ANALYSIS & PDE

Volume 12 No. 2 2019

msp

Analysis & PDE
msp.org/apde

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

The subscription price for 2019 is US $310/year for the electronic version, and $520/year (+$60, 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. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

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

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.259

A UNIFIED FLOW APPROACH TO SMOOTH, EVEN L_p-MINKOWSKI PROBLEMS

PAUL BRYAN, MOHAMMAD N. IVAKI AND JULIAN SCHEUER

We study long-time existence and asymptotic behavior for a class of anisotropic, expanding curvature flows. For this we adapt new curvature estimates, which were developed by Guan, Ren and Wang to treat some stationary prescribed curvature problems. As an application we give a unified flow approach to the existence of smooth, even L_p-Minkowski problems in $\mathbb{R}^{n+1}$ for $p > -n-1$.

1. Introduction

Consider a smooth, closed, strictly convex hypersurface M_0 in Euclidean space $\mathbb{R}^{n+1}$, $n \geq 2$, given by a smooth embedding $F_0 : M \to \mathbb{R}^{n+1}$. Suppose the origin is in the interior of the region enclosed by M_0. We study the long-time behavior of a family of hypersurfaces $\{M_t\}$ given by smooth maps $F : M \times [0, T) \to \mathbb{R}^{n+1}$ satisfying the initial value problem

$$\partial_t F(x,t) = \varphi(\nu(x,t)) \frac{(F(x,t) \cdot \nu(x,t))^{2-p}}{\mathcal{K}(x,t)} \nu(x,t), \quad F(\,\cdot\,,0) = F_0(\,\cdot\,). \tag{1-1}$$

Here $\mathcal{K}(\,\cdot\,,t)$ and $\nu(\,\cdot\,,t)$ are the Gauss curvature and the outer unit normal vector of $M_t = F(M,t)$ and φ is a positive, smooth function on $\mathbb{S}^n$. Furthermore, T is the maximal time for which the solution exists.

For $p = 2$, $\varphi \equiv 1$, flow (1-1) was studied in [Schnürer 2006] in $\mathbb{R}^3$ and in [Gerhardt 2014] in higher dimensions. Both works rely on the reflection principle of [Chow and Gulliver 1996; McCoy 2003]. Their result is as follows: the volume-normalized flow evolves any M_0 in the C^∞-topology to an origin-centered sphere. For $p > 2$, $\varphi \equiv 1$, it follows from [Chow and Gulliver 1996, Theorem 3.1], see also [Tsai 2005, Example 1], that (1-1) evolves M_0, after rescaling to fixed volume, in the C^1-topology to an origin-centered sphere. We refer the reader to [Ivaki 2016] regarding a rather comprehensive list of previous works on this curvature flow. In particular, in either case $\varphi \not\equiv 1$ or $\varphi \equiv 1$, $-n-1 < p < 2$, we are not aware of any result in the literature on the asymptotic behavior of the flow. The following theorem was proved in [Ivaki 2016] regarding the case $p = -n-1$, $\varphi \equiv 1$; in this case the flow belongs to a family of centroaffine normal flows introduced in [Stancu 2012].

Let us write B for the unit ball of $\mathbb{R}^{n+1}$ and put

$$\widetilde{K}_t := \left(\frac{V(B)}{V(K_t)}\right)^{\frac{1}{n+1}} K_t,$$

MSC2010: primary 53C44; secondary 35K55, 52A05, 53A15, 58J35.
Keywords: curvature flow, anisotropic flow, Minkowski problem.

where K_t denotes the convex body enclosed by M_t and $V(\cdot)$ is the $(n+1)$-dimensional Lebesgue measure.

Theorem [Ivaki 2016]. *Let $n \geq 2$, $p = -n-1$, $\varphi \equiv 1$ and suppose K_0 has its Santaló point at the origin; i.e.,*

$$\int_{\mathbb{S}^n} \frac{u}{h_{K_0}(u)^{n+2}}\, d\sigma(u) = 0.$$

Then there exists a unique solution $\{M_t\}$ of flow (1-1) *such that $\widetilde{M}_t$ converges in C^∞ to an origin-centered ellipsoid.*

Here h_{K_0} is the support function of K_0. A closed, convex hypersurface M_0 can be described in terms of its support function $h_{K_0} : \mathbb{S}^n \to \mathbb{R}$ defined by

$$h_{K_0}(u) = \sup\{u \cdot x : x \in M_0\}.$$

If M_0 is smooth and strictly convex, then $h_{K_0}(u) = u \cdot F_0(\nu^{-1}(u))$.

From the evolution equation of $F(\cdot\,, t)$ it follows that

$$h(\cdot\,, t) := h_{K_t}(\cdot) : \mathbb{S}^n \times [0, T) \to \mathbb{R}$$

evolves by

$$\partial_t h(u, t) = \varphi(u)(h^{2-p} S_n)(u, t), \tag{1-2}$$

where $S_n(u, t) = 1/\mathcal{K}(\nu^{-1}(u, t), t)$. A homothetic self-similar solution of this flow satisfies

$$h^{1-p} \det(\overline{\nabla}^2 h + \mathrm{Id}\, h) = \frac{c}{\varphi} \tag{1-3}$$

for some positive constant c. Here $\overline{\nabla}$ is the covariant derivative on $\mathbb{S}^n$. Note that $S_n = \det(\overline{\nabla}^2 h + \mathrm{Id}\, h)$.

We list the main results of the paper extending the previous-mentioned results.

Theorem 1. *Let $-n-1 < p < \infty$ and φ be a positive, smooth function on $\mathbb{S}^n$ that is even, i.e., $\varphi(u) = \varphi(-u)$. Suppose K_0 is origin-symmetric. There exists a unique origin-symmetric solution $\{M_t\}$ of* (1-1) *such that $\{\widetilde{M}_t\}$ converges for a subsequence of times in C^1 to a smooth, origin-symmetric, strictly convex solution of* (1-3). *Also, when $p \leq n+1$ the convergence is in C^∞, and if $p \geq 1$ the convergence holds for the full sequence.*

If $-n-1 < p \leq -n$, we can extend the result of the previous theorem by dropping the assumption that φ is even.

Theorem 2. *Let $-n-1 < p \leq -n$ and K_0 satisfy*

$$\int_{\mathbb{S}^n} \frac{u}{\varphi(u) h_{K_0}(u)^{1-p}}\, d\sigma(u) = 0.$$

There exists a unique solution $\{M_t\}$ of flow (1-1) *such that $\{\widetilde{M}_t\}$ converges for a subsequence of times in C^∞ to a positive, smooth, strictly convex solution of* (1-3).

Given any convex body K_0, there exists a vector $\vec{v}$ such that $K_0 + \vec{v}$ has the origin in its interior and it satisfies the assumption of the second theorem.

For $\varphi \equiv 1$ we prove the following theorem.

Theorem 3. *Let $1 \neq p > -n-1$, $\varphi \equiv 1$ and K_0 satisfy*

$$\int_{\mathbb{S}^n} \frac{u}{h_{K_0}(u)^{1-p}}\, d\sigma(u) = 0.$$

Then there exists a unique solution $\{M_t\}$ of (1-1) such that $\{\widetilde{M}_t\}$ converges in C^1 to the unit sphere. In addition, for $1 \neq p \leq n+1$ the convergence holds in C^∞.

For $p \neq n+1$, self-similar solutions to (1-1) are solutions of the L_p-Minkowski problem (1-4), and for $p = n+1$, a self-similar solution to (1-1) is a solution to the normalized L_{n+1}-Minkowski problem (1-5); we shall introduce them now.

The Minkowski problem deals with existence, uniqueness, regularity, and stability of closed convex hypersurfaces whose Gauss curvature (as a function of the outer normals) is preassigned. Major contributions to this problem were made by Minkowski [1897; 1903], Aleksandrov [1938; 1939; 1942], Fenchel and Jessen [1938], Lewy [1938a; 1938b], Nirenberg [1953], Calabi [1958], Pogorelov [1952; 1971], Cheng and Yau [1976], Caffarelli, Nirenberg, and Spruck [Caffarelli et al. 1984], and others. A generalization of the Minkowski problem known as the L_p-Minkowski problem was introduced in [Lutwak 1993], where for any $1 < p \neq n+1$ and a preassigned even Borel measure on $\mathbb{S}^n$ whose support does not lie in a great sphere of $\mathbb{S}^n$ the existence and uniqueness of the solution were proved. This generalization for $1 < p \neq n+1$ was further studied in [Lutwak and Oliker 1995], where they obtained the $C^{k,\alpha}$ regularity of the solution. Solutions to many cases of these generalized problems followed later in [Ai et al. 2001; Andrews 2000; 2002; 2003; Böröczky et al. 2013; Chen 2006; Chou and Wang 2006; Dou and Zhu 2012; Gage 1993; Gage and Li 1994; Guan and Lin 2000; Huang and Lu 2013; Jiang 2010; Jiang et al. 2011; Lu and Wang 2013; Lutwak et al. 2004; Stancu 1996; 2002; 2003; Umanskiy 2003; Zhu 2014; 2015a; 2015b].

For $p \neq n+1$, in the smooth category, the L_p-Minkowski problem asks, given a smooth, positive function $\varphi : \mathbb{S}^n \to \mathbb{R}$, if there exists a smooth, closed, strictly convex hypersurface $M_0 \subset \mathbb{R}^{n+1}$ such that

$$\frac{h^{1-p}(\nu(x))}{\mathcal{K}(x)} = \frac{1}{\varphi(\nu(x))}, \tag{1-4}$$

where $x \in M_0$, h denotes the support function, $\mathcal{K}$ the Gauss curvature and ν the Gauss map $M_0 \to \mathbb{S}^n$. The *even* L_p-Minkowski problem requires, in addition, that φ is an even function. The case $p = 1$ is the original Minkowski problem.

The special case of $p = n+1$ is troubling since (1-4) might not have a solution. To remedy this, Lutwak, Yang and Zhang [Lutwak et al. 2004] introduced a normalized formulation of the L_{n+1}-Minkowski problem and they proved the existence and uniqueness of the solution for any prescribed even Borel measure on $\mathbb{S}^n$ whose support is not contained in a great sphere of $\mathbb{S}^n$. In the smooth category, the normalized L_{n+1}-Minkowski problem asks for the existence of a smooth, closed, strictly convex hypersurface $M_0 \subset \mathbb{R}^{n+1}$ that solves

$$\frac{1}{h^n(\nu(x))\mathcal{K}(x)} = \frac{V(K_0)}{\varphi(\nu(x))}, \tag{1-5}$$

where K_0 is the convex body with the boundary M_0. In the rest of the paper, the L_p-Minkowski problem refers to either (1-4) or (1-5), and we avoid the word "normalized".

The existence and regularity of solutions to the L_p-Minkowski problem are rather comprehensively discussed in [Chou and Wang 2006] for $p > -n-1$. Our study on (1-1) provides an alternative variational treatment (based on curvature flow) of the even L_p-Minkowski problem. For $p = 1$, Chou and Wang [2000] treated the classical L_1-Minkowski problem in the smooth category by a logarithmic Gauss curvature flow. For $n = 1$ and $1 \neq p > -3$, the existence of solutions to the L_p-Minkowski problems follows from Andrews' results [1998] on the asymptotic behavior of a family of contracting and expanding flows of curves. Also, in higher dimensions, the existence of solutions to the L_p-Minkowski problems follows from [Andrews 2000] when $-n-1 < p \leq -n+1$ (a short proof of this is also given in [Ivaki 2015]) or when φ is even (i.e., $\varphi(u) = \varphi(-u)$) and $-n+1 < p < 1$. See also [Andrews 1999; Andrews et al. 2016; Guan and Ni 2017; Urbas 1998; 1999].

Using our results for the flows above, it is now a simple matter to give a new, unified proof of the smooth, even L_p-Minkowski problem for all ranges of $p > -n-1$.

