcal{O}(1)$, that are even in x_2, in order to show that they satisfy the hypotheses of Proposition 1.8. Section 3D contains a few remarks on the nonsymmetrical case. Finally, in Section 3C, we provide some decay estimates slightly away from the vortices. For the Ginzburg–Landau (stationary) model, such estimates were first given in [35] for minimizing solutions and later generalized in [18] to nonminimizing solutions. They improve some estimates in [14] and are not specific to the way we construct the solutions.

2. Proof of the local uniqueness result (Proposition 1.8)

This section is devoted to the proofs of Proposition 1.8 and Corollary 1.10. The proof of Proposition 1.8 uses arguments from the proof of [15, Theorem 1.14], another local uniqueness result for this problem, but in different spaces. We explain here the core ideas of the proof.

Let us explain schematically the proof of Proposition 1.8. We first pick c', X, γ' in such a way that $Q = Q'_c(\cdot - X)e^{i\gamma}$ has the same vortices as u. This is possible because $c \to d_c$, the position of the vortices, is smooth. We then use the decomposition $u = Qe^{\psi}$, where ψ is the error term. This cannot be done near the zeros of Q, but we focus here on the domain far from the vortices.

The equation satisfied by ψ is then $(\mathrm{TW}_c)(u) = 0 = (\mathrm{TW}_c)(Q) + \mathrm{L}(\psi) + \mathrm{NL}(\psi)$, where we regroup the linear terms in L and the nonlinear terms in NL, and $(\mathrm{TW}_c)(Q) \neq 0$ because $c \neq c'$. We then take the scalar product of this equation with ψ, and we get $0 = \langle(\mathrm{TW}_c)(Q), \psi\rangle + B_Q(\psi) + \langle\mathrm{NL}(\psi), \psi\rangle$. Now, the coercivity of B_Q has been studied in [15]. It holds (for even functions in x_1) up to three orthogonality conditions, which can be satisfied by changing slightly the modulation parameters c', X, γ. We deduce that $B_Q(\psi) \geqslant K\|\psi\|_1^2$ for some norm $\|\cdot\|_1$.

There are two main difficulties at this point. First, since the hypotheses on u in Proposition 1.8 are weak, we simply have $\|\psi\|_1 < +\infty$, but not the fact that it is small. Therefore, an estimate of the form $|\langle \mathrm{NL}(\psi), \psi\rangle| \leqslant K\|\psi\|_1^3$ would not be enough to conclude. Secondly, the norm $\|\cdot\|_1$ is rather weak, and in fact $\langle \mathrm{NL}(\psi), \psi\rangle$ cannot be controlled by powers of $\|\psi\|_1$.

Concerning the term $\langle (TW_c)(Q), \psi\rangle$, we may show that we always have $|c - c'| \leqslant o(1)\|\psi\|_1$, and thus $|\langle (TW_c)(Q), \psi\rangle| \leqslant o(1)\|\psi\|_1^2$. Therefore, we are led to

$$\frac{K}{2}\|\psi\|_1^2 \leqslant \langle (\mathrm{TW}_c)(Q), \psi\rangle + B_Q(\psi) = -\langle \mathrm{NL}(\psi), \psi\rangle.$$

Then, even if $\|\psi\|_1$ is not small, by the hypotheses of Proposition 1.8, ψ will be small in other (nonequivalent) norms. Let us write one of them $\|\cdot\|_2$. Our goal is then to show an estimate of the form $|\langle \mathrm{NL}(\psi), \psi\rangle| \leqslant K\|\psi\|_2\,\|\psi\|_1^2$, which would conclude the proof. This is possible, except for one nonlinear term, which contains two derivatives. We then perform some integrations by parts on it. When both derivatives fall on the same term, we get a term containing $\Delta\psi$, which also appears in the equation $0 = (\mathrm{TW}_c)(Q) + \mathrm{L}(\psi) + \mathrm{NL}(\psi)$ (in $\mathrm{L}(\psi)$). We thus replace it using this equation, which leads to another term containing two derivatives (from $\mathrm{NL}(\psi)$), and other terms that can be successfully estimated. After n such integrations by parts, we have an estimate of the form

$$|\langle \mathrm{NL}(\psi), \psi\rangle| \leqslant K\|\psi\|_2\,\|\psi\|_1^2 + \|\psi\|_3\,\|\psi\|_2^n\,\|\psi\|_1^2,$$

where $\|\cdot\|_3$ is another (semi-)norm in which ψ is not necessarily small. Now, taking n large enough (depending on ψ), since $\|\psi\|_2 \ll 1$, we get $|\langle \mathrm{NL}(\psi), \psi\rangle| \leqslant o(1)\|\psi\|_1^2$, concluding the proof.

The problem is a lot simpler near the vortices. There, we write $u = Q + \phi$ and the coercivity norm is equivalent to the H^1 norm, and the hypotheses of Proposition 1.8 give us $\|\phi\|_{L^\infty} = o(1)$. The estimate of the nonlinear terms then becomes trivial.

As stated in the Introduction, the symmetry condition is necessary to have a coercivity result where the coercivity constant is uniform; see Corollary 2.6 below. This is the only place where the symmetry is used in a crucial way.

2A. ***Some properties of the branch of traveling waves from Theorem 1.6.*** We recall here properties on the branch $c \mapsto Q_c$ from Theorem 1.6, coming mainly from [14; 15]. In this section, we will use the notation

$$\langle f, g\rangle := \int_{\mathbb{R}^2} \mathfrak{Re}(f\bar{g}).$$

2A1. *Properties of vortices.* We start with some estimates on vortices, which compose the traveling wave (see Theorem 1.6).

Lemma 2.1 [12; 26]. *A vortex centered around* 0, $V_1(x) = \rho_1(r)e^{i\theta}$, *satisfies* $V_1(0) = 0$, $E(V_1) = +\infty$ *and there exist constants* $K, \kappa > 0$ *such that*,

$$\textit{for all } r > 0,\ 0 < \rho_1(r) < 1, \quad \rho_1(r) \sim_{r\to 0} \kappa r, \quad \rho_1'(r) \sim_{r\to 0} \kappa,$$

$$\textit{for all } r > 0,\ \rho_1'(r) > 0, \quad \rho_1'(r) = O_{r\to\infty}\left(\frac{1}{r^3}\right), \quad |\rho_1''(r)| + |\rho_1'''(r)| \leqslant K,$$

$$1-|V_1(x)| = \frac{1}{2r^2} + O_{r\to\infty}\left(\frac{1}{r^3}\right),$$
$$|\nabla V_1| \leqslant \frac{K}{1+r}, \quad |\nabla^2 V_1| \leqslant \frac{K}{(1+r)^2},$$
$$\nabla V_1(x) = i V_1(x)\frac{x^\perp}{r^2} + O_{r\to\infty}\left(\frac{1}{r^3}\right),$$

where $x^\perp := (-x_2, x_1)$, $x = re^{i\theta} \in \mathbb{R}^2$. *Furthermore, similar properties hold for* V_{-1}, *since*

$$V_{-1}(x) = \overline{V_1(x)}.$$

2A2. *Toolbox.* We list in this section some results useful for the analysis of traveling waves for not necessarily small speeds.

Theorem 2.2 (uniform L^∞ bound [19]). *Assume that* $U \in L^3_{\mathrm{loc}}(\mathbb{R}^d)$ *solves*

$$\Delta U + ic\partial_2 U + U(1-|U|^2) = 0.$$

Then,

$$\|U\|_{L^\infty(\mathbb{R}^d)} \leqslant 1 + \frac{c^2}{4}.$$

Corollary 2.3. *There exists* $K > 0$ *such that*, *for any* $c \in [-\sqrt{2}, +\sqrt{2}]$ *and any* $U \in L^3_{\mathrm{loc}}(\mathbb{R}^d)$ *satisfying* $(\mathrm{TW}_c)(U) = 0$, *we have*

$$\|\nabla U\|_{L^\infty(\mathbb{R}^d)} + \|\nabla^2 U\|_{L^\infty(\mathbb{R}^d)} \leqslant K. \tag{2-1}$$

The following Pohozaev identity (see [10] for instance) will be useful in our analysis. If $c \in \mathbb{R}$ and $U \in \mathcal{E}$ satisfies (TW_c), then

$$\frac{1}{2}\int_{\mathbb{R}^2} (1-|U|^2)^2\,dx = cP_2(U). \tag{2-2}$$

We shall also make use of the algebraic decay of the traveling waves conjectured in [28] and shown in [24].

Theorem 2.4 (algebraic decay of the traveling waves [24]). *Let* $c \in [0, \sqrt{2}[$. *Assume that* $U \in \mathcal{E}$ *is a solution of* $(\mathrm{TW}_c)(U) = 0$. *Up to a phase shift, we may assume* $U(x) \to 1$ *for* $|x| \to +\infty$. *Then, there exists* M, *depending on* U *and* c *such that*, *for* $x \in \mathbb{R}^2$,

$$|U(x)-1| \leqslant \frac{M}{1+|x|}, \quad |\nabla U(x)| \leqslant \frac{M}{(1+|x|)^2}, \quad ||U(x)|-1| \leqslant \frac{M}{(1+|x|)^2}.$$

2A3. *Symmetries of the traveling waves from Theorem 1.6.* We recall from [14] that the traveling wave Q_c constructed in Theorem 1.6 satisfies, for all $x = (x_1, x_2) \in \mathbb{R}^2$,

$$Q_c(x_1, x_2) = Q_c(-x_1, x_2) = \overline{Q_c(x_1, -x_2)}.$$

This implies that, for all $x = (x_1, x_2) \in \mathbb{R}^2$,

$$\partial_c Q_c(x_1, x_2) = \partial_c Q_c(-x_1, x_2) = \overline{\partial_c Q_c(x_1, -x_2)},$$

$$\begin{aligned}
\partial_{x_1} Q_c(x_1,x_2) &= -\partial_{x_1} Q_c(-x_1,x_2) = \overline{\partial_{x_1} Q_c(x_1,-x_2)},\\
\partial_{x_2} Q_c(x_1,x_2) &= \partial_{x_2} Q_c(-x_1,x_2) = -\overline{\partial_{x_2} Q_c(x_1,-x_2)},\\
\partial_{c^\perp} Q_c(x_1,x_2) &= -\partial_{c^\perp} Q_c(-x_1,x_2) = -\overline{\partial_{c^\perp} Q_c(x_1,-x_2)},
\end{aligned}$$

where $\partial_{c^\perp} Q_c := x^\perp.\nabla Q_c$; see Section 2.2 of [15]. Note that these quantities all have different symmetries.

2A4. *A coercivity result.* From Proposition 1.2 of [15], we recall that Q_c defined in Theorem 1.6 has two zeros, at $\pm\tilde{d}_c\vec{e}_1$, with

$$d_c - \tilde{d}_c = o_{c\to 0}(1). \tag{2-3}$$

We define (as in [15]) the symmetric expended energy space by

$$H^{\exp,s}_{Q_c} := \big\{\varphi \in H^1_{\mathrm{loc}}(\mathbb{R}^2,\mathbb{C}) : \|\varphi\|_{H^{\exp}_{Q_c}} < +\infty \text{ for all } (x_1,x_2)\in\mathbb{R}^2,\ \varphi(-x_1,x_2) = \varphi(x_1,x_2)\big\},$$

where, with $\varphi = Q_c\psi$, $\tilde{r} = \tilde{r}_{\tilde{d}_c} = \min(\tilde{r}_1, \tilde{r}_{-1})$, $\tilde{r}_{\pm 1}$ being the distances to the zeros of Q_c (we use $\tilde{r}$ instead of $\tilde{r}_{\tilde{d}_c}$ to simplify the notation here), we define

$$\|\varphi\|^2_{H^{\exp}_{Q_c}} := \|\varphi\|^2_{H^1(\{\tilde{r}\leqslant 10\})} + \int_{\{\tilde{r}\geqslant 5\}} |\nabla\psi|^2 + \mathfrak{Re}^2(\psi) + \frac{|\psi|^2}{\tilde{r}^2\ln^2\tilde{r}}.$$

By using (2-1), we deduce, for any $R > 0$, $\|\varphi\|_{H^1(\{\tilde{r}\leqslant R\})} \leqslant K(R)\|\varphi\|_{H^{\exp}_{Q_c}}$. The linearized operator around Q_c is

$$L_{Q_c}(\varphi) := -\Delta\varphi - ic\,\partial_{x_2}\varphi - (1-|Q_c|^2)\varphi + 2\,\mathfrak{Re}(\overline{Q}_c\varphi)Q_c.$$

We take a smooth cutoff function $\tilde{\eta}$ such that

$$\tilde{\eta}(x) = \begin{cases} 0 & \text{on } B(\pm\tilde{d}_c\vec{e}_1, 2R),\\ 1 & \text{on } \mathbb{R}^2\backslash B(\pm\tilde{d}_c\vec{e}_1, 2R+1),\end{cases}$$

where $\pm\tilde{d}_c\vec{e}_1$ are the zeros of Q_c and $R > 0$ will be defined later on (it will be a universal constant, independent of any parameters of the problem). We define the quadratic form (as in [15])

$$\begin{aligned}
B^{\exp}_{Q_c}(\varphi) := &\int_{\mathbb{R}^2} (1-\tilde{\eta})\big(|\nabla\varphi|^2 - \mathfrak{Re}(ic\,\partial_{x_2}\varphi\bar{\varphi}) - (1-|Q_c|^2)|\varphi|^2 + 2\,\mathfrak{Re}^2(\overline{Q}_c\varphi)\big)\\
&- \int_{\mathbb{R}^2} \nabla\tilde{\eta}.\big(\mathfrak{Re}(\nabla Q_c\overline{Q}_c)|\psi|^2 - 2\,\mathfrak{Im}(\nabla Q_c\overline{Q}_c)\,\mathfrak{Re}(\psi)\,\mathfrak{Im}(\psi)\big)\\
&+ \int_{\mathbb{R}^2} c\,\partial_{x_2}\tilde{\eta}\,\mathfrak{Re}(\psi)\,\mathfrak{Im}(\psi)|Q_c|^2\\
&+ \int_{\mathbb{R}^2} \tilde{\eta}(|\nabla\psi|^2|Q_c|^2 + 2\,\mathfrak{Re}^2(\psi)|Q_c|^4)\\
&+ \int_{\mathbb{R}^2} \tilde{\eta}\big(4\,\mathfrak{Im}(\nabla Q_c\overline{Q}_c)\,\mathfrak{Im}(\nabla\psi)\,\mathfrak{Re}(\psi) + 2c|Q_c|^2\,\mathfrak{Im}(\partial_{x_2}\psi)\,\mathfrak{Re}(\psi)\big).
\end{aligned} \tag{2-4}$$

We recall from [15] (or by integration by parts) that, for $\varphi\in C^\infty_c(\mathbb{R}^2,\mathbb{C})$, we have $B^{\exp}_{Q_c}(\varphi) = \langle L_{Q_c}(\varphi),\varphi\rangle$ and that $B^{\exp}_{Q_c}(\varphi)$ is well-defined for $\varphi \in H^{\exp,s}_{Q_c}$. This last point is the reason why we write the quadratic form as (2-4), which is equal, up to some integration by parts, to the more natural definition

$$\int_{\mathbb{R}^2} |\nabla\varphi|^2 - (1-|Q_c|^2)|\varphi|^2 + 2\,\mathfrak{Re}^2(\overline{Q}_c\varphi) - \mathfrak{Re}(ic\,\partial_{x_2}\varphi\bar{\varphi}),$$

but this integral is not well-defined for $\varphi \in H^{\exp,s}_{Q_c}$. See [15] for more details on this point. We now quote the following coercivity result:

Theorem 2.5 [15, Theorem 1.13]. *There exist $R, K, c_0 > 0$ such that, for $0 < c \leqslant c_0$, Q_c defined in Theorem 1.6, if a function $\varphi \in H^{\exp,s}_{Q_c}$ satisfies the three orthogonality conditions*

$$\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \partial_c Q_c\,\bar{\varphi} = \mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \partial_{x_2} Q_c\,\bar{\varphi} = 0,$$
$$\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} i\,Q_c\,\bar{\varphi} = 0,$$

then

$$\frac{1}{K}\|\varphi\|^2_{H^{\exp}_{Q_c}} \geqslant B^{\exp}_{Q_c}(\varphi) \geqslant K\|\varphi\|^2_{H^{\exp}_{Q_c}}.$$

We will use a slight variation of this result, given in the next corollary.

Corollary 2.6. *There exist $R, K, c_0 > 0$ such that, for $0 < c \leqslant c_0$, Q_c defined in Theorem 1.6, if a function $\varphi \in H^{\exp,s}_{Q_c}$ satisfies the three orthogonality conditions*

$$\mathfrak{Re}\int_{B(d_c\vec{e}_1,R)\cup B(-d_c\vec{e}_1,R)} \partial_d(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1))_{|d=d_c}\bar{\varphi} = \mathfrak{Re}\int_{B(d_c\vec{e}_1,R)\cup B(-d_c\vec{e}_1,R)} \partial_{x_2} Q_c\,\bar{\varphi} = 0,$$
$$\mathfrak{Re}\int_{B(d_c\vec{e}_1,R)\cup B(-d_c\vec{e}_1,R)} i\,Q_c\,\bar{\varphi} = 0,$$

then

$$\frac{1}{K}\|\varphi\|^2_{H^{\exp}_{Q_c}} \geqslant B^{\exp}_{Q_c}(\varphi) \geqslant K\|\varphi\|^2_{H^{\exp}_{Q_c}}.$$

Note, with Theorem 1.6 (for $p = +\infty$), that $-(1/c^2)\partial_d(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1))_{|d=d_c}$ is the first order of $\partial_c Q_c$ when $c \to 0$ in $L^\infty(\mathbb{R}^2, \mathbb{C})$, and that (with Lemma 2.1) they both have the same symmetries. We need to change the quantity $\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \partial_c Q_c\,\bar{\varphi}$ in the orthogonality conditions because we will differentiate it with respect to c, and

$$c \mapsto \partial_d\big(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1)\big)_{|d=d_c} = -\partial_{x_1}V_1(\cdot - d_c\vec{e}_1)V_{-1}(\cdot + d_c\vec{e}_1) + \partial_{x_1}V_{-1}(\cdot + d_c\vec{e}_1)V_1(\cdot - d_c\vec{e}_1)$$

is a C^1 function ($c \mapsto d_c \in C^1(]0, c_0[, \mathbb{R})$ for $c_0 > 0$ a small constant (see Section 4.6 of [14]), but it is not clear that $c \mapsto \partial_c Q_c$ can be differentiated with respect to c. Precise estimates on

$$\partial_d(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1))_{|d=d_c}$$

can be found in Lemma 2.6 of [14]. Furthermore, we changed, in the area of the integrals, $\tilde{d}_c$ to d_c (they are close when $c \to 0$, see (2-3)).

Proof. Step 1: changing the integrand but not the integration domain. Take a function $\varphi \in H^{\exp,s}_{Q_c}$ satisfying

$$\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \partial_d(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1))_{|d=d_c}\bar{\varphi} = \mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \partial_{x_2}Q_c\,\bar{\varphi} = 0,$$

$$\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} iQ_c\,\bar{\varphi} = 0.$$

Let us show that it satisfies $(1/K)\|\varphi\|^2_{H^{\exp}_{Q_c}} \geqslant B^{\exp}_{Q_c}(\varphi) \geqslant K\|\varphi\|_{H^{\exp}_{Q_c}}$. For $\mu \in \mathbb{R}$, we define

$$\varphi^* = \varphi + c^2\mu\partial_c Q_c.$$

We have that $\partial_c Q_c \in H^{\exp,s}_{Q_c}$. We want to choose $\mu \in \mathbb{R}$ such that φ^* satisfies the hypothesis of Theorem 2.5. By the symmetries of Section 2A3 and the hypotheses on φ, we have that

$$\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} iQ_c\,\overline{\varphi^*} = \mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \partial_{x_2}Q_c\,\overline{\varphi^*} = 0,$$

and we compute, using

$$\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \partial_d\big(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1)\big)_{|d=d_c}\bar{\varphi} = 0,$$

that

$$\begin{aligned}
&\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} c^2\partial_c Q_c\,\overline{\varphi^*}\\
&\quad= \mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} c^2\partial_c Q_c\,\bar{\varphi} + \mu\,\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} c^4|\partial_c Q_c|^2\\
&\quad= \mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} \big(c^2\partial_c Q_c - \partial_d(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1))_{|d=d_c}\big)\bar{\varphi}\\
&\qquad\qquad\qquad + \mu\,\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} c^4|\partial_c Q_c|^2.
\end{aligned}$$

By Theorem 1.6 (for $p = +\infty$) and Lemma 2.6 of [14], we have

$$\big\|c^2\partial_c Q_c - \partial_d\big(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1)\big)_{|d=d_c}\big\|_{L^\infty(\mathbb{R}^2)} = o_{c\to 0}(1),$$

and also that there exists a universal constant $K > 0$ (we recall that $R > 0$ is a universal constant) such that

$$\frac{1}{K} \leqslant \mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} c^4|\partial_c Q_c|^2 \leqslant K.$$

Now, taking

$$\mu = \frac{-\,\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)}\big(c^2\partial_c Q_c - \partial_d(V_1(\cdot - d\vec{e}_1)V_{-1}(\cdot + d\vec{e}_1))_{|d=d_c}\big)\bar{\varphi}}{\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} c^4|\partial_c Q_c|^2},$$

we have

$$\mathfrak{Re}\int_{B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R)} c^2\partial_c Q_c\,\overline{\varphi^*} = 0,$$

with

$$|\mu| \leqslant o_{c\to 0}(1)\|\varphi\|_{L^2(B(\tilde{d}_c\vec{e}_1,R)\cup B(-\tilde{d}_c\vec{e}_1,R))} \leqslant o_{c\to 0}(1)\|\varphi\|_{H^{\exp}_{Q_c}}.$$

Since $\partial_c Q_c \in H^{\exp,s}_{Q_c}$ by Lemma 2.8 of [15], we deduce that φ^* satisfies all the hypotheses of Theorem 2.5; therefore

$$\frac{1}{K}\|\varphi^*\|^2_{H^{\exp}_{Q_c}} \geqslant B^{\exp}_{Q_c}(\varphi^*) \geqslant K\|\varphi^*\|^2_{H^{\exp}_{Q_c}}.$$

Now, from Lemma 6.3 of [15], we have $1/K \leqslant \|c^2\partial_c Q_c\|_{H^{\exp}_{Q_c}} \leqslant K$ for a universal constant $K > 0$. With $|\mu| \leqslant o_{c\to 0}(1)\|\varphi\|_{H^{\exp}_{Q_c}}$, we deduce that, taking $c > 0$ small enough,

$$\frac{1}{K}\|\varphi\|^2_{H^{\exp}_{Q_c}} \geqslant B^{\exp}_{Q_c}(\varphi^*) \geqslant K\|\varphi\|^2_{H^{\exp}_{Q_c}}$$

for some universal constant $K > 0$. Now, we have the decomposition (using Lemmas 6.2 and 6.3 of [15])

$$\begin{aligned} B^{\exp}_{Q_c}(\varphi^*) &= B^{\exp}_{Q_c}(\varphi + c^2\mu\partial_c Q_c) \\ &= B^{\exp}_{Q_c}(\varphi) + 2c^2\mu\langle L_{Q_c}(\partial_c Q_c), \varphi\rangle + c^4\mu^2 B^{\exp}_{Q_c}(\partial_c Q_c), \end{aligned}$$

and by Lemmas 2.8, 5.4 and 6.1 of [15],

$$|\langle L_{Q_c}(\partial_c Q_c), \varphi\rangle| = |\langle i\partial_{x_2} Q_c, \varphi\rangle| \leqslant K\ln\Big(\frac{1}{c}\Big)\|\varphi\|_{H^{\exp}_{Q_c}};$$

hence

$$|2c^2\mu\langle L_{Q_c}(\partial_c Q_c), \varphi\rangle| \leqslant Kc^2\ln\Big(\frac{1}{c}\Big)|\mu|\|\varphi\|_{H^{\exp}_{Q_c}} \leqslant o_{c\to 0}(1)\|\varphi\|^2_{H^{\exp}_{Q_c}}.$$

By Proposition 1.2 of [15], $B^{\exp}_{Q_c}(\partial_c Q_c) = (2\pi + o_{c\to 0}(1))/c^2$; thus

$$|c^4\mu^2 B^{\exp}_{Q_c}(\partial_c Q_c)| \leqslant o_{c\to 0}(1)\|\varphi\|^2_{H^{\exp}_{Q_c}},$$

which concludes the proof of $(1/K)\|\varphi\|^2_{H^{\exp}_{Q_c}} \geqslant B^{\exp}_{Q_c}(\varphi) \geqslant K\|\varphi\|^2_{H^{\exp}_{Q_c}}$ by taking $c > 0$ small enough.

Step 2: Changing the integration domain. To change the conditions

$$\mathfrak{Re}\int_{B(\tilde d_c\vec e_1,R)\cup B(-\tilde d_c\vec e_1,R)} \partial_d(V_1(\cdot - d\vec e_1)V_{-1}(\cdot + d\vec e_1))_{|d=d_c}\bar\varphi = \mathfrak{Re}\int_{B(\tilde d_c\vec e_1,R)\cup B(-\tilde d_c\vec e_1,R)} \partial_{x_2} Q_c\,\bar\varphi = 0,$$

$$\mathfrak{Re}\int_{B(\tilde d_c\vec e_1,R)\cup B(-\tilde d_c\vec e_1,R)} iQ_c\,\bar\varphi = 0$$

to

$$\mathfrak{Re}\int_{B(d_c\vec e_1,R)\cup B(-d_c\vec e_1,R)} \partial_d(V_1(\cdot - d\vec e_1)V_{-1}(\cdot + d\vec e_1))_{|d=d_c}\bar\varphi = \mathfrak{Re}\int_{B(d_c\vec e_1,R)\cup B(-d_c\vec e_1,R)} \partial_{x_2} Q_c\,\bar\varphi = 0,$$

$$\mathfrak{Re}\int_{B(d_c\vec e_1,R)\cup B(-d_c\vec e_1,R)} iQ_c\,\bar\varphi = 0,$$

we use similar arguments, using $|d_c - \tilde d_c| = o_{c\to 0}(1)$ by (2-3). We check for instance that

$$\left|\mathfrak{Re}\int_{B(\tilde d_c\vec e_1,R)\cup B(-\tilde d_c\vec e_1,R)} \partial_{x_2} Q_c\,\bar\varphi - \mathfrak{Re}\int_{B(d_c\vec e_1,R)\cup B(-d_c\vec e_1,R)} \partial_{x_2} Q_c\,\bar\varphi\right| \leqslant K(R)|d_c - \tilde d_c|\,\|\varphi\|_{H^{\exp}_{Q_c}}$$

and $|d_c - \tilde d_c| = o_{c\to 0}(1)$.

Notice that the integration domain remains symmetric with respect to the x_2-axis. □

2B. ***Proof of Proposition 1.8.*** In this subsection, we take $\nu \in]0, 1[$ to be a small but universal constant, which will be fixed at the end of the proof. We take

$$\lambda_* = \max\left(3R + 1, \frac{1}{\nu^2}\right)$$

in the statement of Proposition 1.8 (where $R > 0$ is defined in Corollary 2.6). Then, for any $\lambda \geqslant \lambda_*$, we take

$$\varepsilon(\lambda) = \min\left(\nu, \frac{1}{10\lambda^2 + 100}\right)$$

in the statement of Proposition 1.8. The condition $\varepsilon(\lambda) \leqslant 1/(10\lambda^2 + 100)$ is required only to make sure that the two balls $B(d\vec{e}_1, 2\lambda)$ and $B(-d\vec{e}_1, 2\lambda)$ are disjoint and at a distance at least 1 from one another. This will be used only in the proof of Lemma 2.8.

We take u a function satisfying the hypotheses of Proposition 1.8 for these values of λ_*, λ and $\varepsilon(\lambda)$. In the rest of the subsection, $K, K' > 0$ denote universal constants, independent of any parameters of the problem (in particular, λ, λ_*, $\varepsilon(\lambda)$ and ν).

2B1. *Modulation on the parameters of the branch.* From Theorem 1.1 and the end of Section 4.6 of [14], we have that $Q_c = V_1(\cdot - d_c\vec{e}_1)V_{-1}(\cdot + d_c\vec{e}_1) + \Gamma_c$, with $d_c = (1 + o_{c\to 0}(1))/c$, $\|\Gamma_c\|_{L^\infty} \to 0$, and

$$c \mapsto d_c \in C^1(]0, c_0[, \mathbb{R}),$$

with $\partial_c d_c \sim -1/c^2$ for $c \to 0$ (see Section 4.6 of [14]). In particular, $c \mapsto d_c$ is a smooth decreasing diffeomorphism from $]0, c_0]$ onto $[d_0, +\infty[$, and thus, given $d > 1/\nu > d_0$ (for ν small enough), there exists a unique $c' > 0$ such that $d_{c'} = d$. In addition, $c' \sim_{d\to\infty} 1/d \leqslant K\nu$. Furthermore,

$$\begin{aligned} u(x) - Q_{c'}(x) &= V_1(x - d\vec{e}_1)V_{-1}(x + d\vec{e}_1) + \Gamma(x) - V_1(x - d_{c'}\vec{e}_1)V_{-1}(x + d_{c'}\vec{e}_1) - \Gamma_{c'}(x) \\ &= \Gamma(x) - \Gamma_{c'}(x). \end{aligned}$$

From the hypotheses on Γ, and the fact that $\|\Gamma_{c'}\|_{L^\infty(\mathbb{R}^2)} \leqslant 2\nu$ (since $c' \leqslant 2/d \leqslant 2\nu$), we deduce that (we write $\tilde{r} = \tilde{r}_d = \tilde{r}_{d_{c'}}$ to simplify the notation)

$$\|u - Q_{c'}\|_{L^\infty(\{\tilde{r} \leqslant 2\lambda\})} \leqslant K\nu.$$

Since $(1 + o_{c'\to 0}(1))/c' = d_{c'} = d$ by Theorem 1.6, and $|dc - 1| \leqslant \nu$, we have

$$d|c - c'| \leqslant K\nu. \tag{2-5}$$

We now claim that, for a universal constant $K > 0$,

$$\|u - Q_{c'}\|_{C^1(\{\tilde{r} \leqslant \lambda\})} \leqslant K\nu. \tag{2-6}$$

That is, u is close to $Q_{c'}$ near the vortices (in the region $\{\tilde{r} \leqslant \lambda\}$) in the C^1 norm and not only in L^∞. In order to show this, we use the elliptic equation satisfied by $u - Q_{c'}$, that is,

$$\Delta(u - Q_{c'}) = -ic\,\partial_{x_2}(u - Q_{c'}) - (u - Q_{c'})(1 - |u|^2) + (|u|^2 - |Q_{c'}|^2)Q_{c'}.$$

Let us fix $x \in \{\tilde{r} \leqslant \lambda\}$. We have $\|u - Q_{c'}\|_{L^\infty(\{\tilde{r} \leqslant 2\lambda\})} \leqslant K'\nu$ by hypothesis; thus the right-hand side of the equation is small in $H^{-1}(B(x,4))$. By a standard $H^1 - H^{-1}$ estimate, we deduce

$$\|u - Q_{c'}\|_{H^1(B(x,3))} \leqslant K'\nu.$$

Then, the right-hand side is small in L^2, and standard L^2 elliptic regularity yields first

$$\|u - Q_{c'}\|_{H^2(B(x,2))} \leqslant K'\nu$$

and then

$$\|u - Q_{c'}\|_{H^3(B(x,1))} \leqslant K'\nu,$$

and we conclude by Sobolev imbedding.