Corollary 4. *Let $-n-1 < p < \infty$ and φ be a positive, smooth function on $\mathbb{S}^n$ that is even, i.e., $\varphi(u) = \varphi(-u)$. Then for $p \neq n+1$ there exists an origin-symmetric, smooth, strictly convex body such that* (1-4) *is satisfied. For $p = n+1$, there exists an origin-symmetric, smooth, strictly convex body such that* (1-5) *is satisfied.*

Proof. By the first part of Theorem 1 (only the convergence for a *subsequence* of times is needed), there exists a smooth, strictly convex body K with the volume of the unit ball and a constant $c > 0$ such that

$$\frac{h}{\mathcal{K}} = \frac{ch^p}{\varphi}.$$

Hence $c\int_{\mathbb{S}^n} h^p/\varphi \, d\sigma = (n+1)V(B^n)$. Thus there is a solution to

$$\frac{h^{1-p}(\nu(x))}{\mathcal{K}(x)} = \left(\frac{(n+1)V(B)}{\int_{\mathbb{S}^n} h^p/\varphi \, d\sigma}\right)\frac{1}{\varphi(\nu(x))}.$$

Now let us define

$$\lambda := \begin{cases} \left(\dfrac{\int_{\mathbb{S}^n} h^p/\varphi \, d\sigma}{(n+1)V(B)}\right)^{\frac{1}{n+1-p}}, & p \neq n+1, \\[2ex] \left(\dfrac{(n+1)V(B)}{V(K)\int_{\mathbb{S}^n} h^{n+1}/\varphi \, d\sigma}\right)^{\frac{1}{n+1}}, & p = n+1. \end{cases}$$

Therefore, λK solves the smooth, even L_p-Minkowski problem. □

Let us close this section with a brief outline of this paper. The main difficulty in proving convergence of the normalized solutions is in obtaining long-time existence. The issue arises from the time-dependent anisotropic factor (the support function). We believe in such generality, (1-1) serves as the first example where a time-dependent anisotropic factor is allowed. To prove long-time existence, we first obtain bounds on the Gauss curvature in Section 3.1. Using the well-known standard technique of [Tso 1985] we obtain upper bounds. We obtain lower bounds by applying the same technique to the evolution of the polar body as in [Ivaki 2015]. Controlling the principal curvatures requires estimates of higher derivatives of the speed, which is generally quite difficult due to the nonlinearity of the flow. In Section 3.2 we obtain these crucial estimates by adapting the remarkable C^2 estimates of Guan, Ren and Wang [Guan et al. 2015, (4.2)] for the prescribed curvature problem. Long-time existence then follows readily by standard arguments. Once it is proved that solutions to the flow exist until they expand to infinity uniformly in all directions, the method of [Ivaki 2016, Section 8] applies and yields convergence of the volume-normalized solutions in C^1 to self-similar solutions provided $p \neq 1$. Further work is required to establish convergence of normalized solutions if $p = 1$, and to prove convergence in C^∞ for $p \leq n+1$. This is accomplished in Section 4; see also Remark 10.

2. Basic evolution equations

Let $g = \{g_{ij}\}$ and $W = \{w_{ij}\}$ denote, in order, the induced metric and the second fundamental form of M. At every point in the hypersurface M choose a local orthonormal frame $\{e_1, \dots, e_n\}$.

We use the standard notation

$$w_i^j = g^{mj} w_{im}, \quad (w^2)_i^j = g^{mj} g^{rs} w_{ir} w_{sm}, \quad |W|^2 = g^{ij} g^{kl} w_{ik} w_{lj} = w_{ij} w^{ij}.$$

Here, $\{g^{ij}\}$ is the inverse matrix of $\{g_{ij}\}$.

We use semicolons to denote covariant derivatives. The following geometric formulas are well-known:

$$\begin{aligned} \nu_{;i} &= w_i^k e_k, & h_{;i} &= w_i^k (F \cdot e_k), \\ \nu_{;ij} &= g^{kl} w_{ij;l} e_k - w_i^l w_{lj} \nu, & h_{;ij} &= w_{ij} - h w_i^l w_{lj} + F \cdot \nabla w_{ij}. \end{aligned}$$

Note that above we considered the support function as a function on the boundary of the hypersurface; that is, at the point $x \in M$ we have

$$h(x) = F(x) \cdot \nu(x).$$

For convenience, let $\psi(x) = h^{2-p}(x) \varphi(\nu(x))$. The following evolution can be deduced in a standard manner; see for example [Gerhardt 2006].

Lemma 5. *The following evolution equations hold*:

$$\partial_t \nu = -\nabla\left(\frac{\psi}{\mathcal{K}}\right),$$

$$\begin{aligned} \partial_t w_i^j &= -\left(\frac{\psi}{\mathcal{K}}\right)_{;ik} g^{kj} - \left(\frac{\psi}{\mathcal{K}}\right) w_i^k w_k^j \\ &= \psi \frac{\mathcal{K}^{kl}}{\mathcal{K}^2} w_{i;kl}^j + \psi \frac{\mathcal{K}^{kl}}{\mathcal{K}^2} w_{kr} w_l^r w_i^j - (n+1) \frac{\psi}{\mathcal{K}} w_i^k w_k^j + \psi \frac{\mathcal{K}^{kl,rs}}{\mathcal{K}^2} g^{jm} w_{kl;i} w_{rs;m} \\ &\qquad - \frac{2\psi}{\mathcal{K}^3} g^{jm} \mathcal{K}_{;i} \mathcal{K}_{;m} + \frac{1}{\mathcal{K}^2} g^{jk} \mathcal{K}_{;k} \psi_{;i} + \frac{1}{\mathcal{K}^2} g^{jk} \psi_{;k} \mathcal{K}_{;i} - \frac{1}{\mathcal{K}} g^{jk} \psi_{;ik}, \end{aligned}$$

$$\partial_t h = \psi \frac{\mathcal{K}^{ij}}{\mathcal{K}^2} h_{;ij} + \psi h \frac{\mathcal{K}^{ij}}{\mathcal{K}^2} w_i^l w_{lj} - (n-1) \frac{\psi}{\mathcal{K}} - \frac{1}{\mathcal{K}} F \cdot \nabla \psi.$$

3. Long-time existence

3.1. ***Lower and upper bounds on Gauss curvature.*** The proofs of the following two lemmas are similar to the proofs of [Ivaki 2015, Lemmas 4.1, 4.2]. For completeness, we give the proofs here. In this section we use $\bar{\nabla}$ to denote covariant derivatives on the sphere with respect to the standard metric.

The matrix of the radii of the curvature of a smooth, closed, strictly convex hypersurface is denoted by $\mathfrak{r} = [\mathfrak{r}_{ij}]$ and the entries of $\mathfrak{r}$ are considered as functions on the unit sphere. They can be expressed in terms of the support function as $\mathfrak{r}_{ij} := \bar{\nabla}^2_{ij} h + \bar{g}_{ij} h$, where $[\bar{g}_{ij}]$ is the standard metric on $\mathbb{S}^n$. Additionally, we recall that $S_n = \det[\mathfrak{r}_{ij}]/\det[\bar{g}_{ij}]$.

Lemma 6. *Let $\{M_t\}$ be a solution of* (1-1) *on* $[0, t_1]$. *If $c_2 \le h_{K_t} \le c_1$ on* $[0, t_1]$, *then $\mathcal{K} \le c_4$ on* $[0, t_1]$. *Here c_4 depends on $K_0, c_1, c_2, p, \varphi$ and t_1.*

Proof. We apply the maximum principle to the following auxiliary function defined on the unit sphere:

$$\Theta = \frac{\psi S_n}{2c_1 - h} = \frac{\partial_t h}{2c_1 - h}.$$

At any minimum of Θ we have

$$0 = \bar{\nabla}_i \Theta = \bar{\nabla}_i \left(\frac{\psi S_n}{2c_1 - h} \right) \quad \text{and} \quad \bar{\nabla}^2_{ij} \Theta \ge 0.$$

Therefore, we get

$$\frac{\bar{\nabla}_i (\psi S_n)}{2c_1 - h} = -\frac{\psi S_n \bar{\nabla}_i h}{(2c_1 - h)^2}$$

and

$$\bar{\nabla}^2_{ij} (\psi S_n) + \bar{g}_{ij} \psi S_n \ge \frac{-\psi S_n \mathfrak{r}_{ij} + 2c_1 \psi S_n \bar{g}_{ij}}{2c_1 - h}. \tag{3-1}$$

Differentiating Θ with respect to time yields

$$\partial_t \Theta = \frac{\psi S_n^{ij}}{2c_1 - h} (\bar{\nabla}^2_{ij} (\psi S_n) + \bar{g}_{ij} \psi S_n) + \frac{\psi^2 S_n^2}{(2c_1 - h)^2} (1 + (2 - p) h^{-1} (2c_1 - h)),$$

where S_n^{ij} is the derivative of S_n with respect to the entry $\mathfrak{r}_{ij}$. By applying inequality (3-1) to the preceding identity we deduce

$$\partial_t \Theta \ge \Theta^2 (1 - n + 2c_1 \mathcal{H}) - c\Theta^2, \tag{3-2}$$

where

$$\mathcal{H} = S_n^{-1} S_n^{ij} \bar{g}_{ij}.$$

Therefore, we arrive at

$$\frac{(h^{2-p}/\mathcal{K})\varphi}{2c_1 - h}(t, u) \ge \left(ct + \frac{1}{\min\limits_{u \in \mathbb{S}^n} \frac{(h^{2-p}/\mathcal{K})\varphi}{2c_1 - h}(0, u)} \right)^{-1} \ge \left(ct_1 + \frac{1}{\min\limits_{u \in \mathbb{S}^n} \frac{(h^{2-p}/\mathcal{K})\varphi}{2c_1 - h}(0, u)} \right)^{-1}. \qquad \square$$

Lemma 7. *Let $\{M_t\}$ be a solution of* (1-1) *on* $[0, t_1]$. *If $c_1 \le h_{K_t} \le c_2$ on* $[0, t_1]$, *then $\mathcal{K} \ge 1/(a + bt^{-\frac{n}{n+1}})$ on* $(0, t_1]$, *where a and b depend only on c_1, c_2, p, φ. In particular, $\mathcal{K} \ge c_3$ on* $[0, t_1]$ *for a positive number c_3 that depends on $K_0, c_1, c_2, p, \varphi$ and is independent of t_1.*

Proof. Suppose K_t^* is the polar body[1] of K_t with respect to the origin. We furnish quantities associated with polar bodies with *. The polar bodies evolve by

$$\partial_t h^* = -\psi^* S_n^{*-1}, \quad h^*(\,\cdot\,,t) = h_{K_t^*}(\,\cdot\,),$$

where

$$\psi^* = \frac{(h^{*2}+|\overline{\nabla}h^*|^2)^{\frac{n+1+p}{2}}}{h^{*n+1}}\varphi\biggl(\frac{h^*u+\overline{\nabla}h^*}{\sqrt{h^{*2}+|\overline{\nabla}h^*|^2}}\biggr);$$

see Lemma 11 for the proof. In addition, we have $c_1' = 1/c_2 \le h^* \le 1/c_1 = c_2'$. We will show that the function

$$\Theta = \frac{\psi^* S_n^{*-1}}{h^* - c_1'/2}$$

remains bounded. At any maximal point of Θ,

$$0 = \overline{\nabla}_i\Theta = \overline{\nabla}_i\biggl(\frac{\psi^* S_n^{*-1}}{h^* - c_1'/2}\biggr) \quad \text{and} \quad \overline{\nabla}^2_{ij}\Theta \le 0.$$

Hence, we obtain

$$\frac{\overline{\nabla}_i(\psi^* S_n^{*-1})}{h^* - c_1'/2} = \frac{\psi^* S_n^{*-1}\overline{\nabla}_i h^*}{(h^* - c_1'/2)^2}, \tag{3-3}$$

and consequently,

$$\overline{\nabla}^2_{ij}(\psi^* S_n^{*-1}) + \bar{g}_{ij}\psi^* S_n^{*-1} \le \frac{\psi^* S_n^{*-1}\mathfrak{r}^*_{ij} - (c_1'/2)\psi^* S_n^{*-1}\bar{g}_{ij}}{h^* - c_1'/2}. \tag{3-4}$$

Differentiating Θ with respect to time yields

$$\partial_t\Theta = \frac{\psi^* S_n^{*-2}}{h^* - c_1'/2} S_n^{*ij}(\overline{\nabla}^2_{ij}(\psi^* S_n^{*-1}) + \bar{g}_{ij}\psi^* S_n^{*-1}) + \frac{S_n^{*-1}}{h^* - c_1'/2}\partial_t\psi^* + \Theta^2.$$

On the other hand, in view of

$$|\partial_t h^*| = \psi^* S_n^{*-1}, \quad \|\overline{\nabla}\partial_t h^*\| = \|\overline{\nabla}(\psi^* S_n^{*-1})\| = \frac{\psi^* S_n^{*-1}\|\overline{\nabla}h^*\|}{h^* - c_1'/2}, \quad \|\overline{\nabla}h^*\| \le c_2',$$

where for the second equation we used (3-3), we have

$$\frac{S_n^{*-1}}{h^* - c_1'/2}\partial_t\psi^* \le c(n,p,c_1,c_2,\varphi)\Theta^2.$$

Employing this last inequality and inequality (3-4) we infer that, at any point where the maximum of Θ is reached, we have

$$\partial_t\Theta \le \Theta^2\biggl(c' - \frac{c_1'}{2}\mathcal{H}^*\biggr). \tag{3-5}$$

[1]The polar body of a convex body K with the origin of $\mathbb{R}^{n+1}$ in its interior is the convex body defined by $K^* = \{x \in \mathbb{R}^{n+1} : x\cdot y \le 1 \text{ for all } y \in K\}$.

Moreover, we have

$$\mathcal{H}^* \geq n\left(\frac{h^* - c_1'/2}{\psi^* S_n^{*-1}}\right)^{-\frac{1}{n}} \left(\frac{\psi^*}{h^* - c_1'/2}\right)^{-\frac{1}{n}} \geq n\Theta^{\frac{1}{n}}\left(\frac{c''}{c_1' - c_1'/2}\right)^{-\frac{1}{n}}.$$

Therefore, we can rewrite the inequality (3-5) as

$$\partial_t \Theta \leq \Theta^2 (c - c'\Theta^{\frac{1}{n}})$$

for positive constants c and c' depending only on p, c_1, c_2, φ. Hence,

$$\Theta \leq c + c' t^{-\frac{n}{n+1}} \tag{3-6}$$

for some positive constants depending only on p, c_1, c_2, φ. This follows from the claim below.

Claim. *Suppose f is a positive smooth function of t on $[0, t_1]$ that satisfies*

$$\frac{d}{dt} f \leq c_0 + c_1 f + c_2 f^2 - c_3 f^{2+p}, \tag{3-7}$$

where c_3, p are positive. There exist constant $c, c' > 0$ independent of the solution and depending only on c_0, c_1, c_2, c_3, p, such that $f \leq c + c' t^{-1/(p+1)}$ on $(0, t_1]$.

Proof of claim. Note that there exists $x_0 > 0$ such that $c_0 + c_1 x + c_2 x^2 - c_3 x^{2+p} < -c_3/2 x^{2+p}$ for $x > x_0$. If $f(0) \leq x_0$, then f may increase forward in time, but when f reaches x_0, then f must start decreasing (since the right-hand side of (3-7) becomes negative). Thus we may assume, without loss of generality, that $f(0) > x_0$. Therefore, $f > x_0$ on a maximal time interval $[0, t_0)$. On $[0, t_0)$ we can solve

$$\frac{d}{dt} f \leq -\frac{c_3}{2} f^{2+p}$$

to obtain

$$f \leq \left(c_3 \frac{p+1}{2t}\right)^{-\frac{1}{p+1}}.$$

At t_0 we have $c_0 + c_1 f + c_2 f^2 - c_3 f^{1+p} = -(c_3/2) f^{2+p}$ and $f = x_0$; therefore the right-hand side of (3-7) is still negative. So $f \leq f(t_0)$ on $[t_0, t_1]$. In conclusion,

$$f \leq \max\left\{\left(c_3 \frac{p+1}{2t}\right)^{-\frac{1}{p+1}}, \, x_0 = f(t_0)\right\} \leq c + c' t^{-\frac{1}{1+p}},$$

where c, c' do not depend on solutions. □

The inequality (3-6) implies that

$$S_n^{*-1} \leq a' + b' t^{-\frac{n}{n+1}} \tag{3-8}$$

for some a' and b' depending only on p, c_1, c_2, φ. Now we can use the argument given in [Ivaki and Stancu 2013, Lemma 2.3] to obtain the desired lower bound: For every $u \in \mathbb{S}^n$, there exists a unique $u^* \in \mathbb{S}^n$ such that

$$(S_n h^{n+2})(u)(S_n^* h^{*n+2})(u^*) = 1;$$

see [Hug 1996]. In view of this identity and (3-8) we conclude that on $(0,t_1]$ we have

$$\mathcal{K}\geq \frac{1}{a+bt^{-\frac{n}{n+1}}}$$

for some a and b depending only on p, c_1, c_2, φ. The lower bound for $\mathcal{K}$ on $[0,\delta]$ for a small enough $\delta>0$ follows from the short-time existence of the flow. The lower bound for $\mathcal{K}$ on $[\delta,t_1]$ follows from the inequality $\mathcal{K}\geq 1/(a+b\delta^{-\frac{n}{n+1}})$. □

3.2. ***Upper and lower bounds on principal curvatures.*** To obtain upper and lower bounds on the principal curvatures, denoted by $\{\kappa_i\}_{i=1}^n$, we will consider the auxiliary function used by Guan, Ren and Wang [Guan et al. 2015, (4.2)] for a prescribed curvature problem.

Lemma 8. *Let $\{M_t\}$ be a solution of* (1-1) *on* $[0,t_1]$. *If $c_1\leq h_{K_t}\leq c_2$ on* $[0,t_1]$, *then $c_5\leq \kappa_i\leq c_6$ on* $[0,t_1]$, *where c_5 and c_6 depend on $K_0, c_1, c_2, p, \varphi$ and t_1.*

Proof. In view of Lemmas 6 and 7, it suffices to show that $\|W\|$ remains bounded on $[0,t_1]$. Consider the auxiliary function

$$\Theta=\tfrac{1}{2}\log(\|W\|^2)-\alpha\log h.$$

Assume without loss of generality that $c_1>1$, for otherwise we replace h by $2h/c_1$, which does not effect the evolution equation of Θ. Using the parabolic maximum principle we show that for some α large enough $\Theta(\cdot\,,t)$ is always negative on $[0,t_1]$. If the conclusion of the theorem is false, we may choose (x_0,t_0) with $t_0>0$ and such that $\Theta(x_0,t_0)=0$, $\Theta(x,t_0)\leq 0$, and $\Theta(x,t)<0$ for $t<t_0$. Then,

$$\begin{aligned}
0&\leq\dot{\Theta}-\psi\frac{\mathcal{K}^{kl}}{\mathcal{K}^2}\Theta_{;kl}\\
&=-\frac{\psi}{\|W\|^2}\frac{\mathcal{K}^{kl}}{\mathcal{K}^2}w_{i;k}^{j}w_{j;l}^{i}+\frac{2\psi}{\|W\|^4}\frac{\mathcal{K}^{kl}}{\mathcal{K}^2}w_i^jw_r^sw_{j;k}^iw_{s;l}^r+\psi\frac{\mathcal{K}^{kl}}{\mathcal{K}^2}w_{kr}w_l^r-(n+1)\psi\frac{(w^2)_i^jw_j^i}{\mathcal{K}\|W\|^2}\\
&\quad+\frac{\psi w_j^i}{\|W\|^2}\left(\frac{\mathcal{K}^{kl,rs}}{\mathcal{K}^2}w_{kl;i}g^{jp}w_{rs;p}-2\frac{g^{jp}\mathcal{K}_{;i}\mathcal{K}_{;p}}{\mathcal{K}^3}\right)+\left(\frac{2}{\mathcal{K}^2}g^{jp}\psi_{;i}\mathcal{K}_{;p}-\frac{1}{\mathcal{K}}g^{jp}\psi_{;ip}\right)\frac{w_j^i}{\|W\|^2}\\
&\quad+(n-1)\frac{\alpha\psi}{h\mathcal{K}}+\frac{\alpha}{h\mathcal{K}}(F\cdot\nabla\psi)-\frac{\alpha\psi}{h^2}\frac{\mathcal{K}^{kl}}{\mathcal{K}^2}h_{;k}h_{;l}-\alpha\psi\frac{\mathcal{K}^{kl}}{\mathcal{K}^2}w_{kr}w_l^r.
\end{aligned}$$

Pick normal coordinates around x_0 such that in (x_0,t_0) it holds that

$$g_{ij}=\delta_{ij},\quad w_{ij}=w_{ii}\delta_{ij}.$$

At (x_0,t_0) we may write

$$\mathcal{K}^{kl,rs}w_{kl;i}w_{rs;i}=\mathcal{K}^{kk,ll}w_{kk;i}w_{ll;i}-\mathcal{K}^{kk,ll}w_{kl;i}^2,$$

due to the relation

$$\mathcal{K}^{kl,rs}w_{kl;i}w_{rs;j}w^{ij}=\sum_i w_{ii}\left(\sum_{p,q}\frac{\partial^2\mathcal{K}}{\partial\kappa_p\partial\kappa_q}w_{pp;i}w_{qq;i}+\sum_{p\neq q}\frac{\frac{\partial\mathcal{K}}{\partial\kappa_p}-\frac{\partial\mathcal{K}}{\partial\kappa_q}}{\kappa_p-\kappa_q}w_{pq;i}^2\right);\tag{3-9}$$

see for example [Gerhardt 2006, Lemma 2.1.14]. We obtain after multiplication by $\mathcal{K}^2$ that

$$\begin{aligned}0\le &-\frac{\psi}{\|W\|^2}\mathcal{K}^{ii}\sum_l w^2_{ll;i}-\frac{\psi}{\|W\|^2}\mathcal{K}^{ii}\sum_{p\neq q}w^2_{pq;i}+\frac{2\psi}{\|W\|^4}\mathcal{K}^{ii}\Big(\sum_j w_{jj}w_{jj;i}\Big)^2\\ &+\psi\mathcal{K}^{ii}w^2_{ii}-(n+1)\psi\mathcal{K}\sum_i\frac{w^3_{ii}}{\|W\|^2}+\frac{\psi}{\|W\|^2}\sum_i w_{ii}\Big(\mathcal{K}^{pp,qq}w_{pp;i}w_{qq;i}-\mathcal{K}^{pp,qq}w^2_{pq;i}-2\frac{(\mathcal{K}_{;i})^2}{\mathcal{K}}\Big)\\ &+\sum_i(2\psi_{;i}\mathcal{K}_{;i}-\mathcal{K}\psi_{;ii})\frac{w_{ii}}{\|W\|^2}+(n-1)\frac{\alpha\psi\mathcal{K}}{h}+\frac{\alpha\mathcal{K}}{h}(F\cdot\nabla\psi)-\frac{\alpha\psi}{h^2}\mathcal{K}^{kl}h_{;k}h_{;l}-\alpha\psi\mathcal{K}^{ii}w^2_{ii}.\end{aligned}$$

At (x_0,t_0) we have

$$0=\Theta_{;k}=\sum_i\frac{w_{ii}w_{ii;k}}{\|W\|^2}-\alpha\frac{h_{;k}}{h}. \tag{3-10}$$

We may assume at x_0 that $w_{11}=\max\{w_{ii}:1\le i\le n\}$. Therefore,

$$\Theta(x_0,t_0)=0\quad\Longrightarrow\quad\frac{c_1^\alpha}{\sqrt{n}}\le w_{11}\le c_2^\alpha. \tag{3-11}$$

On the other hand, since ψ is bounded above and below in view of the hypotheses of the lemma, we obtain

$$\begin{aligned}\psi_{;i}\le C_0w_{ii}\quad\Longrightarrow\quad 2\psi_{;i}\mathcal{K}_{;i}&\le\frac{\varepsilon\psi}{c_4}(\mathcal{K}_{;i})^2+\frac{c_4C_0^2}{\psi\varepsilon}w^2_{ii}\\ &\le\varepsilon\psi\frac{(\mathcal{K}_{;i})^2}{\mathcal{K}}+C(\varepsilon,K_0,\varphi,t_1)\psi w^2_{ii},\end{aligned} \tag{3-12}$$

where c_4 (depending on t_1) is from Lemma 6, and

$$\psi_{;ii}\ge -C-Cw_{ii}-Cw^2_{ii}+\sum_k w_{ii;k}d_\nu\psi(\partial_k). \tag{3-13}$$

Using (3-10) in (3-13) we obtain

$$\begin{aligned}-\frac{\mathcal{K}}{\|W\|^2}\sum_i w_{ii}\psi_{;ii}&\le\frac{\mathcal{K}}{\|W\|^2}\sum_i w_{ii}\Big(C+Cw_{ii}+Cw^2_{ii}-\sum_k w_{ii;k}d_\nu\psi(\partial_k)\Big)\\ &\le\frac{\mathcal{K}}{\|W\|^2}\sum_i w_{ii}(C+Cw_{ii}+Cw^2_{ii})-\frac{\alpha\mathcal{K}}{h}\sum_k h_{;k}d_\nu\psi(\partial_k)\\ &=\frac{\mathcal{K}}{\|W\|^2}\sum_i w_{ii}(C+Cw_{ii}+Cw^2_{ii})-\frac{\alpha\mathcal{K}}{h}\sum_i w_{ii}(\partial_i\cdot F)d_\nu\psi(\partial_i)\\ &\le\frac{\psi}{\|W\|^2}\sum_i w_{ii}(C+Cw^2_{ii})-\frac{\alpha\mathcal{K}}{h}\sum_i w_{ii}(\partial_i\cdot F)d_\nu\psi(\partial_i).\end{aligned} \tag{3-14}$$

For the last inequality, we used that $\mathcal{K}$ is bounded above and ψ is bounded below (so the constant C depends on K_0,φ,t_1).