Outside of this domain, u and $Q_{c'}$ are close only in modulus. Indeed, by equation (2.6) of [15] $\big($for $\sigma = \frac{1}{2}\big)$ and the hypotheses on u, we have for a universal constant $K > 0$ that on $\{\tilde{r} \geqslant \lambda\}$,

$$\big||u| - |Q_{c'}|\big| \leqslant \big||u| - 1\big| + \big||Q_{c'}| - 1\big| \leqslant \nu + \frac{K}{\lambda^{3/2}} \leqslant K'\nu.$$

Now, we modulate on the parameters of the family of traveling waves to get the orthogonality conditions of Corollary 2.6. For $c'' > 0$ close enough to c' and $X, \gamma \in \mathbb{R}$, we define

$$Q := Q_{c''}(\cdot - X\vec{e}_2)\mathrm{e}^{i\gamma}. \tag{2-7}$$

Lemma 2.7. *There exist $K > 0$, $\nu_0 > 0$ universal constants such that, for u satisfying the hypotheses of Proposition 1.8 for values of $\lambda_*, \lambda, \varepsilon(\lambda), \nu$ described above, if $\nu \leqslant \nu_0$, then there exists $c'' > 0$, $X, \gamma \in \mathbb{R}$ such that, for $R > 0$ defined in Corollary 2.6, and $\vec{d}_\pm := \pm d_{c''}\vec{e}_1 + X\vec{e}_2$,*

$$\begin{aligned}\mathfrak{Re}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} &\partial_d\big(V_1(\cdot - d\vec{e}_1 - X\vec{e}_2)V_{-1}(\cdot + d\vec{e}_1 - X\vec{e}_2)\mathrm{e}^{i\gamma}\big)_{|d=d_{c''}}\overline{(u-Q)}\\ &= \mathfrak{Re}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} \partial_{x_2}Q\overline{(u-Q)} = \mathfrak{Re}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} iQ\overline{(u-Q)} = 0.\end{aligned}$$

Furthermore,

$$\frac{|c'' - c'|}{c'^2} + |X| + |\gamma| \leqslant K\nu.$$

Proof. To simplify the notation, in this proof, we define

$$\partial_d V := \partial_d\big(V_1(\cdot - d\vec{e}_1 - X\vec{e}_2)V_{-1}(\cdot + d\vec{e}_1 + X\vec{e}_2)\mathrm{e}^{i\gamma}\big)_{|d=d_{c''}}.$$

We will keep the notation $\tilde{r}$ for the minimum of the distance to the zeros of Q.

First, from equation (7.5) of [15], there exists a universal constant $K > 0$ such that, for $c'' < c_0$, $c'/2 \leqslant c'' \leqslant 2c'$,

$$\|Q - Q_{c'}\|_{L^\infty(\mathbb{R}^2)} \leqslant K\left(|X| + \frac{|c'' - c'|}{c'^2} + |\gamma|\right). \tag{2-8}$$

Now, we follow closely the proof of Lemma 7.6 of [15], which is done in Appendix C.3 there. We define

$$G\begin{pmatrix} X \\ c'' \\ \gamma \end{pmatrix} := \begin{pmatrix} \mathfrak{Re} \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} \partial_{x_2} Q\overline{(u-Q)} \\ \mathfrak{Re} \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} \partial_d V\overline{(u-Q)} \\ \mathfrak{Re} \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} iQ\overline{(u-Q)} \end{pmatrix}.$$

Note that Q, $\partial_d V$ and $\vec{d}_\pm$ all depend on X and c'', and Q depends also on γ. From (2-6) and the fact that $\lambda \geqslant \lambda_* > 2R$, we have $\|u - Q_{c'}\|_{L^\infty(\{\tilde{r}\leqslant R\})} \leqslant K\nu$, and from Theorem 1.6 with $p = +\infty$, as well as Lemma 2.6 of [14],

$$\|\partial_{x_2} Q_{c'}\|_{L^\infty(\mathbb{R}^2)} + \|\partial_d V\|_{L^\infty(\mathbb{R}^2)} + \|iQ_{c'}\|_{L^\infty(\mathbb{R}^2)} \leqslant K \tag{2-9}$$

for some universal constant $K > 0$. Therefore, since $Q = Q_{c'}$ for $X = \gamma = 0,\ c'' = c'$, we obtain

$$\left| G\begin{pmatrix} 0 \\ c' \\ 0 \end{pmatrix} \right| \leqslant K\|u - Q_{c'}\|_{L^\infty(\{\tilde{r}\leqslant\lambda\})} \leqslant K\nu.$$

We want to show that G is invertible in a vicinity of $(0\ c'\ 0)^\top$. With (2-6) and (2-8), we check that (we recall that $\tilde{r} = \min(|x - \vec{d}_+|, |x - \vec{d}_-|)$)

$$\begin{aligned} \|u - Q\|_{L^\infty(\{\tilde{r}\leqslant 2R\})} &\leqslant \|u - Q_{c'}\|_{L^\infty(\{\tilde{r}\leqslant 2R\})} + \|Q - Q_{c'}\|_{L^\infty(\mathbb{R}^2)} \\ &\leqslant K\nu + K\left(|X| + \frac{|c'' - c'|}{c'^2} + |\gamma|\right), \end{aligned}$$

and as in Lemma 7.1 of [15], this implies

$$\|u - Q\|_{C^1(\{\tilde{r}\leqslant R\})} \leqslant K\nu + K\left(|X| + \frac{|c'' - c'|}{c'^2} + |\gamma|\right). \tag{2-10}$$

Now, we compute

$$\begin{aligned} &\left| \partial_X\left(\mathfrak{Re} \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} \partial_{x_2} Q\overline{(u-Q)} \right) - \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} |\partial_{x_2} Q|^2 \right| \\ &\qquad \leqslant \int_{\partial B(\vec{d}_+,R)\cup \partial B(\vec{d}_-,R)} |\partial_{x_2} Q\overline{(u-Q)}| + \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} |\partial_{x_2}^2 Q\overline{(u-Q)}|. \end{aligned}$$

Therefore, with (2-1) and (2-10), we check that

$$\int_{\partial B(\vec{d}_+,R)\cup \partial B(\vec{d}_-,R)} |\partial_{x_2} Q\overline{(u-Q)}| + \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} |\partial_{x_2}^2 Q\overline{(u-Q)}| \leqslant K\nu + K\left(|X| + \frac{|c'' - c'|}{c'^2} + |\gamma|\right);$$

hence

$$\left| \partial_X\left(\mathfrak{Re} \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} \partial_{x_2} Q\overline{(u-Q)} \right) - \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)} |\partial_{x_2} Q|^2 \right| \leqslant K\nu + K\left(|X| + \frac{|c''-c'|}{c'^2} + |\gamma|\right).$$

With similar computations, using Lemma 2.6 of [14], (2-1) and (2-10), we infer that

$$\left|\partial_X G-\begin{pmatrix}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}|\partial_{x_2}Q|^2\\ \mathfrak{Re}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}\partial_d V\overline{\partial_{x_2}Q}\\ \mathfrak{Re}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}iQ\overline{\partial_{x_2}Q}\end{pmatrix}\right|\leqslant K\nu+K\biggl(|X|+\frac{|c''-c'|}{c'^2}+|\gamma|\biggr).$$

By the symmetries of $Q(\cdot+X\vec{e}_2)\mathrm{e}^{-i\gamma}$ and $\partial_d V(\cdot+X\vec{e}_2)\mathrm{e}^{-i\gamma}$, we have that

$$\mathfrak{Re}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}\partial_d V\overline{\partial_{x_2}Q}=0,$$

and from Theorem 1.6 (with $p=+\infty$), with the symmetries of Q_c and V_1 (see Sections 2A1 and 2A3), we have

$$\left|\mathfrak{Re}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}iQ\overline{\partial_{x_2}Q}-2\,\mathfrak{Re}\int_{B(0,R)}iV_1\overline{\partial_{x_2}V_1}\right|\leqslant K\biggl(|X|+\frac{|c''-c'|}{c'^2}\biggr).$$

By decomposition in harmonics and Lemma 2.1, we check easily that $\mathfrak{Re}\int_{B(0,R)}iV_1\overline{\partial_{x_2}V_1}=0$; thus

$$\left|\partial_X G-\begin{pmatrix}\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}|\partial_{x_2}Q|^2\\ 0\\ 0\end{pmatrix}\right|\leqslant K\nu+K\biggl(|X|+\frac{|c''-c'|}{c'^2}+|\gamma|\biggr).$$

Similarly, we check that (using $\partial_c(d_c)=(-1+o_{c\to 0}(1))/c^2$ from Section 4.6 and Lemma 2.6 of [14])

$$\left|c'^2\partial_{c''}G-\begin{pmatrix}0\\ \int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}|\partial_d V|^2\\ 0\end{pmatrix}\right|\leqslant K\nu+K\biggl(|X|+\frac{|c''-c'|}{c'^2}+|\gamma|\biggr)$$

(we use here the fact that $c\mapsto\partial_d V$ and $c\mapsto\vec{d}_\pm$ are differentiable) and

$$\left|\partial_\gamma G-\begin{pmatrix}0\\ 0\\ -\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}|Q|^2\end{pmatrix}\right|\leqslant K\nu+K\biggl(|X|+\frac{|c''-c'|}{c'^2}+|\gamma|\biggr).$$

From (2-1) and Theorem 1.6 (for $p=+\infty$) as well as Lemma 2.6 of [14], there exists a universal constant $K>0$ such that

$$\frac{1}{K}\leqslant\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}|\partial_{x_2}Q|^2\leqslant K,$$
$$\frac{1}{K}\leqslant\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}|\partial_d V|^2\leqslant K,$$
$$\frac{1}{K}\leqslant\int_{B(\vec{d}_+,R)\cup B(\vec{d}_-,R)}|Q|^2\leqslant K,$$

provided $|X|+c''$ is small enough. We deduce that there exists K_1, K_2, $\nu_0>0$ such that, for $0<\nu\leqslant\nu_0$ and u satisfying the hypotheses of Proposition 1.8 with the parameters λ, ν, dG is invertible in the ball

$$\biggl\{(X,c'',\gamma)\in\mathbb{R}^3:|X|+\frac{|c''-c'|}{c'^2}+|\gamma|\leqslant K_1\nu\biggr\},$$

and there exists $X, c'', \gamma \in \mathbb{R}$ such that

$$G\begin{pmatrix} X \\ c'' \\ \gamma \end{pmatrix} = 0,$$

with

$$\frac{|c''-c'|}{c'^2} + |X| + |\gamma| \leqslant K_2 \nu. \qquad \square$$

2B2. *Construction and properties of the perturbation term.* We define η a smooth cutoff function with

$$\eta(x) = \begin{cases} 0 & \text{for } x \in B(\vec{d}_\pm, 2R), \\ 1 & \text{for } x \in \mathbb{R}^2 \backslash B(\pm\vec{d}_\pm, 2R+1), \end{cases}$$

which is even in x_1. We infer the following result, where the space $H_Q^{\exp,s}$ is simply defined by

$$H_Q^{\exp,s} := \{\varphi \in H^1_{\mathrm{loc}}(\mathbb{R}^2, \mathbb{C}) : \|\varphi\|_{H_Q^{\exp}} < +\infty \text{ for all } (x_1, x_2) \in \mathbb{R}^2,\ \varphi(-x_1, x_2) = \varphi(x_1, x_2)\},$$

with, for $\tilde{r}$ the minimum of the distances to the zeros of Q, $\varphi = Q\psi$,

$$\|\varphi\|^2_{H_Q^{\exp}} := \|\varphi\|^2_{H^1(\{\tilde{r}\leqslant 10\})} + \int_{\{\tilde{r}\geqslant 5\}} |\nabla\psi|^2 + \mathfrak{Re}^2(\psi) + \frac{|\psi|^2}{\tilde{r}^2 \ln^2 \tilde{r}},$$

and $B_Q^{\exp}$ has the same definition as $B_{Q_c}^{\exp}$, replacing $\tilde{\eta}$ by η and Q_c by Q.

Lemma 2.8. *There exist $K_1, K_2 > 0$, $\nu_0 > \nu_1 > 0$ universal constants such that, for u satisfying the hypotheses of Proposition 1.8 for values of $\lambda_*, \lambda, \varepsilon(\lambda), \nu$ described above, if $\nu \leqslant \nu_1$, then there exists a function $\varphi = Q\psi \in H_Q^{\exp,s} \cap C^1(\mathbb{R}^2, \mathbb{C})$ such that, for Q defined in* (2-7) *with the values of $c'', X, \gamma \in \mathbb{R}$ from Lemma 2.7*,

$$u - Q = (1-\eta)\varphi + \eta Q(\mathrm{e}^{\psi} - 1).$$

Furthermore,

$$B_Q^{\exp}(\varphi) \geqslant K_1 \|\varphi\|^2_{H_Q^{\exp}}$$

and

$$\|\varphi\|_{C^1(\{\tilde{r}\leqslant\lambda\})} + \|\,\mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde{r}\geqslant\lambda\})} \leqslant K_2 \nu.$$

The goal of this lemma is to decompose the error $u - Q$ into a particular form. In the area $\{\eta = 1\}$, that is, far from the zeros of Q, the error is written in an exponential form: $u = Q\mathrm{e}^{\psi}$. This form was already used in [14; 15], and it is useful to have a particular form on the cubic error terms. Furthermore, we fix the parameters of Q such that φ satisfies the orthogonality conditions of Corollary 2.6, yielding the coercivity.

Note that we have no smallness on $\mathfrak{Im}(\psi)$ in $\{\tilde{r} \geqslant \lambda\}$, where $\varphi = Q\psi$. We will simply be able to show that it is bounded (see (2-11) below), with no a priori bound on it. This lack of smallness is one of the main difficulties in the proof of Proposition 1.8. Analogously, we show that $\varphi \in H_Q^{\exp,s}$, but we have no good control on $\|\varphi\|_{H_Q^{\exp}}$: this quantity might be a priori very large at this point.

Proof. This proof follows some ideas of the proofs of Lemmas 7.2 and 7.3 of [15]. First, in the area $\{\tilde{r} \leqslant \lambda\}$, the proof is identical to that of Lemma 7.2 of [15] for the existence of $\varphi = Q\psi \in C^1(\{\tilde{r} \leqslant \lambda\}, \mathbb{C})$ such that $u - Q = (1-\eta)\varphi + \eta Q(\mathrm{e}^{\psi} - 1)$ in $\{\tilde{r} \leqslant \lambda\}$, with $\|\varphi\|_{C^1(\{\tilde{r}\leqslant\lambda\})} \leqslant K\nu$ (this is a consequence of the estimate $\|u-Q\|_{C^1(\{\tilde r\leqslant\lambda\})}\leqslant K\nu$, obtained using Lemma 2.7). The main idea is that $u-Q$ is small there (in $C^1(\{\tilde r\leqslant\lambda\},\mathbb{C})$), and the equation on φ is a perturbation of the identity for functions φ that are small in $C^1(\{\tilde r\leqslant\lambda\},\mathbb{C})$. In particular, since u and Q are symmetric with respect to the x_2-axis, φ and ψ are also symmetric with respect to the x_2-axis.

We then focus our attention in the area $\{\tilde r\geqslant\lambda\}$, where $\eta\equiv 1$, so that the problem reduces to the equation

$$u=Q\mathrm{e}^{\psi}.$$

By Theorem 1.6 and the hypotheses of Proposition 1.8, there exists $\nu_1>0$ such that, if $\nu\leqslant\nu_1$, then, as a consequence of

$$\varepsilon(\lambda)\leqslant\min\left(\nu_1,\frac{1}{10\lambda^2+100}\right),$$

the domain $\{\tilde r\geqslant\lambda\}$ consists of the complement of the two disjointed disks $B(\vec d_\pm,\lambda)$, with

$$|Q|\geqslant\tfrac12,\quad |u|\geqslant\tfrac12\quad\text{in }\{\tilde r\geqslant\lambda\}$$

and

$$\deg(Q,\partial B(\vec d_\pm,\lambda))=\deg(u,\partial B(\vec d_\pm,\lambda))=\pm1,$$

so that u/Q is smooth in $\{\tilde r\geqslant\lambda\}=\mathbb{R}^2\setminus(B(\vec d_+,\lambda)\cup B(\vec d_-,\lambda))$, does not vanish and has degree zero on the two circles $\partial B(\vec d_\pm,\lambda)$. It then follows from standard lifting theorems (even though $\{\tilde r\geqslant\lambda\}$ is not simply connected) that there exists $\psi^\dagger\in C^1(\{\tilde r\geqslant\lambda\})$ such that $\mathrm{e}^{\psi^\dagger}=u/Q$, as wished. We then notice that u and Q are symmetric with respect to the x_2-axis; thus $x\mapsto\psi^\dagger(-x_1,x_2)$ is also a lifting of u/Q in the connected set $\{\tilde r\geqslant\lambda\}$, which implies that there exists $q\in\mathbb{Z}$ such that $\psi^\dagger(-x_1,x_2)=\psi^\dagger(x_1,x_2)+2iq\pi$ in $\{\tilde r\geqslant\lambda\}$. Letting $x_1=0$, we obtain $q=0$; $\psi^\dagger$ is also symmetric with respect to the x_2-axis.

Recalling that $\psi:=\varphi/Q$ in the set $\{\lambda\leqslant\tilde r\leqslant 2\lambda\}$ (where Q does not vanish), we see that, since $\eta\equiv1$ there, the equation $u-Q=(1-\eta)\varphi+\eta Q(\mathrm{e}^{\psi}-1)$ becomes $u=Q\mathrm{e}^{\psi}$. We then infer that there exists $m\in\mathbb{Z}$ such that $\psi=\psi^\dagger+2im\pi$ in the connected annulus $B(\vec d_+,2\lambda)\setminus B(\vec d_+,\lambda)$. By symmetry in x_1, this is also true in the annulus $B(\vec d_-,2\lambda)\setminus B(\vec d_-,\lambda)$. It then suffices to extend ψ by the formula $\psi=\psi^\dagger+2im\pi$ in $\{\tilde r\geqslant\lambda\}$ to obtain the formula $u-Q=(1-\eta)\varphi+\eta Q(\mathrm{e}^{\psi}-1)$. In the region $\{\tilde r\geqslant\lambda\}$, the relation $u=Q\mathrm{e}^{\psi}$ yields

$$\mathrm{e}^{\mathfrak{Re}(\psi)}=\left|\frac{u}{Q}\right|;$$

thus, taking the decomposition

$$\left|\frac{u}{Q}\right|=1+|u|-1+\frac{(|u|-1)-(|Q|-1)}{|Q|},$$

since there exists a universal constant $K'>0$ such that in this region

$$\left||u|-1+\frac{(|u|-1)-(|Q|-1)}{|Q|}\right|\leqslant K'\nu,$$

we deduce that, for $\nu\leqslant\nu_1$ with ν_1 small enough,

$$\|\mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})}\leqslant K\nu.$$

Since u is a traveling wave and $E(u) < +\infty$, u converges to a constant at infinity (uniformly in all directions) by [24]. Therefore, u/Q converges to a constant at infinity, and the function ψ converges to a constant, and thus it is bounded near infinity, that is,

$$\|\psi\|_{L^\infty(\{\tilde{r} \geqslant \lambda\})} < +\infty. \tag{2-11}$$

Now, we want to show that $\varphi \in H_Q^{\exp,s}$. We already know that φ satisfies the symmetry,

$$\text{for all } (x_1, x_2) \in \mathbb{R}^2, \quad \varphi(-x_1, x_2) = \varphi(x_1, x_2).$$

Furthermore, to check that $\|\varphi\|_{H_Q^{\exp}} < +\infty$, since $\varphi \in C^1(\mathbb{R}^2, \mathbb{C})$, we only have to check the integrability in $\{\tilde{r} \geqslant \lambda\}$, where $\mathrm{e}^\psi = u/Q$. We check that there, with (2-11),

$$\int_{\{\tilde{r} \geqslant \lambda\}} \frac{|\psi|^2}{\tilde{r}^2 \ln^2(\tilde{r})} < +\infty.$$

Now, using Theorem 11 of [24] (we recall that $E(u) < +\infty,\ E(Q) < +\infty$),

$$|\mathrm{e}^{\mathfrak{Re}(\psi)} - 1| = \frac{\big||u| - |Q|\big|}{|Q|} \leqslant 2\big(\big||u| - 1\big| + \big||Q| - 1\big|\big) \leqslant \frac{K(u, c, Q, c'')}{(1 + r)^2},$$

where $K(u, c, Q, c'') > 0$ is a constant depending on u, c, c'' and Q; hence

$$|\mathfrak{Re}(\psi)| \leqslant \frac{K(u, c, Q, c'')}{(1 + r)^2}$$

and

$$\int_{\{\tilde{r} \geqslant \lambda\}} \mathfrak{Re}^2(\psi) \leqslant \int_{\{\tilde{r} \geqslant \lambda\}} \frac{K(u, c, Q, c'')}{(1 + r)^4} < +\infty.$$

We finally compute

$$\nabla\psi = \frac{\nabla u}{u} - \frac{\nabla Q}{Q},$$

and with Theorem 11 of [24], in $\{\tilde{r} \geqslant \lambda\}$, we deduce that

$$(1 + r)^2 |\nabla\psi| \leqslant (1 + r)^2 \left|\frac{\nabla u}{u}\right| + (1 + r)^2 \left|\frac{\nabla Q}{Q}\right| \leqslant K(u, c, Q, c'');$$

therefore

$$\int_{\{\tilde{r} \geqslant \lambda\}} |\nabla\psi|^2 < +\infty.$$

This concludes the proof that $\varphi = Q\psi \in H_Q^{\exp,s}$. The fact that $B_Q^{\exp}(\varphi) \geqslant K\|\varphi\|^2_{H_Q^{\exp}}$ is a consequence of Corollary 2.6 and Lemma 2.7, using in particular that

$$B_Q^{\exp}(\varphi) = B_{Q_{c''}}^{\exp}(\varphi(\cdot + X\vec{e}_2)\mathrm{e}^{-i\gamma}) \quad \text{and} \quad \|\varphi\|_{H_Q^{\exp}} = \|\varphi(\cdot + X\vec{e}_2)\mathrm{e}^{-i\gamma}\|_{H_{Q_{c''}}^{\exp}}. \qquad \square$$

We now compute the equation satisfied by φ. By Lemma 2.8, in $\{0 < \eta < 1\} = \{2R < \tilde{r} < 2R + 1\}$, we have $|\mathfrak{Re}(\psi)| = |\mathfrak{Re}(\varphi/Q)| \leqslant K\nu$ uniformly; thus $|\mathrm{e}^{\mathfrak{Re}(\psi)} - 1| \leqslant K\nu$ uniformly in this region and then $|(1 - \eta) + \eta\mathrm{e}^\psi| \geqslant \frac{1}{2}$ for $\nu \leqslant \nu_1$, possibly diminishing ν_1 of Lemma 2.8.

Lemma 2.9. *For u satisfying the hypotheses of Proposition 1.8 for values of λ_*, λ, $\varepsilon(\lambda)$, ν described above, if $\nu \leqslant \nu_1$ (where ν_1 is defined in Lemma 2.8), then the function $\varphi = Q\psi$ defined in Lemma 2.8 satisfies the equation*

$$L_Q(\varphi) - i(c-c'')\vec{e}_2.H(\psi) + \mathrm{NL}_{\mathrm{loc}}(\psi) + F(\psi) = 0,$$

with L_Q the linearized operator around Q: $L_Q(\varphi) = -\Delta\varphi - ic''\partial_{x_2}\varphi - (1-|Q|^2)\varphi + 2\,\mathfrak{Re}(\overline{Q}\varphi)Q$,

$$S(\psi) := \mathrm{e}^{2\,\mathfrak{Re}(\psi)} - 1 - 2\,\mathfrak{Re}(\psi),$$
$$F(\psi) := Q\eta(-\nabla\psi.\nabla\psi + |Q|^2 S(\psi)),$$
$$H(\psi) := \nabla Q + \frac{\nabla(Q\psi)(1-\eta) + Q\nabla\psi\eta\mathrm{e}^{\psi}}{(1-\eta)+\eta\mathrm{e}^{\psi}}$$

and $\mathrm{NL}_{\mathrm{loc}}(\psi)$ is a sum of terms at least quadratic in ψ, localized in the area where $\eta \neq 1$. Furthermore,

$$|\langle \mathrm{NL}_{\mathrm{loc}}(\psi), Q\psi\rangle| \leqslant K\|\mathrm{NL}_{\mathrm{loc}}(\psi)\|_{L^2(\{\eta<1\})}\|\varphi\|_{L^\infty(\{\eta<1\})} \leqslant K\nu\|\varphi\|^2_{H^1(\{\eta\neq 1\})}.$$

Notice that $F(\psi)$ (the notation $X.Y$ for complex vector fields stands for $X_1Y_1 + X_2Y_2$) contains all the nonlinear terms far from the zeros of Q, and its structure relies on the fact that the error is written in an exponential form far from the vortices. Close to the zeros of Q, this particular form does not hold, but it will not be necessary, since there the error φ is small in the C^1 norm, whereas, at infinity, it is small only in a weaker norm.

Proof. The proof is identical to the proof of Lemma 7.5 of [15], and it is in the particular case where all the speeds are along $\vec{e}_2$. The proof consists simply of decomposing the equation

$$0 = (\mathrm{TW}_c)(u) = \mathrm{TW}_c(Q + (1-\eta)\varphi + \eta Q(\mathrm{e}^{\psi}-1))$$

into the different terms.

The last estimate uses Lemmas 2.8 and 2.7. □

This result shows in particular that $\psi \in C^2(\{\eta \neq 0\}, \mathbb{C})$, and we can check with it, as in Lemma 7.3 of [15], that $\|\Delta\psi(1+r)^2\|_{L^\infty(\{\tilde{r}\geqslant\lambda\})} \leqslant K(u, Q, c, c'')$.

We now infer a critical estimate on the differences of the speeds of the problem, namely c (the speed of u) and c'' (the speed of Q). The method for the estimate has been used in [15] (we take the scalar product of the equation of Lemma 2.9 with $\partial_c Q$), but since we have worse estimates on the error term, we need to be more careful ($\|\varphi\|_{H_Q^{\exp}}$ is not a priori small at this point).

Lemma 2.10. *There exist universal constants $K > 0$, $\nu_1 \geqslant \nu_2 > 0$ (where ν_1 is defined in Lemma 2.8), such that, for u satisfying the hypotheses of Proposition 1.8 for values of λ_*, λ, $\varepsilon(\lambda)$, ν described above, if $\nu \leqslant \nu_2$, then, with $\varphi = Q\psi$ defined in Lemma 2.8, we have*

$$|c''-c| \leqslant K\sqrt{c''}\|\varphi\|_{H_Q^{\exp}}.$$

Proof. First, from (2-5) and Lemma 2.7, taking $\nu > 0$ small enough, we have

$$|c''-c| \leqslant |c''-c'| + |c'-c| \leqslant Kc''. \tag{2-12}$$

We will show the estimate

$$|c''-c| \leqslant K\Big(c''^2 \ln\Big(\frac{1}{c''}\Big)\|\varphi\|_{H_Q^{\exp}} + \|\varphi\|^2_{H_Q^{\exp}}\Big) + K|c''-c|\,\|\varphi\|_{H_Q^{\exp}}. \tag{2-13}$$

This is related to equation (7.13) of [15] (its proof is in Step 1 in Section 7.3.1 of [15]). With both estimates, we can conclude the proof of this lemma. Indeed, either $\|\varphi\|_{H_Q^{\exp}} \geqslant \sqrt{c''}$, and in that case

$$|c''-c| \leqslant Kc'' \leqslant K\sqrt{c''}\|\varphi\|_{H_Q^{\exp}},$$

or $\|\varphi\|_{H_Q^{\exp}} \leqslant \sqrt{c''}$, and then with (2-13),

$$\begin{aligned}|c''-c| &\leqslant K\Big(c''^2 \ln\Big(\frac{1}{c''}\Big)\|\varphi\|_{H_Q^{\exp}} + \|\varphi\|^2_{H_Q^{\exp}}\Big) + K|c''-c|\,\|\varphi\|_{H_Q^{\exp}}\\ &\leqslant K\sqrt{c''}\|\varphi\|_{H_Q^{\exp}} + C_2\sqrt{c''}|c''-c|.\end{aligned}$$

Therefore, for $c''>0$ small enough such that $C_2\sqrt{c''}<\frac{1}{2}$ (which is implied by taking $\nu>0$ small enough, independently of λ), we have $|c''-c| \leqslant K\sqrt{c''}\|\varphi\|_{H_Q^{\exp}}$.