Combining (3-10), (3-12) and (3-14) implies

$$
\begin{aligned}
0 \le{}& -\frac{\psi}{\|W\|^2}\mathcal{K}^{ii}\sum_l w_{ll;i}^2 - \frac{\psi}{\|W\|^2}\mathcal{K}^{ii}\sum_{p\neq q} w_{pq;i}^2 + \frac{2\psi}{\|W\|^4}\mathcal{K}^{ii}\Big(\sum_j w_{jj}w_{jj;i}\Big)^2 + \psi\mathcal{K}^{ii}w_{ii}^2\\
&-(n+1)\psi\mathcal{K}\sum_i \frac{w_{ii}^3}{\|W\|^2} + \frac{\psi}{\|W\|^2}\sum_l w_{ll}\Big(\mathcal{K}^{pp,qq}w_{pp;l}w_{qq;l} - \mathcal{K}^{pp,qq}w_{pq;l}^2 - (2-\varepsilon)\frac{(\mathcal{K}_{;l})^2}{\mathcal{K}}\Big)\\
&+\frac{\psi}{\|W\|^2}\sum_i w_{ii}(C + Cw_{ii}^2) - \frac{\alpha\mathcal{K}}{h}\sum_i w_{ii}(\partial_i\cdot F)d_\nu\psi(\partial_i) + (n-1)\frac{\alpha\psi\mathcal{K}}{h}\\
&+\frac{\alpha\mathcal{K}}{h}\sum_s(\partial_s\cdot F)d_F\psi(\partial_s) + \frac{\alpha\mathcal{K}}{h}\sum_i w_{ii}(\partial_i\cdot F)d_\nu\psi(\partial_i) - \frac{\alpha\psi}{h^2}\mathcal{K}^{ii}w_{ii}^2(\partial_i\cdot F)^2 - \alpha\psi\mathcal{K}^{ii}w_{ii}^2\\
\le{}& \frac{\psi}{\|W\|^2}\Big(\sum_l w_{ll}(C+Cw_{ll}^2) - n\mathcal{K}\sum_l w_{ll}^3 + \mathcal{K}^{ii}w_{ii}^2\|W\|^2\Big)\\
&+\alpha\psi\Big(\frac{n\mathcal{K}}{h} - \mathcal{K}^{ii}w_{ii}^2 - \frac{\mathcal{K}^{ii}w_{ii}^2(\partial_i\cdot F)^2}{h^2} + \frac{\mathcal{K}}{h\psi}\sum_s(\partial_s\cdot F)d_F\psi(\partial_s)\Big)\\
&-\psi\sum_i(A_i+B_i+C_i+D_i-E_i) - \frac{\alpha\psi\mathcal{K}}{h} - \psi\mathcal{K}\sum_i\frac{w_{ii}^3}{\|W\|^2},
\end{aligned}
\tag{3-15}
$$

where C depends on $\varepsilon, K_0, \varphi, t_1$, and

$$A_i = \frac{2-\varepsilon}{\|W\|^2\mathcal{K}}w_{ii}(\mathcal{K}_{;i})^2 - \frac{w_{ii}}{\|W\|^2}\sum_{p,q}\mathcal{K}^{pp,qq}w_{pp;i}w_{qq;i},$$

$$B_i = \frac{2}{\|W\|^2}\sum_j w_{jj}\mathcal{K}^{jj,ii}w_{jj;i}^2, \quad C_i = \frac{2}{\|W\|^2}\sum_{j\neq i}\mathcal{K}^{jj}w_{jj;i}^2,$$

$$D_i = \frac{1}{\|W\|^2}\mathcal{K}^{ii}\sum_j w_{jj;i}^2, \qquad E_i = \frac{2}{\|W\|^4}\mathcal{K}^{ii}\Big(\sum_j w_{jj}w_{jj;i}\Big)^2.$$

The terms B_i and C_i deserve some explanation. C_i comes from the second term in (3-15), which is given by

$$-\frac{\psi}{\|W\|^2}\sum_i\mathcal{K}^{ii}\sum_{p\neq q}w_{pq;i}^2 \le -\frac{\psi}{\|W\|^2}\sum_{p\neq q}\mathcal{K}^{pp}w_{pq;p}^2 - \frac{\psi}{\|W\|^2}\sum_{p\neq q}\mathcal{K}^{qq}w_{pq;q}^2,$$

which is exactly C_i due to the Codazzi equation.

The third line of (3-15) arises from (3-9). Since the second term in the bracket of (3-9) is negative and the hypersurface is convex, we can proceed in the same way as we derived C_i and just throw away all indices i which are neither p nor q. This gives term B_i. The first term in the big bracket goes into A_i.

In Corollary 14 of the Appendix we will present an adaption of the method developed in [Guan et al. 2015] to deal with the curvature derivative terms A_i, B_i, C_i, D_i, E_i. There we prove that we obtain the following alternative: there exist positive numbers $\delta_2, \dots, \delta_n$, which only depend on the dimension and bounds on the Gauss curvature, such that either

$$w_{ii} > \delta_i w_{11} \quad \text{for all } 2 \le i \le n$$

or

$$A_i + B_i + C_i + D_i - E_i \ge 0 \quad \text{for all } 1 \le i \le n.$$

By taking α large in (3-11), in the first case we get a contradiction to the bound on the Gauss curvature. In the second case, using also $\mathcal{K}^{ii} w_{ii}^2 = \mathcal{K} \sum_i w_{ii}$, (3-15) yields

$$\begin{aligned} 0 \le \frac{\psi}{\|W\|^2}\Big(\sum_l w_{ll}(C + C w_{ll}^2) - n\mathcal{K}\sum_l w_{ll}^3\Big) - (\alpha-1)\mathcal{K}\psi \sum_i w_{ii} \\ + \alpha\psi\Big((n-1)\frac{\mathcal{K}}{h} - \frac{\mathcal{K}}{h^2}\sum_i w_{ii}(\partial_i \cdot F)^2 + \frac{\mathcal{K}}{h\psi}\sum_l (\partial_l \cdot F) d_F\psi(\partial_l)\Big). \end{aligned}$$

Consequently we obtain

$$0 \le \frac{C(\varepsilon, K_0, \varphi, t_1) w_{11}^3}{\|W\|^2} - (\alpha-1)\mathcal{K}\psi w_{11} + C(K_0, \varphi, t_1)\alpha,$$

where we discarded $-(\alpha-1)\mathcal{K}\psi \sum_{i\ne 1} w_{ii} \le 0$ and used the bounds on h, ψ and $\mathcal{K}$ to bound w_{11} in terms of w_{11}^3.

Now take α such that $(\alpha-1)\mathcal{K}\psi \ge C(\varepsilon, K_0, \varphi, t_1) + 1$. Therefore, in view of (3-11)

$$\begin{aligned} 0 &\le \frac{C(\varepsilon, K_0, \varphi, t_1) w_{11}^3}{\|W\|^2} - (\alpha-1)\mathcal{K}\psi w_{11} + C(K_0, \varphi, t_1)\alpha \\ &\le C(\varepsilon, K_0, \varphi, t_1)\Big(\frac{w_{11}^2}{\|W\|^2} - 1\Big) w_{11} - w_{11} + C(K_0, \varphi, t_1)\alpha \\ &\le -\frac{c_1^\alpha}{\sqrt{n}} + C(K_0, \varphi, t_1)\alpha. \end{aligned} \tag{3-16}$$

Taking α large enough yields a contradiction. ☐

Proposition 9. *The solution to* (1-1) *satisfies* $\lim_{t\to T} \max h_{K_t} = \infty$.

Proof. First, let $p \ge n+1$. In this case, by comparing with suitable outer balls, the flow exists on $[0, \infty)$. For $p > n+1$, consider an origin-centered ball B_r such that $K_0 \supseteq B_r$. Then $K_t \supseteq B_{r(t)}$, where

$$r(t) = ((\min h_{K_0})^{p-n-1} + t(p-n-1)\min\varphi)^{\frac{1}{p-n-1}}$$

and $B_{r(t)}$ expands to infinity as t approaches ∞. For $p = n+1$, $K_t \supseteq B_{r(t)}$ with $r(t) = e^{t\min\varphi}\min h_{K_0}$ and $B_{r(t)}$ expands to infinity as t approaches ∞.

Second, if $p < n+1$, then the flow exists only on a finite-time interval. If $\max h_{K_t} < \infty$, then by Lemmas 6, 7 and 8, the evolution equation (1-1) is uniformly parabolic on $[0, T)$. Thus, the result of [Krylov and Safonov 1980] and standard parabolic theory allow us to extend the solution smoothly past time T, contradicting its maximality. ☐

4. Convergence of normalized solutions

4.1. ***Convergence in C^1, $1 \neq p > -n-1$.*** By the proof of [Ivaki 2016, Corollary 7.5], there exist r, R such that

$$0 < r \leq h_{\widetilde{K}_t} \leq R < \infty. \tag{4-1}$$

Therefore, a subsequence of $\{\widetilde{K}_{t_k}\}$ converges in the Hausdorff distance to a limiting shape $\widetilde{K}_\infty$ with the origin in its interior. The argument of [Ivaki 2016, Section 8.1] implies

$$\varphi h^{1-p}_{\widetilde{K}_\infty} f_{\widetilde{K}_\infty} = c,$$

where $f_{\widetilde{K}_\infty}$ is the positive continuous curvature function of $\widetilde{K}_\infty$ and c is some positive constant. By [Chou and Wang 2006, Proposition 1.2], $\widetilde{K}_\infty$ is smooth and strictly convex. The C^1 convergence follows, which is purely geometric and does not depend on the evolution equation, from [Andrews 1997, Lemma 13].

Remark 10. Section 4.1 completes the discussion on the existence of solutions to the smooth, even L_p-Minkowski problems in $\mathbb{R}^{n+1}$ for $1 \neq p > -n-1$. The next section discusses the C^∞ convergence when $1 \neq p \leq n+1$, and also when $p = 1$ and solutions are origin-symmetric. We mention that in the latter case, by the proof of [Ivaki 2016, Corollary 7.5], the estimate (4-1) still holds.

4.2. ***Convergence in C^∞.*** By [Ivaki 2016, Lemma 9.2], there is a uniform upper bound on the Gauss curvature of the normalized solution when $p \leq n+1$. In the following, we first obtain a uniform lower bound on the Gauss curvature of the normalized solution $\widetilde{K}_t$.

Let $h : \mathbb{S}^n \times [0, T) \to \mathbb{R}^{n+1}$ be a solution of (1-2). Then for each $\lambda > 0$, $\bar{h}$ defined by

$$\bar{h} : \mathbb{S}^n \times [0, T/\lambda^{\frac{1+n-p}{n+1}}) \to \mathbb{R}^{n+1},$$
$$\bar{h}(u,t) = \lambda^{\frac{1}{n+1}} h(u, \lambda^{\frac{1+n-p}{n+1}} t),$$

is also a solution of evolution equation (1-2) but with the initial data $\lambda^{\frac{1}{n+1}} h(\cdot, 0)$.

For each *fixed* time $t \in [0, T)$, define $\bar{h}$ a solution of (1-2) as follows:

$$\bar{h}(u,\tau) = \left(\frac{V(B)}{V(K_t)}\right)^{\frac{1}{n+1}} h\left(u, t + \left(\frac{V(B)}{V(K_t)}\right)^{\frac{1+n-p}{n+1}} \tau\right).$$

Note that $\bar{h}(\cdot, 0)$ is the support function of $(V(B)/V(K_t))^{\frac{1}{n+1}} K_t$; therefore,

$$r \leq \bar{h}(u, 0) \leq R.$$

Write $\overline{K}_\tau$ for the convex body associated with $\bar{h}(\cdot, \tau)$ and let B_c denote the ball of radius c centered at the origin. Since B_R encloses $\overline{K}_0$, the comparison principle implies that B_{2R} will enclose $\overline{K}_\tau$ for $\tau \in [0, \delta]$, where δ depends only on p, R, ψ. By the first statement of Lemma 7 applied to $\bar{h}$, there is a uniform lower bound (depending only on r, R, p, φ) on the Gauss curvature of $\overline{K}_{\frac{\delta}{2}}$.

On the other hand, the volume of $\overline{K}_{\frac{\delta}{2}}$ is bounded above by $V(B_{2R})$; therefore,

$$\frac{V(B)}{V(B_{2R})} \le c_t := \frac{V(K_t)}{V(K_{t+(\frac{V(B)}{V(K_t)})^{(1+n-p)/(n+1)}\frac{\delta}{2}})} \le 1$$

for all $t \in [0, T)$. Consequently,

$$\left(\frac{V(B)}{V(K_{t+(\frac{V(B)}{V(K_t)})^{(1+n-p)/(n+1)}\frac{\delta}{2}})}\right)^{\frac{1}{n+1}} h\left(u, t + \left(\frac{V(B)}{V(K_t)}\right)^{\frac{1+n-p}{n+1}} \frac{\delta}{2}\right) = c_t^{\frac{1}{n+1}} \bar{h}\left(\cdot, \frac{\delta}{2}\right)$$

has Gauss curvature bounded below for all $t \in [0, T)$.

Now we show that for every $\tilde{t} \in \big[(V(B)/V(K_0))^{\frac{1+n-p}{n+1}} \frac{\delta}{2}, T\big)$, we can find $t \in [0, T)$ such that

$$\tilde{t} = t + \left(\frac{V(B)}{V(K_t)}\right)^{\frac{1+n-p}{n+1}} \frac{\delta}{2}.$$

Define $f(t) = t + (V(B)/V(K_t))^{\frac{1+n-p}{n+1}} \frac{\delta}{2} - \tilde{t}$ on $[0, T)$. Then f is continuous, and

$$\begin{cases} f(T) = T - \tilde{t} > 0, & p < n+1, \\ f(\infty) = \infty, & p = n+1, \\ f(0) \le 0, & p \le n+1. \end{cases}$$

The claim follows.

Next we obtain uniform lower and upper bounds on the principal curvatures of the normalized solution. Consider the convex bodies $\widetilde{K}_\tau := (V(B)/V(K_t))^{\frac{1}{n+1}} K_t$, where

$$\tau(t) := \int_0^t \left(\frac{V(K_s)}{V(B)}\right)^{\frac{1+n-p}{n+1}} ds.$$

Let us furnish all geometric quantities associated with $\widetilde{K}_\tau$ by a tilde. The evolution equation of $\tilde{h}_\tau$ is given by

$$\partial_\tau \tilde{h}_\tau = \varphi \tilde{h}^{2-p} \widetilde{S}_n - \frac{\int_{\mathbb{S}^n} \varphi \tilde{h}^{2-p} \widetilde{S}_n^2 \, d\sigma}{(n+1)V(B)} \tilde{h}.$$

Since $(\int_{\mathbb{S}^n} \varphi \tilde{h}^{2-p} \widetilde{S}_n^2 \, d\sigma)/((n+1)V(B))$ is uniformly bounded above, applying the maximum principle to $\Theta = \frac{1}{2}\log(\|\widetilde{W}\|^2) - \alpha \log \tilde{h}$, and arguing as in the proof of Lemma 8, we see that $\|\widetilde{W}\|$ has a uniform upper bound. This in turn, in view of our lower and upper bounds on the Gauss curvature of $\widetilde{K}_\tau$, implies that we have uniform lower and upper bounds on the principal curvatures of $\widetilde{K}_\tau$. Higher-order regularity estimates and convergence in C^∞ for a subsequence of $\{\widetilde{K}_\tau\}$ follow from [Krylov and Safonov 1980], standard parabolic theory and the Arzelà-Ascoli theorem. The convergence for the full sequence when $p \ge 1$ follows from the uniqueness of the self-similar solutions to (1-3); see [Lutwak 1993; Chou and Wang 2006]. Moreover, note that when $\varphi \equiv 1$ and $-n-1 < p < 1$, by the result of [Brendle et al. 2017], the limit is the unit sphere.

Appendix

Evolution of polar bodies. Let K be a smooth, strictly convex body with the origin in its interior. Suppose ∂K, the boundary of K, is parametrized by the radial function $r = r(u) : \mathbb{S}^n \to \mathbb{R}$. The metric $[g_{ij}]$, unit normal ν, support function h, and the second fundamental form $[w_{ij}]$ of ∂K can be written in terms of r and its partial derivatives as follows:

(a) $g_{ij} = r^2\bar{g}_{ij} + \bar{\nabla}_i r\bar{\nabla}_j r$.

(b) $\nu = (ru - \bar{\nabla} r)/\sqrt{r^2 + \|\bar{\nabla} r\|^2}$.

(c) $h = r^2/\sqrt{r^2 + \|\bar{\nabla} r\|^2}$.

(d) $w_{ij} = (-r\bar{\nabla}^2_{ij} r + 2\bar{\nabla}_i r\bar{\nabla}_j r + r^2\bar{g}_{ij})/\sqrt{r^2 + \|\bar{\nabla} r\|^2}$.

Since $1/r$ is the support function of K^*, see, e.g., [Schneider 2014, p. 57], we can calculate the entries of $[\mathfrak{r}^*_{ij}]$:

$$\mathfrak{r}^*_{ij} = \bar{\nabla}^2_{ij}\frac{1}{r} + \frac{1}{r}\bar{g}_{ij} = \frac{-r\bar{\nabla}^2_{ij} r + 2\bar{\nabla}_i r\bar{\nabla}_j r + r^2\bar{g}_{ij}}{r^3}.$$

Thus, using (d) we get

$$\mathfrak{r}^*_{ij} = \frac{\sqrt{r^2 + \|\bar{\nabla} r\|^2}}{r^3} w_{ij}.$$

Lemma 11. *As K_t evolve by* (1-2), *their polars K_t^* evolve as follows*:

$$\partial_t h^* = -\varphi\left(\frac{h^* u + \bar{\nabla} h^*}{\sqrt{h^{*2} + |\bar{\nabla} h^*|^2}}\right)\frac{(h^{*2} + |\bar{\nabla} h^*|^2)^{\frac{n+1+p}{2}}}{h^{*n+1} S_n^*}, \quad h^*(\cdot\,, t) := h_{K_t^*}(\cdot).$$

Proof. To obtain the evolution equation of $h_{K_t^*}$, we first need to parametrize M_t over the unit sphere

$$F = r(u(\cdot\,, t), t)u(\cdot\,, t) : \mathbb{S}^n \to \mathbb{R}^{n+1},$$

where $r(u(\cdot\,, t), t)$ is the radial function of M_t in the direction $u(\cdot\,, t)$. Note that

$$\partial_t r = \varphi\frac{h^{2-p}}{\mathcal{K}}\frac{\sqrt{r^2 + \|\bar{\nabla} r\|^2}}{r},$$

and

$$\mathcal{K} = \frac{\det w_{ij}}{\det g_{ij}}, \quad \frac{1}{S_n^*} = \frac{\det \bar{g}_{ij}}{\det \mathfrak{r}^*_{ij}}, \quad \frac{\det \bar{g}_{ij}}{\det g_{ij}} = \frac{1}{r^{2n-2}(r^2 + \|\bar{\nabla} r\|^2)}, \quad h = \frac{1}{\sqrt{h^{*2} + \|\bar{\nabla} h^*\|^2}}.$$

Now we calculate

$$\begin{aligned}\partial_t h^* = \partial_t\frac{1}{r} &= -\frac{h^{2-p}}{\mathcal{K}}\frac{\sqrt{r^2 + \|\bar{\nabla} r\|^2}}{r^3}\varphi(\nu)\\ &= -h^{2-p}\frac{\sqrt{r^2 + \|\bar{\nabla} r\|^2}}{r^3}\frac{\det g_{ij}}{\det w_{ij}}\varphi(\nu)\end{aligned}$$

$$= -h^{2-p} \frac{\sqrt{r^2 + \|\bar{\nabla} r\|^2}}{r^3} \frac{\det \bar{g}_{ij}}{\det \mathfrak{r}^*_{ij}} \frac{\det g_{ij}}{\det \bar{g}_{ij}} \frac{\det \mathfrak{r}^*_{ij}}{\det w_{ij}} \varphi(\nu)$$
$$= -\left(\frac{\sqrt{r^2 + \|\bar{\nabla} r\|^2}}{r^3} \right)^{n+1} \frac{r^{2n-2}(r^2 + \|\bar{\nabla} r\|^2)}{(h^{*2} + \|\bar{\nabla} h^*\|^2)^{\frac{2-p}{2}}} \frac{\varphi(\nu)}{S^*_n}.$$

Replacing r by $1/h^*$ and taking into account (b) finishes the proof. $\square$

Estimates for curvature derivatives. For convenience we present some of the main ideas regarding how one can prove the alternative in Lemma 8 about balancing the curvature derivatives. This method was used in [Guan et al. 2015] for a similar stationary prescribed curvature equation. Recall that

$$A_i = \frac{2-\varepsilon}{\|W\|^2 \mathcal{K}} w_{ii} (\mathcal{K}_{;i})^2 - \frac{w_{ii}}{\|W\|^2} \sum_{p,q} \mathcal{K}^{pp,qq} w_{pp;i} w_{qq;i},$$

$$B_i = \frac{2}{\|W\|^2} \sum_j w_{jj} \mathcal{K}^{jj,ii} w^2_{jj;i}, \quad C_i = \frac{2}{\|W\|^2} \sum_{j \neq i} \mathcal{K}^{jj} w^2_{jj;i},$$

$$D_i = \frac{1}{\|W\|^2} \mathcal{K}^{ii} \sum_j w^2_{jj;i}, \qquad E_i = \frac{2}{\|W\|^4} \mathcal{K}^{ii} \left(\sum_j w_{jj} w_{jj;i} \right)^2.$$

Note that the term A_i looks slightly different from the term A_i in [Guan et al. 2015, p. 1309], where the $\mathcal{K}$ is not present in the denominator. We have to define A_i in the way we did, because due to the inverse nature of the curvature flow equation we obtain an extra good derivative term. This allows us to choose the constant in A_i as $2-\varepsilon$, whereas a large constant was required in [Guan et al. 2015] (denoted by K there). Fortunately the proofs of Lemmas 4.2 and 4.3 in that paper also work for sufficiently small ε. The remaining terms B_i, C_i, D_i, E_i are all identical to those in [Guan et al. 2015].

In the following σ_k denotes the k-th elementary symmetric function of principal curvatures. We begin by recalling the following special case ($k = n$) of inequality (2.4) from [Guan et al. 2015, Lemma 2.2], which can be deduced easily by differentiating

$$G = \left(\frac{\sigma_n}{\sigma_l} \right)^{\frac{1}{n-l}}$$

twice, using the concavity of G and applying the Schwarz inequality. For any $\delta > 0$, $1 \le i \le n$ and $1 \le l < n$ we have

$$-\mathcal{K}^{pp,qq} w_{pp;i} w_{qq;i} + \left(1 - \frac{1}{n-l} + \frac{1}{(n-l)\delta} \right) \frac{(\mathcal{K}_{;i})^2}{\mathcal{K}} \ge \left(1 + \frac{1-\delta}{n-l} \right) \frac{\mathcal{K}((\sigma_l)_{;i})^2}{\sigma_l^2} - \frac{\mathcal{K}}{\sigma_l} \sigma_l^{pp,qq} w_{pp;i} w_{qq;i}.$$

In particular, by taking $\delta = 1/(2-\varepsilon)$, we have

$$(2-\varepsilon) \frac{(\mathcal{K}_{;i})^2}{\mathcal{K}} - \mathcal{K}^{pp,qq} w_{pp;i} w_{qq;i} \ge \left[1 + \frac{1-\varepsilon}{(n-1)(2-\varepsilon)} \right] \frac{\mathcal{K}((\sigma_l)_{;i})^2}{\sigma_l^2} - \frac{\mathcal{K} \sigma_l^{pp,qq} w_{pp;i} w_{qq;i}}{\sigma_l}, \quad \text{(A-1)}$$

provided $(2-\varepsilon) > 1$, i.e., $0 < \varepsilon < 1$.

Lemma 12. *For each $i \neq 1$, if $\sqrt{3}\kappa_i \leq \kappa_1$, we have*

$$A_i + B_i + C_i + D_i - E_i \geq 0.$$

Proof. Note that from (A-1) with $l = 1$, it follows that $A_i \geq 0$ since $\sigma_1^{pp,qq} = 0$. The proof that $B_i + C_i + D_i - E_i \geq 0$ can literally be taken from [Guan et al. 2015, Lemma 4.2], starting with (4.10) of that paper. □

In the following proof we will write $\sigma_n = \mathcal{K}$ for better comparability with [Guan et al. 2015, Lemma 4.3]. Also denote by $\sigma_k(\kappa \mid i)$ the k-th elementary symmetric polynomial in the variables $\kappa_1, \ldots, \kappa_{i-1}, \kappa_{i+1}, \ldots, \kappa_n$ and $\sigma_k(\kappa \mid ij)$ accordingly.

Lemma 13. *For $\lambda = 1, \ldots, n-1$ suppose there exists some $\delta \leq 1$ such that $\kappa_\lambda/\kappa_1 \geq \delta$. There exists a sufficiently small positive constant δ' depending on δ, ε and the bounds for $\mathcal{K}$ such that if $\kappa_{\lambda+1}/\kappa_1 \leq \delta'$, we have*

$$A_i + B_i + C_i + D_i - E_i \geq 0 \quad \text{for } i = 1, \ldots, \lambda.$$

Proof. This corresponds to [Guan et al. 2015, Lemma 4.3]. We highlight the main estimates in this proof. First of all, from [Guan et al. 2015, (4.16), (4.17)] one can extract the following estimate:

$$\begin{aligned}\|W\|^4(B_i + C_i + D_i - E_i) &\geq \|W\|^2 \sum_{j\neq i}(\sigma_{n-1}(\kappa \mid j) - 2\sigma_{n-1}(\kappa \mid ij))w_{jj;i}^2 - w_{ii}^2\sigma_n^{ii}w_{ii;i}^2 \\ &= \|W\|^2 \sum_{j\neq i}\sigma_{n-1}(\kappa \mid j)w_{jj;i}^2 - w_{ii}^2\sigma_n^{ii}w_{ii;i}^2, \qquad \text{(A-2)}\end{aligned}$$

since $\sigma_{n-1}(\kappa \mid ij) = 0$.