We now focus on the proof of (2-13). We take the scalar product of the equation

$$L_Q(\varphi) - i(c-c'')\vec{e}_2.H(\psi) + \mathrm{NL}_{\mathrm{loc}}(\psi) + F(\psi) = 0,$$

with $c''^2\partial_{c''}Q$. We estimate, as in Section 7.3.1 of [15], that

$$|\langle L_Q(\varphi), c''^2\partial_{c''}Q\rangle| = c''^2|\langle\varphi, L_Q(\partial_{c''}Q)\rangle| = c''^2|\langle\varphi, i\partial_{x_2}Q\rangle| \leqslant Kc''^2\ln\Big(\frac{1}{c''}\Big)\|\varphi\|_{H_Q^{\exp}}.$$

We recall that

$$i\vec{e}_2.H(\psi) = i\partial_{x_2}Q + i\,\frac{\partial_{x_2}(Q\psi)(1-\eta) + Q\partial_{x_2}\psi\,\eta e^{\psi}}{(1-\eta)+\eta e^{\psi}},$$

and we check that (estimating the local terms in the area where $\eta\neq 1$ by Cauchy–Schwarz and $\|c''^2\partial_{c''}Q\|_{L^\infty(\mathbb{R}^2)} \leqslant K$ from Theorem 1.6 for $p=+\infty$ and Lemma 2.6 of [14])

$$\begin{aligned}&|(c-c'')\langle i\vec{e}_2.H(\psi), c''^2\partial_{c''}Q\rangle - (c-c'')\langle i\partial_{x_2}Q, c''^2\partial_{c''}Q\rangle|\\ &\qquad\leqslant K(|c-c''|\,\|\varphi\|_{H^1(\{\eta\neq 1\})} + |(c-c'')\langle\eta Qi\partial_{x_2}\psi, c''^2\partial_{c''}Q\rangle|)\\ &\qquad\leqslant K(|c-c''|\,\|\varphi\|_{H_Q^{\exp}} + |(c-c'')\langle\eta Qi\partial_{x_2}\psi, c''^2\partial_{c''}Q\rangle|).\end{aligned}$$

We recall from Section 7.3.1 of [15] (using decay estimates on $c''^2\partial_{c''}Q\bar{Q}$ and integrations by parts), that

$$|(c-c'')\langle\eta Qi\partial_{x_2}\psi, c''^2\partial_{c''}Q\rangle| \leqslant K|c-c''|\,\|\varphi\|_{H_Q^{\exp}}$$

and, from Proposition 1.2 of [15] (we check easily that the translation and phase on Q instead of $Q_{c''}$ do not change the computation),

$$(c-c'')\langle i\partial_{x_2}Q, c''^2\partial_{c''}Q\rangle = (2\pi + o_{c''\to 0}(1))(c-c'') = (2\pi + o_{\nu\to 0}(1))(c-c'').$$

We deduce that, taking $\nu>0$ small enough (independently of λ), that

$$|c-c''| \leqslant Kc''^2\ln\Big(\frac{1}{c''}\Big)\|\varphi\|_{H_Q^{\exp}} + K|c-c''|\,\|\varphi\|_{H_Q^{\exp}} + K|\langle\mathrm{NL}_{\mathrm{loc}}(\psi) + F(\psi), c''^2\partial_{c''}Q\rangle|.$$

We take $\nu_2 > 0$ with $\nu_2 \leqslant \nu_1$ such that all the above conditions on the smallness of ν are satisfied if $\nu \leqslant \nu_2$. Since $\mathrm{NL}_{\mathrm{loc}}(\psi)$ contains terms at least quadratic in φ, $\|\varphi\|_{C^1(\{\eta\neq 1\})} \leqslant C_3\nu$ from Lemma 2.8 and $\|c''^2\partial_{c''}Q\|_{L^\infty(\mathbb{R}^2)} \leqslant K$, we obtain that for $\nu \leqslant \nu_2$, diminishing ν_2 if necessary so that $\|\varphi\|_{C^1(\{\eta\neq 1\})} \leqslant K\nu \leqslant 1$,

$$|\langle \mathrm{NL}_{\mathrm{loc}}(\psi), c''^2\partial_{c''}Q\rangle| \leqslant K\|\varphi\|^2_{H^1(\{\eta\neq 1\})} \leqslant K\|\varphi\|^2_{H_Q^{\exp}}.$$

Finally, we estimate, using $\|c''^2\partial_{c''}Q\|_{L^\infty(\mathbb{R}^2)} \leqslant K$,

$$|\langle Q\eta\nabla\psi.\nabla\psi, c''^2\partial_{c''}Q\rangle| \leqslant K\int_{\mathbb{R}^2}\eta|\nabla\psi|^2\|c''^2\partial_{c''}Q\|_{L^\infty(\mathbb{R}^2)} \leqslant K\|\varphi\|^2_{H_Q^{\exp}}.$$

Similarly, since $\|\eta\,\mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})} \leqslant K\nu$ by Lemma 2.8, diminishing ν_2 if necessary, for $\nu \leqslant \nu_2$, we have $\|\eta\,\mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})} \leqslant 1$, and hence

$$|Q\eta|Q|^2S(\psi)| = |Q\eta|Q|^2(\mathrm{e}^{2\mathfrak{Re}(\psi)} - 1 - 2\,\mathfrak{Re}(\psi))| \leqslant K\eta\,\mathfrak{Re}^2(\psi).$$

Therefore

$$|\langle Q\eta|Q|^2S(\psi), c''^2\partial_{c''}Q\rangle| \leqslant K\int_{\mathbb{R}^2}\eta\,\mathfrak{Re}^2(\psi)\|c''^2\partial_{c''}Q\|_{L^\infty(\mathbb{R}^2)} \leqslant K\|\varphi\|^2_{H_Q^{\exp}}.$$

This concludes the proof of (2-13), and therefore of the lemma. □

2B3. *Proof of Proposition 1.8 completed.* We take u satisfying the hypotheses of Proposition 1.8 for values of λ_*, λ, $\varepsilon(\lambda)$, ν described above, with $\nu \leqslant \nu_2$, where ν_2 is defined in Lemma 2.10. We want to take the scalar product of the equation of Lemma 2.9 with φ. It is however not clear at this point that every term is integrable. In Section 7.3 of [15], we took the scalar product of the equation with $\varphi + i\gamma Q$ for some $\gamma \in \mathbb{R}$, using a decay estimate $\|\,\mathfrak{Im}(\psi + i\gamma)(1+r)\|_{L^\infty(\{\tilde r\leqslant\lambda\})} \leqslant K(u, Q, c, c'')$ to justify that some terms are well-defined, and to do some integration by parts. Here, we need to change our approach a little. We first require better decay estimates on ψ. At this stage, we know (see Theorem 11 of [24] and the proof of Lemma 2.8) that

$$\|\Delta\psi(1+r)^2\|_{L^\infty(\{\tilde r\geqslant\lambda\})} + \|(1+r)^2\nabla\psi\|_{L^\infty(\{\tilde r\geqslant\lambda\})} + \|\psi\|_{L^\infty(\{\tilde r\geqslant\lambda\})} + \|(1+r)^2\,\mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})} \leqslant K(u, Q, c, c'').$$

Now, let us show the following improvements:

$$\|\,\mathfrak{Im}(\Delta\psi)(1+r)^3\|_{L^\infty(\{\tilde r\geqslant\lambda\})} + \|(1+r)^3\,\mathfrak{Re}(\nabla\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})} \leqslant K(u, Q, c, c''). \tag{2-14}$$

The proof of $\|(1+r)^3|\,\mathfrak{Re}(\nabla\psi)|\|_{L^\infty(\{\tilde r\geqslant\lambda\})} \leqslant K(u, Q, c, c'')$ is identical to the one for the same result in Lemma 7.3 of [15] (see the penultimate estimate of its proof). We focus on the estimate on $\mathfrak{Im}(\Delta\psi)$. In $\{\tilde r \geqslant \lambda\}$, we have $u = Q\mathrm{e}^{\psi}$; therefore,

$$\Delta\psi = -\frac{\Delta Q}{Q} + \frac{\Delta u}{u} - 2\frac{\nabla Q}{Q}.\nabla\psi - \nabla\psi.\nabla\psi.$$

With the previous estimates and Theorem 11 of [24], we have

$$\left\|\left(-2\frac{\nabla Q}{Q}.\nabla\psi - \nabla\psi.\nabla\psi\right)(1+r)^4\right\|_{L^\infty(\{\tilde r\geqslant\lambda\})} \leqslant K(u, Q, c, c''),$$

and since $(\mathrm{TW}_{c''})(Q)=0$,

$$\frac{\Delta Q}{Q}=ic''\frac{\partial_{x_2}Q}{Q}-(1-|Q|^2);$$

therefore, with [24] ($E(Q)<+\infty$),

$$\left|\mathfrak{Im}\left(\frac{\Delta Q}{Q}\right)\right|\leqslant c''\left|\mathfrak{Re}\left(\frac{\partial_{x_2}Q}{Q}\right)\right|\leqslant\frac{K(Q,c'')}{(1+r)^3}.$$

Similarly, since $(\mathrm{TW}_c)(u)=0$ and $E(u)<+\infty$,

$$\left|\mathfrak{Im}\left(\frac{\Delta u}{u}\right)\right|\leqslant c\left|\mathfrak{Re}\left(\frac{\partial_{x_2}u}{u}\right)\right|\leqslant\frac{K(u,c)}{(1+r)^3};$$

thus

$$\|\,\mathfrak{Im}(\Delta\psi)(1+r)^3\|_{L^\infty(\{\tilde r\geqslant\lambda\})}\leqslant K(u,Q,c,c'').$$

We infer, with these two additional estimates on ψ, that we can do the same computations as in the proof of [15, Lemma 7.4], with $\gamma=0$. The only difference is that where we used $\|\,\mathfrak{Im}(\psi+i\gamma)(1+r)\|_{L^\infty(\{\tilde r\geqslant\lambda\})}\leqslant K(u,Q)$ we can use (2-14) instead to get the same decay for these terms, with $\|\,\mathfrak{Im}(\psi)\|_{L^\infty(\{\tilde r\leqslant\lambda\})}\leqslant K(u,Q)$. The only two terms where this change is needed are

$$\begin{aligned}\left|\int_{\mathbb{R}}\eta|Q|^2\,\mathfrak{Re}(\Delta\psi\bar\psi)\right|&\leqslant\left|\int_{\mathbb{R}}\eta|Q|^2\,\mathfrak{Re}(\Delta\psi)\,\mathfrak{Re}(\psi)\right|+\left|\int_{\mathbb{R}}\eta|Q|^2\,\mathfrak{Im}(\Delta\psi)\,\mathfrak{Im}(\psi)\right|\\&\leqslant K(\|\,\mathfrak{Re}(\Delta\psi)(1+r)^2\|_{L^\infty(\{\tilde r\geqslant\lambda\})}\|\,\mathfrak{Re}(\psi)(1+r)^2\|_{L^\infty(\{\tilde r\geqslant\lambda\})})\\&\qquad+K(\|\,\mathfrak{Im}(\Delta\psi)(1+r)^3\|_{L^\infty(\{\tilde r\geqslant\lambda\})}\|\,\mathfrak{Im}(\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})})\end{aligned}$$

and

$$\begin{aligned}\left|\int_{\mathbb{R}}\eta|Q|^2\,\mathfrak{Re}(i\partial_{x_2}\psi\bar\psi)\right|&\leqslant\left|\int_{\mathbb{R}}\eta|Q|^2\,\mathfrak{Re}(\partial_{x_2}\psi)\,\mathfrak{Im}(\psi)\right|+\left|\int_{\mathbb{R}}\eta|Q|^2\,\mathfrak{Im}(\partial_{x_2}\psi)\,\mathfrak{Re}(\psi)\right|\\&\leqslant K(\|\,\mathfrak{Re}(\partial_{x_2}\psi)(1+r)^3\|_{L^\infty(\{\tilde r\geqslant\lambda\})}\|\,\mathfrak{Im}(\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})})\\&\qquad+K(\|\,\mathfrak{Im}(\partial_{x_2}\psi)(1+r)^2\|_{L^\infty(\{\tilde r\geqslant\lambda\})}\|\,\mathfrak{Re}(\psi)(1+r)^2\|_{L^\infty(\{\tilde r\geqslant\lambda\})}).\end{aligned}$$

We deduce, taking the scalar product of the equation of Lemma 2.9 with φ, that

$$B_Q^{\exp}(\varphi)-\langle i(c-c'')\vec e_2.H(\psi),\varphi\rangle+\langle\mathrm{NL}_{\mathrm{loc}}(\psi),\varphi\rangle+\langle F(\psi),\varphi\rangle=0.\tag{2-15}$$

From Lemma 2.8,

$$B_Q^{\exp}(\varphi)\geqslant K\|\varphi\|^2_{H^{\exp}_{Q_c}},\tag{2-16}$$

and from Lemma 2.9,

$$|\langle\mathrm{NL}_{\mathrm{loc}}(\psi),\varphi\rangle|\leqslant K\nu\|\varphi\|^2_{H^1(\{\eta\neq1\})}\leqslant K\nu\|\varphi\|^2_{H^{\exp}_{Q_c}}.\tag{2-17}$$

Let us now show that

$$|\langle i(c-c'')\vec e_2.H(\psi),\varphi\rangle|\leqslant K\nu\|\varphi\|^2_{H^{\exp}_{Q_c}}.\tag{2-18}$$

We recall that

$$i\vec e_2.H(\psi)=i\partial_{x_2}Q+i\frac{\partial_{x_2}(Q\psi)(1-\eta)+Q\partial_{x_2}\psi\eta e^{\psi}}{(1-\eta)+\eta e^{\psi}}.$$

We compute, with Lemma 2.10 and Lemma 5.4 of [15],

$$|(c-c'')\langle i\partial_{x_2}Q,\varphi\rangle| \leqslant K\sqrt{c''}\|\varphi\|_{H_Q^{\exp}}|\langle i\partial_{x_2}Q,\varphi\rangle| \leqslant K\sqrt{c''}\ln\Big(\frac{1}{c''}\Big)\|\varphi\|^2_{H_Q^{\exp}} \leqslant K\nu\|\varphi\|^2_{H_Q^{\exp}}.$$

Indeed, although $Q = Q_{c''}(\cdot - X\vec{e}_2)e^{i\gamma}$ has a phase that is not present in Lemma 5.4 of [15], since $\varphi = Q\psi$, we have $\partial_{x_2}Q\bar{\varphi} = \partial_{x_2}Q\bar{Q}\bar{\psi}$, which no longer depends on γ.

Now, with $\|\varphi\|_{H^1(\{\eta\neq 1\})} \leqslant K\nu$ from Lemmas 2.7 and 2.8, we compute easily that

$$\left|\left\langle i\frac{\partial_{x_2}(Q\psi)(1-\eta)+Q\partial_{x_2}\psi\eta e^{\psi}}{(1-\eta)+\eta e^{\psi}},\varphi\right\rangle - \langle iQ\partial_{x_2}\psi\eta,\varphi\rangle\right| \leqslant K\nu\|\varphi\|_{H_Q^{\exp}}$$

since the left-hand side is supported in $\{\eta\neq 1\}$; therefore

$$|\langle i(c-c'')\vec{e}_2.H(\psi),\varphi\rangle| \leqslant K\nu\|\varphi\|^2_{H_{Q_c}^{\exp}} + |(c-c'')\langle iQ\partial_{x_2}\psi\eta,\varphi\rangle|.$$

With the same computations as in Section 7.3.2 of [15] (taking $\gamma'=0$), we check that

$$|\langle iQ\partial_{x_2}\psi\eta,\varphi\rangle| \leqslant K\|\varphi\|^2_{H_Q^{\exp}};$$

therefore, using Lemma 2.7 and (2-12), for $\nu>0$ small enough,

$$|(c-c'')\langle iQ\partial_{x_2}\psi\eta,\varphi\rangle| \leqslant K|c-c''|\|\varphi\|^2_{H_Q^{\exp}} \leqslant K\nu\|\varphi\|^2_{H_Q^{\exp}}.$$

This completes the proof of (2-18). We focus now on the proof of

$$|\langle F(\psi),\varphi\rangle| \leqslant K\nu\|\varphi\|^2_{H_Q^{\exp}}. \tag{2-19}$$

We compute

$$\int_{\mathbb{R}^2}\mathfrak{Re}(Q\eta(|Q|^2S(\psi))\bar{\varphi}) = \int_{\mathbb{R}^2}|Q|^4\eta(e^{2\mathfrak{Re}(\psi)}-1-2\mathfrak{Re}(\psi))\mathfrak{Re}(\psi),$$

and since, as already seen at the end of the proof of Lemma 2.10, we have $\|\mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde{r}\geqslant\lambda\})} \leqslant 1$ if $\nu\leqslant\nu_2$, we deduce

$$|e^{2\mathfrak{Re}(\psi)}-1-2\mathfrak{Re}(\psi)| \leqslant K\,\mathfrak{Re}^2(\psi)$$

and

$$\left|\int_{\mathbb{R}^2}\mathfrak{Re}\big(Q\eta(|Q|^2S(\psi))\bar{\varphi}\big)\right| \leqslant K\int_{\mathbb{R}^2}\eta\,\mathfrak{Re}^3(\psi) \leqslant K\nu\int_{\mathbb{R}^2}\eta\,\mathfrak{Re}^2(\psi) \leqslant K\nu\|\varphi\|^2_{H_Q^{\exp}}.$$

We are left with the estimation of $\int_{\mathbb{R}^2}\mathfrak{Re}(Q\eta(-\nabla\psi.\nabla\psi)\bar{\varphi})$, which will be slightly more delicate. First, we compute, using $\varphi = Q\psi$

$$\begin{aligned}\int_{\mathbb{R}^2}\mathfrak{Re}\big(Q\eta(-\nabla\psi.\nabla\psi)\bar{\varphi}\big) &= -\int_{\mathbb{R}^2}|Q|^2\eta\,\mathfrak{Re}(\nabla\psi.\nabla\psi\bar{\psi})\\ &= -\int_{\mathbb{R}^2}|Q|^2\eta\,\mathfrak{Re}(\nabla\psi.\nabla\psi)\,\mathfrak{Re}(\psi) - \int_{\mathbb{R}^2}|Q|^2\eta\,\mathfrak{Im}(\nabla\psi.\nabla\psi)\,\mathfrak{Im}(\psi)\\ &= -\int_{\mathbb{R}^2}|Q|^2\eta\,\mathfrak{Re}(\nabla\psi.\nabla\psi)\,\mathfrak{Re}(\psi) - 2\int_{\mathbb{R}^2}|Q|^2\eta\,\mathfrak{Re}(\nabla\psi).\,\mathfrak{Im}(\nabla\psi)\,\mathfrak{Im}(\psi).\end{aligned}$$

Note that there exists a universal constant $K > 0$ such that $\| \mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde{r}\geqslant R\})} \leqslant K\nu$ by Lemma 2.8 (considering the regions $\{\tilde{r} \geqslant \lambda\}$ with ψ and $\{\tilde{r} \leqslant \lambda\}$ with φ). Then, we estimate

$$\left| -\int_{\mathbb{R}^2} |Q|^2 \eta\, \mathfrak{Re}(\nabla\psi.\nabla\psi)\, \mathfrak{Re}(\psi) \right| \leqslant K\nu \int_{\mathbb{R}^2} \eta |\nabla\psi|^2 \leqslant K\nu \|\varphi\|^2_{H_Q^{\exp}}.$$

Now, by integration by parts (that can be justified as in [15]), we have

$$\begin{aligned}
&\int_{\mathbb{R}^2} |Q|^2 \eta\, \mathfrak{Re}(\nabla\psi).\, \mathfrak{Im}(\nabla\psi)\, \mathfrak{Im}(\psi) \\
&\quad = -\int_{\mathbb{R}^2} \nabla(|Q|^2)\eta\, \mathfrak{Re}(\psi).\, \mathfrak{Im}(\nabla\psi)\, \mathfrak{Im}(\psi) - \int_{\mathbb{R}^2} |Q|^2 \nabla\eta\, \mathfrak{Re}(\psi).\, \mathfrak{Im}(\nabla\psi)\, \mathfrak{Im}(\psi) \\
&\qquad - \int_{\mathbb{R}^2} |Q|^2 \eta\, \mathfrak{Re}(\psi)\, \mathfrak{Im}(\Delta\psi)\, \mathfrak{Im}(\psi) - \int_{\mathbb{R}^2} |Q|^2 \eta\, \mathfrak{Re}(\psi)\, \mathfrak{Im}(\nabla\psi).\, \mathfrak{Im}(\nabla\psi),
\end{aligned}$$

and with $|\nabla(|Q|^2)| \leqslant K/(1+\tilde{r})^{5/2}$ from equation (2.9) of [15] $\left(\text{for } \sigma = \frac{1}{2}\right)$ with $K > 0$ a universal constant, we have by Cauchy–Schwarz

$$\left| \int_{\mathbb{R}^2} \nabla(|Q|^2)\eta\, \mathfrak{Re}(\psi).\, \mathfrak{Im}(\nabla\psi)\, \mathfrak{Im}(\psi) \right| \leqslant K\nu \sqrt{\int_{\mathbb{R}^2} \eta|\nabla\psi|^2 \int_{\mathbb{R}^2} \eta \frac{|\psi|^2}{(1+\tilde{r})^5}} \leqslant K\nu\|\varphi\|^2_{H_Q^{\exp}},$$

$$\left| \int_{\mathbb{R}^2} |Q|^2 \eta\, \mathfrak{Re}(\psi)\, \mathfrak{Im}(\nabla\psi).\, \mathfrak{Im}(\nabla\psi) \right| \leqslant K\nu \int_{\mathbb{R}^2} \eta|\nabla\psi|^2 \leqslant K\nu\|\varphi\|^2_{H_{Q_c}^{\exp}}.$$

Since $\nabla\eta$ is supported in $\{0 < \eta < 1\}$, we check easily that

$$\left| \int_{\mathbb{R}^2} |Q|^2 \nabla\eta\, \mathfrak{Re}(\psi).\, \mathfrak{Im}(\nabla\psi)\, \mathfrak{Im}(\psi) \right| \leqslant K\nu\|\varphi\|^2_{H_Q^{\exp}}.$$

We focus now on the estimation of the last remaining term, $\int_{\mathbb{R}^2} |Q|^2\eta\, \mathfrak{Re}(\psi)\, \mathfrak{Im}(\Delta\psi)\, \mathfrak{Im}(\psi)$. For that purpose, we define more generally for $n \geqslant 1$

$$A_n := \int_{\mathbb{R}^2} |Q|^2 \eta^n\, \mathfrak{Re}^n(\psi)\, \mathfrak{Im}(\Delta\psi)\, \mathfrak{Im}(\psi).$$

Note that we want to estimate A_1.

We compute, using $(\mathrm{TW}_{c''})(Q) = 0$, that

$$L_Q(\varphi) = Q\left(-\Delta\psi - ic''\partial_{x_2}\psi - 2\frac{\nabla Q}{Q}.\nabla\psi + 2\,\mathfrak{Re}(\psi)|Q|^2\right);$$

therefore, by Lemma 2.9, in $\{\eta \neq 0\}$,

$$\begin{aligned}
\mathfrak{Im}(\Delta\psi) &= \mathfrak{Im}\left(-ic''\partial_{x_2}\psi - 2\frac{\nabla Q}{Q}.\nabla\psi + 2\,\mathfrak{Re}(\psi)|Q|^2 + \frac{-i(c-c'')\vec{e}_2.H(\psi) + \mathrm{NL}_{\mathrm{loc}}(\psi) + F(\psi)}{Q}\right) \\
&= -c''\,\mathfrak{Re}(\partial_{x_2}\psi) - 2\,\mathfrak{Im}\left(\frac{\nabla Q}{Q}.\nabla\psi\right) + \mathfrak{Im}\left(\frac{-i(c-c'')\vec{e}_2.H(\psi) + \mathrm{NL}_{\mathrm{loc}}(\psi) + F(\psi)}{Q}\right).
\end{aligned}$$

We compute, by integration by parts, with $\mathfrak{Re}^n(\psi)\,\mathfrak{Re}(\partial_{x_2}\psi) = (1/(n+1))\partial_{x_2}(\mathfrak{Re}^{n+1}(\psi))$, that

$$\begin{aligned}\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)c''\,\mathfrak{Re}(\partial_{x_2}\psi)\,\mathfrak{Im}(\psi)\\ = -\frac{1}{n+1}\int_{\mathbb{R}^2}(\partial_{x_2}|Q|^2)\eta^n\,\mathfrak{Re}^{n+1}(\psi)c''\,\mathfrak{Im}(\psi)\\ -\frac{n}{n+1}\int_{\mathbb{R}^2}|Q|^2\partial_{x_2}\eta\eta^{n-1}\,\mathfrak{Re}^{n+1}(\psi)c''\,\mathfrak{Im}(\psi) - \frac{1}{n+1}\int_{\mathbb{R}^2}|Q|^2\eta^n\,\mathfrak{Re}^{n+1}(\psi)c''\,\mathfrak{Im}(\partial_{x_2}\psi).\end{aligned}$$

Since $|c''| \leqslant \nu$ by (2-5) (diminishing ν_2 if necessary), Lemma 2.7 and the hypotheses of Proposition 1.8, $\|\varphi\|_{C^1(\{\tilde r\leqslant\lambda\})} + \|\,\mathfrak{Re}(\psi)\|_{L^\infty(\{\tilde r\geqslant\lambda\})} \leqslant K\nu$ by Lemma 2.8 and $|\nabla(|Q|^2)| \leqslant K/(1+\tilde r)^{5/2}$ from equation (2.9) of [15], we infer by Cauchy–Schwarz that

$$\begin{aligned}\left|\int_{\mathbb{R}^2}(\partial_{x_2}|Q|^2)\eta^n\,\mathfrak{Re}^{n+1}(\psi)c''\,\mathfrak{Im}(\psi)\right| &\leqslant Kc''\nu^n\sqrt{\int_{\mathbb{R}^2}\eta\,\mathfrak{Im}^2(\psi)(\partial_{x_2}|Q|^2)^2\int_{\mathbb{R}^2}\eta\,\mathfrak{Re}^2(\psi)}\\ &\leqslant K\nu^n\|\varphi\|^2_{H_Q^{\exp}},\end{aligned} \tag{2-20}$$

$$\left|\int_{\mathbb{R}^2}|Q|^2\partial_{x_2}\eta\eta^{n-1}\,\mathfrak{Re}^{n+1}(\psi)c''\,\mathfrak{Im}(\psi)\right| \leqslant K\nu^n\|\varphi\|^2_{H_Q^{\exp}}, \tag{2-21}$$

$$\left|\int_{\mathbb{R}^2}|Q|^2\eta^n\,\mathfrak{Re}^{n+1}(\psi)c''\,\mathfrak{Im}(\partial_{x_2}\psi)\right| \leqslant K\nu^n\sqrt{\int_{\mathbb{R}^2}\eta|\nabla\psi|^2\int_{\mathbb{R}^2}\eta\,\mathfrak{Re}^2(\psi)} \leqslant K\nu^n\|\varphi\|^2_{H_Q^{\exp}}. \tag{2-22}$$

We deduce that

$$\left|\int_{\mathbb{R}^2}|Q|^2\eta^n\,\mathfrak{Re}^n(\psi)c''\,\mathfrak{Re}(\partial_{x_2}\psi)\,\mathfrak{Im}(\psi)\right| \leqslant (K\nu)^n\|\varphi\|^2_{H_Q^{\exp}}. \tag{2-23}$$

For

$$\int_{\mathbb{R}^2}|Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Im}\left(\frac{\nabla Q}{Q}.\nabla\psi\right)\mathfrak{Im}(\psi),$$

we compute

$$\mathfrak{Im}\left(\frac{\nabla Q}{Q}.\nabla\psi\right) = \mathfrak{Re}\left(\frac{\nabla Q}{Q}\right).\,\mathfrak{Im}(\nabla\psi) + \mathfrak{Re}(\nabla\psi).\,\mathfrak{Im}\left(\frac{\nabla Q}{Q}\right),$$

and with previous estimates, we check easily that

$$\begin{aligned}\left|\int_{\mathbb{R}^2}|Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Re}\left(\frac{\nabla Q}{Q}\right).\,\mathfrak{Im}(\nabla\psi)\,\mathfrak{Im}(\psi)\right|\\ \leqslant (K\nu)^n\sqrt{\int_{\mathbb{R}^2}\eta|\nabla\psi|^2\int_{\mathbb{R}^2}\eta\,\mathfrak{Im}^2(\psi)\,\mathfrak{Re}^2\left(\frac{\nabla Q}{Q}\right)} \leqslant (K\nu)^n\|\varphi\|^2_{H_Q^{\exp}},\end{aligned} \tag{2-24}$$

and by integration by parts, with computations similar to those for the proof of (2-23), using

$$\left|\nabla.\,\mathfrak{Im}\left(\frac{\nabla Q}{Q}\right)\right| \leqslant \frac{K}{(1+\tilde r)^{3/2}}$$

from (2.9) to (2.11) of [15] $\left(\text{for } \sigma = \frac{1}{2}\right)$ for a universal constant $K > 0$ and Lemma 2.1, we infer that

$$\left|\int_{\mathbb{R}^2}|Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Re}(\nabla\psi).\,\mathfrak{Im}\left(\frac{\nabla Q}{Q}\right)\mathfrak{Im}(\psi)\right| \leqslant (K\nu)^n\|\varphi\|^2_{H_Q^{\exp}}, \tag{2-25}$$

and we check easily that

$$\left|\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Im}\left(\frac{\mathrm{NL}_{\mathrm{loc}}(\psi)}{Q}\right)\mathfrak{Im}(\psi)\right| \leqslant (K\nu)^n\|\varphi\|^2_{H_Q^{\exp}}. \tag{2-26}$$

Now, we look at

$$\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Im}\left(\frac{-i(c-c'')\vec{e}_2.H(\psi)}{Q}\right)\mathfrak{Im}(\psi)$$

for the part of $\vec{e}_2.H(\psi)$ related to the cutoff, the estimation can be done as previously, and we are left with the estimation of

$$\begin{aligned}(c-c'')\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Im}\left(-i\frac{\partial_{x_2}Q}{Q}-i\partial_{x_2}\psi\right)\mathfrak{Im}(\psi)&\\ =(c-c'')\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Re}\left(\frac{\partial_{x_2}Q}{Q}+\partial_{x_2}\psi\right)\mathfrak{Im}(\psi).&\end{aligned}$$

From (2-5) and Lemma 2.7, we have $|c-c''|\leqslant\nu$ (diminishing ν_2 if necessary), and from equation (2.9) of [15],

$$\left|\mathfrak{Re}\left(\frac{\partial_{x_2}Q}{Q}\right)\right| \leqslant \frac{K}{(1+\tilde{r})^{5/2}}.$$

Therefore

$$\begin{aligned}\left|(c-c'')\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Re}\left(\frac{\partial_{x_2}Q}{Q}\right)\mathfrak{Im}(\psi)\right|&\\ \leqslant (K\nu)^n\sqrt{\int_{\mathbb{R}^2}\eta\,\mathfrak{Re}^2(\psi)\int_{\mathbb{R}^2}\eta\,\mathfrak{Re}^2\left(\frac{\partial_{x_2}Q}{Q}\right)\mathfrak{Im}^2(\psi)} \leqslant (K\nu)^n\|\varphi\|^2_{H_Q^{\exp}},&\end{aligned} \tag{2-27}$$

and we estimate

$$\left|(c-c'')\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Re}(\partial_{x_2}\psi)\,\mathfrak{Im}(\psi)\right| \leqslant (K\nu)^n\|\varphi\|^2_{H_Q^{\exp}} \tag{2-28}$$

by (2-23). For the last remaining term, since

$$\mathfrak{Im}\left(\frac{F(\psi)}{Q}\right) = \mathfrak{Im}(-\eta\nabla\psi.\nabla\psi),$$

we have

$$\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Im}\left(\frac{F(\psi)}{Q}\right)\mathfrak{Im}(\psi) = -2\int_{\mathbb{R}^2}|Q|^2\eta^{n+1}\,\mathfrak{Re}^n(\psi)\,\mathfrak{Im}(\nabla\psi).\mathfrak{Re}(\nabla\psi)\,\mathfrak{Im}(\psi).$$

In particular,

$$\begin{aligned}\left|\int_{\mathbb{R}^2} |Q|^2\eta^n\,\mathfrak{Re}^n(\psi)\,\mathfrak{Im}\left(\frac{F(\psi)}{Q}\right)\mathfrak{Im}(\psi)\right| &\leqslant (K\nu)^n\|\eta\,\mathfrak{Im}(\psi)\|_{L^\infty(\mathbb{R}^2)}\int_{\mathbb{R}^2}\eta|\nabla\psi|^2\\ &\leqslant (K\nu)^n\|\eta\,\mathfrak{Im}(\psi)\|_{L^\infty(\mathbb{R}^2)}\|\varphi\|^2_{H_Q^{\exp}}.\end{aligned} \tag{2-29}$$

Combining this result with the previous estimates, this implies that

$$|A_n| \leqslant (C_6\nu)^n(1+\|\eta\,\mathfrak{Im}(\psi)\|_{L^\infty(\mathbb{R}^2)})\|\varphi\|^2_{H_Q^{\exp}} \tag{2-30}$$

for some universal constant $C_6 > 0$, but that is not enough to show that we have

$$\left|\int_{\mathbb{R}^2} |Q|^2 \eta^n \,\mathfrak{Re}^n(\psi)\, \mathfrak{Im}\left(\frac{F(\psi)}{Q}\right) \mathfrak{Im}(\psi)\right| \leqslant (K\nu)^n \|\varphi\|^2_{H_Q^{\exp}},$$

since we have no control on $\|\eta\, \mathfrak{Im}(\psi)\|_{L^\infty(\mathbb{R}^2)}$ other than the fact that it is a finite quantity. By integration by parts (integrating $\mathfrak{Re}(\nabla\psi)$), with computations similar to those for the proof of (2-23), we infer that

$$\begin{aligned}\left|2\int_{\mathbb{R}^2} |Q|^2 \eta^{n+1} \,\mathfrak{Re}^n(\psi)\, \mathfrak{Im}(\nabla\psi).\, \mathfrak{Re}(\nabla\psi)\, \mathfrak{Im}(\psi)\right| \\ \leqslant \left|2\int_{\mathbb{R}^2} |Q|^2 \eta^{n+1} \,\mathfrak{Re}^n(\psi)\, \mathfrak{Im}(\Delta\psi)\, \mathfrak{Re}(\psi)\, \mathfrak{Im}(\psi)\right| + (K\nu)^n \|\varphi\|^2_{H_Q^{\exp}} \\ \leqslant 2|A_{n+1}| + (K\nu)^n \|\varphi\|^2_{H_Q^{\exp}}.\end{aligned}$$

Combining this result with estimates (2-20) to (2-29), we deduce that, for some universal constant $C_7 > 0$,

$$|A_n| \leqslant 2|A_{n+1}| + (C_7\nu)^n \|\varphi\|^2_{H_Q^{\exp}}.$$

Therefore, by induction,

$$|A_1| \leqslant 2^n |A_n| + \sum_{k=1}^{n-1} (2C_7\nu)^k \|\varphi\|^2_{H_Q^{\exp}}.$$

Hence, with (2-30),

$$|A_1| \leqslant \left((2C_6\nu)^n (1 + \|\eta\, \mathfrak{Im}(\psi)\|_{L^\infty(\mathbb{R}^2)}) + \sum_{k=1}^{n-1} (2C_7\nu)^k\right) \|\varphi\|^2_{H_Q^{\exp}}.$$

Taking $\nu > 0$ such that $\nu \leqslant \nu_2$ and $2C_6\nu < \frac{1}{2}$ and $2C_7\nu < \frac{1}{2}$, then $n \geqslant 1$ large enough (depending on $\|\eta\, \mathfrak{Im}(\psi)\|_{L^\infty(\mathbb{R}^2)}$) such that

$$\frac{1}{2^{n-1}} (1 + \|\eta\, \mathfrak{Im}(\psi)\|_{L^\infty(\mathbb{R}^2)}) \leqslant 1,$$

we conclude that

$$|A_1| \leqslant \left(2C_6 + 2C_7 \sum_{k=0}^{n-2} \frac{1}{2^k}\right) \nu \|\varphi\|^2_{H_Q^{\exp}} \leqslant 2(C_6 + 2C_7)\nu \|\varphi\|^2_{H_Q^{\exp}}.$$

This concludes the proof of (2-19).

Combining estimates (2-16) to (2-19) in (2-15), we deduce that

$$(1 - C_8\nu)\|\varphi\|^2_{H_Q^{\exp}} \leqslant 0$$

for some universal constant $C_8 > 0$; therefore, taking $\nu > 0$ small enough such that the previous constraints are satisfied and $C_8\nu < \frac{1}{2}$, we have $\|\varphi\|_{H_Q^{\exp}} = 0$. From Lemma 2.10, we deduce $c'' = c$. The proof is complete.