Now we show the right-hand side of (A-2) is dominated by $\|W\|^4 A_i$. From (A-1) we get for all $1 \leq \lambda < n$ and for all $1 \leq i \leq n$:

$$\begin{aligned}A_i &= \frac{(2-\varepsilon)w_{ii}}{\|W\|^2\sigma_n}((\sigma_n)_{;i})^2 - \frac{w_{ii}}{\|W\|^2}\sum_{p,q}\sigma_n^{pp,qq}w_{pp;i}w_{qq;i} \\ &\geq \frac{w_{ii}}{\|W\|^2}\left(1+\frac{1-\varepsilon}{(n-1)(2-\varepsilon)}\right)\frac{\sigma_n((\sigma_\lambda)_{;i})^2}{\sigma_\lambda^2} - \frac{w_{ii}}{\|W\|^2}\frac{\sigma_n\sum_{p,q}\sigma_\lambda^{pp,qq}w_{pp;i}w_{qq;i}}{\sigma_\lambda} \\ &= \frac{w_{ii}\sigma_n}{\|W\|^2\sigma_\lambda^2}\Bigg[\left(1+\frac{1-\varepsilon}{(n-1)(2-\varepsilon)}\right)\sum_a(\sigma_\lambda^{aa}w_{aa;i})^2 + \frac{1-\varepsilon}{(n-1)(2-\varepsilon)}\sum_{a\neq b}\sigma_\lambda^{aa}\sigma_\lambda^{bb}w_{aa;i}w_{bb;i} \\ &\qquad + \sum_{a\neq b}(\sigma_\lambda^{aa}\sigma_\lambda^{bb} - \sigma_\lambda\sigma_\lambda^{aa,bb})w_{aa;i}w_{bb;i}\Bigg]. \qquad \text{(A-3)}\end{aligned}$$

For sufficiently small δ' and $\lambda = 1$ the simple estimates [Guan et al. 2015, (4.19), (4.20)] give

$$\|W\|^4 A_i \geq w_{ii}^2\sigma_n^{ii}w_{11;i}^2 - C_\epsilon w_{ii}\sum_{a\neq 1}w_{aa;i}^2. \qquad \text{(A-4)}$$

Combining this with (A-2) for $i = 1$ yields,

$$\begin{aligned}\|W\|^2(A_1+B_1+C_1+D_1-E_1) &\geq \sum_{j\neq 1}\sigma_{n-1}(\kappa\mid j)w^2_{jj;1}-\frac{C_\epsilon}{w_{11}}\sum_{j\neq 1}w^2_{jj;1}\\ &=\sum_{j\neq 1}\left(\frac{\sigma_n}{w_{jj}}-\frac{C_\epsilon}{w_{11}}\right)w^2_{jj;1}\\ &\geq \sum_{j\neq 1}\left(\frac{\sigma_n}{\delta' w_{11}}-\frac{C_\epsilon}{w_{11}}\right)w^2_{jj;1},\end{aligned}\tag{A-5}$$

which is nonnegative for δ' sufficiently small. Hence the lemma is true in the case $\lambda = 1$.

For $\lambda > 1$ the series of elementary estimates [Guan et al. 2015, (4.22)–(4.27)] gives

$$\|W\|^4 A_i \geq w^2_{ii}\sigma_n^{ii}\sum_{a\leq\lambda}w^2_{aa;i}-\frac{w_{ii}C_\epsilon}{\delta^2}\sum_{a>\lambda}w^2_{aa;i},$$

after adapting ϵ if necessary and choosing δ' sufficiently small again. Combining this last inequality with (A-2) for $1 \leq i \leq \lambda$ yields

$$\begin{aligned}\|W\|^2(A_i+B_i+C_i+D_i-E_i) &\geq \sum_{j\neq i}\sigma_{n-1}(\kappa\mid j)w^2_{jj;i}-\frac{C_\epsilon}{w_{ii}\delta^2}\sum_{j>\lambda}w^2_{jj;i}\\ &\geq \sum_{j>\lambda}\left(\sigma_{n-1}(\kappa\mid j)-\frac{C_\epsilon}{w_{ii}\delta^2}\right)w^2_{jj;i}\\ &\geq \sum_{j>\lambda}\left(\frac{\sigma_n}{w_{11}\delta'}-\frac{C_\epsilon}{w_{ii}\delta^2}\right)w^2_{jj;i},\end{aligned}\tag{A-6}$$

which is nonnegative for small δ' for the same reason as in (A-5). □

Corollary 14. *There exist positive numbers $\delta_2, \dots, \delta_n$, depending only on the dimension, on ϵ and on the bounds for the Gauss curvature, such that either*

$$\kappa_i > \delta_i\kappa_1 \quad \text{for all } 2 \leq i \leq n \tag{A-7}$$

or

$$A_i + B_i + C_i + D_i - E_i \geq 0 \quad \text{for all } 1 \leq i \leq n. \tag{A-8}$$

Proof. Choosing $\lambda = 1$ and $\delta = 1$ in Lemma 13 yields the existence of δ' with the following property: if $\kappa_2/\kappa_1 \leq \delta'$, then

$$A_1 + B_1 + C_1 + D_1 - E_1 \geq 0.$$

Note that $\kappa_i \leq \kappa_2$ for $i \geq 2$. Choose $\delta_2 = \min\{\delta', 1/\sqrt{3}\}$. Therefore, in view of Lemma 12, $\kappa_2/\kappa_1 \leq \delta_2$ implies

$$A_i + B_i + C_i + D_i - E_i \geq 0 \quad \text{for all } i \geq 2.$$

We now apply induction, assuming we have constructed $\delta_2, \dots, \delta_j$. We may assume $\kappa_i > \delta_i \kappa_1$ for $2 \le i \le j$; otherwise $A_i + B_i + C_i + D_i - E_i \ge 0$ is already true for $2 \le i \le n$. Choose $\delta = \delta_j$ and $\lambda = j$ in Lemma 13 to get a δ' so that if $\kappa_{j+1} \le \delta' \kappa_1$, then $A_i + B_i + C_i + D_i - E_i \ge 0$ holds for $1 \le i \le j$. Now in view of Lemma 12, taking $\delta_{j+1} = \min\{\delta', 1/\sqrt{3}\}$ gives $A_i + B_i + C_i + D_i - E_i \ge 0$ for $j \le i \le n$. □

Acknowledgment

The work of Bryan was supported by the EPSRC on a Programme Grant entitled "Singularities of Geometric Partial Differential Equations", reference number EP/K00865X/1. The work of Ivaki was supported by Austrian Science Fund (FWF) Project M1716-N25 and the European Research Council (ERC) Project 306445.

References

[Ai et al. 2001] J. Ai, K.-S. Chou, and J. Wei, "Self-similar solutions for the anisotropic affine curve shortening problem", *Calc. Var. Partial Differential Equations* **13**:3 (2001), 311–337. MR Zbl

[Aleksandrov 1938] A. D. Aleksandrov, "On the theory of mixed volumes, III: Extensions of two theorems of Minkowski on convex polyhedra to arbitrary convex bodies", *Mat. Sb.* (*N.S.*) **3**:1 (1938), 27–47. In Russian. Zbl

[Aleksandrov 1939] A. D. Aleksandrov, "On the surface area measure of convex bodies", *Mat. Sb.* (*N.S.*) **6**:1 (1939), 167–174. In Russian. JFM

[Aleksandrov 1942] A. Alexandroff, "Smoothness of the convex surface of bounded Gaussian curvature", *C. R.* (*Doklady*) *Acad. Sci. URSS* (*N.S.*) **36** (1942), 195–199. In Russian. MR Zbl

[Andrews 1997] B. Andrews, "Monotone quantities and unique limits for evolving convex hypersurfaces", *Internat. Math. Res. Notices* **1997**:20 (1997), 1001–1031. MR Zbl

[Andrews 1998] B. Andrews, "Evolving convex curves", *Calc. Var. Partial Differential Equations* **7**:4 (1998), 315–371. MR Zbl

[Andrews 1999] B. Andrews, "Gauss curvature flow: the fate of the rolling stones", *Invent. Math.* **138**:1 (1999), 151–161. MR Zbl

[Andrews 2000] B. Andrews, "Motion of hypersurfaces by Gauss curvature", *Pacific J. Math.* **195**:1 (2000), 1–34. MR Zbl

[Andrews 2002] B. Andrews, "Singularities in crystalline curvature flows", *Asian J. Math.* **6**:1 (2002), 101–121. MR Zbl

[Andrews 2003] B. Andrews, "Classification of limiting shapes for isotropic curve flows", *J. Amer. Math. Soc.* **16**:2 (2003), 443–459. MR Zbl

[Andrews et al. 2016] B. Andrews, P. Guan, and L. Ni, "Flow by powers of the Gauss curvature", *Adv. Math.* **299** (2016), 174–201. MR Zbl

[Böröczky et al. 2013] K. J. Böröczky, E. Lutwak, D. Yang, and G. Zhang, "The logarithmic Minkowski problem", *J. Amer. Math. Soc.* **26**:3 (2013), 831–852. MR Zbl

[Brendle et al. 2017] S. Brendle, K. Choi, and P. Daskalopoulos, "Asymptotic behavior of flows by powers of the Gaussian curvature", *Acta Math.* **219**:1 (2017), 1–16. MR Zbl

[Caffarelli et al. 1984] L. Caffarelli, L. Nirenberg, and J. Spruck, "The Dirichlet problem for nonlinear second-order elliptic equations, I: Monge–Ampère equation", *Comm. Pure Appl. Math.* **37**:3 (1984), 369–402. MR Zbl

[Calabi 1958] E. Calabi, "Improper affine hyperspheres of convex type and a generalization of a theorem by K. Jörgens", *Michigan Math. J.* **5** (1958), 105–126. MR Zbl

[Chen 2006] W. Chen, "L_p Minkowski problem with not necessarily positive data", *Adv. Math.* **201**:1 (2006), 77–89. MR Zbl

[Cheng and Yau 1976] S. Y. Cheng and S. T. Yau, "On the regularity of the solution of the n-dimensional Minkowski problem", *Comm. Pure Appl. Math.* **29**:5 (1976), 495–516. MR Zbl

[Chou and Wang 2000] K.-S. Chou and X.-J. Wang, "A logarithmic Gauss curvature flow and the Minkowski problem", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **17**:6 (2000), 733–751. MR Zbl

[Chou and Wang 2006] K.-S. Chou and X.-J. Wang, "The L_p-Minkowski problem and the Minkowski problem in centroaffine geometry", *Adv. Math.* **205**:1 (2006), 33–83. MR Zbl

[Chow and Gulliver 1996] B. Chow and R. Gulliver, "Aleksandrov reflection and nonlinear evolution equations, I: The n-sphere and n-ball", *Calc. Var. Partial Differential Equations* **4**:3 (1996), 249–264. MR Zbl

[Dou and Zhu 2012] J. Dou and M. Zhu, "The two dimensional L_p Minkowski problem and nonlinear equations with negative exponents", *Adv. Math.* **230**:3 (2012), 1209–1221. MR Zbl

[Fenchel and Jessen 1938] W. Fenchel and B. Jessen, "Mengenfunktionen und konvexe Körper", *Danske Vid. Selsk. Mat. Fys. Medd.* **16**:1 (1938), 1–31. Zbl

[Gage 1993] M. E. Gage, "Evolving plane curves by curvature in relative geometries", *Duke Math. J.* **72**:2 (1993), 441–466. MR Zbl

[Gage and Li 1994] M. E. Gage and Y. Li, "Evolving plane curves by curvature in relative geometries, II", *Duke Math. J.* **75**:1 (1994), 79–98. MR Zbl

[Gerhardt 2006] C. Gerhardt, *Curvature problems*, Series in Geometry and Topology **39**, International, Somerville, MA, 2006. MR Zbl

[Gerhardt 2014] C. Gerhardt, "Non-scale-invariant inverse curvature flows in Euclidean space", *Calc. Var. Partial Differential Equations* **49**:1-2 (2014), 471–489. MR Zbl

[Guan and Lin 2000] P. Guan and C. S. Lin, "On equation $\det(u_{ij} + \delta_{ij} u) = u^p f$ on $\mathbb{S}^n$", preprint, 2000.

[Guan and Ni 2017] P. Guan and L. Ni, "Entropy and a convergence theorem for Gauss curvature flow in high dimension", *J. Eur. Math. Soc.* (*JEMS*) **19**:12 (2017), 3735–3761. MR Zbl

[Guan et al. 2015] P. Guan, C. Ren, and Z. Wang, "Global C^2-estimates for convex solutions of curvature equations", *Comm. Pure Appl. Math.* **68**:8 (2015), 1287–1325. MR Zbl

[Huang and Lu 2013] Y. Huang and Q. Lu, "On the regularity of the L_p Minkowski problem", *Adv. in Appl. Math.* **50**:2 (2013), 268–280. MR Zbl

[Hug 1996] D. Hug, "Curvature relations and affine surface area for a general convex body and its polar", *Results Math.* **29**:3-4 (1996), 233–248. MR Zbl

[Ivaki 2015] M. N. Ivaki, "An application of dual convex bodies to the inverse Gauss curvature flow", *Proc. Amer. Math. Soc.* **143**:3 (2015), 1257–1271. MR Zbl

[Ivaki 2016] M. N. Ivaki, "Deforming a hypersurface by Gauss curvature and support function", *J. Funct. Anal.* **271**:8 (2016), 2133–2165. MR Zbl

[Ivaki and Stancu 2013] M. N. Ivaki and A. Stancu, "Volume preserving centro-affine normal flows", *Comm. Anal. Geom.* **21**:3 (2013), 671–685. MR Zbl

[Jiang 2010] M.-Y. Jiang, "Remarks on the 2-dimensional L_p-Minkowski problem", *Adv. Nonlinear Stud.* **10**:2 (2010), 297–313. MR Zbl

[Jiang et al. 2011] M. Jiang, L. Wang, and J. Wei, "2π-periodic self-similar solutions for the anisotropic affine curve shortening problem", *Calc. Var. Partial Differential Equations* **41**:3-4 (2011), 535–565. MR Zbl

[Krylov and Safonov 1980] N. V. Krylov and M. V. Safonov, "A property of the solutions of parabolic equations with measurable coefficients", *Izv. Akad. Nauk SSSR Ser. Mat.* **44**:1 (1980), 161–175, 239. In Russian; translated in *Math. USSR Izv.* **16**:1 (1981), 151–164. MR Zbl

[Lewy 1938a] H. Lewy, "On differential geometry in the large, I: Minkowski's problem", *Trans. Amer. Math. Soc.* **43**:2 (1938), 258–270. MR Zbl

[Lewy 1938b] H. Lewy, "On the existence of a closed convex surface realizing a given Riemannian metric", *Proc. Nat. Acad. Sci. USA* **24**:2 (1938), 104–106. Zbl

[Lu and Wang 2013] J. Lu and X.-J. Wang, "Rotationally symmetric solutions to the L_p-Minkowski problem", *J. Differential Equations* **254**:3 (2013), 983–1005. MR Zbl

[Lutwak 1993] E. Lutwak, "The Brunn–Minkowski–Firey theory, I: Mixed volumes and the Minkowski problem", *J. Differential Geom.* **38**:1 (1993), 131–150. MR Zbl

[Lutwak and Oliker 1995] E. Lutwak and V. Oliker, "On the regularity of solutions to a generalization of the Minkowski problem", *J. Differential Geom.* **41**:1 (1995), 227–246. MR Zbl

[Lutwak et al. 2004] E. Lutwak, D. Yang, and G. Zhang, "On the L_p-Minkowski problem", *Trans. Amer. Math. Soc.* **356**:11 (2004), 4359–4370. MR Zbl

[McCoy 2003] J. McCoy, "The surface area preserving mean curvature flow", *Asian J. Math.* **7**:1 (2003), 7–30. MR Zbl

[Minkowski 1897] H. Minkowski, "Allgemeine Lehrsätze über die convexen Polyeder", *Nachrichten Akad. Wiss. Göttingen Math.-Phys. Kl.* **1897** (1897), 198–219. JFM

[Minkowski 1903] H. Minkowski, "Volumen und Oberfläche", *Math. Ann.* **57**:4 (1903), 447–495. MR Zbl

[Nirenberg 1953] L. Nirenberg, "The Weyl and Minkowski problems in differential geometry in the large", *Comm. Pure Appl. Math.* **6** (1953), 337–394. MR Zbl

[Pogorelov 1952] A. V. Pogorelov, "Regularity of a convex surface with given Gaussian curvature", *Mat. Sbornik N.S.* **31(73)** (1952), 88–103. In Russian. MR Zbl

[Pogorelov 1971] A. V. Pogorelov, "A regular solution of the n-dimensional Minkowski problem", *Soviet Math. Dokl.* **12** (1971), 1192–1196. In Russian. MR Zbl

[Schneider 2014] R. Schneider, *Convex bodies: the Brunn–Minkowski theory*, Encyclopedia of Mathematics and its Applications **151**, Cambridge University Press, 2014. MR Zbl

[Schnürer 2006] O. C. Schnürer, "Surfaces expanding by the inverse Gauß curvature flow", *J. Reine Angew. Math.* **600** (2006), 117–134. MR Zbl

[Stancu 1996] A. Stancu, "Uniqueness of self-similar solutions for a crystalline flow", *Indiana Univ. Math. J.* **45**:4 (1996), 1157–1174. MR Zbl

[Stancu 2002] A. Stancu, "The discrete planar L_0-Minkowski problem", *Adv. Math.* **167**:1 (2002), 160–174. MR Zbl

[Stancu 2003] A. Stancu, "On the number of solutions to the discrete two-dimensional L_0-Minkowski problem", *Adv. Math.* **180**:1 (2003), 290–323. MR Zbl

[Stancu 2012] A. Stancu, "Centro-affine invariants for smooth convex bodies", *Int. Math. Res. Not.* **2012**:10 (2012), 2289–2320. MR Zbl

[Tsai 2005] D.-H. Tsai, "Behavior of the gradient for solutions of parabolic equations on the circle", *Calc. Var. Partial Differential Equations* **23**:3 (2005), 251–270. MR Zbl

[Tso 1985] K. Tso, "Deforming a hypersurface by its Gauss–Kronecker curvature", *Comm. Pure Appl. Math.* **38**:6 (1985), 867–882. MR Zbl

[Umanskiy 2003] V. Umanskiy, "On solvability of two-dimensional L_p-Minkowski problem", *Adv. Math.* **180**:1 (2003), 176–186. MR Zbl

[Urbas 1998] J. Urbas, "Complete noncompact self-similar solutions of Gauss curvature flows, I: Positive powers", *Math. Ann.* **311**:2 (1998), 251–274. MR Zbl

[Urbas 1999] J. Urbas, "Complete noncompact self-similar solutions of Gauss curvature flows, II: Negative powers", *Adv. Differential Equations* **4**:3 (1999), 323–346. MR Zbl

[Zhu 2014] G. Zhu, "The logarithmic Minkowski problem for polytopes", *Adv. Math.* **262** (2014), 909–931. MR Zbl

[Zhu 2015a] G. Zhu, "The centro-affine Minkowski problem for polytopes", *J. Differential Geom.* **101**:1 (2015), 159–174. MR Zbl

[Zhu 2015b] G. Zhu, "The L_p Minkowski problem for polytopes for $0 < p < 1$", *J. Funct. Anal.* **269**:4 (2015), 1070–1094. MR Zbl

Received 9 Aug 2016. Accepted 1 May 2018.

PAUL BRYAN: paul.bryan@mq.edu.au
Department of Mathematics, Macquarie University, Sydney, Australia

MOHAMMAD N. IVAKI: m.ivaki@utoronto.ca
Department of Mathematics, University of Toronto, Toronto, ON, Canada

JULIAN SCHEUER: julian.scheuer@math.uni-freiburg.de
Mathematisches Institut, Albert-Ludwigs-Universität, Freiburg, Germany

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.281

THE MUSKAT PROBLEM IN TWO DIMENSIONS: EQUIVALENCE OF FORMULATIONS, WELL-POSEDNESS, AND REGULARITY RESULTS

BOGDAN-VASILE MATIOC

We consider the Muskat problem describing the motion of two unbounded immiscible fluid layers with equal viscosities in vertical or horizontal two-dimensional geometries. We first prove that the mathematical model can be formulated as an evolution problem for the sharp interface separating the two fluids, which turns out to be, in a suitable functional-analytic setting, quasilinear and of parabolic type. Based upon these properties, we then establish the local well-posedness of the problem for arbitrary large initial data and show that the solutions become instantly real-analytic in time and space. Our method allows us to choose the initial data in the class H^s, $s \in \left(\frac{3}{2}, 2\right)$, when neglecting surface tension, respectively in H^s, $s \in (2, 3)$, when surface-tension effects are included. Besides, we provide new criteria for the global existence of solutions.

1. Introduction and the main results

The Muskat problem [1934] is a classical model describing the motion of two immiscible fluids in a porous medium or a Hele-Shaw cell. We consider here the particular case when the fluids have equal viscosities and we assume that the flows are two-dimensional. Furthermore, we consider an unbounded geometry corresponding to fluid layers that occupy the entire space, the fluid motion being localized and the fluid system close to the rest state far away from the origin. We further assume that the fluids are separated by a sharp interface which flattens out at infinity, evolves in time, and is unknown. We consider two different scenarios for this unconfined Muskat problem:

MSC2010: 35R37, 35K59, 35K93, 35Q35, 42B20.
Keywords: Muskat problem, surface tension, singular integral.

(a) In the absence of surface-tension effects at the free boundary, the Hele-Shaw cell is vertical and the fluid located below is more dense.

(b) In the presence of surface-tension effects, the Hele-Shaw cell is either vertical or horizontal and we make no restrictions on the densities of the fluids.

One big advantage of considering this setting is that the equations of motion can be very elegantly formulated as a single evolution equation for the interface between the fluids. Indeed, parametrizing this interface as the graph $[y=f(t,x)]$, the Muskat problem is equivalent in this setting to an evolution problem for the unknown function f, see Section 2, and it can be written as

$$\begin{cases} \partial_t f(t,x) = \dfrac{\sigma k}{2\pi\mu} f'(t,x)\,\mathrm{PV}\displaystyle\int_{\mathbb{R}} \frac{f(t,x)-f(t,x-y)}{y^2+(f(t,x)-f(t,x-y))^2}(\kappa(f))'(t,x-y)\,dy \\ \qquad\qquad + \dfrac{\sigma k}{2\pi\mu}\,\mathrm{PV}\displaystyle\int_{\mathbb{R}} \frac{y}{y^2+(f(t,x)-f(t,x-y))^2}(\kappa(f))'(t,x-y)\,dy \\ \qquad\qquad + \dfrac{\Delta_\rho k}{2\pi\mu}\,\mathrm{PV}\displaystyle\int_{\mathbb{R}} \frac{y(f'(t,x)-f'(t,x-y))}{y^2+(f(t,x)-f(t,x-y))^2}\,dy \qquad \text{for } t>0, \quad x\in\mathbb{R}, \\ f(0,\cdot)=f_0. \end{cases} \tag{1-1}$$

For brevity we write f' for the spatial derivative $\partial_x f$. We let k denote the permeability of the homogeneous porous medium, μ is the viscosity coefficient of the fluids, σ is the surface-tension coefficient at the free boundary, and

$$\Delta_\rho := g(\rho_- - \rho_+),$$

where g is the Earth's gravity and $\rho_\pm$ is the density of the fluid which occupies the domain $\Omega_\pm(t)$ defined by

$$\Omega_-(t) := [y<f(t,x)] \quad \text{and} \quad \Omega_+(t) := [y>f(t,x)].$$

Furthermore, $\kappa(f(t))$ is the curvature of the graph $[y=f(t,x)]$ and PV denotes the principal value which, depending on the regularity of the functions under the integral, is taken at zero and/or at infinity. Our analysis covers the following scenarios

$$\text{(a) } \sigma=0,\ \Delta_\rho>0 \quad \text{and} \quad \text{(b) } \sigma>0,\ \Delta_\rho\in\mathbb{R},$$

meaning that (a) corresponds to the stable case when the denser fluid is located below.

Due to its physical relevance [Bear 1972], the Muskat problem has been widely studied in the last decades in several geometries and physical settings and with various methods. When neglecting surface-tension effects the well-posedness of the Muskat problem is in strong relationship with the Rayleigh–Taylor condition, being implied by the latter. The Rayleigh–Taylor condition, which appears first in [Saffman and Taylor 1958], is a sign restriction on the jump of the pressure gradient in normal direction at the free boundary. For fluids with equal viscosities moving in a vertical geometry, it reduces to the simple relation

$$\Delta_\rho > 0;$$

see, e.g., [Córdoba and Gancedo 2010; Escher et al. 2018] and also (2-1a)–(2-1b). The first local existence result was established in [Yi 1996] by using Newton's iteration method; the analysis in [Ambrose 2004;

Berselli et al. 2014; Cheng et al. 2016; Gómez-Serrano and Granero-Belinchón 2014; Constantin et al. 2017; Córdoba et al. 2011; 2013; 2014; Córdoba and Gancedo 2007; 2010] is based on energy estimates and the energy method; the authors of [Siegel et al. 2004] use methods from complex analysis and a version of the Cauchy–Kowalewski theorem; a fixed-point argument is employed in [Bazaliy and Vasylyeva 2014] for nonregular initial data, and the approach in [Escher and Matioc 2011; Escher et al. 2012; 2018] relies on the formulation of the problem as a nonlinear and nonlocal parabolic equation together with an abstract well-posedness result from [Da Prato and Grisvard 1979] based on continuous maximal regularity. Other papers study the qualitative aspects of solutions to the Muskat problem for fluids with equal viscosities, such as global existence of strong and weak solutions [Constantin et al. 2013; 2017; Granero-Belinchón 2014], existence of initial data for which solutions turn over [Castro et al. 2011; 2012; 2013], and the absence of squirt or splash singularities [Córdoba and Gancedo 2010; Gancedo and Strain 2014].