2C. ***Proof of Corollary 1.10.*** Take a function u satisfying the hypotheses of Corollary 1.10. Then, u is even in x_1 and it has finite energy. Furthermore, by Theorem 1.6 (for $p = +\infty$),

$$\begin{aligned}\|u - V_1(\cdot - d_c\vec{e}_1)V_{-1}(\cdot + d_c\vec{e}_1)\|_{L^\infty(\mathbb{R}^2)} &\leqslant \|u - Q_c\|_{L^\infty(\mathbb{R}^2)} + \|Q_c - V_1(\cdot - d_c\vec{e}_1)V_{-1}(\cdot + d_c\vec{e}_1)\|_{L^\infty(\mathbb{R}^2)} \\ &\leqslant \varepsilon + o_{c\to 0}(1).\end{aligned}$$

Next,

$$\| |u| - 1 \|_{L^\infty(\{\tilde{r}_d \geqslant \lambda\})} \leqslant \|u - Q_c\|_{L^\infty(\{\tilde{r}_d \geqslant \lambda\})} + \| |Q_c| - 1 \|_{L^\infty(\{\tilde{r}_d \geqslant \lambda\})} \leqslant \varepsilon + \frac{K}{\lambda}$$

by equation (2.6) of [15]. We now fix the parameters. We first choose $\lambda \geqslant \lambda_*$ large enough so that $K/\lambda \leqslant 1/(2\lambda_*)$. Then, we fix $c_0 > 0$ and $\varepsilon > 0$ so small that $\varepsilon \leqslant 1/(2\lambda_*)$, $|cd_c - 1| \leqslant \varepsilon(\lambda)$, $d_c \geqslant 1/\varepsilon(\lambda)$ and $\varepsilon + o_{c\to 0}(1) \leqslant \varepsilon(\lambda)$ for $c < c_0$. Therefore, u satisfies the hypotheses of Proposition 1.8 with $d = d_c$, and this concludes the proof.

3. Properties of quasiminimizers of the energy and proof of Theorem 1.11

3A. ***Tools for the vortex analysis.*** We list in this section some results useful for the analysis of traveling waves for small speeds or, equivalently, large momentum, with vorticity. We shall denote by $\langle u \mid v\rangle = \mathrm{Re}(u\bar{v})$ the real scalar product of the complex numbers u, v. The Jacobian (or vorticity)

$$Jv := \langle i\partial_1 v \mid \partial_2 v\rangle = \tfrac{1}{2}\partial_1\langle iv \mid \partial_2 v\rangle - \tfrac{1}{2}\partial_2\langle iv \mid \partial_1 v\rangle$$

is then relevant, and we shall use the following concentration property of the Jacobian. We define

$$E_\varepsilon(u, \Omega) := \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{2\varepsilon^2}(1 - |u|^2)^2\, dx.$$

Theorem 3.1 (concentration of the Jacobian [2; 27]). *Let $M_0 > 0$, $R > 0$ and $\beta \in]0, 1]$. Then, for every $\delta > 0$, there exists $\varepsilon_0 > 0$ (depending only on β, δ, R and M_0) such that, for any $0 < \varepsilon < \varepsilon_0$, and for any $u \in H^1(B(0, 4R))$ such that $E_\varepsilon(u, B(0, 4R)) \leqslant M_0|\ln\varepsilon|$ and $|u| \geqslant \frac{1}{2}$ in $B(0, 4R) \setminus B(0, R)$, there exist $N \in \mathbb{N}$, $y_1, \dots, y_N \in \bar{B}(0, R)$, $d_1, \dots, d_N \in \mathbb{Z}$ such that*

$$\Big\| Ju - \pi\sum_{k=1}^N d_k\delta_{y_k}\Big\|_{[\mathcal{C}_c^{0,\beta}(B(0,4R))]^*} \leqslant \delta$$

and

$$\pi\sum_{k=1}^N |d_k| \leqslant \frac{E_\varepsilon(u, B(0, 4R))}{|\ln\varepsilon|} + \delta.$$

Finally, we may choose the points y_k, $1 \leqslant k \leqslant N$, in $\{|u| \leqslant \frac{1}{2}\}$.

Here, we recall that the space $[\mathcal{C}_c^{0,\beta}(B(0, R))]^*$ is endowed with the dual norm associated with

$$\|\zeta\|_{\mathcal{C}_c^{0,\beta}(B(0,R))} = \sup_{x\neq y\in B(0,R)} \frac{|\zeta(x) - \zeta(y)|}{|x - y|^\beta}$$

for $\zeta \in \mathcal{C}^{0,\beta}(B(0, R))$ compactly supported.

Remark 3.2. The above-mentioned theorem is actually Lemma 3.3 in [8]. It is related to the works [2; 27], which both correspond to the limit $\varepsilon \to 0$, whereas we have here a statement (obtained by compactness) at fixed ε. The hypothesis "$|u| \geqslant \frac{1}{2}$ in $B(0, 4R) \setminus B(0, R)$" ensures that the vortices do not approach the boundary $\partial B(0, 4R)$.

Theorem 3.3 (clearing-out theorem [8]). *Let $M_0 > 0$ and $\sigma > 0$ be given. Then there exist $\epsilon_0 > 0$ and $\eta > 0$, depending only on M_0 and σ, such that, if $R_0 = 1/(1+M_0)$, if $U : B(0, R_0) \to \mathbb{C}$ solves*

$$\Delta U + i\mathfrak{c}\partial_2 U + \frac{1}{\epsilon^2} U(1-|U|^2) = 0 \tag{3-1}$$

in $B(0, R_0) \subset \mathbb{R}^2$, with $\epsilon < \epsilon_0$, $|\mathfrak{c}| \leqslant M_0 |\ln \epsilon|$, and

$$E_\epsilon(U, B(0, R_0)) \leqslant \eta |\ln \epsilon|,$$

then

$$|U(0)| \geqslant 1 - \sigma.$$

For the elliptic PDE

$$\Delta \mathcal{U} + \frac{1}{\varepsilon^2} \mathcal{U}(1-|\mathcal{U}|^2) = 0, \tag{3-2}$$

that is, without the transport term $i\partial_2 U$, this result has been shown in two dimensions in [6] for minimizing maps, and in [4] for the Ginzburg–Landau equation with magnetic field. In higher dimensions, see [7; 31] for (3-2) and [8] for an equation including the Ginzburg–Landau equation with magnetic field and (3-1). One may use the change of unknown

$$\mathcal{U}(x) := (1+\mathfrak{c}^2\epsilon^2/4)^{-1/2} \mathrm{e}^{i\mathfrak{c}x_2/2} U(x), \quad \varepsilon = \epsilon(1+\mathfrak{c}^2\epsilon^2/4)^{-1/2},$$

to transform (3-2) without the transport term into (3-1) with the transport term. However, the assumptions $E_\epsilon(U, B(0, R_0)) \leqslant \eta|\ln \epsilon|$ and $E_\varepsilon(\mathcal{U}, B(0, R_0)) \leqslant \eta |\ln \varepsilon|$ are not equivalent (due to the extra phase term).

3B. ***Vortex structure for quasiminimizers of E at fixed P.*** In this section, some $\Lambda_0 > 0$ is fixed and we consider a large momentum $\mathfrak{p}$ and $u_\mathfrak{p}$ such that

$$E(u_\mathfrak{p}) \leqslant 2\pi \ln \mathfrak{p} + \Lambda_0 \tag{3-3}$$

and such that there exists $c_\mathfrak{p} > 0$ (depending on $u_\mathfrak{p}$) such that

$$0 = (\mathrm{TW}_{c_\mathfrak{p}})(u_\mathfrak{p}) = -ic_\mathfrak{p}\partial_{x_2} u_\mathfrak{p} - \Delta u_\mathfrak{p} - (1-|u_\mathfrak{p}|^2) u_\mathfrak{p}.$$

It then follows from [24] (see Theorem 2.4) that we may assume, using the phase-shift invariance, that $u_\mathfrak{p} \to 1$ at spatial infinity. In particular, we have

$$\mathfrak{p} = P_2(u_\mathfrak{p}) = \frac{1}{2} \int_{\mathbb{R}^2} \langle i\partial_2 u_\mathfrak{p} | u_\mathfrak{p} - 1 \rangle \, dx.$$

Our goal is to show that $u_\mathfrak{p}$ satisfies the hypotheses of Proposition 1.8. We shall follow [5; 8] in order to analyze the vortex structure of $u_\mathfrak{p}$.

3B1. *Localizing the vorticity set at scale $x/\mathfrak{p}$.* We define the following rescaling $\hat{u}_\mathfrak{p}$ of $u_\mathfrak{p}$:

$$\hat{u}_\mathfrak{p}(\hat{x}) = u_\mathfrak{p}(\mathfrak{p}\hat{x}). \tag{3-4}$$

Therefore, $\hat{u}_\mathfrak{p}$ solves

$$\Delta \hat{u}_\mathfrak{p} + ic_\mathfrak{p}\mathfrak{p}\partial_2 \hat{u}_\mathfrak{p} + \mathfrak{p}^2 \hat{u}_\mathfrak{p}(1-|\hat{u}_\mathfrak{p}|^2) = 0, \tag{3-5}$$

which is a particular case of (3-1) with

$$\epsilon = 1/\mathfrak{p}, \quad \mathfrak{c} = c_\mathfrak{p}\mathfrak{p}.$$

The universal L^∞ bound on the gradient of Corollary 2.3 reads now

$$\|\nabla \hat{u}_\mathfrak{p}\|_{L^\infty(\mathbb{R}^2)} \leqslant K\mathfrak{p}. \tag{3-6}$$

We shall have, in the end, $c_\mathfrak{p} \sim 1/\mathfrak{p}$. The first step provides a rough upper bound for the speed $c_\mathfrak{p}$ (the Lagrange multiplier for the minimization problem $E_{\min}(\mathfrak{p})$).

<u>Step 1</u>: There exists $\mathfrak{p}_1 = \mathfrak{p}_1(\Lambda_0)$ such that, for $\mathfrak{p} \geqslant \mathfrak{p}_1$, we have

$$0 < c_\mathfrak{p} \leqslant \frac{2E(u_\mathfrak{p})}{\mathfrak{p}} \leqslant 13\frac{\ln \mathfrak{p}}{\mathfrak{p}}.$$

In particular, $c_\mathfrak{p} \leqslant \frac{1}{2}$ and $\ln \mathfrak{p} \leqslant 2|\ln c_\mathfrak{p}|$.

We shall use the Pohozaev identity (2-2), that is,

$$\frac{1}{2}\int_{\mathbb{R}^2} (1-|u_\mathfrak{p}|^2)^2\,dx = c_\mathfrak{p}\mathfrak{p}.$$

At this stage, we only have the rough upper bound $0 \leqslant \frac{1}{4}\int_{\mathbb{R}^2}(1-|u_\mathfrak{p}|^2)^2\,dx \leqslant E(u_\mathfrak{p}) \leqslant 2\pi\ln\mathfrak{p} + \Lambda_0$, which concludes this step.

Another argument we could use for minimizers is that we know from [10] (see also [13]) that $0 \leqslant c_\mathfrak{p} \leqslant d^+E_{\min}(\mathfrak{p}) \leqslant E_{\min}(\mathfrak{p})/\mathfrak{p}$.

<u>Step 2</u>: There exists $\mathfrak{p}_2 > \mathfrak{p}_1$, $R_* \geqslant \frac{1}{8}$ and $n_* \in \mathbb{N}$, depending only on Λ_0, such that, if $\mathfrak{p} > \mathfrak{p}_2$, there exist $n_\mathfrak{p}$ points $\hat{z}_{\mathfrak{p},j}$, $1 \leqslant j \leqslant n_\mathfrak{p}$, with $n_\mathfrak{p} \leqslant n_*$ such that $\left\{|\hat{u}_\mathfrak{p}(\hat{x})| \leqslant \frac{1}{2}\right\} \subset \bigcup_{j=1}^{n_\mathfrak{p}} B(\hat{z}_{\mathfrak{p},j}, R_*)$ and the disks $\bar{B}(\hat{z}_{\mathfrak{p},j}, 4R_*)$, $1 \leqslant j \leqslant n_\mathfrak{p}$, are mutually disjoint.

We apply Theorem 3.3 with $\epsilon = 1/\mathfrak{p}$, $\mathfrak{c} = c_\mathfrak{p}\mathfrak{p}$ and $\sigma = \frac{1}{2}$ to $\hat{u}_\mathfrak{p}$. This is possible in view of the upper bound on $0 \leqslant c_\mathfrak{p}\mathfrak{p} \leqslant 13\ln\mathfrak{p}$ of Step 1 (that is, $M_0 = 13$). We then let $R_0 := 1/(1+13) = \frac{1}{14}$ for $\mathfrak{p} \geqslant \mathfrak{p}_1$ and denote by $\eta_{1/2}$ the positive constant η given by Theorem 3.3.

We now proceed in this way: we choose (if it exists) some $\hat{z}_{\mathfrak{p},1} \in \mathbb{R}^2$ such that $|\hat{u}_\mathfrak{p}(\hat{z}_{\mathfrak{p},1})| < \frac{1}{2}$. If $\left\{|\hat{u}_\mathfrak{p}| \leqslant \frac{1}{2}\right\} \subset \bar{B}(\hat{z}_{\mathfrak{p},1}, 2R_0)$, then we stop. If not, we choose $\hat{z}_{\mathfrak{p},2} \in \mathbb{R}^2 \setminus \bar{B}(\hat{z}_{\mathfrak{p},1}, 2R_0)$ such that $|\hat{u}_\mathfrak{p}(\hat{z}^{\mathfrak{p},2})| < \frac{1}{2}$. If $\left\{|\hat{u}_\mathfrak{p}| \leqslant \frac{1}{2}\right\} \subset \bigcup_{j=1}^{2} \bar{B}(\hat{z}_{\mathfrak{p},j}, 2R_0)$, then we stop, if not, we continue. This process ends in a finite number of steps (depending only on K_0) since, by construction, the disks $\bar{B}(\hat{z}_{\mathfrak{p},j}, R_0)$, $1 \leqslant j \leqslant n$, are pairwise disjoint. Hence, by Theorem 3.3, we have

$$2\pi\ln\mathfrak{p} + K_0 \geqslant E(u_\mathfrak{p}) = E_{1/\mathfrak{p}}(\hat{u}_\mathfrak{p}) \geqslant \sum_{j=1}^{n} E_{1/\mathfrak{p}}(\hat{u}_\mathfrak{p}, B(\hat{z}_{\mathfrak{p},j}, R_0)) \geqslant n \times \eta_{1/2}\ln\mathfrak{p},$$

which implies

$$n \leqslant \frac{2\pi\ln\mathfrak{p} + K_0}{\eta_{1/2}\ln\mathfrak{p}} \leqslant \frac{7}{\eta_{1/2}}$$

for $\mathfrak{p}$ large enough, say $\mathfrak{p} \geqslant \mathfrak{p}_2$.

At this stage, the disks $B(\hat{z}_{\mathfrak{p},j}, 2R_0)$, $1 \leqslant j \leqslant n_\mathfrak{p}$, cover the vorticity set $\left\{|\hat{u}_\mathfrak{p}| \leqslant \frac{1}{2}\right\}$, but the disks $\bar{B}(\hat{z}_{\mathfrak{p},j}, 8R_0)$ may not be pairwise disjoint. To get this property, we argue as in [6, Theorem IV.1]. Let us recall the idea: if the disks $\bar{B}(\hat{z}_{\mathfrak{p},j}, 8R_0)$, $1 \leqslant j \leqslant n_\mathfrak{p}$, are pairwise disjoint, then we are done with $R_* = 2R_0$. If not, then we have, for instance, $|\hat{z}_{\mathfrak{p},1} - \hat{z}_{\mathfrak{p},2}| \leqslant 16R_0$. We then remove the disk $B(\hat{z}_{\mathfrak{p},1}, 8R_0)$

from the list and set $R_1 := 17R_0$. The disks $B(\hat{z}_{\mathfrak{p},j}, R_1)$, $2 \leqslant j \leqslant n_\mathfrak{p}$, cover $\bigcup_{1\leqslant j\leqslant n_\mathfrak{p}} B(\hat{z}_{\mathfrak{p},j}, 2R_0)$, and hence the vorticity set $\{|\hat{u}_\mathfrak{p}| \leqslant \frac{1}{2}\}$, and their number has decreased. In a finite number of steps (depending only on K_0), we obtain the conclusion. The radius R_* is necessarily $\leqslant R_0 \times 17^{n_\mathfrak{p}} \leqslant R_0 \times 17^{n_*}$.

Similar arguments are given in [8], whereas in [5] the vorticity set is included in some disks of radii of order $c_\mathfrak{p}^\gamma$, which requires some extra work.

Step 3: We have

$$\mathfrak{p}^2 \int_{\mathbb{R}^2} (1-|\hat{u}_\mathfrak{p}|^2)^2\, d\hat{x} = o_{\mathfrak{p}\to+\infty}(\ln \mathfrak{p}).$$

This follows exactly as in [8] (see Proposition A.1 in the Appendix there). Notice that the result in [8] is stated for the potential on a compact set in a domain Ω, but it holds as well in the entire plane.

We then define, as in [8], the function $\hat{u}'_\mathfrak{p} : \mathbb{R}^2 \to \mathbb{C}$ by

$$\hat{u}'_\mathfrak{p}(\hat{x}) := \begin{cases} \hat{u}_\mathfrak{p}(\hat{x}) & \text{if } \hat{x} \in \bigcup_{j=1}^{n_\mathfrak{p}} \bar{B}(\hat{z}_{\mathfrak{p},j}, 2R_*), \\ \dfrac{\hat{u}_\mathfrak{p}(\hat{x})}{|\hat{u}_\mathfrak{p}(\hat{x})|} & \text{if } \hat{x} \notin \bigcup_{j=1}^{n_\mathfrak{p}} \bar{B}(\hat{z}_{\mathfrak{p},j}, 3R_*), \\ \Big(3 - \dfrac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\Big)\hat{u}_\mathfrak{p}(\hat{x}) + \Big(-2 + \dfrac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\Big)\dfrac{\hat{u}_\mathfrak{p}(\hat{x})}{|\hat{u}_\mathfrak{p}(\hat{x})|} & \text{if } \hat{x} \in \bar{B}(\hat{z}_{\mathfrak{p},j}, 3R_*) \setminus \bar{B}(\hat{z}_{\mathfrak{p},j}, 2R_*) \end{cases}$$

for some $1 \leqslant j \leqslant n_\mathfrak{p}$ (this last formula is valid since the disks $\bar{B}(\hat{z}_{\mathfrak{p},j}, 4R_*)$, $1 \leqslant j \leqslant n_\mathfrak{p}$, are mutually disjoint).

Step 4: We have, as $\mathfrak{p} \to +\infty$,

$$E_{1/\mathfrak{p}}(\hat{u}'_\mathfrak{p}) \leqslant 2\pi \ln \mathfrak{p} + o(\ln \mathfrak{p}).$$

Letting $\Omega_R := \bigcup_{j=1}^{n_\mathfrak{p}} \bar{B}(\hat{z}_{\mathfrak{p},j}, R)$, we have

$$\int_{\mathbb{R}^2} (1-|\hat{u}'_\mathfrak{p}|^2)^2\, d\hat{x} = \int_{\Omega_{2R_*}} (1-|\hat{u}_\mathfrak{p}|^2)^2\, d\hat{x} + \int_{\Omega_{3R_*}\setminus\Omega_{2R_*}} (1-|\hat{u}'_\mathfrak{p}|^2)^2\, d\hat{x}.$$

We notice that in $\Omega_{3R_*} \setminus \Omega_{2R_*}$, say for $\hat{x} \in \bar{B}(\hat{z}_{\mathfrak{p},j}, 3R_*) \setminus \bar{B}(\hat{z}_{\mathfrak{p},j}, 2R_*)$, we have

$$|\hat{u}'_\mathfrak{p}(\hat{x})| = \left(3 - \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right)|\hat{u}_\mathfrak{p}(\hat{x})| + \left(-2 + \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right) \in [|\hat{u}_\mathfrak{p}(\hat{x})|, 1];$$

hence $|1-|\hat{u}'_\mathfrak{p}(\hat{x})|^2| \leqslant |1-|\hat{u}_\mathfrak{p}(\hat{x})|^2|$ and thus

$$\begin{aligned} \int_{\mathbb{R}^2} (1-|\hat{u}'_\mathfrak{p}|^2)^2\, d\hat{x} &\leqslant \int_{\Omega_{2R_*}} (1-|\hat{u}_\mathfrak{p}|^2)^2\, d\hat{x} + \int_{\Omega_{3R_*}\setminus\Omega_{2R_*}} (1-|\hat{u}_\mathfrak{p}|^2)^2\, d\hat{x} \\ &= \int_{\Omega_{3R_*}} (1-|\hat{u}_\mathfrak{p}|^2)^2\, d\hat{x}. \end{aligned} \tag{3-7}$$

For the kinetic term, we have

$$|\nabla \hat{u}'_\mathfrak{p}(\hat{x})|^2 = |\nabla \hat{u}_\mathfrak{p}(\hat{x})|^2$$

if $\hat{x} \in \Omega_{2R_*}$. Outside $\bigcup_{j=1}^{n_\mathfrak{p}} \bar{B}(\hat{z}_{\mathfrak{p},j}, R_*)$ we have $|\hat{u}_\mathfrak{p}| \geqslant \frac{1}{2}$ and we may then lift, at least locally, $\hat{u}_\mathfrak{p} = A e^{i\phi}$ and get

$$|\nabla \hat{u}_\mathfrak{p}|^2 = A^2 |\nabla\phi|^2 + |\nabla A|^2.$$

If $\hat{x} \notin \Omega_{3R_*}$, then, by (3-6),

$$|\nabla \hat{u}'_\mathfrak{p}|^2 = |\nabla\phi|^2 = A^2|\nabla\phi|^2 + \frac{1-A^2}{A^2} \times A^2 |\nabla\phi|^2 \leqslant |\nabla \hat{u}_\mathfrak{p}|^2 + 4K\mathfrak{p}|1-A^2| \times |\nabla \hat{u}_\mathfrak{p}|$$

since $A = |\hat{u}_\mathfrak{p}| \geqslant \frac{1}{2}$ outside Ω_{R_*}. Finally, in $\bar{B}(\hat{z}_{\mathfrak{p},j}, 3R_*) \setminus \bar{B}(\hat{z}_{\mathfrak{p},j}, 2R_*)$ (for some unique $1 \leqslant j \leqslant n_\mathfrak{p}$), we have

$$|\nabla \hat{u}'_\mathfrak{p}|^2 = |\nabla\phi|^2 \left(\left(3 - \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right) A + \left(-2 + \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right)\right)^2 + \left| \nabla \left[\left(3 - \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right) A + \left(-2 + \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right)\right]\right|^2.$$

We then use that, since $|\hat{u}_\mathfrak{p}(\hat{x})| \geqslant \frac{1}{2}$ and letting $\theta = 3 - |\hat{x} - \hat{z}_{\mathfrak{p},j}|/R_* \in [0, 1]$,

$$\begin{aligned} |\nabla\phi|^2 \left[\left(3 - \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right) A + \left(-2 + \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right)\right]^2 & \\ &= A^2|\nabla\phi|^2 \times \frac{1}{A^2}[1 + \theta(A-1)]^2 \leqslant A^2 |\nabla\phi|^2 \times (1 + K|A^2 - 1|) \\ &\leqslant A^2|\nabla\phi|^2 + K\mathfrak{p}|\nabla \hat{u}_\mathfrak{p}| \times |A^2 - 1|, \end{aligned}$$

by Corollary 2.3. On the other hand, since $|\cdot|$ is 1-Lipschitz continuous,

$$\begin{aligned} \left| \nabla \left[\left(3 - \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right) A + \left(-2 + \frac{|\hat{x}-\hat{z}_{\mathfrak{p},j}|}{R_*}\right)\right]\right|^2 &\leqslant \frac{1}{R_*^2}|1-A|^2 + |\nabla A|^2 + \frac{2}{R_*}|1-A| \times |\nabla A| \\ &\leqslant |\nabla A|^2 + K(A^2-1)^2 + K|\nabla A| \times |A^2 - 1|. \end{aligned}$$

Therefore, by the Cauchy–Schwarz inequality, for some absolute constant $K > 0$,

$$\int_{\mathbb{R}^2} |\nabla \hat{u}'_\mathfrak{p}|^2 \, d\hat{x} \leqslant \int_{\mathbb{R}^2} |\nabla \hat{u}_\mathfrak{p}|^2 \, d\hat{x} + K \left(\int_{\mathbb{R}^2} \mathfrak{p}^2 (1 - |\hat{u}_\mathfrak{p}|^2)^2 \, d\hat{x} \right)^{1/2} \left(\int_{\mathbb{R}^2} |\nabla \hat{u}_\mathfrak{p}|^2 \, d\hat{x} \right)^{1/2} + K \int_{\mathbb{R}^2} (1 - |\hat{u}_\mathfrak{p}|^2)^2 \, d\hat{x}.$$

Combining this with (3-7) yields

$$E_{1/\mathfrak{p}}(\hat{u}'_\mathfrak{p}) \leqslant E_\mathfrak{p}(\hat{u}_\mathfrak{p}) + K\sqrt{E_\mathfrak{p}(\hat{u}_\mathfrak{p})} \left(\int_{\mathbb{R}^2} \mathfrak{p}^2 (1 - |\hat{u}_\mathfrak{p}|^2)^2 \, d\hat{x} \right)^{1/2} + K \frac{E_\mathfrak{p}(\hat{u}_\mathfrak{p})}{\mathfrak{p}^2} \leqslant 2\pi \ln \mathfrak{p} + o(\ln \mathfrak{p}),$$

by the upper bound (3-3) and the estimate for the potential term of Step 3.

<u>Step 5</u>: We claim that for any $\delta \in \left]0, \frac{\pi}{2}\right[$, there exist $\mathfrak{p}_\delta^\dagger > \mathfrak{p}_2$ such that, for all $\mathfrak{p} \geqslant \mathfrak{p}_\delta^\dagger$, we are in one of the following cases:

(I) For any $1 \leqslant j \leqslant n_\mathfrak{p}$,

$$\|J\hat{u}'_\mathfrak{p}\|_{[\mathcal{C}_c^{0,1}(B(\hat{z}_{\mathfrak{p},j}, 4R_*))]^*} \leqslant \delta.$$

(II) There exist (up to a relabeling) two points $\hat{y}_{\mathfrak{p},\pm} \in \mathbb{R}^2$, depending on $\hat{u}_\mathfrak{p}$, such that

$$\max_{1 \leqslant j \leqslant n_\mathfrak{p}} \|J\hat{u}'_\mathfrak{p} - \pi(\delta_{\hat{y}_{\mathfrak{p},+}} - \delta_{\hat{y}_{\mathfrak{p},-}})\|_{[\mathcal{C}_c^{0,1}(B(\hat{z}_{\mathfrak{p},j}, 4R_*))]^*} \leqslant \delta.$$

We apply Theorem 3.1 to $\hat{u}'_{\mathfrak{p}}$ on each disk $B(\hat{z}_{\mathfrak{p},j}, 4R_*)$, $1 \leqslant j \leqslant n_{\mathfrak{p}}$. This yields points $\hat{y}_{\mathfrak{p},j,k} \in \left\{|\hat{u}_{\mathfrak{p}}| \leqslant \frac{1}{2}\right\} \subset B(\hat{z}_{\mathfrak{p},j}, R_*) \subset B(\hat{z}_{\mathfrak{p},j}, 4R_*)$ and integers $d_{\mathfrak{p},j,k} \in \mathbb{Z}$, $1 \leqslant k \leqslant N_{\mathfrak{p},j}$, such that

$$\left\| J\hat{u}'_{\mathfrak{p}} - \pi \sum_{k=1}^{N_{\mathfrak{p},j}} d_{\mathfrak{p},j,k} \delta_{\hat{y}_{\mathfrak{p},j,k}} \right\|_{[C_c^{0,1}(B(\hat{z}_{\mathfrak{p},j},4R_*))]^*} \leqslant \delta \tag{3-8}$$

and

$$\pi \sum_{k=1}^{N_{\mathfrak{p},j}} |d_{\mathfrak{p},j,k}| \leqslant \frac{E_{1/\mathfrak{p}}(\hat{u}'_{\mathfrak{p}}, B(\hat{z}_{\mathfrak{p},j}, 4R_*))}{\ln \mathfrak{p}} + \delta. \tag{3-9}$$

By summing the inequalities (3-9) over $1 \leqslant j \leqslant n_{\mathfrak{p}}$, we infer

$$\pi \sum_{j=1}^{n_{\mathfrak{p}}} \sum_{k=1}^{N_{\mathfrak{p},j}} |d_{\mathfrak{p},j,k}| \leqslant \frac{E_{1/\mathfrak{p}}(\hat{u}'_{\mathfrak{p}}, \Omega_{4R_*})}{\ln \mathfrak{p}} + \delta \leqslant 2.5\pi$$

by using $\delta < \frac{\pi}{2}$ and Step 3, and for $\mathfrak{p}$ large enough. Therefore,

$$\sum_{j=1}^{n_{\mathfrak{p}}} \sum_{k=1}^{N_{\mathfrak{p},j}} |d_{\mathfrak{p},j,k}| \leqslant 2 \tag{3-10}$$

and two cases may occur: all the integers $d_{\mathfrak{p},j,k}$ are zero (this is case (I)) or at least one of the integers $d_{\mathfrak{p},j,k}$ is not zero.

In addition, we have, for $1 \leqslant j \leqslant n_{\mathfrak{p}}$,

$$\sum_{k=1}^{N_{\mathfrak{p},j}} d_{\mathfrak{p},j,k} = \deg(\hat{u}_{\mathfrak{p}}, \partial B(\hat{z}_{\mathfrak{p},j}, 3R_*)). \tag{3-11}$$

Indeed, since $|\hat{u}'_{\mathfrak{p}}| = 1$ on $B(\hat{z}_{\mathfrak{p},j}, 4R_*) \setminus B(\hat{z}_{\mathfrak{p},j}, 3R_*)$, we have $J\hat{u}'_{\mathfrak{p}} = 0$ there. Therefore, by fixing $\chi \in \mathcal{C}_c^{\infty}(B(0, 4R_*))$ such that $\chi \equiv 1$ on $\bar{B}(0, 3R_*)$, we deduce

$$\begin{aligned}\left|\sum_{k=1}^{N_{\mathfrak{p},j}} d_{\mathfrak{p},j,k} - \deg(\hat{u}_{\mathfrak{p}}, \partial B(\hat{z}_{\mathfrak{p},j}, 3R_*))\right| &= \left| \int_{B(\hat{z}_{\mathfrak{p},j},3R_*)} \sum_{k=1}^{N_{\mathfrak{p},j}} d_{\mathfrak{p},j,k} \delta_{\hat{y}_{\mathfrak{p},j,k}} \, d\hat{x} - \frac{1}{\pi} \int_{B(\hat{z}_{\mathfrak{p},j},4R_*)} J\hat{u}'_{\mathfrak{p}} \, d\hat{x} \right| \\ &= \frac{1}{\pi} \left| \int_{B(\hat{z}^j_{\mathfrak{p}},4R_*)} \chi(\hat{x} - \hat{z}_{\mathfrak{p},j}) \left(\sum_{k=1}^{N_{\mathfrak{p},j}} d_{\mathfrak{p},j,k} \delta_{\hat{y}_{\mathfrak{p},j,k}} - J\hat{u}'_{\mathfrak{p}} \right) d\hat{x} \right| \\ &\leqslant \frac{1}{\pi} \|\chi\| \times \left\| J\hat{u}'_{\mathfrak{p}} - \pi \sum_{k=1}^{N_{\mathfrak{p},j}} d_{\mathfrak{p},j,k} \delta_{\hat{y}_{\mathfrak{p},j,k}} \right\|_{[C_c^{0,1}(\bar{D}(\hat{z}_{\mathfrak{p},j},4R_*))]^*}\end{aligned}$$

by (3-8). Since the left-hand side is an integer and the right-hand side is $\leqslant \frac{1}{2}$ provided $\mathfrak{p} \geqslant \mathfrak{p}_{2,1}(\delta, \Lambda_0)$, (3-11) follows.

We finally notice that the degree of $\hat{u}'_{\mathfrak{p}}$ on some large circle $\partial B(0, R)$ (with $R \gg \max_{1 \leqslant j \leqslant n_{\mathfrak{p}}} |\hat{z}_{\mathfrak{p},j}|$) is zero, for otherwise $\hat{u}'_{\mathfrak{p}}$ (and $\hat{u}_{\mathfrak{p}}$) would have infinite kinetic energy. Therefore,

$$0 = \sum_{j=1}^{n_{\mathfrak{p}}} \deg(\hat{u}_{\mathfrak{p}}, \partial B(\hat{z}_{\mathfrak{p},j}, 3R_*)) = \sum_{j=1}^{n_{\mathfrak{p}}} \sum_{k=1}^{N_{\mathfrak{p},j}} d_{\mathfrak{p},j,k}.$$

Combining this with (3-10), we deduce that if we are not in case (I), then one of the $d_{\mathfrak{p},j,k}$ must be equal to $+1$ and another one must be equal to -1, which is case (II).

Notice that for case (II), if $B(\hat{z}_{\mathfrak{p},j}, 4R_*)$ contains neither $y_{\mathfrak{p},+}$ nor $y_{\mathfrak{p},-}$, then $\| J\hat{u}'_{\mathfrak{p}}\|_{[C_c^{0,1}(B(\hat{z}_{\mathfrak{p},j},4R_*))]^*} \leqslant \delta$.

As in [5], we now relate the location of the points $\hat{y}_{\mathfrak{p},\pm}$ to the momentum $P(\hat{u}_{\mathfrak{p}})$.

<u>Step 6</u>: Case (I) does not occur for $\mathfrak{p}$ sufficiently large, say $\mathfrak{p} \geqslant \mathfrak{p}_3$. In addition, we have

$$1 = P(\hat{u}_{\mathfrak{p}}) = \pi((\hat{y}_{\mathfrak{p},+})_1 - (\hat{y}_{\mathfrak{p},-})_1) + o(1).$$

First, we have, by computations similar to those of Step 3, $\hat{u}_{\mathfrak{p}} = A e^{i\varphi}$ locally outside Ω_{R_*}; hence $\langle i\hat{u}_{\mathfrak{p}} \mid \nabla \hat{u}_{\mathfrak{p}}\rangle = A^2 \nabla\varphi$ and then, outside Ω_{3R_*},

$$\langle i\hat{u}_{\mathfrak{p}} \mid \nabla \hat{u}_{\mathfrak{p}}\rangle - \langle i\hat{u}'_{\mathfrak{p}} \mid \nabla \hat{u}'_{\mathfrak{p}}\rangle = A^2\nabla\varphi - \nabla\varphi = \frac{A^2-1}{A} \times A\nabla\varphi.$$

In $B(\hat{z}_{\mathfrak{p},j}, 3R_*) \setminus B(\hat{z}_{\mathfrak{p},j}, 2R_*)$, we obtain

$$|\langle i\hat{u}_{\mathfrak{p}} \mid \nabla \hat{u}_{\mathfrak{p}}\rangle - \langle i\hat{u}'_{\mathfrak{p}} \mid \nabla \hat{u}'_{\mathfrak{p}}\rangle| = |A^2\nabla\varphi - |\hat{u}'_{\mathfrak{p}}|^2\nabla\varphi| \leqslant \frac{|A^2-1|}{A} \times |A\nabla\varphi|,$$

since $|\hat{u}'_{\mathfrak{p}}| \in [|\hat{u}_{\mathfrak{p}}|, 1]$. Therefore,

$$\|\langle i\hat{u}_{\mathfrak{p}} \mid \nabla \hat{u}_{\mathfrak{p}}\rangle - \langle i\hat{u}'_{\mathfrak{p}} \mid \nabla \hat{u}'_{\mathfrak{p}}\rangle\|_{L^1(\mathbb{R}^2)} \leqslant K \int_{\mathbb{R}^2\setminus\Omega_{2R_*}} |1-|\hat{u}_{\mathfrak{p}}|^2| \times |\nabla\hat{u}_{\mathfrak{p}}|\, d\hat{x} \leqslant \frac{K}{\mathfrak{p}} E_{1/\mathfrak{p}}(\hat{u}_{\mathfrak{p}}) \leqslant K\frac{\ln\mathfrak{p}}{\mathfrak{p}}. \quad (3\text{-}12)$$

Following [5; 8], we write

$$\begin{aligned} 1 = \frac{P(u_{\mathfrak{p}})}{\mathfrak{p}} = P(\hat{u}_{\mathfrak{p}}) &= \frac{1}{2}\int_{\mathbb{R}^2} \langle i\partial_2\hat{u}_{\mathfrak{p}} \mid \hat{u}_{\mathfrak{p}} - 1\rangle\, d\hat{x} \\ &= \frac{1}{2}\int_{\mathbb{R}^2} \langle i\partial_2\hat{u}'_{\mathfrak{p}} \mid \hat{u}'_{\mathfrak{p}} - 1\rangle\, d\hat{x} + \frac{1}{2}\int_{\mathbb{R}^2} (\langle i\partial_2\hat{u}_{\mathfrak{p}} \mid \hat{u}_{\mathfrak{p}} - 1\rangle - \langle i\partial_2\hat{u}'_{\mathfrak{p}} \mid \hat{u}'_{\mathfrak{p}} - 1\rangle)\, d\hat{x}. \end{aligned}$$

For the second integral, we write that, on the one hand,

$$\left|\int_{\mathbb{R}^2} (\langle i\hat{u}_{\mathfrak{p}} \mid \partial_2\hat{u}_{\mathfrak{p}}\rangle - \langle i\hat{u}'_{\mathfrak{p}} \mid \partial_2\hat{u}'_{\mathfrak{p}}\rangle)\, d\hat{x}\right| \leqslant \|\langle i\hat{u}_{\mathfrak{p}} \mid \nabla \hat{u}_{\mathfrak{p}}\rangle - \langle i\hat{u}'_{\mathfrak{p}} \mid \nabla \hat{u}'_{\mathfrak{p}}\rangle\|_{L^1(\mathbb{R}^2)} \leqslant K\frac{\ln\mathfrak{p}}{\mathfrak{p}} \to 0$$

when $\mathfrak{p} \to +\infty$; on the other hand, by the decays given in Theorem 2.4,

$$\begin{aligned} \left|\int_{\mathbb{R}^2} (\langle i\partial_2\hat{u}_{\mathfrak{p}} \mid 1\rangle - \langle i\partial_2\hat{u}'_{\mathfrak{p}} \mid 1\rangle)\, d\hat{x}\right| &= \lim_{r\to+\infty}\left|\int_{\partial B(0,r)} \nu_2\, \mathfrak{Im}(\hat{u}_{\mathfrak{p}} - \hat{u}'_{\mathfrak{p}})\, d\ell\right| \\ &\leqslant \lim_{r\to+\infty}\int_{\partial B(0,r)} |A-1|\, d\ell = \lim_{r\to+\infty} \mathcal{O}(1/r) = 0. \end{aligned}$$

We then integrate by parts to get

$$\frac{1}{2}\int_{\mathbb{R}^2} \langle i\partial_2\hat{u}'_{\mathfrak{p}} \mid \hat{u}'_{\mathfrak{p}} - 1\rangle\, d\hat{x} = \frac{1}{2}\int_{\mathbb{R}^2} \partial_1\hat{x}_1 \langle i\partial_2\hat{u}'_{\mathfrak{p}} \mid \hat{u}'_{\mathfrak{p}} - 1\rangle - \partial_2\hat{x}_1\langle i\partial_1\hat{u}'_{\mathfrak{p}} \mid \hat{u}'_{\mathfrak{p}} - 1\rangle\, d\hat{x} = \int_{\mathbb{R}^2} J\hat{u}'_{\mathfrak{p}}\hat{x}_1\, d\hat{x}.$$

The integration by parts is justified by the algebraic decay at infinity given in Theorem 2.4:

$$\hat{x}_1\langle i\partial_2\hat{u}'_{\mathfrak{p}} \mid \hat{u}'_{\mathfrak{p}}-1\rangle = \mathcal{O}\left(\frac{1}{|x|^2}\right).$$

Then, since $J\hat{u}'_{\mathfrak{p}}$ is supported in Ω_{R_*}, we obtain

$$\begin{aligned}\int_{\mathbb{R}^2}\hat{x}_1 J\hat{u}'_{\mathfrak{p}}\,d\hat{x} &= \sum_{j=1}^{n_{\mathfrak{p}}}\int_{B(\hat{z}_{\mathfrak{p},j},3R_*)}\hat{x}_1 J\hat{u}'_{\mathfrak{p}}\,d\hat{x}\\ &= \sum_{j=1}^{n_{\mathfrak{p}}}\int_{B(\hat{z}_{\mathfrak{p},j},3R_*)}(\hat{x}_1-(\hat{z}_{\mathfrak{p},j})_1)J\hat{u}'_{\mathfrak{p}}\,d\hat{x}+\sum_{j=1}^{n_{\mathfrak{p}}}\hat{z}_{\mathfrak{p},j,1}\int_{B(\hat{z}_{\mathfrak{p},j},3R_*)}J\hat{u}'_{\mathfrak{p}}\,d\hat{x}.\end{aligned}$$

We then fix $\chi\in\mathcal{C}_c^\infty(B(0,4R_*))$ such that $\chi\equiv 1$ on $\bar{B}(0,3R_*)$. Next, for any $1\leqslant j\leqslant n_{\mathfrak{p}}$, we write

$$\begin{aligned}&\int_{B(\hat{z}_{\mathfrak{p},j},3R_*)}(\hat{x}_1-(\hat{z}_{\mathfrak{p},j})_1)J\hat{u}'_{\mathfrak{p}}\,d\hat{x}\\ &=\int_{B(\hat{z}_{\mathfrak{p},j},4R_*)}(\hat{x}_1-(\hat{z}_{\mathfrak{p},j})_1)\chi(\hat{x}-\hat{z}_{\mathfrak{p},j})J\hat{u}'_{\mathfrak{p}}\,d\hat{x}\\ &=\int_{B(\hat{z}_{\mathfrak{p},j},4R_*)}(\hat{x}_1-(\hat{z}_{\mathfrak{p},j})_1)\chi(\hat{x}-\hat{z}_{\mathfrak{p},j})\left(J\hat{u}'_{\mathfrak{p}}-\pi\sum_{k=1}^{N_{\mathfrak{p},j}}d_{\mathfrak{p},j,k}\delta_{y_{\mathfrak{p},j,k}}\right)d\hat{x}+\pi\sum_{k=1}^{N_{\mathfrak{p},j}}d_{\mathfrak{p},j,k}((y_{\mathfrak{p},j,k})_1-(\hat{z}_{\mathfrak{p},j})_1).\end{aligned}$$

We now estimate the first integral (actually, a duality bracket) by using Step 5:

$$\begin{aligned}&\left|\int_{B(\hat{z}_{\mathfrak{p},j},2R_*)}(\hat{x}_1-(\hat{z}_{\mathfrak{p},j})_1)\chi(\cdot-\hat{z}_{\mathfrak{p},j})\left(J\hat{u}'_{\mathfrak{p}}-\pi\sum_{k=1}^{N_{\mathfrak{p},j}}d_{\mathfrak{p},j,k}\delta_{y_{\mathfrak{p},j,k}}\right)d\hat{x}\right|\\ &\qquad\leqslant\|(\hat{x}_1-(\hat{z}_{\mathfrak{p},j})_1)\chi(\cdot-\hat{z}_{\mathfrak{p},j})\|_{\mathcal{C}_c^{0,1}(B(\hat{z}_{\mathfrak{p},j},2R_*))}\left\|J\hat{u}'_{\mathfrak{p}}-\pi\sum_{k=1}^{N_{\mathfrak{p},j}}d_{\mathfrak{p},j,k}\delta_{y_{\mathfrak{p},j,k}}\right\|_{[\mathcal{C}_c^{0,1}(B(\hat{z}_{\mathfrak{p},j},2R_*))]^*}\\ &\qquad\leqslant Ko(1).\end{aligned}$$

As a consequence of (3-11), which implies, for each $1\leqslant j\leqslant n_{\mathfrak{p}}$,

$$\sum_{k=1}^{N_{\mathfrak{p},j}}d_{\mathfrak{p},j,k}=\deg(\hat{u}_{\mathfrak{p}},\partial B(\hat{z}_{\mathfrak{p},j},3R_*))=\deg(\hat{u}'_{\mathfrak{p}},\partial B(\hat{z}_{\mathfrak{p},j},3R_*))=\int_{B(\hat{z}_{\mathfrak{p},j},3R_*)}J\hat{u}'_{\mathfrak{p}}\,d\hat{x},$$

we infer, after some cancellation,

$$\left|P(\hat{u}_{\mathfrak{p}})-\pi\sum_{j=1}^{n_{\mathfrak{p}}}\sum_{k=1}^{N_{\mathfrak{p},j}}d_{\mathfrak{p},j,k}(y_{\mathfrak{p},j,k})_1\right|\leqslant K\frac{\ln\mathfrak{p}}{\mathfrak{p}}+n_*Ko(1). \tag{3-13}$$

Since $P(\hat{u}_{\mathfrak{p}})=1$, it follows that for $\mathfrak{p}$ large enough, we cannot be in Case (I), and the conclusion is a recasting of (3-13).

<u>Step 7</u>: There exists $\mathfrak{p}_4$ large such that, for $\mathfrak{p}\geqslant\mathfrak{p}_4$, we have $\{|\hat{u}_{\mathfrak{p}}|\leqslant\frac{1}{2}\}\subset B\big(\hat{y}_{\mathfrak{p},+},\frac{3}{20}\big)\cup B\big(\hat{y}_{\mathfrak{p},-},\frac{3}{20}\big)$ and $\deg\big(u,\partial B\big(\hat{y}_{\mathfrak{p},\pm},\frac{3}{20}\big)\big)=\pm 1$.

From Step 6, we know that $1=P(\hat{u}_{\mathfrak{p}})=\pi((\hat{y}_{\mathfrak{p},+})_1-(\hat{y}_{\mathfrak{p},-})_1)+o(1)$; hence the two points $\hat{y}_{\mathfrak{p},\pm}$ are far away from each other:

$$|\hat{y}_{\mathfrak{p},+}-\hat{y}_{\mathfrak{p},-}|\geqslant\tfrac{4}{10}$$

$\big($since $\frac{1}{\pi} \approx 0.318 < \frac{4}{10}\big)$ for $\mathfrak{p}$ large enough (but they may be, at this stage, very far away from each other). By applying Theorem 1.1(i) of [2] or Theorem 3.1 of [27] (this is not very far from Theorem 3.1), since $J\hat{u}_{\mathfrak{p}}(\hat{y}_{\mathfrak{p},\pm} + \cdot\,) \to \pm\pi\delta_0$ weakly, we deduce

$$E_{1/\mathfrak{p}}\big(\hat{u}_{\mathfrak{p}}, B\big(\hat{y}_{\mathfrak{p},\pm}, \tfrac{1}{10}\big)\big) \geqslant (\pi + o(1)) \ln \mathfrak{p};$$

hence, by the upper bound (3-3),

$$E_{1/\mathfrak{p}}\big(\hat{u}_{\mathfrak{p}}, \mathbb{R}^2 \setminus \big(B\big(\hat{y}_{\mathfrak{p},+}, \tfrac{1}{10}\big) \cup B\big(\hat{y}_{\mathfrak{p},-}, \tfrac{1}{10}\big)\big)\big) \leqslant o(\ln \mathfrak{p}),$$

and this in turn implies, by the clearing-out theorem (Theorem 3.3), that if $\mathfrak{p}$ is large enough, say $\mathfrak{p} \geqslant \mathfrak{p}_4$, then,

$$\text{for all } \hat{x} \in \mathbb{R}^2 \setminus \big(B\big(\hat{y}_{\mathfrak{p},+}, \tfrac{3}{20}\big) \cup B\big(\hat{y}_{\mathfrak{p},-}, \tfrac{3}{20}\big)\big), \quad |\hat{u}_{\mathfrak{p}}(\hat{x})| \geqslant \tfrac{3}{4},$$

as wished. In particular, $\hat{z}_{\mathfrak{p},\pm} \in B\big(\hat{y}_{\mathfrak{p},+}, \frac{3}{20}\big) \cup B\big(\hat{y}_{\mathfrak{p},-}, \frac{3}{20}\big)$.

We emphasize that at this stage, we have $|\hat{y}_{\mathfrak{p},+} - \hat{y}_{\mathfrak{p},-}| \gtrsim 1$, but we do not know whether $|\hat{y}_{\mathfrak{p},+} - \hat{y}_{\mathfrak{p},-}| \lesssim 1$ or $|\hat{y}_{\mathfrak{p},+} - \hat{y}_{\mathfrak{p},-}| \gg 1$. We may now take advantage of the fact that $\hat{u}_{\mathfrak{p}}$ is by hypothesis symmetric with respect to the x_2-axis (i.e., $\hat{u}_{\mathfrak{p}}(-\hat{x}_1, \hat{x}_2) = \hat{u}_{\mathfrak{p}}(\hat{x}_1, \hat{x}_2)$), so that, possibly translating along the x_2-axis, we may assume

$$(\hat{y}_{\mathfrak{p},-})_2 = (\hat{y}_{\mathfrak{p},+})_2 = 0 \quad \text{and} \quad -(\hat{y}_{\mathfrak{p},-})_1 = (\hat{y}_{\mathfrak{p},+})_1 \to \tfrac{1}{2\pi}. \tag{3-14}$$

If we do not assume a priori the symmetry in x_1, then we may remove the translation invariance by imposing $\hat{y}_{\mathfrak{p},+} + \hat{y}_{\mathfrak{p},-} = 0$, and then we may still show that $\hat{y}_{\mathfrak{p},+} = -\hat{y}_{\mathfrak{p},-} \to \big(\frac{1}{2\pi}, 0\big)$ by using the Hopf differential as in [6, Chapter VII].

3B2. *Strong convergence outside the vorticity set at scale $x/\mathfrak{p}$.* We start with a $W^{1,p}_{\mathrm{loc}}$ bound at scale $\hat{x}$ for $1 \leqslant p < 2$.

Step 1: For any $1 \leqslant p < 2$, there exists C_p such that, for any $\widehat{X} \in \mathbb{R}^2$, we have

$$\int_{B(\widehat{X},1)} |\nabla \hat{u}_{\mathfrak{p}}|^p \, d\hat{x} \leqslant C_p.$$

We shall adapt the proof of [8] (see the proof of Theorem 4, Step 3, p. 83) to the two-dimensional case. Actually, the only modification to make in the estimate is to replace (C.26) there by the standard convolution

$$\psi_{0,i}(\hat{x}) = -\frac{\ln r}{2\pi} \star \omega_{0,i}(\hat{x}) = -\frac{1}{2\pi} \int_{\mathrm{Supp}(\omega_{0,i})} \omega_{0,i}(\hat{y}) \ln |\hat{x} - \hat{y}| \, d\hat{y},$$

and then use, for $|\hat{x} - \hat{y}_{\mathfrak{p},\pm}| \geqslant 3R_*$, that

$$\begin{aligned} |\nabla \psi_{0,\pm}(\hat{x})| &= \left| \frac{1}{2\pi} \int_{\mathrm{Supp}(\omega_{0,\pm})} \omega_{0,i}(\hat{y}) \nabla_{\hat{x}} \ln |\hat{x} - \hat{y}| \, d\hat{y} \right| \\ &\leqslant \frac{1}{2\pi} \|\omega_{0,\pm}\|_{[C^{0,1}_c(B(\hat{y}_{\mathfrak{p},\pm}, 2R_*))]^*} \|(\hat{x} - \hat{y})/|\hat{x} - \hat{y}|^2\|_{C^{0,1}(B(\hat{y}_{\mathfrak{p},\pm}, 3R_*))} \leqslant K \end{aligned}$$

(the estimate $\|\psi_{0,\pm}\|_{\mathcal{C}^k(\mathbb{R}^2\setminus B(\hat{y}_{\mathfrak{p},\pm},3R_*))} \leqslant C_k$ does not hold since the two-dimensional fundamental solution $(\ln r)/(2\pi)$ goes to $+\infty$ at spatial infinity, but $\|\nabla\psi_{0,\pm}\|_{\mathcal{C}^k(\mathbb{R}^2\setminus B(\hat{y}_{\mathfrak{p},\pm},3R_*))} \leqslant C_k$ is true). The rest of the proof remains unchanged.

Step 2: For any $\widehat{X} \in \mathbb{R}^2 \setminus \big(B\big(\hat{y}_{\mathfrak{p},+}, \frac{2}{10}\big) \cup B\big(\hat{y}_{\mathfrak{p},-}, \frac{2}{10}\big)\big)$, we may write $\hat{u}_{\mathfrak{p}} = A\mathrm{e}^{i\phi}$ in $B\big(\widehat{X}, \frac{1}{20}\big)$, with, for any $k \in \mathbb{N}$,

$$\left\| 2(1-A) - \frac{c_{\mathfrak{p}}}{\mathfrak{p}}\partial_2\phi \right\|_{\mathcal{C}^k(B(\widehat{X},1/20))} \leqslant \frac{C_k}{\mathfrak{p}^2}, \quad \|\nabla\phi\|_{\mathcal{C}^k(B(\widehat{X},1/20))} \leqslant C_k, \tag{3-15}$$

for some constant C_k independent of $\widehat{X}$.

The proof (relying on Step 1) follows the lines of the proof of Step 7 (p. 48) of Theorem 1 in [8] and is omitted.

In view of the upper bound of Step 1 of Section 3B1, we infer the uniform estimate

$$\|1 - |\hat{u}_{\mathfrak{p}}|\|_{\mathcal{C}^k(B(\widehat{X},1/20))} \leqslant C_k \frac{\ln \mathfrak{p}}{\mathfrak{p}^2} \tag{3-16}$$

for $\widehat{X} \in \mathbb{R}^2 \setminus \big(B\big(\hat{y}_{\mathfrak{p},+}, \frac{2}{10}\big) \cup B\big(\hat{y}_{\mathfrak{p},-}, \frac{2}{10}\big)\big)$.

3B3. *Lower bound for the energy and upper bound for the potential energy.*

Step 1: Upper bound for the potential. We claim that

$$\int_{\mathbb{R}^2} |\nabla|\hat{u}_{\mathfrak{p}}||^2 + \frac{\mathfrak{p}^2}{2}(1-|\hat{u}_{\mathfrak{p}}|^2)^2 \, d\hat{x} \leqslant C(\Lambda_0),$$

$$\int_{\mathbb{R}^2\setminus(B(\hat{y}_{\mathfrak{p},+},2/10)\cup B(\hat{y}_{\mathfrak{p},-},2/10))} |\nabla\hat{u}_{\mathfrak{p}}|^2 + \frac{\mathfrak{p}^2}{2}(1-|\hat{u}_{\mathfrak{p}}|^2)^2 \, d\hat{x} \leqslant C(\Lambda_0).$$

The proof of this upper bound will be a direct consequence of the lower bounds established in [43] (see Theorems 2 and 3 there).

Theorem 3.4 [43]. *Let $\Omega \subset \mathbb{R}^2$ be a bounded smooth domain. Assume that $u \in H^1(\Omega, \mathbb{C})$ and that $u_{|\partial\Omega} \in \mathcal{C}^1(\partial\Omega, \mathcal{S}^1)$. Let $\delta \in]0, 1[$.*

(i) *There exists a constant Λ_1, depending on Ω and $\|u_{|\partial\Omega}\|_{\mathcal{C}^1}$, such that*

$$\frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{2\delta^2}(1-|u|^2)^2 \geqslant \pi|\deg(u_{|\partial\Omega}, \partial\Omega)| \ln(1/\delta) - \Lambda_1.$$

(ii) *If, moreover, for some constant Λ_2, we have*

$$\frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{2\delta^2}(1-|u|^2)^2 \leqslant \pi|\deg(u_{|\partial\Omega}, \partial\Omega)| \ln(1/\delta) + \Lambda_2,$$

then

$$\frac{1}{2}\int_\Omega |\nabla|u||^2 + \frac{1}{2\delta^2}(1-|u|^2)^2 \leqslant C(\Omega, \Lambda_2, \|u_{|\partial\Omega}\|_{\mathcal{C}^1}).$$

We shall apply this result with $\delta = 1/\mathfrak{p} \ll 1$, $\Omega = B\big(\hat{y}_{\mathfrak{p},\pm}, \frac{2}{10}\big)$ and $u = \hat{u}_{\mathfrak{p}}$. In view of the upper bound (3-3) on the energy of $\hat{u}_{\mathfrak{p}}$ and since $\deg\big(\hat{u}_{\mathfrak{p}}, \partial B\big(\hat{y}_{\mathfrak{p},\pm}, \frac{2}{10}\big)\big) = \pm 1$, this yields

$$\int_{B(\hat{y}_{\mathfrak{p},\pm},2/10)} |\nabla\hat{u}_{\mathfrak{p}}|^2 + \frac{\mathfrak{p}^2}{2}(1-|\hat{u}_{\mathfrak{p}}|^2)^2 \, d\hat{x} \geqslant \pi \ln \mathfrak{p} - \Lambda_1,$$

$$\int_{B(\hat{y}_{\mathfrak{p},\pm},2/10)} |\nabla|\hat{u}_{\mathfrak{p}}||^2 + \frac{\mathfrak{p}^2}{2}(1-|\hat{u}_{\mathfrak{p}}|^2)^2\,d\hat{x} \leqslant C(\Lambda_0).$$

We conclude by using once again the upper bound (3-3). Actually, $\hat{u}_{\mathfrak{p}}$ does not belong to $\mathcal{C}^1\big(\partial B\big(\hat{y}_{\mathfrak{p},\pm}, \frac{2}{10}\big)\big)$, but it is easy, using (3-15), to construct an extension of $\hat{u}_{\mathfrak{p}}$ on $B\big(\hat{y}_{\mathfrak{p},\pm}, \frac{3}{10}\big)$ with the required properties by linear interpolation (see, for instance the lemma on p. 395-396 in [43]).

Step 2: There exists $\sigma_0 > 0$ such that we have, for $R \geqslant 1$,

$$\int_{\mathbb{R}^2\setminus B(0,R)} |\nabla\hat{u}_{\mathfrak{p}}|^2 + \frac{\mathfrak{p}^2}{2}(1-|\hat{u}_{\mathfrak{p}}|^2)^2\,d\hat{x} \leqslant \frac{C(\Lambda_0)}{R^{\sigma_0}}.$$

The proof is similar to that of Lemma 5.1 (p. 50) in [8], and relies on the fact that $|\hat{u}_{\mathfrak{p}}| \geqslant \frac{1}{2}$ in $\mathbb{R}^2 \setminus B(0,1)$ (hence we may write the PDE in terms of modulus and phase), and the upper bound in $\mathbb{R}^2 \setminus \big(B\big(\hat{y}_{\mathfrak{p},+}, \frac{2}{10}\big) \cup B\big(\hat{y}_{\mathfrak{p},-}, \frac{2}{10}\big)\big) \supset \mathbb{R}^2 \setminus B(0,1)$ of the energy of $\hat{u}_{\mathfrak{p}}$ (in [8], this last upper bound was derived differently).

3B4. *Convergence on the scale $x/\mathfrak{p}$.* By Step 1 of Section 3B3 and (3-14), we have, as $\mathfrak{p} \to +\infty$,

$$\hat{y}_{\mathfrak{p},\pm} \to \hat{y}_{\infty,\pm} := \pm(1/(2\pi), 0) \in \mathbb{R}^2. \tag{3-17}$$

We then define (identifying $\mathbb{R}^2$ and $\mathbb{C}$)

$$\hat{u}_\infty(\hat{x}) := \frac{\hat{x}-\hat{y}_{\infty,+}}{|\hat{x}-\hat{y}_{\infty,+}|} \times \overline{\frac{\hat{x}+\hat{y}_{\infty,-}}{|\hat{x}+\hat{y}_{\infty,-}|}}.$$

Step 1: For any $p \in [1,2[$, there holds, in $W^{1,p}_{\text{loc}}(\mathbb{R}^2)$,

$$\hat{u}_{\mathfrak{p}} \rightharpoonup \hat{u}_\infty.$$

From the $W^{1,p}_{\text{loc}}$ upper bound of Step 1 in Section 3B2 and by weak compactness, there exists $\widehat{U} \in W^{1,p}_{\text{loc}}(\mathbb{R}^2)$ such that $\hat{u}_{\mathfrak{p}} \rightharpoonup \widehat{U}$ in $W^{1,p}_{\text{loc}}(\mathbb{R}^2)$. Moreover, $\widehat{U} \in \mathcal{C}^\infty_{\text{loc}}(\mathbb{R}^2 \setminus \{\hat{y}_{\infty,+}, \hat{y}_{\infty,-}\})$ and the convergence holds in $\mathcal{C}^k_{\text{loc}}(\mathbb{R}^2 \setminus \{\hat{y}_{\infty,+}, \hat{y}_{\infty,-}\})$ by Step 2 of Section 3B2 (for any $k \in \mathbb{N}$). In order to determine $\widehat{U}$, we shall pass to the limit in the system

$$\begin{cases} \nabla\cdot(\hat{u}_{\mathfrak{p}} \wedge \nabla\hat{u}_{\mathfrak{p}}) = -\frac{1}{2}c_{\mathfrak{p}}\mathfrak{p}\partial_2(|\hat{u}_{\mathfrak{p}}|^2-1), \\ \nabla^\perp\cdot(\hat{u}_{\mathfrak{p}} \wedge \nabla\hat{u}_{\mathfrak{p}}) = 2J\hat{u}_{\mathfrak{p}} \end{cases}$$

obtained from (3-5) and the definition of the Jacobian. From (3-3) (implying $c_{\mathfrak{p}}\mathfrak{p}\partial_2(|\hat{u}_{\mathfrak{p}}|^2-1) \to 0$ in the distributional or the H^{-1} sense) and Step 5 of Section 3B1, we then infer

$$\begin{cases} \nabla\cdot(\widehat{U} \wedge \nabla\widehat{U}) = 0, \\ \nabla^\perp\cdot(\widehat{U} \wedge \nabla\widehat{U}) = 2\pi(\delta_{\hat{y}_{\infty,+}} - \delta_{\hat{y}_{\infty,-}}). \end{cases}$$

It then follows that $\widehat{U} \wedge \nabla\widehat{U} = \hat{u}_\infty \wedge \nabla\hat{u}_\infty$; hence we have the existence of $\Theta \in \mathbb{R}$ such that $\widehat{U} = \mathrm{e}^{i\Theta}\hat{u}_\infty$. We finally use the x_1-symmetry to infer $\Theta = 0$.

Step 2: As $\mathfrak{p} \to +\infty$, we have

$$\mathfrak{p}c_{\mathfrak{p}} = \frac{\mathfrak{p}^2}{2}\int_{\mathbb{R}^2} (1-|\hat{u}_{\mathfrak{p}}|^2)^2 \, d\hat{x} \to 2\pi.$$

This is claimed in [5, Proposition VI.7], but the proof is not clearly given.

One way to prove this point is to use the Hopf differential as in [6, Chapter VII]. We shall follow the alternative proof of Theorem VII.2 given in Section VII.1 there. The first equality is the Pohozaev identity (2-2).

First, notice that

$$W_{\mathfrak{p}} := \frac{\mathfrak{p}^2}{2}(1-|\hat{u}_{\mathfrak{p}}|^2)^2$$

is a nonnegative function which is bounded in $L^1(\mathbb{R}^2)$ by Step 1 of Section 3B3 and enjoys the decay estimate of Step 2 of Section 3B3. In addition, by (3-16) (see Step 2 of Section 3B2), we have $W_{\mathfrak{p}} \to 0$ locally uniformly in $\mathbb{R}^2 \setminus \{\pm(1/(2\pi), 0)\}$. Up to a subsequence, we may then assume that

$$W_{\mathfrak{p}} \rightharpoonup \mu_+ \delta_{\hat{y}_{\infty,+}} + \mu_- \delta_{\hat{y}_{\infty,-}}$$

in the weak $*$ topology of $\mathcal{C}_b(\mathbb{R}^2)$ for some reals $\mu_\pm \geqslant 0$, with $\mu_+ + \mu_- = \lim_{\mathfrak{p}\to+\infty} \int_{\mathbb{R}^2} W_{\mathfrak{p}}$.

We shall now compute μ_+ (the case of μ_- is similar). First, we write, for some $R_5 \leqslant \frac{2}{10}$, the Pohozaev identity for $\hat{u}_{\mathfrak{p}}$ on $B(\hat{y}_{\infty,+}, R_5)$ (obtained by multiplying the equation by the conjugate of $(\hat{x} - \hat{y}_{\infty,+}) \cdot \nabla \hat{u}_{\mathfrak{p}}$ and integrating the real part over $B(\hat{y}_{\infty,+}, R_5)$), which yields

$$\int_{B(\hat{y}_{\infty,+},R_5)} \frac{\mathfrak{p}^2}{2}(1-|\hat{u}_{\mathfrak{p}}|^2)^2 + c_{\mathfrak{p}}\mathfrak{p}\int_{B(\hat{y}_{\infty,+},R_5)} (\hat{x}_1 - \hat{y}_{\infty,+,1})\langle i\partial_2\hat{u}_{\mathfrak{p}} \mid \partial_1\hat{u}_{\mathfrak{p}}\rangle$$
$$= \frac{R_5}{2}\int_{\partial B(\hat{y}_{\infty,+},R_5)} |\partial_\tau \hat{u}_{\mathfrak{p}}|^2 - |\partial_\nu \hat{u}_{\mathfrak{p}}|^2 + \frac{\mathfrak{p}^2}{4}(1-|\hat{u}_{\mathfrak{p}}|^2)^2.$$

We then pass to the limit $\mathfrak{p} \to +\infty$. For the boundary term, we use the strong convergences outside the vorticity set; for the second term of the first line, we prove that it tends to zero by following the arguments given for Step 6 in Section 3B1. We then get

$$\mu_+ = \frac{R_5}{2}\int_{\partial B(\hat{y}_{\infty,+},R_5)} |\partial_\tau \hat{u}_\infty|^2 - |\partial_\nu \hat{u}_\infty|^2.$$

By Step 1, we know that $\hat{u}_\infty = \exp(i \operatorname{Arg}(\hat{x} - \hat{y}_{\infty,+}) - i \operatorname{Arg}(\hat{x} - \hat{y}_{\infty,-}))$ on $\partial B(\hat{y}_{\infty,+}, R_5)$, and the second term $\operatorname{Arg}(\hat{x} - \hat{y}_{\infty,-})$ is smooth and harmonic in $\overline{D}(\hat{y}_{\infty,+}, R_5)$. As a consequence, we have the Pohozaev identity for $\operatorname{Arg}(\cdot - \hat{y}_{\infty,-})$

$$0 = \frac{R_5}{2}\int_{\partial B(\hat{y}_{\infty,+},R_5)} |\partial_\tau \operatorname{Arg}(\hat{x} - \hat{y}_{\infty,-})|^2 - |\partial_\nu \operatorname{Arg}(\hat{x} - \hat{y}_{\infty,-})|^2,$$

$\partial_\tau \operatorname{Arg}(\hat{x} - \hat{y}_{\infty,+}) = 1/R_5$, $\partial_\nu \operatorname{Arg}(\hat{x} - \hat{y}_{\infty,+}) = 0$, and thus by expansion

$$\mu_+ = \frac{R_5}{2}\int_{\partial B(\hat{y}_{\infty,+},R_5)} |\partial_\tau \hat{u}_\infty|^2 - |\partial_\nu \hat{u}_\infty|^2 = \frac{R_5}{2}\int_{\partial B(\hat{y}_{\infty,+},R_5)} \frac{1}{R_5^2} + \frac{2\partial_\tau \operatorname{Arg}(\hat{x} - \hat{y}_{\infty,-})}{R_5} = \pi.$$

This concludes the proof.

3B5. *Convergence on the scale x.* We shall now focus on verifying hypothesis (2) of Proposition 1.8. The main tool is the following result. We now work on the scale x.

Proposition 3.5. *Assume that $\hat{z}_{\mathfrak{p}} \in \mathbb{R}^2$ is such that*

$$\limsup_{\mathfrak{p}\to+\infty} |\hat{u}_{\mathfrak{p}}(\hat{z}_{\mathfrak{p}})| < 1$$

and consider the rescaled mapping

$$U_{\mathfrak{p}}(y) := \hat{u}_{\mathfrak{p}}(\hat{z}_{\mathfrak{p}} + y/\mathfrak{p}).$$

Then, there exists a sign $\pm$ and $\beta \in \mathbb{R}$ (depending on the choice of the family $(\hat{z}_{\mathfrak{p}})$) such that, up to a subsequence, we have, in $C^k_{\mathrm{loc}}(\mathbb{R}^2)$ for any $k \in \mathbb{N}$,

$$U_{\mathfrak{p}} \to \mathrm{e}^{i\beta} V_{\pm}.$$

Proof. The rescaling $U_{\mathfrak{p}}$ solves

$$\Delta U_{\mathfrak{p}} + i c_{\mathfrak{p}} \partial_2 U_{\mathfrak{p}} + U_{\mathfrak{p}}(1 - |U_{\mathfrak{p}}|^2) = 0$$

and satisfies $\limsup_{\mathfrak{p}\to+\infty} |U_{\mathfrak{p}}(0)| < 1$ and, by Step 2 of Section 3B4,

$$\int_{\mathbb{R}^2} (1 - |U_{\mathfrak{p}}|^2)^2 \, dy = 4\pi + o_{\mathfrak{p}\to+\infty}(1).$$

Then, from the uniform bounds of Theorem 2.2 and Corollary 2.3, we may assume, up to a subsequence,

$$U_{\mathfrak{p}} \to U_\infty \tag{3-18}$$

in $C^k_{\mathrm{loc}}(\mathbb{R}^2)$ with $|U_\infty(0)| < 1$,

$$\Delta U_\infty + U_\infty(1 - |U_\infty|^2) = 0$$

and, by Fatou's lemma,

$$\int_{\mathbb{R}^2} (1 - |U_\infty|^2)^2 \, dy \leqslant 4\pi.$$

By [11], we know that $\int_{\mathbb{R}^2} (1 - |U_\infty|^2)^2 \, dy = 2\pi d^2$, where $d \in \mathbb{Z}$ is the degree of U_∞ at infinity. It follows that $|d| \leqslant 1$, and that the case $d = 0$ is excluded since $|U_\infty(0)| < 1$; hence $|U_\infty| \not\equiv 1$. Therefore $d = \pm 1$. It then follows from [36] that $U_\infty = \mathrm{e}^{i\beta} V_d$ for some $\beta \in \mathbb{R}$. □

We may now localize the set $\{|\hat{u}_{\mathfrak{p}}| \leqslant 1 - 1/\lambda_*\}$, where λ_* is as in Proposition 1.8, rather precisely.

Step 1: There exists $\mathfrak{p}_6$ large such that, for $\mathfrak{p} \geqslant \mathfrak{p}_6$, $\hat{u}_{\mathfrak{p}}$ has exactly two zeros $\hat{z}_{\mathfrak{p},\pm}$. Up to a translation in the x_2-direction, we may assume

$$\mathbb{R} \times \{0\} \ni \hat{z}_{\mathfrak{p},\pm} \to \left(\pm\tfrac{1}{2\pi}, 0\right) \in \mathbb{R}^2.$$

Moreover, there exists $R_0 > 0$ such that $\{|\hat{u}_{\mathfrak{p}}| \leqslant 1 - 1/\lambda_*\} \subset B(\hat{z}_{\mathfrak{p},+}, R_0/\mathfrak{p}) \cup B(\hat{z}_{\mathfrak{p},-}, R_0/\mathfrak{p})$. Here, $\lambda_* > 0$ is the large universal constant appearing in Proposition 1.8.

By Step 8 of Section 3B1, we know (due to the nonzero degree) that $\hat{u}_{\mathfrak{p}}$ has at least two zeros, one in each disk $B\left(\hat{y}_{\mathfrak{p},\pm}, \tfrac{3}{20}\right)$.

Now, if $\hat{z}_{\mathfrak{p}}$ is a zero of $\hat{u}_{\mathfrak{p}}$, we know by Proposition 3.5 that, for some $\beta \in \mathbb{R}$ (depending on the sequence $(\hat{z}_{\mathfrak{p}})_{\mathfrak{p}}$) and $d_0 = \pm 1$, we have

$$\hat{u}_{\mathfrak{p}}(\hat{z}_{\mathfrak{p}} + \mathfrak{p} y) \to \mathrm{e}^{i\beta} V_{d_0}(y) \tag{3-19}$$

in $\mathcal{C}^k_{\mathrm{loc}}(\mathbb{R}^2)$. As noticed in [41], since $V_\pm : \mathbb{R}^2 \to \mathbb{C} \approx \mathbb{R}^2$ has nonzero Jacobian at the origin, we deduce that for any $R > 0$, and for $\mathfrak{p} \geqslant \mathfrak{p}_R$ large enough, 0 is the only zero of $U_{\mathfrak{p}}$ in $B(0, R)$. Roughly speaking, there do not exist zeros $\hat{z}$, $\hat{z}'$ of $\hat{u}_{\mathfrak{p}}$ such that $0 < |\hat{z} - \hat{z}'| = \mathcal{O}(1/\mathfrak{p})$.

We now fix $R_0 > 0$ sufficiently large so that

$$\int_{\{|y| \leqslant R_0/2\}} (1 - |V_1(y)|^2)^2 \, dy \geqslant \tfrac{3\pi}{2}.$$

and we assume that (for any large $\mathfrak{p}$) $\{|\hat{u}_{\mathfrak{p}}| \leqslant 1 - 1/\lambda_*\}$ (where $\lambda_* > 0$ is the one appearing in Proposition 1.8) is not included in $B(\hat{z}_{\mathfrak{p},+}, R_0/\mathfrak{p}) \cup B(\hat{z}_{\mathfrak{p},-}, R_0/\mathfrak{p})$. This means that there exists $\widehat{Z}_{\mathfrak{p}} \in B\big(\hat{z}_{\mathfrak{p},+}, \tfrac{3}{20}\big) \setminus B(\hat{z}_{\mathfrak{p},+}, R_0/\mathfrak{p})$ (say) with $|\hat{u}_{\mathfrak{p}}(\widehat{Z}_{\mathfrak{p}})| \leqslant 1 - 1/\lambda_*$. By Proposition 3.5, the rescaled mapping $U_{\mathfrak{p}}(y) := \hat{u}_{\mathfrak{p}}(\widehat{Z}_{\mathfrak{p}} + \mathfrak{p} y)$ converges (up to a subsequence) in $\mathcal{C}^k_{\mathrm{loc}}(\mathbb{R}^2)$ to $U_\infty \in \mathbb{S}^1 V_\pm$ and we know (from [11]) that $\int_{\mathbb{R}^2} (1 - |U_\infty|^2)^2 \, dy = 2\pi$. As a consequence, since $|\hat{z}_{\mathfrak{p},+} - \widehat{Z}_{\mathfrak{p}}| \geqslant R_0/\mathfrak{p}$,

$$\begin{aligned}
2\pi + o(1) &= \mathfrak{p}^2 \int_{B(\hat{y}_{\mathfrak{p},+}, 3/20)} (1 - |\hat{u}_{\mathfrak{p}}|^2)^2 \, d\hat{x} \\
&\geqslant \mathfrak{p}^2 \int_{B(\hat{z}_{\mathfrak{p},+}, R_0/(2\mathfrak{p}))} (1 - |\hat{u}_{\mathfrak{p}}|^2)^2 \, d\hat{x} + \mathfrak{p}^2 \int_{B(\widehat{Z}_{\mathfrak{p}}, R_0/(2\mathfrak{p}))} (1 - |\hat{u}_{\mathfrak{p}}|^2)^2 \, d\hat{x} \\
&\geqslant \int_{\{|y| \leqslant R_0/2\}} (1 - |V_1|^2)^2 \, dy + \int_{\{|y| \leqslant R_0/2\}} (1 - |U_\infty|^2)^2 \, dy + o(1) \\
&\geqslant \tfrac{3\pi}{2} + \tfrac{3\pi}{2} + o(1),
\end{aligned}$$

which is absurd. We then conclude $\||u_{\mathfrak{p}}| - 1\|_{L^\infty(\{\tilde{r}_d \geqslant R_0\})} \leqslant 1/\lambda_*$ for $\mathfrak{p}$ sufficiently large, then proving hypothesis (3) of Proposition 1.8 with $\lambda = \max(R_0, \lambda_*)$. Another consequence of this fact is that $\hat{u}_{\mathfrak{p}}$ possesses at most two (simple) zeros $\hat{z}_{\mathfrak{p},\pm}$.

We then define $d = d_{\mathfrak{p}}$ such that the unique zero $\hat{z}_{\mathfrak{p},+}$ of $\hat{u}_{\mathfrak{p}}$ in the right half-plane is

$$\hat{z}_{\mathfrak{p},+} = \frac{d_{\mathfrak{p}}}{\mathfrak{p}} \vec{e}_1 \to \big(\tfrac{1}{2\pi}, 0\big) \in \mathbb{R}^2.$$

We deduce from Step 2 of Section 3B4 that

$$d_{\mathfrak{p}} \sim \frac{\mathfrak{p}}{2\pi} \sim \frac{1}{c_{\mathfrak{p}}},$$

so that hypothesis (4) of Proposition 1.8 is satisfied for $\mathfrak{p}$ large enough (still for $\lambda = \max(R_0, \lambda_*)$). Furthermore, hypothesis (2) of Proposition 1.8 is satisfied by taking $\mathfrak{p}$ large enough, associated with the choice $\lambda = \max(R_0, \lambda_*)$.

Step 2: Conclusion. Applying Proposition 1.8 to $\mathrm{e}^{-i\beta} u_{\mathfrak{p}}$, we infer that there exists $\gamma_{\mathfrak{p}} \in \mathbb{R}$ such that (for large $\mathfrak{p}$)

$$u_{\mathfrak{p}} = \mathrm{e}^{i\gamma_{\mathfrak{p}}} Q_{c_{\mathfrak{p}}}$$

(no translation is needed in the x_2-direction at this stage since the zeros of $\hat{u}_{\mathfrak{p}}$ are on the x_1-axis).

3C. *Decay slightly away from the vortices.* In this section, we provide some estimates for $\hat{u}_{\mathfrak{p}}$ in the region $B(\hat{z}_{\mathfrak{p},+}, 2R_0) \cup B(\hat{z}_{\mathfrak{p},-}, 2R_0)$. For the Ginzburg–Landau (stationary) model, such estimates were first given in [35] for minimizing solutions and later generalized in [18] to nonminimizing solutions. However, since the paper [35] is difficult to find, we give here a proof of these estimates that includes the transport term. They improve some estimates in [14] and are not specific to the way we construct the solutions.

Proposition 3.6. *We have*, *for* $|\hat{y}| \leqslant \frac{3}{20}$,

$$\big||\hat{u}_{\mathfrak{p}}(\hat{z}_{\mathfrak{p},\pm} + \hat{y})| - 1\big| \leqslant \frac{C}{\mathfrak{p}^2 |\hat{y}|^2}, \quad |\nabla |\hat{u}_{\mathfrak{p}}|(\hat{z}_{\mathfrak{p},\pm} + \hat{y})| \leqslant \frac{C}{\mathfrak{p}^2 |\hat{y}|^3}, \quad |\nabla \hat{u}_{\mathfrak{p}}(\hat{z}_{\mathfrak{p},\pm} + \hat{y})| \leqslant \frac{C}{|\hat{y}|}.$$

Proof. We work near $\hat{z}_{\mathfrak{p},+}$ (the minus sign is similar), say in the annulus $B\big(\hat{z}_{\mathfrak{p},+}, \frac{1}{10}\big) \setminus B(\hat{z}_{\mathfrak{p},+}, 1/\mathfrak{p})$ and set

$$\hat{u}_{\mathfrak{p}}(\hat{z}_{\mathfrak{p},+} + \hat{y}) = \hat{A}_{\mathfrak{p}}(\hat{y}) e^{i\theta + i\hat{\varphi}_{\mathfrak{p}}(\hat{y})},$$

with $\hat{A}_{\mathfrak{p}}$ and $\hat{\varphi}_{\mathfrak{p}}$ real-valued and smooth in the annulus (θ is the polar angle centered at $\hat{z}_{\mathfrak{p},+}$). Then, we obtain the system

$$\begin{cases} \Delta \hat{A}_{\mathfrak{p}} - \hat{A}_{\mathfrak{p}} |\nabla \hat{\varphi}_{\mathfrak{p}}|^2 + \mathfrak{p}^2 \hat{A}_{\mathfrak{p}} |V_1|^2 (1 - \hat{A}_{\mathfrak{p}}^2) - 2\hat{A}_{\mathfrak{p}} \dfrac{\partial_\theta \varphi}{r^2} - c_{\mathfrak{p}} \mathfrak{p} \hat{A}_{\mathfrak{p}} \partial_2 \hat{\varphi}_{\mathfrak{p}} - c_{\mathfrak{p}} \mathfrak{p} \dfrac{\cos\theta}{r} \hat{A}_{\mathfrak{p}} = 0, \\ \hat{A}_{\mathfrak{p}} \Delta \hat{\varphi}_{\mathfrak{p}} + 2\nabla \hat{A}_{\mathfrak{p}} \cdot \nabla \hat{\varphi}_{\mathfrak{p}} + 2 \dfrac{\partial_\theta \hat{A}_{\mathfrak{p}}}{r^2} + c_{\mathfrak{p}} \mathfrak{p} \partial_2 \hat{A}_{\mathfrak{p}} = 0. \end{cases}$$

The second equation may be recast as

$$\nabla \cdot (\hat{A}_{\mathfrak{p}}^2 \nabla \hat{\varphi}_{\mathfrak{p}}) + \frac{\partial_\theta \hat{A}_{\mathfrak{p}}^2}{r^2} = -\frac{c_{\mathfrak{p}} \mathfrak{p}}{2} \partial_2 (\hat{A}_{\mathfrak{p}}^2 - 1). \tag{3-20}$$

Multiplying by $\hat{\varphi}_{\mathfrak{p}}$ and integrating over $B\big(0, \frac{3}{20}\big) \setminus B(0, R_0/\mathfrak{p})$, we obtain

$$\begin{aligned} \int_{B(0,3/20)\setminus B(0,R_0/\mathfrak{p})} \hat{A}_{\mathfrak{p}}^2 |\nabla \hat{\varphi}_{\mathfrak{p}}|^2 \, d\hat{y} = \int_{B(0,3/20)\setminus B(0,R_0/\mathfrak{p})} (1 - \hat{A}_{\mathfrak{p}}^2) \frac{\partial_\theta \hat{\varphi}_{\mathfrak{p}}}{r^2} + \frac{c_{\mathfrak{p}} \mathfrak{p}}{2} (1 - \hat{A}_{\mathfrak{p}}^2) \partial_2 \hat{\varphi}_{\mathfrak{p}} \, d\hat{y} \\ + \int_{\partial B(0,3/20)} \hat{A}_{\mathfrak{p}}^2 \frac{\partial \hat{\varphi}_{\mathfrak{p}}}{\partial \nu} + \frac{c_{\mathfrak{p}} \mathfrak{p}}{2} (\hat{A}_{\mathfrak{p}}^2 - 1) \hat{\varphi}_{\mathfrak{p}} \nu_2 \, d\ell. \end{aligned}$$

By the Cauchy–Schwarz inequality, (3-3) and Step 1 of Section 3B3, we infer

$$\|\nabla \hat{\varphi}_{\mathfrak{p}}\|_{L^2(B(0,3/20)\setminus B(0,R_0/\mathfrak{p}))}^2 \leqslant C(1 + c_{\mathfrak{p}}) \|\nabla \hat{\varphi}_{\mathfrak{p}}\|_{L^2(B(0,3/20)\setminus B(0,R_0/\mathfrak{p}))} + C,$$

where, for the contribution of the integral over $\partial B\big(0, \frac{3}{20}\big)$, we have used (3-16) and (3-15) (see Step 2 of Section 3B2). This implies

$$\|\nabla \hat{\varphi}_{\mathfrak{p}}\|_{L^2(B(0,3/20)\setminus B(0,R_0/\mathfrak{p}))} \leqslant C. \tag{3-21}$$

We fix $\hat{y} \in \mathbb{R}^2$ such that $2R_0/\mathfrak{p} \leqslant |\hat{y}| \leqslant \frac{3}{20}$. Then, since $|\hat{u}_{\mathfrak{p}}| \geqslant \frac{1}{2}$ in the annulus $B\big(0, \frac{3}{20}\big) \setminus B(0, R_0/\mathfrak{p}) \supset B(\hat{y}, |\hat{y}|/2)$, we deduce

$$\int_{B(\hat{y},|\hat{y}|/2)} \hat{A}_{\mathfrak{p}}^2 |\nabla \hat{\varphi}_{\mathfrak{p}} + \vec{e}_\theta / r|^2 \, d\hat{x} \leqslant C \int_{B(\hat{y},|\hat{y}|/2)} |\nabla \hat{\varphi}_{\mathfrak{p}}|^2 + \frac{1}{r^2} \, d\hat{x} \leqslant C$$

by (3-21) and the fact that $r = |\hat{x}| \geqslant |\hat{y}|/2$. By Step 1 of Section 3B3, we then infer the upper bound (also shown in [35])

$$E_{1/\mathfrak{p}}(\hat{u}_{\mathfrak{p}}, B(\hat{y}, |\hat{y}|/2)) \leqslant C. \tag{3-22}$$

We now make some rescaling and consider

$$v(X) := \hat{u}_{\mathfrak{p}}\left(\hat{y} + \frac{|\hat{y}|}{2} X\right)$$

in $B(0, 1)$ (v depends on $\hat{y}$ and $\mathfrak{p}$), which solves

$$\Delta v + i\frac{c_{\mathfrak{p}}}{\delta}\partial_2 v + \frac{1}{\delta^2} v(1 - |v|^2) = 0$$

in $B(0, 1)$, with $\delta := 2/(\mathfrak{p}|\hat{y}|)$. This equation is of the type (3-1) with "$\epsilon = \delta$" and "$\mathfrak{c} = c_{\mathfrak{p}}/\delta$". Let us check that the assumption $|\mathfrak{c}| \leqslant M_0|\ln\epsilon|$ is satisfied with $M_0 = 1$. As a matter of fact, we have

$$\delta = \frac{2}{\mathfrak{p}|\hat{y}|} \in \left]\frac{40}{3\mathfrak{p}}, \frac{1}{2}\right];$$

thus

$$M_0\,\delta|\ln\delta| \geqslant \frac{40}{3\mathfrak{p}} \ln 2 \geqslant c_{\mathfrak{p}} = \frac{2\pi}{\mathfrak{p}} + o(1)$$

by Step 2 of Section 3B4 (note $40(\ln 2)/3 \approx 9.24(1) > 2\pi$). Furthermore, the upper bound (3-22) reads now

$$E_\delta(v, B(0, 1)) \leqslant C.$$

It then follows from the proof of Step 7 (p. 48) of Theorem 1 in [8] that, for δ sufficiently small,

$$\|2\delta^{-2}(1 - |v|) - c_{\mathfrak{p}}\delta^{-1}\partial_2 \arg(v)\|_{C^1(B(0,1/2))} \leqslant C, \quad \|\nabla \arg(v)\|_{C^1(B(0,1/2))} \leqslant C.$$

Therefore, by Step 2 of Section 3B3,

$$\big|1 - |v(0)|\big| + \big|\nabla|v|(0)\big| \leqslant Cc_{\mathfrak{p}}\delta + C\delta^2 \leqslant \frac{C}{\mathfrak{p}^2|\hat{y}|^2}, \quad |\nabla \arg(v)(0)| \leqslant C,$$

and scaling this back yields the conclusion, at least for $\delta = 2/(\mathfrak{p}|\hat{y}|)$ sufficiently small, say $\mathfrak{p}|\hat{y}| \geqslant \delta_0/2$, but the estimate is easy to show if $\mathfrak{p}|\hat{y}| \leqslant \delta_0/2$. □

3D. ***Some remarks on the nonsymmetrical case.*** In the case where we do not assume the x_1-symmetry for $u_{\mathfrak{p}}$, the location of the vortices $\hat{y}_{\mathfrak{p},\pm}$ is more delicate. Indeed, we can no longer assume (3-14), that is,

$$(\hat{y}_{\mathfrak{p},-})_2 = (\hat{y}_{\mathfrak{p},+})_2 = 0 \quad \text{and} \quad -(\hat{y}_{\mathfrak{p},-})_1 = (\hat{y}_{\mathfrak{p},+})_1 \to \tfrac{1}{2\pi}.$$

Up to a translation, we may assume $\hat{y}_{\mathfrak{p},+} + \hat{y}_{\mathfrak{p},-} = 0$, and it remains true that $\hat{y}_{\mathfrak{p},+,1} - \hat{y}_{\mathfrak{p},-,1} \to \frac{1}{\pi}$, but we may have $|\hat{y}_{\mathfrak{p},+} - \hat{y}_{\mathfrak{p},-}| \gg 1$. By carefully following the proof in [43], one could show that

$$|\hat{y}_{\mathfrak{p},+} - \hat{y}_{\mathfrak{p},-}| \leqslant C.$$

Then, the location of the limiting vortices $\hat{y}_{\infty,\pm} = \lim_{\mathfrak{p}\to+\infty} \hat{y}_{\mathfrak{p},\pm}$ can be obtained through the use of the Hopf differential as in [6] (Chapter VII), and would lead as before to $\hat{y}_{\infty,\pm} = \left(\pm\frac{1}{2\pi}, 0\right)$. This is of course related to the fact that the only critical point of the action functional

$$\mathcal{F}(\hat{y}_{\infty,+},\hat{y}_{\infty,-}) := 2\pi\big(2\ln|\hat{y}_{\infty,+}-\hat{y}_{\infty,-}| - 2\pi[(\hat{y}_{\infty,+})_1-(\hat{y}_{\infty,-})_1]\big)$$

associated with the action of the Kirchhoff energy is (up to translation) $(\hat{y}_{\infty,+},\hat{y}_{\infty,-}) = \big(\frac{1}{2\pi},-\frac{1}{2\pi}\big)\in\mathbb{C}^2$.

Next, Step 1 of Section 3B4 becomes, for any $p\in[1,2[$, and in $W^{1,p}_{\mathrm{loc}}(\mathbb{R}^2)$,

$$\hat{u}_{\mathfrak{p}} \rightharpoonup \mathrm{e}^{i\Theta}\hat{u}_\infty.$$

The term Θ is somewhat the phase at infinity, even though we do not claim some uniformity at infinity in space. Next, for the local convergences, there are two phases $\beta_\pm\in\mathbb{R}$ such that

$$\hat{u}_{\mathfrak{p}}(\hat{z}_{\mathfrak{p},\pm}+\mathfrak{p}\,\cdot\,) \to \mathrm{e}^{i\beta_\pm}V_\pm \tag{3-23}$$

in $C^k_{\mathrm{loc}}(\mathbb{R}^2)$ for any $k\in\mathbb{N}$. We are then simply able to show that $\beta_\pm=\Theta$, but this is not enough for the uniqueness result. This follows from the arguments given in [44], as we explain.

We work for the $+$ sign. Integrating (3-20) over the disk $B(0,R)$ yields

$$\int_{\partial B(0,R)} \hat{A}_{\mathfrak{p}}^2\frac{\partial\hat{\varphi}_{\mathfrak{p}}}{\partial\nu}\,d\ell + c_{\mathfrak{p}}\mathfrak{p}\int_{\partial B(0,R)}\nu_2(\hat{A}_{\mathfrak{p}}^2-1)\,d\ell = 0.$$

We now consider the average

$$\beta_{\mathfrak{p}}(r) := \frac{1}{2\pi r}\int_{\partial B(0,r)}\hat{\varphi}_{\mathfrak{p}}\,d\ell,$$

which satisfies, for $1/\mathfrak{p}\leqslant r_0\leqslant r_1\leqslant\frac{3}{20}$,

$$\begin{aligned}\beta_{\mathfrak{p}}(r_0)-\beta_{\mathfrak{p}}(r_1) &= \int_{r_0}^{r_1}\partial_r\beta_{\mathfrak{p}}(r)\,dr = \int_{r_0}^{r_1}\frac{1}{2\pi r}\int_{\partial B(0,r)}\partial_r\hat{\varphi}_{\mathfrak{p}}\,d\ell\,dr\\ &= \int_{r_0}^{r_1}\frac{1}{2\pi r}\int_{\partial B(0,r)}(1-\hat{A}_{\mathfrak{p}}^2)\partial_r\hat{\varphi}_{\mathfrak{p}}\,d\ell\,dr + c_{\mathfrak{p}}\mathfrak{p}\int_{r_0}^{r_1}\frac{1}{2\pi r}\int_{\partial B(0,r)}\nu_2(\hat{A}_{\mathfrak{p}}^2-1)\,d\ell\,dr.\end{aligned}$$

Therefore, by Step 5,

$$|\beta_{\mathfrak{p}}(r_0)-\beta_{\mathfrak{p}}(r_1)| \leqslant C\int_{r_0}^{r_1}\frac{dr}{\mathfrak{p}^2r^3} + C\int_{r_0}^{r_1}\frac{dr}{\mathfrak{p}^2r^2} \leqslant \frac{C}{(r_0\mathfrak{p})^2}+\frac{C}{\mathfrak{p}}.$$

We now fix $\eta\in\,]0,1]$. Taking $r_0=1/(\sqrt{\eta}\mathfrak{p})$ and $r_1=\frac{3}{20}$, we infer

$$|\beta_{\mathfrak{p}}(r_0)-\beta_{\mathfrak{p}}(r_1)| \leqslant C\eta+\frac{C}{\mathfrak{p}}.$$

Moreover, by (3-23), we have

$$\beta_{\mathfrak{p}}(r_0) = \beta_{\mathfrak{p}}(1/(\sqrt{\eta}\mathfrak{p})) \to \beta_+$$

as $\mathfrak{p}\to+\infty$, and by Step 1 of Section 3B4, we deduce

$$\beta_{\mathfrak{p}}(r_1)\to\Theta.$$

As a consequence,

$$|\beta_+-\Theta|\leqslant C\eta,$$

and the conclusion follows by letting $\eta\to 0$.

Acknowledgement

Pacherie is supported by Tamkeen under the NYU Abu Dhabi Research Institute grant CG002. We would like to thank the referee for a careful reading of the manuscript and for suggestions that have helped and clarified the presentation.

References

[1] M. Abid, C. Huepe, S. Metens, C. Nore, C. T. Pham, L. S. Tuckerman, and M. E. Brachet, "Gross–Pitaevskii dynamics of Bose–Einstein condensates and superfluid turbulence", *Fluid Dynam. Res.* **33**:5-6 (2003), 509–544. MR Zbl

[2] G. Alberti, S. Baldo, and G. Orlandi, "Variational convergence for functionals of Ginzburg–Landau type", *Indiana Univ. Math. J.* **54**:5 (2005), 1411–1472. MR Zbl

[3] J. Bellazzini and D. Ruiz, "Finite energy traveling waves for the Gross–Pitaevskii equation in the subsonic regime", *Amer. J. Math.* **145**:1 (2023), 109–149. MR Zbl

[4] F. Bethuel and T. Rivière, "Vortices for a variational problem related to superconductivity", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **12**:3 (1995), 243–303. MR Zbl

[5] F. Bethuel and J.-C. Saut, "Travelling waves for the Gross–Pitaevskii equation, I", *Ann. Inst. H. Poincaré Phys. Théor.* **70**:2 (1999), 147–238. MR Zbl

[6] F. Bethuel, H. Brezis, and F. Hélein, *Ginzburg–Landau vortices*, Progr. Nonlinear Differ. Eq. Appl. **13**, Birkhäuser, Boston, 1994. MR Zbl

[7] F. Bethuel, H. Brézis, and G. Orlandi, "Asymptotics for the Ginzburg–Landau equation in arbitrary dimensions", *J. Funct. Anal.* **186**:2 (2001), 432–520. Correction in **188**:2 (2002), 548–549. Zbl

[8] F. Bethuel, G. Orlandi, and D. Smets, "Vortex rings for the Gross–Pitaevskii equation", *J. Eur. Math. Soc.* **6**:1 (2004), 17–94. MR Zbl

[9] F. Béthuel, P. Gravejat, and J.-C. Saut, "Ondes progressives pour l'équation de Gross–Pitaevskii", pp. exposé XV in *Séminaire*: *Équations aux Dérivées Partielles* (Palaiseau, 2007–2008), École Polytech., Palaiseau, France, 2009. MR Zbl

[10] F. Béthuel, P. Gravejat, and J.-C. Saut, "Travelling waves for the Gross–Pitaevskii equation, II", *Comm. Math. Phys.* **285**:2 (2009), 567–651. MR Zbl

[11] H. Brezis, F. Merle, and T. Rivière, "Quantization effects for $-\Delta u = u(1-|u|^2)$ in $\mathbb{R}^2$", *Arch. Ration. Mech. Anal.* **126**:1 (1994), 35–58. MR Zbl

[12] X. Chen, C. M. Elliott, and T. Qi, "Shooting method for vortex solutions of a complex-valued Ginzburg–Landau equation", *Proc. Roy. Soc. Edinburgh Sect. A* **124**:6 (1994), 1075–1088. MR Zbl

[13] D. Chiron and M. Mariş, "Traveling waves for nonlinear Schrödinger equations with nonzero conditions at infinity", *Arch. Ration. Mech. Anal.* **226**:1 (2017), 143–242. MR Zbl

[14] D. Chiron and E. Pacherie, "Smooth branch of travelling waves for the Gross–Pitaevskii equation in $\mathbb{R}^2$ for small speed", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **22**:4 (2021), 1937–2038. MR Zbl

[15] D. Chiron and E. Pacherie, "Coercivity for travelling waves in the Gross–Pitaevskii equation in $\mathbb{R}^2$ for small speed", *Publ. Mat.* **67**:1 (2023), 277–410. MR Zbl

[16] D. Chiron and C. Scheid, "Travelling waves for the nonlinear Schrödinger equation with general nonlinearity in dimension two", *J. Nonlinear Sci.* **26**:1 (2016), 171–231. MR Zbl

[17] D. Chiron and C. Scheid, "Multiple branches of travelling waves for the Gross–Pitaevskii equation", *Nonlinearity* **31**:6 (2018), 2809–2853. MR Zbl

[18] M. Comte and P. Mironescu, "The behavior of a Ginzburg–Landau minimizer near its zeroes", *Calc. Var. Partial Differential Equations* **4**:4 (1996), 323–340. MR Zbl

[19] A. Farina, "From Ginzburg–Landau to Gross–Pitaevskii", *Monatsh. Math.* **139**:4 (2003), 265–269. MR Zbl

[20] C. Gallo, "The Cauchy problem for defocusing nonlinear Schrödinger equations with non-vanishing initial data at infinity", *Comm. Partial Differential Equations* **33**:4-6 (2008), 729–771. MR Zbl

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

[22] P. Gérard, "The Gross–Pitaevskii equation in the energy space", pp. 129–148 in *Stationary and time dependent Gross–Pitaevskii equations* (Vienna, 2006), edited by A. Farina and J.-C. Saut, Contemp. Math. **473**, Amer. Math. Soc., Providence, RI, 2008. MR Zbl

[23] V. L. Ginzburg and L. P. Pitaevskii, "On the theory of superfluidity", *Zh. Éksper. Teoret. Fiz.* **34** (1958), 1240–1245. In Russian; translated in *Soviet Phys. JETP* **34(7)**:5 (1958), 858–861.

[24] P. Gravejat, "Decay for travelling waves in the Gross–Pitaevskii equation", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **21**:5 (2004), 591–637. MR Zbl

[25] P. Gravejat, "Asymptotics for the travelling waves in the Gross–Pitaevskii equation", *Asymptot. Anal.* **45**:3-4 (2005), 227–299. MR Zbl

[26] R.-M. Hervé and M. Hervé, "Étude qualitative des solutions réelles d'une équation différentielle liée à l'équation de Ginzburg–Landau", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **11**:4 (1994), 427–440. MR Zbl

[27] R. L. Jerrard and H. M. Soner, "The Jacobian and the Ginzburg–Landau energy", *Calc. Var. Partial Differential Equations* **14**:2 (2002), 151–191. MR Zbl

[28] C. A. Jones and P. H. Roberts, "Motions in a Bose condensate, IV: Axisymmetric solitary waves", *J. Phys. A* **15**:8 (1982), 2599–2619.

[29] C. A. Jones, S. J. Putterman, and P. H. Roberts, "Motions in a Bose condensate, V: Stability of solitary wave solutions of nonlinear Schrödinger equations in two and three dimensions", *J. Phys. A* **19**:15 (1986), 2991–3011.

[30] Y. S. Kivshar and B. Luther-Davies, "Dark optical solitons: physics and applications", *Phys. Rep.* **298**:2-3 (1998), 81–197.

[31] F. Lin and T. Rivière, "Complex Ginzburg–Landau equations in high dimensions and codimension two area minimizing currents", *J. Eur. Math. Soc.* **1**:3 (1999), 237–311. MR Zbl

[32] Z. Lin and C. Zeng, *Instability, index theorem, and exponential trichotomy for linear Hamiltonian PDEs*, Mem. Amer. Math. Soc. **1347**, Amer. Math. Soc., Providence, RI, 2022. MR Zbl

[33] Y. Liu and J. Wei, "Multivortex traveling waves for the Gross–Pitaevskii equation and the Adler–Moser polynomials", *SIAM J. Math. Anal.* **52**:4 (2020), 3546–3579. MR Zbl

[34] M. Mariş, "Traveling waves for nonlinear Schrödinger equations with nonzero conditions at infinity", *Ann. of Math.* (2) **178**:1 (2013), 107–182. MR Zbl

[35] P. Mironescu, "Explicit bounds for solutions to a Ginzburg–Landau type equation", *Rev. Roumaine Math. Pures Appl.* **41**:3-4 (1996), 263–271. MR Zbl

[36] P. Mironescu, "Les minimiseurs locaux pour l'équation de Ginzburg–Landau sont à symétrie radiale", *C. R. Acad. Sci. Paris Sér. I Math.* **323**:6 (1996), 593–598. MR Zbl

[37] J. C. Neu, "Vortices in complex scalar fields", *Phys. D* **43**:2-3 (1990), 385–406. MR Zbl

[38] M. del Pino, P. Felmer, and M. Kowalczyk, "Minimality and nondegeneracy of degree-one Ginzburg–Landau vortex as a Hardy's type inequality", *Int. Math. Res. Not.* **2004**:30 (2004), 1511–1527. MR Zbl

[39] M. del Pino, M. Kowalczyk, and M. Musso, "Variational reduction for Ginzburg–Landau vortices", *J. Funct. Anal.* **239**:2 (2006), 497–541. MR Zbl

[40] L. M. Pismen, *Vortices in nonlinear fields*: *from liquid crystals to superfluids*, *from non-equilibrium patterns to cosmic strings*, Int. Ser. Monogr. Phys. **100**, Oxford Univ. Press, 1999. Zbl

[41] J. Qing, "Zeros of wave functions in Ginzburg–Landau model for small ϵ", *Commun. Contemp. Math.* **3**:2 (2001), 187–199. MR Zbl

[42] P. H. Roberts and N. G. Berloff, "The nonlinear Schrödinger equation as a model of superfluidity", pp. 235–257 in *Quantized vortex dynamics and superfluid turbulence*, edited by C. F. Barenghi et al., Lect. Notes in Phys. **571**, Springer, 2001. Zbl

[43] E. Sandier, "Lower bounds for the energy of unit vector fields and applications", *J. Funct. Anal.* **152**:2 (1998), 379–403. Correction in **171**:1 (2000), 233. MR Zbl

[44] I. Shafrir, "L^∞-approximation for minimizers of the Ginzburg–Landau functional", *C. R. Acad. Sci. Paris Sér. I Math.* **321**:6 (1995), 705–710. MR Zbl

[45] M. I. Weinstein, "On the vortex solutions of some nonlinear scalar field equations", *Rocky Mountain J. Math.* **21**:2 (1991), 821–827. MR Zbl

Received 16 Sep 2021. Revised 26 Feb 2022. Accepted 9 Apr 2022.

DAVID CHIRON: david.chiron@univ-cotedazur.fr
Université Côte d'Azur, CNRS, LJAD, Nice, France

ELIOT PACHERIE: ep2699@nyu.edu
NYUAD Research Institute, New York University, Abu Dhabi, United Arab Emirates

ANALYSIS AND PDE
Vol. 16 (2023), No. 9, pp. 2225–2240
DOI: 10.2140/apde.2023.16.2225

CLASSIFICATION OF CONVEX ANCIENT FREE-BOUNDARY CURVE-SHORTENING FLOWS IN THE DISC

THEODORA BOURNI AND MAT LANGFORD

Using a combination of direct geometric methods and an analysis of the linearization of the flow about the horizontal bisector, we prove that there exists a unique (modulo rotations about the origin) convex ancient curve-shortening flow in the disc with free boundary on the circle. This appears to be the first result of its kind in the free-boundary setting.

1. Introduction

Curve-shortening flow is the gradient flow of length for regular curves. It models the evolution of grain boundaries [Mullins 1956; von Neumann 1952] and the shapes of worn stones [Firey 1974] in two dimensions, and has been exploited in a multitude of further applications; see, for example, [Sapiro 2001].

The evolution of closed planar curves by curve-shortening was initiated by Mullins [1956] and was later taken up by Gage [1984] and Gage and Hamilton [1986], who proved that closed convex curves remain convex and shrink to "round" points in finite time. Soon after, Grayson showed that closed embedded planar curves become convex in finite time under the flow, thereafter shrinking to round points according to the Gage–Hamilton theorem. Different proofs of these results were discovered later by others [Andrews 2012; Andrews and Bryan 2011a; 2011b; Hamilton 1995b; Huisken 1998]. Ancient solutions to geometric flows (that is, solutions defined on backwards-infinite time-intervals) are important from an analytical standpoint as they model singularity formation [Hamilton 1995a]. They also arise in quantum field theory, where they model the ultraviolet regime in certain Dirichlet sigma models [Bakas and Sourdis 2007]. They have generated a great deal of interest from a purely geometric standpoint due to their symmetry and rigidity properties. For example, ancient solutions to curve-shortening flow of convex planar curves have been classified through the work of Daskalopoulos, Hamilton and Sesum [Daskalopoulos et al. 2010] and the authors in collaboration with Tinaglia [Bourni et al. 2020]. Bryan and Louie [2016] proved that the shrinking parallel is the only convex ancient solution to curve-shortening flow on the two-sphere, and Choi and Mantoulidis [2022] showed that it is the only embedded ancient solution on the two-sphere with uniformly bounded length.

The natural Neumann boundary value problem for curve-shortening flow, called the *free-boundary problem*, asks for a family of curves whose endpoints lie on (but are free to move on) a fixed barrier curve which is met by the solution curve orthogonally. Study of the free-boundary problem was initiated by Huisken [1989] and further developed by Stahl [1996a; 1996b]. In particular, Stahl proved that convex

MSC2020: 53E10.
Keywords: curve-shortening flow, free-boundary, ancient solutions.

curves with free boundary on a smooth, convex, locally uniformly convex barrier remain convex and shrink to a point on the barrier curve.

The analysis of ancient solutions to free-boundary curve-shortening flow remains in its infancy. Indeed, to our knowledge, the only examples previously known seem to be those inherited from closed or complete examples (one may restrict the shrinking circle, for example, to the upper halfplane). We provide here a classification of convex[1] ancient free-boundary curve-shortening flows in the disc.

Theorem 1.1. *Modulo rotation about the origin and translation in time, there exists exactly one convex, locally uniformly convex ancient solution to free-boundary curve-shortening flow in the unit disc $D \subset \mathbb{R}^2$. It converges to the point $(0, 1)$ as $t \to 0$ and smoothly to the segment $[-1, 1] \times \{0\}$ as $t \to -\infty$. It is invariant under reflection across the y-axis. As a graph over the x-axis, it satisfies*

$$e^{\lambda^2 t} y(x, t) \to A \cosh(\lambda x) \quad \textit{uniformly in } x \textit{ as } t \to -\infty$$

for some $A > 0$, where λ is the solution to $\lambda \tanh \lambda = 1$.

Theorem 1.1 is a consequence of Propositions 2.8, 3.4, and 3.5 proved below. Note that it is actually a classification of all convex ancient solutions, since the strong maximum principle and the Hopf boundary point lemma imply that any convex solution to the flow is either a stationary segment (and hence a bisector of the disc by the free-boundary condition) or is locally uniformly convex at interior times.

A higher-dimensional counterpart of Theorem 1.1 will be treated in a forthcoming paper.

Another natural setting in which to seek ancient solutions is within the class of *soliton* solutions. Since free-boundary curve-shortening flow in the disc is invariant under ambient rotations, one might expect to find rotating solutions. In Section 4, we provide a short proof that none exist.

Theorem 1.2. *There exist no proper rotating solutions to free-boundary curve-shortening flow in the disc.*

2. Existence

Our first goal is the explicit construction of a nontrivial ancient free-boundary curve-shortening flow in the disc. It will be clear from the construction that the solution is reflection-symmetric about the vertical axis, emerges at time negative infinity from the horizontal bisector, and converges at time zero to the point (0, 1). We shall also prove an estimate for the height of the constructed solution (which will be needed to prove its uniqueness).

2A. *Barriers.* Given $\theta \in \left(0, \frac{\pi}{2}\right)$, denote by C_θ the circle centered on the y-axis which meets ∂B^2 orthogonally at $(\cos\theta, \sin\theta)$. That is,

$$\mathrm{C}_\theta \doteqdot \{(x, y) \in \mathbb{R}^2 : x^2 + (\csc\theta - y)^2 = \cot^2\theta\}. \tag{1}$$

If we set

$$\theta^-(t) \doteqdot \arcsin e^t \quad \text{and} \quad \theta^+(t) \doteqdot \arcsin e^{2t},$$

[1] A free-boundary curve in the open disc B^2 is *convex* if it bounds a convex region in B^2 and *locally uniformly convex* if it is of class C^2 and its curvature is positive.

then $C_{\theta^\pm(t)}$ is defined for $t \in (-\infty, 0)$ and flows from the x-axis to the point $(0, 1)$. Moreover, since the inward normal speed of $\mathrm{C}_{\theta^-(t)}$ is no greater than its curvature κ^- and the inward normal speed of $\mathrm{C}_{\theta^+(t)}$ is no less than its curvature κ^+, the maximum principle and the Hopf boundary point lemma imply that:

Proposition 2.1. *A solution to free-boundary curve-shortening flow in B^2 which lies below* (*resp. above*) *the circle C_{θ_0} at time t_0 lies below $\mathrm{C}_{\theta^+(t_0^+ + t - t_0)}$* (*resp. above $\mathrm{C}_{\theta^-(t_0^- + t - t_0)}$*) *for all $t > t_0$, where $2t_0^+ = \log\sin\theta_0$* (*resp. $t_0^- = \log\sin\theta_0$*).

Consider now the shifted scaled Angenent oval $\{\mathrm{A}_t^\lambda\}_{t\in(-\infty,0)}$, where

$$\mathrm{A}_t^\lambda \doteqdot \left\{(x, y) \in \mathbb{R} \times \left(0, \frac{\pi}{2\lambda}\right) : \sin(\lambda y) = \mathrm{e}^{\lambda^2 t}\cosh(\lambda x)\right\}.$$

This evolves by curve-shortening flow, passes through the point $(\cos\theta, \sin\theta) \in \partial B^2$ at a time t given by

$$t = \lambda^{-2}\log\left(\frac{\sin(\lambda\sin\theta)}{\cosh(\lambda\cos\theta)}\right)$$

and at that point, the normal satisfies

$$\nu_\lambda(\cos\theta, \sin\theta) \cdot (\cos\theta, \sin\theta) = \frac{\cos\theta\tanh(\lambda\cos\theta) - \sin\theta\cot(\lambda\sin\theta)}{\sqrt{\tanh^2(\lambda\cos\theta) + \cot^2(\lambda\sin\theta)}}.$$

Lemma 2.2. *For each $\theta \in \left(0, \frac{\pi}{2}\right)$, there is a unique $\lambda(\theta) \in (0, \pi/(2\sin\theta))$ such that*

$$\nu_{\lambda(\theta)}(\cos\theta, \sin\theta) \cdot (\cos\theta, \sin\theta) = 0.$$

Given $\theta, \theta_0 \in \left(0, \frac{\pi}{2}\right)$ with $\theta > \theta_0$,

$$\nu_{\lambda(\theta_0)}(\cos\theta, \sin\theta) \cdot (\cos\theta, \sin\theta) < 0.$$

Proof. Define

$$f(\lambda, \theta) \doteqdot \cos\theta\tanh(\lambda\cos\theta) - \sin\theta\cot(\lambda\sin\theta)).$$

Observe that

$$\lim_{\lambda\searrow 0} f(\lambda, \theta) = -\infty, \qquad \lim_{\lambda\nearrow\pi/(2\sin\theta)} f(\lambda, \theta) = \cos\theta\tanh\left(\tfrac{\pi}{2}\cot\theta\right) > 0$$

and

$$\frac{\partial f}{\partial\lambda} = \cos^2\theta(1 - \tanh^2(\lambda\cos\theta)) + \sin^2\theta(1 + \cot^2(\lambda\sin\theta)) > 0. \tag{2}$$

The first claim follows.

Next observe that

$$\frac{\partial f}{\partial\theta} = -\sin\theta\tanh(\lambda\cos\theta) - \lambda\cos\theta\sin\theta\operatorname{sech}^2(\lambda\cos\theta) - \cos\theta\cot(\lambda\sin\theta) + \lambda\sin\theta\cos\theta\csc^2(\lambda\sin\theta).$$

Given $\theta \in \left(0, \frac{\pi}{2}\right)$, we obtain, at the unique zero $\lambda \in (0, \pi/(2\sin\theta))$ of $f(\,\cdot\,, \theta)$,

$$\begin{aligned}\frac{\partial f}{\partial\theta} &= -\sin\theta\tan\theta\cot(\lambda\sin\theta) - \lambda\cos\theta\sin\theta(1 - \tan^2\theta\cot^2(\lambda\sin\theta)) \\ &\qquad\qquad - \cos\theta\cot(\lambda\sin\theta) + \lambda\sin\theta\cos\theta\csc^2(\lambda\sin\theta) \\ &= -\sec\theta\cot(\lambda\sin\theta)(1 - \lambda\sin\theta\cot(\lambda\sin\theta)).\end{aligned}$$

Since $Y\cot Y < 1$ for $Y \in \left(0, \frac{\pi}{2}\right)$, this is less than zero. The second claim follows. $\square$

The maximum principle and the Hopf boundary point lemma now imply the following.

Proposition 2.3. *Let $\{\Gamma_t\}_{t\in[\alpha,\omega)}$ be a solution to free-boundary curve-shortening flow in B^2. Suppose that $\lambda \le \lambda(\theta_\alpha)$, where θ_α denotes the smaller, in absolute value, of the two turning angles to Γ_α at its boundary. If Γ_α lies above A^{λ}_{s}, then Γ_t lies above $\mathrm{A}^{\lambda}_{s+t-\alpha}$ for all $t \in (\alpha, \omega)\cap(-\infty, \alpha - s)$.*

Proof. By the strong maximum principle, the two families of curves can never develop contact at an interior point. Since the families are monotonic, they cannot develop boundary contact at a boundary point $(\cos\theta, \sin\theta)$ with $|\theta| \le \theta_\alpha$. On the other hand, since $\lambda \le \lambda(\theta_\alpha)$, (2) implies that

$$f(\lambda, \theta_\alpha) \le f(\lambda_\alpha, \theta_\alpha) = 0,$$

and hence, by the argument of Lemma 2.2,

$$f(\lambda, \theta) \le 0 \quad \text{for } \theta \ge \theta_\alpha.$$

So the Hopf boundary point lemma implies that no boundary contact can develop for $\theta \ge \theta_\alpha$ either. $\square$

Remark 2.4. Since $s\cot s \to 1$ as $s \to 0$, we have that $f(\lambda, \theta)$ is nonnegative at $\theta = 0$ so long as $\lambda \ge \lambda_0$, where $\lambda_0 \tanh\lambda_0 = 1$.

2B. *Old-but-not-ancient solutions.* For each $\rho > 0$, choose a curve Γ^ρ in $\bar{B}^2$ with the following properties:

- Γ^ρ meets ∂B^2 orthogonally at $(\cos\rho, \sin\rho)$.
- Γ^ρ is reflection-symmetric about the y-axis.
- $\Gamma^\rho \cap B^2$ is the relative boundary of a convex region $\Omega^\rho \subset B^2$.
- $\kappa^\rho_s > 0$ in $B^2 \cap \{x > 0\}$.

For example, we could take $\Gamma^\rho \doteqdot \mathrm{A}^{\lambda_\rho}_{t_\rho} \cap B^2$, where $\lambda_\rho > \lambda_0$ and t_ρ are (uniquely) chosen so that

$$\cos\rho \tanh(\lambda_\rho \cos\rho) - \sin\rho \cot(\lambda_\rho \sin\rho)) = 0$$

and

$$-t_\rho = \lambda_\rho^{-2} \log\left(\frac{\cosh(\lambda_\rho \cos\rho)}{\sin(\lambda_\rho \sin\rho)}\right).$$

Observe that the circle C_{θ_ρ} defined by

$$\sin\theta_\rho = \frac{2\sin\rho}{1+\sin^2\rho}$$

is tangent to the line $y = \sin\rho$, and hence lies above Γ^ρ.

Work of Stahl [1996b; 1996a] now yields the following *old-but-not-ancient solutions*.

Lemma 2.5. *For each $\rho \in \left(0, \frac{\pi}{2}\right)$, there exists a smooth solution*[2] *$\{\Gamma^\rho_t\}_{t\in[\alpha_\rho,0)}$ to curve-shortening flow with $\Gamma^\rho_{\alpha_\rho} = \Gamma^\rho$ which satisfies the following properties*:

- *Γ^ρ_t meets ∂B^2 orthogonally for each $t \in (\alpha_\rho, 0)$.*

[2]Given by a one parameter family of immersions $X : [-1, 1] \times [\alpha_\rho, 0) \to \bar{B}^2$ satisfying $X \in C^\infty([-1, 1] \times (\alpha_\rho, 0)) \cap C^{2+\beta,1+\beta/2}([-1, 1] \times [\alpha_\rho, 0))$ for some $\beta \in (0, 1)$.

- Γ_t^ρ *is convex and locally uniformly convex for each* $t \in (\alpha_\rho, 0)$.
- Γ_t^ρ *is reflection-symmetric about the* y*-axis for each* $t \in (\alpha_\rho, 0)$.
- $\Gamma_t^\rho \to (0, 1)$ *uniformly as* $t \to 0$.
- $\kappa_s^\rho > 0$ *in* $B^2 \cap \{x > 0\}$.
- $\alpha_\rho < \frac{1}{2}\log(2\sin\rho/(1+\sin^2\rho)) \to -\infty$ *as* $\rho \to 0$.

Proof. Existence of a maximal solution to curve-shortening flow out of Γ^ρ which meets ∂B^2 orthogonally was proved by Stahl [Stahl 1996b, Theorem 2.1]. Stahl [1996a, Proposition 1.4] also proved that this solution remains convex and locally uniformly convex and shrinks to a point on the boundary of B^2 at the final time (which is finite). We obtain $\{\Gamma_t^\rho\}_{t\in[\alpha_\rho,0)}$ by time-translating Stahl's solution.

By uniqueness of solutions Γ_t^ρ remains reflection-symmetric about the y-axis for $t \in (\alpha_\rho, 0)$, so the final point is $(0, 1)$.

The reflection symmetry also implies that $\kappa_s^\rho = 0$ at the point $p_t \doteqdot \Gamma_t^\rho \cap \{x = 0\}$ for all $t \in [\alpha_\rho, 0)$. By [Stahl 1996a, Proposition 2.1], $\kappa_s^\rho = \kappa^\rho > 0$ at the boundary point $q_t \doteqdot \partial\Gamma_t^\rho \cap \{x > 0\}$ for all $t \in (\alpha_\rho, 0)$. Applying Sturm's theorem [Angenent 1988] to κ_s^ρ, we thus find that $\kappa_s^\rho > 0$ on $\Gamma_t^\rho \cap B^2 \cap \{x > 0\}$ for all $t \in (\alpha_\rho, 0)$.

Since $\mathrm{C}_{\theta_\rho} \subset \Omega^\rho$, the final property follows from Proposition 2.1. □

We now fix $\rho > 0$ and drop the super/subscript ρ. Set

$$\underline{\kappa}(t) \doteqdot \min_{\Gamma_t} \kappa = \kappa(p_t) \quad \text{and} \quad \bar{\kappa}(t) \doteqdot \max_{\Gamma_t} \kappa = \kappa(q_t),$$

and define $\underline{y}(t)$, $\bar{y}(t)$ and $\bar{\theta}(t)$ by

$$p_t = (0, \underline{y}(t)), \quad q_t = (\cos\bar{\theta}(t), \sin\bar{\theta}(t)), \quad \text{and} \quad \bar{y}(t) = \sin\bar{\theta}(t).$$

Lemma 2.6. *Each old-but-not-ancient solution satisfies*

$$\underline{\kappa} \le \tan\bar{\theta} \le \bar{\kappa}, \tag{3}$$

$$\sin\bar{\theta} \le e^t, \tag{4}$$

$$\frac{\sin\bar{\theta}}{1+\cos\bar{\theta}} \le \underline{y} \le \sin\bar{\theta}. \tag{5}$$

Proof. To prove the lower bound for $\bar{\kappa}$, it suffices to show that the circle $\mathrm{C}_{\bar{\theta}(t)}$ (see (1)) lies locally below Γ_t near q_t. If this is not the case, then, locally around q_t, Γ_t lies below $\mathrm{C}_{\bar{\theta}(t)}$ and hence $\kappa(q_t) \le \tan\bar{\theta}(t)$. But then we can translate $\mathrm{C}_{\bar{\theta}(t)}$ downwards until it touches Γ_t from below in an interior point at which the curvature must satisfy $\kappa \ge \tan\bar{\theta}(t)$. This contradicts the unique maximization of the curvature at q_t.

The estimate (4) now follows by integrating the inequality

$$\frac{d}{dt}\sin\bar{\theta} = \cos\bar{\theta}\,\bar{\kappa} \ge \sin\bar{\theta}$$

between any initial time t and the final time 0 (at which $\bar{\theta} = \frac{\pi}{2}$ since the solution contracts to the point $(0, 1)$).

The upper bound for $\underline{y}$ follows from convexity and the boundary condition $\bar{y} = \sin\bar{\theta}$. To prove the lower bound, we will show that the circle $\mathrm{C}_{\bar{\theta}(t)}$ lies nowhere above Γ_t. Suppose that this is not the case. Then, since $\mathrm{C}_{\bar{\theta}(t)}$ lies locally below Γ_t near q_t, we can move $\mathrm{C}_{\bar{\theta}(t)}$ downwards until it is tangent from below to a point p_t' on $\Gamma_t \cap \{x \geq 0\}$, at which we must have $\kappa \geq \tan\bar{\theta}(t)$. But then, since $\kappa_s \geq 0$ in $\{x > 0\}$, we find that $\kappa \geq \tan\bar{\theta}(t)$ for all points between p_t' and q_t. But this implies that this whole arc (including p_t') lies above $\mathrm{C}_{\bar{\theta}(t)}$, a contradiction. To prove the upper bound for $\underline{\kappa}$, fix t and consider the circle C centered on the y-axis through the points p_t and q_t. Its radius is $r(t)$, where

$$r \doteq \frac{\cos^2\bar{\theta} + (\sin\bar{\theta} - \underline{y})^2}{2(\sin\bar{\theta} - \underline{y})}.$$

We claim that Γ_t lies locally below C near p_t. Suppose that this is not the case. Then, by the symmetry of Γ_t and C across the y-axis, Γ_t lies locally above C near p_t. This implies two things: first, that

$$\kappa(p_t) \geq r^{-1},$$

and second, that, by moving C vertically upwards, we can find a point p_t' (the final point of contact) which satisfies

$$\kappa(p_t') \leq r^{-1}.$$

These two inequalities contradict the (unique) minimization of κ at p_t. We conclude that

$$\underline{\kappa} \leq \frac{2(\sin\bar{\theta} - \underline{y})}{\cos^2\bar{\theta} + (\sin\bar{\theta} - \underline{y})^2} \leq \tan\bar{\theta}$$

due to the lower bound for $\underline{y}$. □

Remark 2.7. If we parametrize by turning angle $\theta \in [-\bar{\theta}, \bar{\theta}]$, so that

$$\tau = (\cos\theta, \sin\theta),$$

then the estimates (3) are also easily obtained from the monotonicity of κ and the formulas

$$x(\theta) = x_0 + \int_0^\theta \frac{\cos u}{\kappa(u)}\,du \quad \text{and} \quad y(\theta) = y_0 + \int_0^\theta \frac{\sin u}{\kappa(u)}\,du. \tag{6}$$

2C. *Taking the limit.*

Proposition 2.8. *There exists a convex, locally uniformly convex ancient curve-shortening flow in the disc with free boundary on the circle.*

Proof. For each $\rho > 0$, consider the old-but-not-ancient solution $\{\Gamma_t^\rho\}_{t\in[\alpha_\rho,0)}$, $\Gamma_t^\rho = \partial\Omega_t^\rho$, constructed in Lemma 2.5. By (4), Ω_t^ρ contains $\mathrm{C}_{\omega(t)} \cap B^2$, where $\omega(t) \in \left(0, \frac{\pi}{2}\right)$ is uniquely defined by

$$\frac{1 - \cos\omega(t)}{\sin\omega(t)} = \mathrm{e}^t.$$

If we represent Γ_t^ρ as a graph $x \mapsto y^\rho(x,t)$ over the x-axis, then convexity and the boundary condition imply that $|y_x^\rho| \leq \tan\omega$. Since $\omega(t)$ is independent of ρ, the (global-in-space, interior-in-time) Ecker–Huisken-type estimates in [Stahl 1996b] imply uniform-in-ρ bounds for the curvature and its derivatives.

So the limit

$$\{\Gamma_t^\rho\}_{t\in[\alpha_\rho,0)} \to \{\Gamma_t\}_{t\in(-\infty,0)}$$

exists in C^∞ (globally in space on compact subsets of time) and the limit $\{\Gamma_t\}_{t\in(-\infty,0)}$ satisfies curve-shortening flow with free boundary in B^2. On the other hand, since $\{\Gamma_t^\rho\}_{t\in(\alpha_\rho,0)}$ contracts to $(0, 1)$ as $t \to 0$, (the contrapositive of) Proposition 2.1 implies that Γ_t^ρ must intersect the closed region enclosed by $\mathrm{C}_{\theta^+(t)}$ for all $t < 0$. It follows that Γ_t must intersect the closed region enclosed by $\mathrm{C}_{\theta^+(t)}$ for all $t < 0$. Since each Γ_t is the limit of convex boundaries, each is convex. It follows that Γ_t converges to $(0, 1)$ as $t \to 0$ and, by [Stahl 1996b, Corollary 4.5], that Γ_t is locally uniformly convex for each t. □

2D. ***Asymptotics for the height.*** For the purposes of this section, we fix an ancient solution $\{\Gamma_t\}_{(-\infty,0)}$ obtained as in Proposition 2.8 by taking a sublimit as $\lambda \searrow \lambda_0$ of the specific old-but-not ancient solutions $\{\Gamma_t^\lambda\}_{t\in[\alpha_\lambda,0)}$ corresponding to $\Gamma^\lambda_{\alpha_\lambda} = \mathrm{A}^\lambda_{t_\lambda} \cap B^2$, t_λ being the time at which $\{\mathrm{A}_t^\lambda\}_{t\in(-\infty,0)}$ meets ∂B^2 orthogonally. The asymptotics we obtain for this solution will be used to prove its uniqueness.

We will need to prove that the limit $\lim_{t\to-\infty} \mathrm{e}^{-\lambda_0^2 t}\underline{y}(t)$ exists in $(0, \infty)$. The following speed bound will imply that it exists in $[0, \infty)$.

Lemma 2.9. *The ancient solution $\{\Gamma_t\}_{(-\infty,0)}$ satisfies*

$$\frac{\kappa}{\cos\theta} \geq \lambda_0 \tan(\lambda_0 y). \tag{7}$$

Proof. It suffices to prove that $\kappa/\cos\theta \geq \lambda\tan(\lambda y)$ on each of the old-but-not-ancient solutions $\{\Gamma_t^\lambda\}_{t\in[\alpha_\lambda,0)}$. Note that equality holds on the initial timeslice $\Gamma^\lambda_{\alpha_\lambda} = \mathrm{A}^\lambda_{t_\lambda}$.

Given any $\mu < \lambda$, set $u \doteqdot \mu\tan(\mu y)$ and $v \doteqdot x_s = \cos\theta = \langle \nu, e_2\rangle$. Observe that

$$u_s = \mu^2 \sec^2(\mu y)\sin\theta, \quad (\partial_t - \Delta)u = -2\mu^2\sec^2(\mu y)\sin^2\theta u,$$
$$v_s = -\kappa\sin\theta \quad \text{and} \quad (\partial_t - \Delta)v = \kappa^2 v.$$

At an interior maximum of uv/κ we observe that

$$\frac{\nabla\kappa}{\kappa} = \frac{\nabla u}{u} + \frac{\nabla v}{v}$$

and hence

$$0 \leq (\partial_t - \Delta)\frac{uv}{\kappa} = \frac{uv}{\kappa}\left(\frac{(\partial_t - \Delta)u}{u} - 2\left\langle\frac{\nabla u}{u}, \frac{\nabla v}{v}\right\rangle\right) = 2\mu^2\sec^2(\mu y)\sin^2\theta\left(1 - \frac{uv}{\kappa}\right). \tag{8}$$

At a (without loss of generality right) boundary maximum of uv/κ, we have $y_s = y$ and $\kappa_s = \kappa$, and hence

$$\begin{aligned}\left(\frac{uv}{\kappa}\right)_s &= \frac{uv}{\kappa}\left(\frac{u_s}{u} + \frac{v_s}{v} - \frac{\kappa_s}{\kappa}\right) = \frac{uv}{\kappa}\left(\frac{\sec^2(\mu y)\mu y}{\tan\mu y} - \kappa\frac{y}{v} - 1\right)\\ &= \left(\frac{\mu y}{\tan(\mu y)} - 1\right)\frac{uv}{\kappa} + \left(\frac{uv}{\kappa} - 1\right)\tan(\mu y)\mu y\\ &\leq \left(\frac{uv}{\kappa} - 1\right)\tan(\mu y)\mu y.\end{aligned} \tag{9}$$

We may now conclude that $\max_{\overline{\Gamma}_t^\lambda} uv/\kappa$ remains less than 1. Indeed, if uv/κ ever reaches 1, then there must be a first time $t_0 > 0$ and a point $x_0 \in \overline{\Gamma}_t$ at which this occurs (note that uv/κ is continuous on $\overline{\Gamma}_t$ up to the initial time). The point x_0 cannot be an interior point, due to (8), and it cannot be a boundary point, due to (9) and the Hopf boundary point lemma. We conclude that

$$\frac{\kappa}{\cos\theta} \geq \mu\tan(\mu y)$$

on $\{\Gamma_t^\lambda\}_{t\in[\alpha_\lambda,0)}$ for all $\mu < \lambda$. Now take $\mu \to \lambda$. □

If we parametrize Γ_t as a graph $x \mapsto y(x,t)$ over the x-axis, then (7) yields

$$(\sin(\lambda_0 y))_t = \lambda_0\cos(\lambda_0 y)\kappa\sqrt{1+|y_x|^2} = \lambda_0\cos(\lambda_0 y)\frac{\kappa}{\cos\theta} \geq \lambda_0^2\sin(\lambda_0 y)$$

and hence

$$(\mathrm{e}^{-\lambda_0^2 t}\sin(\lambda_0 y(x,t)))_t \geq 0. \tag{10}$$

In particular, the limit

$$A(x) \doteqdot \lim_{t\to-\infty} \mathrm{e}^{-\lambda_0^2 t} y(x,t)$$

exists in $[0,\infty)$ for each $x \in (-1,1)$, as claimed.

We want next to prove that the above limit is positive. We will achieve this through a suitable upper bound for the speed. Recall that

$$(\partial_t - \Delta)\kappa_s = 4\kappa^2\kappa_s \quad \text{and} \quad (\partial_t - \Delta)\langle\gamma,\nu\rangle = \kappa^2\langle\gamma,\nu\rangle - 2\kappa, \tag{11}$$

where γ denotes the position and s is an arc-length parameter. The good -2κ-term in the second equation may be exploited to obtain the following crude speed bound.

Lemma 2.10. *There exist $T > -\infty$ and $C < \infty$ such that*

$$\bar{\kappa} \leq Ce^t \quad \textit{for all } t < T. \tag{12}$$

Proof. We will prove the estimate for each old-but-not-ancient solution $\{\Gamma_t^\lambda\}_{t\in(\alpha_\lambda,0)}$. We first prove a crude gradient estimate of the form

$$|\kappa_s| \leq 2\kappa \tag{13}$$

for t sufficiently negative. It will suffice to prove that

$$|\kappa_s| - \kappa + \langle\gamma,\nu\rangle \leq 0, \tag{14}$$

where γ denotes the position. Indeed, since $\langle\gamma,\nu\rangle_s = \kappa\langle\gamma,\tau\rangle$ has the same sign as the x-coordinate, we may estimate, as in (7),

$$|\langle\gamma,\nu\rangle| \leq |\langle\gamma,\nu\rangle|_{x=0} \leq \lambda^{-2}\kappa|_{x=0} = \lambda^{-2}\min_{\Gamma_t}\kappa \leq \kappa. \tag{15}$$

For λ sufficiently close to λ_0, we have $\kappa|_{t=\alpha_\lambda} < \frac{1}{2}$. Denote by T^λ the first time at which κ reaches $\frac{1}{2}$. Since κ is continuous up to the initial time α_λ, we have $T^\lambda > \alpha_\lambda$. We claim that (14) holds for $t < T^\lambda$.

Indeed, it is satisfied on the initial timeslice $\Gamma^\lambda_{\alpha_\lambda} = \mathrm{A}^\lambda_{t_\lambda}$ since

$$\kappa_s^2 - \kappa^2 = \lambda^2(\cos^2\theta\sin^2\theta - \sin^2\theta - a_\lambda^2) = -\lambda^2(\sin^4\theta + a_\lambda^2) \le 0,$$

whereas $\langle\gamma, \nu\rangle \le 0$. We will show that

$$f_\varepsilon \doteqdot |\kappa_s| - \kappa + \langle\gamma, \nu\rangle - \varepsilon \mathrm{e}^{t-\alpha_\lambda}$$

remains negative up to time T^λ. Suppose, to the contrary, that f_ε reaches zero at some time $t < T^\lambda$ at some point $p \in \overline{\Gamma}^\lambda_t$. Since $|\kappa_s| - \kappa + \langle\gamma, \nu\rangle$ vanishes at the boundary, p must be an interior point. Since κ_s vanishes at the y-axis, and the curve is symmetric, we may assume that $x(p) > 0$. At such a point, using the evolution equations (11), we have

$$\begin{aligned} 0 \le (\partial_t - \Delta) f_\varepsilon &= \kappa^2(4\kappa_s - \kappa + \langle\gamma, \nu\rangle) - 2\kappa - \varepsilon\mathrm{e}^{t-\alpha_\lambda} \\ &= \kappa^2(3[\kappa - \langle\gamma, \nu\rangle] + 4\varepsilon\mathrm{e}^{t-\alpha_\lambda}) - 2\kappa - \varepsilon\mathrm{e}^{t-\alpha_\lambda}. \end{aligned}$$

Recalling (15) and estimating $\kappa \le \frac{1}{2}$ yields

$$0 \le 6\kappa^3 - 2\kappa + (4\kappa^2 - 1)\varepsilon\mathrm{e}^{t-\alpha_\lambda} < 0,$$

which is absurd. So f_ε does indeed remain negative, and taking $\varepsilon \to 0$ yields (13) for $t < T^\lambda$.

Since $\mathrm{Length}(\Gamma^\lambda_t \cap \{x \ge 0\}) \le 1$, integrating (13) yields

$$\bar{\kappa} \le \mathrm{e}^2 \underline{\kappa} \quad \text{for } t < T^\lambda.$$

Recalling (3) and (4), this implies that

$$\bar{\kappa} \le \mathrm{e}^2 \frac{\mathrm{e}^t}{\sqrt{1 - \mathrm{e}^{2t}}} \quad \text{for } t < T^\lambda.$$

Taking $t = T^\lambda$ we find that $T^\lambda \ge T$, where T is independent of λ, so we conclude that

$$\bar{\kappa} \le C\mathrm{e}^t \quad \text{for } t < T,$$

where C and T do not depend on λ. □

We now bootstrap (12) to obtain the desired speed bound.

Lemma 2.11. *There exist $C < \infty$ and $T > -\infty$ such that*

$$\frac{\kappa}{y} \le \lambda_0^2 + C\mathrm{e}^{2t} \quad \textit{for } t < T.$$

Proof. Consider the old-but-not-ancient solutions $\{\Gamma^\lambda_t\}_{t\in(-\infty,0)}$. By (12), we can find $C < \infty$ and $T > -\infty$ such that

$$(\partial_t - \Delta)\frac{\kappa}{y} = \kappa^2\frac{\kappa}{y} + 2\left\langle \nabla\frac{\kappa}{y}, \frac{\nabla y}{y}\right\rangle \le C\mathrm{e}^{2t}\frac{\kappa}{y} + 2\left\langle \nabla\frac{\kappa}{y}, \frac{\nabla y}{y}\right\rangle \quad \text{for } t < T.$$

Since, at a boundary point,

$$\left(\frac{\kappa}{y}\right)_s = \frac{\kappa_s}{y} - \frac{\kappa}{y}\frac{y_s}{y} = 0,$$

the Hopf boundary point lemma and the ODE comparison principle yield

$$\max_{\Gamma_t^\lambda} \frac{\kappa}{y} \leq C \max_{\Gamma_{\alpha_\lambda}^\lambda} \frac{\kappa}{y} \quad \text{for } t \in (\alpha_\lambda, T).$$

But now

$$(\partial_t - \Delta)\frac{\kappa}{y} \leq C\mathrm{e}^{2t} \max_{\Gamma_{\alpha_\lambda}^\lambda} \frac{\kappa}{y} + 2\left\langle \nabla \frac{\kappa}{y}, \frac{\nabla y}{y} \right\rangle \quad \text{for } t < T,$$

and hence, by ODE comparison,

$$\max_{\Gamma_t^\lambda} \frac{\kappa}{y} \leq \max_{\Gamma_{\alpha_\lambda}^\lambda} \frac{\kappa}{y}(1 + C\mathrm{e}^{2t}) \quad \text{for } t \in (\alpha_\lambda, T).$$

Since, on the initial timeslice $\Gamma_{\alpha_\lambda}^\lambda = \mathrm{A}_{t_\lambda}^\lambda$,

$$\frac{\kappa}{y} = \frac{\lambda \tan(\lambda y)}{y} \cos\theta,$$

the claim follows upon taking $\lambda \to \lambda_0$. □

It follows that

$$(\log \underline{y}(t) - \lambda_0^2 t)_t \leq C\mathrm{e}^{2t} \quad \text{for } t < T,$$

and hence, integrating from time t up to time T,

$$\log \underline{y}(t) - \lambda_0^2 t \geq \log \underline{y}(T) - \lambda_0^2 T - C \quad \text{for } t < T.$$

So we indeed find that:

Lemma 2.12. *The limit*

$$A \doteqdot \lim_{t \to -\infty} \mathrm{e}^{-\lambda_0^2 t} \underline{y}(t) \tag{16}$$

exists in $(0, \infty)$ *on the particular ancient solution* $\{\Gamma_t\}_{(-\infty,0)}$.

3. Uniqueness

Now let $\{\Gamma_t\}_{t\in(-\infty,0)}$, $\Gamma_t = \partial_{\mathrm{rel}}\Omega_t$, be *any* convex, locally uniformly convex ancient free-boundary curve-shortening flow in the disc. By Stahl's theorem [1996a], we may assume that Γ_t contracts to a point on the boundary as $t \to 0$.

3A. *Backwards convergence.* We first show that $\overline{\Gamma}_t$ converges to a bisector as $t \to -\infty$.

Lemma 3.1. *Up to a rotation of the plane*,

$$\overline{\Gamma}_t \xrightarrow[C^\infty]{} [-1, 1] \times \{0\} \quad \text{as } t \to -\infty.$$

Proof. Set $A(t) \doteqdot \mathrm{area}(\Omega_t)$. Integrating the variational formula for area yields

$$A(t) = \int_t^0 \int_{\Gamma_t} d\theta,$$

where θ is the turning angle. Since convexity ensures that the total turning angle $\int_{\Gamma_t} d\theta$ is increasing and $A(t) \le \pi$ for all t, we find that

$$\int_{\Gamma_t} d\theta \to 0 \quad \text{as } t \to -\infty.$$

Monotonicity of the flow, the free-boundary condition and convexity now imply that the enclosed regions Ω_t satisfy

$$\overline{\Omega}_t \to B^2 \cap \{y \ge 0\} \quad \text{as } t \to -\infty$$

in the Hausdorff topology.

If we now represent Γ_t graphically over the x-axis, then convexity and the boundary condition ensure that the height and gradient are bounded by the height at the boundary. Stahl's estimates [1996b] now give bounds for κ and its derivatives up to the boundary depending only on the height at the boundary. We then get smooth subsequential convergence along any sequence of times $t_j \to -\infty$. The claim follows since any sublimit is the horizontal segment. □

We henceforth assume, without loss of generality, that the backwards limit is the horizontal bisector.

3B. *Reflection symmetry.* We can now prove that the solution is reflection-symmetric using Alexandrov reflection across lines through the origin; see [Chow and Gulliver 2001].

Lemma 3.2. *Γ_t is reflection-symmetric about the y-axis for all t.*

Proof. Given any $\omega \in \left(0, \frac{\pi}{2}\right)$, we define the halfspace

$$H_\omega = \{(x, y) : (x, y) \cdot (-\sin\omega, \cos\omega) > 0\}$$

and denote by R_ω the reflection about ∂H_ω. We first claim that, for every ω, there exists $t = t_\omega$ such that

$$(R_\omega \cdot \Gamma_t) \cap (\Gamma_t \cap H_\omega) = \varnothing \quad \text{for all } t < t_\omega. \tag{17}$$

Assume that the claim is not true. Then there exists $\omega \in \left(0, \frac{\pi}{2}\right)$, a sequence of times $t_i \to -\infty$, and a sequence of pairs of points $p_i, q_i \in \Gamma_{t_i}$ such that $R_\omega(p_i) = q_i$. This implies that the line passing through p_i and q_i is parallel to the vector $(\sin\omega, -\cos\omega)$, so the mean value theorem yields for each i a point r_i on Γ_{t_i} where the normal is parallel to $(\cos\omega, \sin\omega)$. This contradicts Lemma 3.1.

The strong maximum principle now implies that (17) holds for all $t < 0$ (note that $R_\omega \cdot \Gamma_t$ also intersects ∂B^2 orthogonally). In fact, $(R_\omega \cdot \Gamma_t) \cap H_\omega$ lies above $\Gamma_t \cap H_\omega$ for all $t < 0$ and all $\omega \in \left(0, \frac{\pi}{2}\right)$ and by continuity the same holds for $\omega = \frac{\pi}{2}$. Repeating the argument on the "other side" with the halfspaces

$$H_\omega = \{(x, y) : (x, y) \cdot (\sin\omega, -\cos\omega) > 0\}, \quad \omega \in \left(\tfrac{\pi}{2}, \pi\right),$$

implies the reflection symmetry. □

3C. *Asymptotics for the height.* We begin with a lemma.

Lemma 3.3. *For all $t < 0$,*

$$\kappa_s > 0 \quad \textit{in } \{x > 0\} \cap \Gamma_t$$

and hence

$$\frac{\sin\bar{\theta}}{1+\cos\bar{\theta}} \leq \underline{y}. \tag{18}$$

Proof. Choose $T > -\infty$ so that $\kappa < \frac{2}{7}$ for $t < T$ and, given $\varepsilon > 0$, set

$$v_\varepsilon \doteqdot \kappa_s + \varepsilon(1 - \langle\gamma, \nu\rangle).$$

We claim that $v_\varepsilon \geq 0$ in $\{x \geq 0\} \cap (-\infty, T)$. Suppose that this is not the case. Since at the spatial boundary $v_\varepsilon > \varepsilon$, and $v_\varepsilon \to \varepsilon$ as $t \to -\infty$, there must exist a first time in $(-\infty, T)$ and an interior point at which $v_\varepsilon = 0$. But, at such a point,

$$\begin{aligned} 0 \geq (\partial_t - \Delta)v_\varepsilon &= \kappa^2(\kappa_s - \varepsilon\langle\gamma, \nu\rangle) + 3\kappa^2\kappa_s + 2\varepsilon\kappa \\ &= -\varepsilon\kappa^2 - 3\varepsilon\kappa^2(1 - \langle\gamma, \nu\rangle) + 2\varepsilon\kappa \\ &\geq \varepsilon(2 - 7\kappa)\kappa > 0, \end{aligned}$$

which is absurd. Now take $\varepsilon \to 0$ to obtain $\kappa_s \geq 0$ in $\{x \geq 0\} \cap \Gamma_t$ for $t \in (-\infty, T]$. Since $\kappa_s = 0$ at the y-axis and $\kappa_s = \kappa > 0$ at the right boundary point, the strong maximum principle and the Hopf boundary point lemma imply that $\kappa_s > 0$ in $\{x > 0\} \cap \Gamma_t$ for $t \in (-\infty, T]$. But then Sturm's theorem implies that κ_s does not develop additional zeroes up to time 0.

Having established the first claim, the second follows as in Lemma 2.6. □

Proposition 3.4. *If we define $A \in (0, \infty)$ as in* (16), *then*

$$e^{\lambda_0^2 t} y(x, t) \to A\cosh(\lambda_0 x) \quad \textit{uniformly as } t \to -\infty.$$

Proof. Given $\tau < 0$, consider the rescaled height function

$$y^\tau(x, t) \doteqdot e^{-\lambda_0^2 \tau} y(x, t + \tau),$$

which is defined on the time-translated flow $\{\Gamma_t^\tau\}_{t \in (-\infty, -\tau)}$, where $\Gamma_t^\tau \doteqdot \Gamma_{t+\tau}$. Note that

$$\begin{cases} (\partial_t - \Delta^\tau)y^\tau = 0 & \text{in } \{\Gamma_t^\tau\}_{t \in (-\infty, -\tau)}, \\ \langle\nabla^\tau y^\tau, N\rangle = y & \text{on } \{\partial\Gamma_t^\tau\}_{t \in (-\infty, -\tau)}, \end{cases} \tag{19}$$

where ∇^τ and Δ^τ are the gradient and Laplacian on $\{\Gamma_t^\tau\}_{t \in (-\infty, -\tau)}$, respectively, and N is the outward unit normal to ∂B^2.

Since $\{\Gamma_t\}_{t \in (-\infty, 0)}$ reaches the origin at time zero, it must intersect the constructed solution for all $t < 0$. In particular, the value of $\underline{y}$ on the former can at no time exceed the value of $\bar{y}$ on the latter. But then (16) and (18) yield

$$\limsup_{t \to -\infty} e^{-\lambda_0^2 t} \bar{y} < \infty. \tag{20}$$

This implies a uniform bound for y^τ on $\{\Gamma_t^\tau\}_{t \in (-\infty, T]}$ for any $T \in \mathbb{R}$. So Alaoglu's theorem yields a sequence of times $\tau_j \to -\infty$ such that y^{τ_j} converges in the weak* topology as $j \to \infty$ to some $y^\infty \in L^2_{\text{loc}}([-1, 1] \times (-\infty, \infty))$. Since convexity and the boundary condition imply a uniform bound for $\nabla^\tau y^\tau$ on any time interval of the form $(-\infty, T]$, we may also arrange that the convergence is uniform in space at time zero, say.

Weak* convergence ensures that y^∞ satisfies the problem

$$\begin{cases} y_t = y_{xx} & \text{in } [-1,1]\times(-\infty,\infty), \\ y_x(\pm 1) = \pm y(\pm 1). \end{cases} \tag{21}$$

Indeed, a smooth function y^τ satisfies the boundary value problem (19) (and analogously for (21)) if and only if

$$\int_{-\infty}^{-\tau}\int_{\Gamma_t^\tau} y^\tau(\partial_t - \Delta^\tau)^*\eta = 0$$

for all smooth η which are compactly supported in time and satisfy

$$\nabla^\tau\eta\cdot N = \eta \quad \text{on } \partial\Gamma_t^\tau,$$

where $(\partial_t-\Delta^\tau)^* \doteqdot -(\partial_t+\Delta^\tau)$ is the formal L^2-adjoint of the heat operator. Since $\{\Gamma_t^\tau\}_{t\in(-\infty,-\tau)}$ converges uniformly in the smooth topology to the stationary interval $\{[-1,1]\times\{0\}\}_{t\in(-\infty,\infty)}$ as $\tau\to-\infty$, we conclude that the limit y^∞ must satisfy (21) in the L^2 sense (and hence in the classical sense due to the L^2 theory for the heat equation). Indeed, by the definition of smooth convergence, we may (after possibly applying a diffeomorphism) parametrize each flow $\{\bar\Gamma_t^{\tau_j}\}_{t\in(-\infty,-\tau_j)}$ over $I \doteqdot [-1,1]$ by a family of embeddings $\gamma_t^j : I\times(-\infty,-\tau_j)\to \bar B^2$ which converge in $C^\infty_{\mathrm{loc}}(I\times(-\infty,\infty))$ as $j\to\infty$ to the stationary embedding $(x,t)\mapsto xe_1$. Given $\eta\in C_0^\infty(I\times(-\infty,\infty))$ satisfying $\eta_\zeta(\pm1)=\pm\eta$, set $\eta^j \doteqdot \varphi^j\eta$, where $\varphi^j : [-1,1]\times(-\infty,-\tau^j)\to\mathbb{R}$ is defined by

$$\varphi_\zeta^j + (1-|\gamma_\zeta^j|)\varphi^j = 0, \quad \varphi^j(0,t)=1.$$

That is, $\varphi^j(\zeta,t)=\mathrm{e}^{s^j(\zeta,t)-\zeta}$, where $s^j(\zeta,t)\doteqdot\int_0^\zeta|\gamma_\zeta^j(\xi,t)|\,d\xi$. This ensures that $\nabla^{\tau^j}\eta^j\cdot N=\eta^j$ at the boundary, and hence

$$0=\int_{-\infty}^{\infty}\int_I y^{\tau_j}(\partial_t-\Delta^{\tau_j})^*\eta^j\,ds^j\,dt.$$

Since $\varphi^j\to 1$ in $C^\infty_{\mathrm{loc}}(I\times(-\infty,\infty))$, a short computation reveals that

$$0=\int_{-\infty}^{\infty}\int_I y^{\infty}(\partial_t-\Delta)^*\eta\,d\zeta\,dt.$$

Finally, we characterize the limit (uniqueness of which implies full convergence, completing the proof). Separation of variables leads us to consider the problem

$$\begin{cases} -\phi_{xx} = \mu\phi & \text{in } [-1,1], \\ \phi_x(\pm1) = \pm\phi(\pm1). \end{cases}$$

There is only one negative eigenspace, and its frequency turns out to be λ_0, with the corresponding mode given by

$$\phi_{-1}(x) \doteqdot \cosh(\lambda_0 x).$$

Thus, recalling (20), we are able to conclude that

$$y^\infty(x,t) = A\mathrm{e}^{\lambda_0^2 t}\cosh(\lambda_0 x)$$

for some $A \geq 0$. In particular,

$$e^{-\lambda_0^2 \tau_j} y(x, \tau_j) = y^{\tau_j}(x, 0) \to A\cosh(\lambda_0 x) \quad \text{uniformly as } j \to \infty.$$

Now, if A is not equal to the corresponding value on the constructed solution (note that the full limit exists for the latter), then one of the two solutions must lie above the other at time τ_j for j sufficiently large. But this violates the avoidance principle. $\square$

3D. *Uniqueness.* Uniqueness of the constructed ancient solution now follows directly from the avoidance principle.

Proposition 3.5. *Modulo time translation and rotation about the origin, there is only one convex, locally uniformly convex ancient solution to free-boundary curve-shortening flow in the disc.*

Proof. Denote by $\{\Gamma_t\}_{t\in(-\infty,0)}$ the constructed ancient solution and let $\{\Gamma'_t\}_{t\in(-\infty,0)}$ be a second ancient solution which, without loss of generality, contracts to the point (0, 1) at time 0. Given any $\tau > 0$, consider the time-translated solution $\{\Gamma^\tau_t\}_{t\in(-\infty,-\tau)}$ defined by $\Gamma^\tau_t = \Gamma'_{t+\tau}$. By Proposition 3.4,

$$e^{-\lambda_0^2 t} y^\tau(x, t) \to A e^{\lambda_0^2 \tau} \cosh(\lambda_0 x) \quad \text{as } t \to -\infty$$

uniformly in x. So Γ^τ_t lies above Γ_t for $-t$ sufficiently large. The avoidance principle then ensures that Γ^τ_t lies above Γ_t for all $t \in (-\infty, 0)$. Taking $\tau \to 0$, we find that Γ'_t lies above Γ_t for all $t < 0$. Since the two curves reach the point (0, 1) at time zero, they intersect for all $t < 0$ by the avoidance principle. The strong maximum principle then implies that the two solutions coincide for all t. $\square$

4. Supplement: nonexistence of rotators

Free-boundary curve-shortening flow in B^2 is invariant under rotations about the origin, so it is natural to seek solutions which move by rotation, that is, solutions $\gamma : (-L/2, L/2) \times (-\infty, \infty) \to \bar{B}^2$ satisfying

$$\gamma(\,\cdot\,, t) = e^{iBt}\gamma(\,\cdot\,, 0)$$

for some $B > 0$. Differentiating yields the *rotator equation*

$$\kappa = -B\langle \gamma, \tau \rangle. \tag{22}$$

It turns out, however, that there are no solutions to (22) in B^2 satisfying the free-boundary condition.

Proof of Theorem 1.2. Following [Halldorsson 2012], we rewrite the rotator equation as the pair of ordinary differential equations

$$x' = B + xy \quad \text{and} \quad y' = -x^2, \tag{23}$$

where

$$x \doteqdot B\langle \gamma, \tau \rangle \quad \text{and} \quad y \doteqdot B\langle \gamma, \nu \rangle.$$

Arc-length parametrized solutions γ to the rotator equation (22) can be recovered from solutions to the system (23) via

$$\gamma \doteqdot B^{-1}(x + iy)e^{i\theta}, \quad \theta(s) \doteqdot -\int_0^s x(\sigma)\, d\sigma,$$

and this parametrization is unique up to an ambient rotation and a unit linear reparametrization, i.e., $(\theta, s) \mapsto (\pm\theta + \theta_0, \pm s + s_0)$.

Note that

$$|\gamma| = B^{-1}\sqrt{x^2 + y^2}.$$

So we seek solutions $(x, y) : (-L/2, L/2) \to B^2$ to (23) satisfying the free-boundary condition

$$\left(x\left(\pm\frac{L}{2}\right), y\left(\pm\frac{L}{2}\right)\right) = (\pm B, 0).$$

Let γ be such a solution. Since (23) can be uniquely solved with initial condition $(x(s_0), y(s_0)) = (B, 0)$ (which corresponds to $\gamma(s_0) \in \partial B^2$ with $\langle\gamma, \tau\rangle|_{s_0} = 1$), we find that γ must be invariant under rotation by π about the origin. In particular, the points $\gamma(-L/2)$ and $\gamma(L/2)$ are diametrically opposite. It follows that $\gamma(0)$ is the origin. Indeed, for topological reasons, γ must cross the line orthogonally bisecting the segment joining its endpoints an odd number of times (with multiplicity). But since the rotational invariance pairs each crossing above the origin with one below, we are forced to include the origin in the set of crossings. We conclude that

$$0 = y\left(\frac{L}{2}\right) = \int_0^{L/2} y' = -\int_0^{L/2} x^2\, ds,$$

which is impossible since $x(L/2) = B > 0$. □

Acknowledgements

We wish to thank Jonathan Zhu for sharing his thoughts on the problem.

Bourni was supported through grant 707699 of the Simons Foundation and grant DMS-2105026 of the National Science Foundation. Langford was supported through an Australian Research Council DECRA fellowship (grant DE200101834).

References

[Andrews 2012] B. Andrews, "Noncollapsing in mean-convex mean curvature flow", *Geom. Topol.* **16**:3 (2012), 1413–1418. MR Zbl

[Andrews and Bryan 2011a] B. Andrews and P. Bryan, "A comparison theorem for the isoperimetric profile under curve-shortening flow", *Comm. Anal. Geom.* **19**:3 (2011), 503–539. MR Zbl

[Andrews and Bryan 2011b] B. Andrews and P. Bryan, "Curvature bound for curve shortening flow via distance comparison and a direct proof of Grayson's theorem", *J. Reine Angew. Math.* **653** (2011), 179–187. MR Zbl

[Angenent 1988] S. Angenent, "The zero set of a solution of a parabolic equation", *J. Reine Angew. Math.* **390** (1988), 79–96. MR Zbl

[Bakas and Sourdis 2007] I. Bakas and C. Sourdis, "Dirichlet sigma models and mean curvature flow", *J. High Energy Phys.* **2007**:6 (2007), art. id. 057. MR

[Bourni et al. 2020] T. Bourni, M. Langford, and G. Tinaglia, "Convex ancient solutions to curve shortening flow", *Calc. Var. Partial Differential Equations* **59**:4 (2020), art. id. 133. MR Zbl

[Bryan and Louie 2016] P. Bryan and J. Louie, "Classification of convex ancient solutions to curve shortening flow on the sphere", *J. Geom. Anal.* **26**:2 (2016), 858–872. MR Zbl

[Choi and Mantoulidis 2022] K. Choi and C. Mantoulidis, “Ancient gradient flows of elliptic functionals and Morse index”, *Amer. J. Math.* **144**:2 (2022), 541–573. MR Zbl

[Chow and Gulliver 2001] B. Chow and R. Gulliver, “Aleksandrov reflection and geometric evolution of hypersurfaces”, *Comm. Anal. Geom.* **9**:2 (2001), 261–280. MR Zbl

[Daskalopoulos et al. 2010] P. Daskalopoulos, R. Hamilton, and N. Sesum, “Classification of compact ancient solutions to the curve shortening flow”, *J. Differential Geom.* **84**:3 (2010), 455–464. MR Zbl

[Firey 1974] W. J. Firey, “Shapes of worn stones”, *Mathematika* **21** (1974), 1–11. MR Zbl

[Gage 1984] M. E. Gage, “Curve shortening makes convex curves circular”, *Invent. Math.* **76**:2 (1984), 357–364. MR Zbl

[Gage and Hamilton 1986] M. Gage and R. S. Hamilton, “The heat equation shrinking convex plane curves”, *J. Differential Geom.* **23**:1 (1986), 69–96. MR Zbl

[Halldorsson 2012] H. P. Halldorsson, “Self-similar solutions to the curve shortening flow”, *Trans. Amer. Math. Soc.* **364**:10 (2012), 5285–5309. MR Zbl

[Hamilton 1995a] R. S. Hamilton, “The formation of singularities in the Ricci flow”, pp. 7–136 in *Surveys in differential geometry*, *II* (Cambridge, 1993), edited by S.-T. Yau, Int. Press, Cambridge, MA, 1995. MR Zbl

[Hamilton 1995b] R. S. Hamilton, “Isoperimetric estimates for the curve shrinking flow in the plane”, pp. 201–222 in *Modern methods in complex analysis* (Princeton, 1992), edited by T. Bloom et al., Ann. of Math. Stud. **137**, Princeton Univ. Press, 1995. MR Zbl

[Huisken 1989] G. Huisken, “Nonparametric mean curvature evolution with boundary conditions”, *J. Differential Equations* **77**:2 (1989), 369–378. MR Zbl

[Huisken 1998] G. Huisken, “A distance comparison principle for evolving curves”, *Asian J. Math.* **2**:1 (1998), 127–133. MR Zbl

[Mullins 1956] W. W. Mullins, “Two-dimensional motion of idealized grain boundaries”, *J. Appl. Phys.* **27** (1956), 900–904. MR

[von Neumann 1952] J. von Neumann, “Discussion: ‘Shape of metal grains”’, pp. 108–110 in *Metal interfaces*, edited by C. Herring, Amer. Soc. Metals, Cleveland, OH, 1952.

[Sapiro 2001] G. Sapiro, *Geometric partial differential equations and image analysis*, Cambridge Univ. Press, 2001. MR Zbl

[Stahl 1996a] A. Stahl, “Convergence of solutions to the mean curvature flow with a Neumann boundary condition”, *Calc. Var. Partial Differential Equations* **4**:5 (1996), 421–441. MR Zbl

[Stahl 1996b] A. Stahl, “Regularity estimates for solutions to the mean curvature flow with a Neumann boundary condition”, *Calc. Var. Partial Differential Equations* **4**:4 (1996), 385–407. MR Zbl

Received 10 Nov 2021. Revised 23 Feb 2022. Accepted 29 Mar 2022.

THEODORA BOURNI: tbourni@utk.edu
Department of Mathematics, University of Tennessee, Knoxville, TN, United States

MAT LANGFORD: mathew.langford@anu.edu.au
Mathematical Sciences Institute, Australian National University, Canberra, Australia

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 16 No. 9 2023