Compared to the zero-surface-tension case, the Muskat problem with surface tension is less well-studied. When allowing for surface tension, the Rayleigh–Taylor condition is no longer needed and the problem is well-posed for general initial data. While some of the references require quite high regularity from the initial data, see [Ambrose 2014; Friedman and Tao 2003; Hong et al. 1997; Tofts 2017], optimal results are established in bounded or periodic geometries under the observation that the Muskat problem with surface tension can be formulated as a quasilinear parabolic evolution problem; see [Escher et al. 2018; Prüss and Simonett 2016a].

The stability properties of equilibria which are, depending on the physical scenario, horizontal lines [Cheng et al. 2016; Ehrnström et al. 2013; Escher et al. 2012; Escher and Matioc 2011], finger-shaped [Ehrnström et al. 2013; Escher et al. 2012; Escher and Matioc 2011], circular [Friedman and Tao 2003], or a union of disjoint circles/spheres [Prüss and Simonett 2016b] have been also addressed in the references just mentioned.

In this paper we first rigorously prove in Section 2 that the Muskat problem in the classical formulation (2-1) and the system (1-1) are equivalent for a certain class of solutions. Thereafter, the analysis of (1-1) starts from the obvious observation that the right-hand side of the first equation of (1-1) is linear with respect to the highest-order spatial derivative of f; that is, this particular Muskat problem has a quasilinear structure (also when neglecting surface tension). This property is not obvious in the particular geometry considered in [Escher et al. 2018; Yi 1996] (when $\sigma = 0$). In a suitable functional-analytic setting we then prove that (1-1) is additionally parabolic for general initial data. The parabolic character was established previously for bounded geometries [Escher et al. 2012; 2018; Escher and Matioc 2011; Prüss and Simonett 2016a; 2016b] (in the absence of surface-tension effects only when the Rayleigh–Taylor condition holds), but for (1-1) only for small initial data; see [Constantin et al. 2013; Córdoba and Gancedo 2007]. These two aspects, that is, the quasilinearity and the parabolicity, enable us to use abstract results for quasilinear parabolic problems due to H. Amann [1993, Section 12] to prove, by similar strategies, the well-posedness of the Muskat problem with and without surface tension.

It is worth emphasizing that for this particular Muskat problem the local well-posedness is established, in the zero-surface-tension case, only for initial data that are twice-weakly differentiable and which belong to $W^2_p(\mathbb{R})$ for some $p \in (1, \infty]$; see [Constantin et al. 2017]. Our first main result, i.e., Theorem 1.1, extends the local well-posedness to general initial data in $H^s(\mathbb{R})$ with $s \in \left(\frac{3}{2}, 2\right)$. For the unconfined Muskat problem with surface tension, the well-posedness is considered only in [Ambrose 2014; Tofts 2017] and in both papers the authors require that $f_0 \in H^s(\mathbb{X})$, with $\mathbb{X} \in \{\mathbb{R}, \mathbb{T}\}$ and $s \geq 6$. In our well-posedness result, i.e., Theorem 1.2, the curvature of the initial data may be even unbounded as we allow for general initial data $f_0 \in H^s(\mathbb{R})$ with $s \in (2, 3)$. Additionally, we also obtain new criteria for the existence of global solutions to the Muskat problem with and without surface tension and, as a consequence of the parabolic character of the equations, we show that the fluid interfaces become instantly real-analytic.

Our strategy is the following: we formulate, in a suitable functional-analytic setting, (1-1) as a quasilinear evolution problem of the form[1]

$$\dot{f} = \Phi_\sigma(f)[f], \quad t > 0, \quad f(0) = f_0,$$

and then we study the properties of the operator Φ_σ. We differentiate between the case $\sigma = 0$, studied in Sections 3–5, when we simply write $\Phi_\sigma =: \Phi$, and the case $\sigma > 0$, as in the first case $\Phi(f)$ is a nonlocal operator of order 1 and in the second case $\Phi_\sigma(f)$ has order 3 (for f appropriately chosen). At the core of our estimates lies the following deep result from harmonic analysis: given a Lipschitz function $a : \mathbb{R} \to \mathbb{R}$, the singular integral operator

$$h \mapsto \left[x \mapsto \mathrm{PV}\int_{\mathbb{R}} \frac{h(x-y)}{y} \exp\left(i \frac{a(x) - a(x-y)}{y} \right) dy \right] \tag{1-2}$$

belongs to $\mathcal{L}(L_2(\mathbb{R}))$ and its norm is bounded by $C(1 + \|a'\|_\infty)$, see [Murai 1986], with C denoting a universal constant independent of a. Relying on (1-2), we study the mapping properties of Φ_σ and show, for suitable f, that $\Phi_\sigma(f)$ is the generator of a strongly continuous and real-analytic semigroup. The main results of this paper, that is, Theorems 1.1–1.3, are then obtained by employing abstract results presented in [Amann 1993, Section 12], and which we briefly recall at the end of this section. The line of approach is close to the one we followed in [Escher et al. 2018]; however, the functional-analytic setting and the methods used to establish the needed estimates are substantially different. We expect that our method extends to the general case when $\mu_- \neq \mu_+$ and we believe to obtain, for periodic flows, a similar stability behavior of the — flat and finger shaped — equilibria, as in [Escher and Matioc 2011].

Our first main result is the following well-posedness theorem for the Muskat problem without surface-tension effects.

Theorem 1.1 (well-posedness: no surface tension). *Let $\sigma = 0$ and $\Delta_\rho > 0$. The problem* (1-1) *possesses for each $f_0 \in H^s(\mathbb{R})$, $s \in \left(\frac{3}{2}, 2\right)$, a unique maximal classical solution*

$$f := f(\cdot\,; f_0) \in C([0, T_+(f_0)), H^s(\mathbb{R})) \cap C((0, T_+(f_0)), H^2(\mathbb{R})) \cap C^1((0, T_+(f_0)), H^1(\mathbb{R})),$$

with $T_+(f_0) \in (0, \infty]$, and $[(t, f_0) \mapsto f(t; f_0)]$ defines a semiflow on $H^s(\mathbb{R})$. Additionally, if

$$\sup_{[0,T_+(f_0))\cap[0,T]} \|f(t)\|_{H^s} < \infty \quad \text{for all } T > 0,$$

then $T_+(f_0) = \infty$.

[1] We write $\dot{f}$ to denote the derivative df/dt.

The quasilinear character of the problem is enhanced by the presence of surface tension. For this reason we may consider, when $\sigma > 0$, initial data with unbounded curvature. We show in Theorem 1.3, however, that the curvature becomes instantly real-analytic and bounded.

Theorem 1.2 (well-posedness: with surface tension). *Let $\sigma > 0$ and $\Delta_\rho \in \mathbb{R}$. The problem* (1-1) *possesses for each $f_0 \in H^s(\mathbb{R})$, $s \in (2, 3)$, a unique maximal classical solution*

$$f := f(\cdot\,; f_0) \in C([0, T_+(f_0)), H^s(\mathbb{R})) \cap C((0, T_+(f_0)), H^3(\mathbb{R})) \cap C^1((0, T_+(f_0)), L_2(\mathbb{R})),$$

with $T_+(f_0) \in (0, \infty]$, and $[(t, f_0) \mapsto f(t; f_0)]$ defines a semiflow on $H^s(\mathbb{R})$. Additionally, if

$$\sup_{[0,T_+(f_0))\cap[0,T]} \|f(t)\|_{H^s} < \infty \quad \textit{for all } T > 0,$$

then $T_+(f_0) = \infty$.

These results reflect the fact that the Muskat problem without surface tension is a first-order evolution problem, while the Muskat problem with surface tension is of third order. The solutions obtained in Theorems 1.1 and 1.2 become instantly real-analytic.

Theorem 1.3. *Let $s \in \left(\frac{3}{2}, 2\right)$ if $\sigma = 0$ and $\Delta_\rho > 0$, and let $s \in (2, 3)$ if $\sigma > 0$. Given $f_0 \in H^s(\mathbb{R})$, let $f = f(\cdot\,; f_0)$ denote the unique maximal solution to* (1-1) *found in Theorems 1.1 and 1.2, respectively. Then*

$$[(t, x) \mapsto f(t, x)] : (0, T_+(f_0)) \times \mathbb{R} \to \mathbb{R}$$

is a real-analytic function. In particular, $f(t, \cdot)$ is real-analytic for each $t \in (0, T_+(f_0))$. Moreover, given $k \in \mathbb{N}$, it holds that

$$f \in C^\omega((0, T_+(f_0)), H^k(\mathbb{R})),$$

where C^ω denotes real-analyticity.

As a direct consequence of Theorems 1.1 and 1.3 and of [Constantin et al. 2013, Theorem 3.1], see also [Constantin et al. 2016, Remark 6.2], we obtain a global existence result for solutions to the Muskat problem without surface tension that correspond to initial data of medium size in $H^s(\mathbb{R})$, $s \in \left(\frac{3}{2}, 2\right)$. In the following $\mathcal{F}$ denotes the Fourier transform.

Corollary 1.4. *There exists a constant $c_0 \geq \frac{1}{5}$ such that for all $f_0 \in H^s(\mathbb{R})$, $s \in \left(\frac{3}{2}, 2\right)$, with*

$$|||f_0||| := \int_{\mathbb{R}} |\xi|\,|\mathcal{F} f_0(\xi)|\, d\xi < c_0,$$

the solution found in Theorem 1.1 exists globally.

Proof. The claim follows from the inequality

$$|||f||| = \int_{\mathbb{R}} |\xi|\,|\mathcal{F} f(\xi)|\, d\xi \leq \|f\|_{H^s} \int_{\mathbb{R}} \frac{1}{(1 + |\xi|^2)^{s-1}}\, d\xi \leq C\|f\|_{H^s}$$

for $s \in \left(\frac{3}{2}, 2\right)$ and $f \in H^s(\mathbb{R})$. ☐

An abstract setting for quasilinear parabolic evolution equations. In Theorem 1.5 we collect abstract results from [Amann 1993, Section 12] for a general class of abstract quasilinear parabolic evolution equations, which we use in an essential way in our analysis.

Given Banach spaces $\mathbb{E}_0, \mathbb{E}_1$ with dense embedding $\mathbb{E}_1 \hookrightarrow \mathbb{E}_0$, we define $\mathcal{H}(\mathbb{E}_1, \mathbb{E}_0)$ as the subset of $\mathcal{L}(\mathbb{E}_1, \mathbb{E}_0)$ consisting of negative generators of strongly continuous analytic semigroups. More precisely, $\mathbb{A} \in \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0)$ if $-\mathbb{A}$, considered as an unbounded operator in $\mathbb{E}_0$ with domain $\mathbb{E}_1$, generates a strongly continuous and analytic semigroup in $\mathcal{L}(\mathbb{E}_0)$.

Theorem 1.5. *Let $\mathbb{E}_0, \mathbb{E}_1$ be Banach spaces with dense embedding $\mathbb{E}_1 \hookrightarrow \mathbb{E}_0$ and let $\mathbb{E}_\theta := [\mathbb{E}_0, \mathbb{E}_1]_\theta$ for $0 < \theta < 1$ be endowed with the $\|\cdot\|_\theta$-norm. Let further $0 < \beta < \alpha < 1$ and assume that*

$$-\Phi \in C^{1-}(\mathcal{O}_\beta, \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0)), \tag{1-3}$$

where $\mathcal{O}_\beta$ denotes an open subset of $\mathbb{E}_\beta$ and C^{1-} stands for local Lipschitz continuity. The following assertions hold for the quasilinear evolution problem

$$\dot{f} = \Phi(f)[f], \quad t > 0, \quad f(0) = f_0. \tag{QP}$$

Existence: *given $f_0 \in \mathcal{O}_\alpha := \mathcal{O}_\beta \cap \mathbb{E}_\alpha$, the problem* (QP) *possesses a maximal solution*

$$f := f(\cdot\,; f_0) \in C([0, T_+(f_0)), \mathcal{O}_\alpha) \cap C((0, T_+(f_0)), \mathbb{E}_1) \cap C^1((0, T_+(f_0)), \mathbb{E}_0) \cap C^{\alpha-\beta}([0, T], \mathbb{E}_\beta)$$

for all $T \in (0, T_+(f_0))$, with $T_+(f_0) \in (0, \infty]$.

Uniqueness: *if $\widetilde{T} \in (0, \infty]$, $\eta \in (0, \alpha - \beta]$, and $\tilde{f} \in C((0, \widetilde{T}), \mathbb{E}_1) \cap C^1((0, \widetilde{T}), \mathbb{E}_0)$ satisfies*

$$\tilde{f} \in C^\eta([0, T], \mathbb{E}_\beta) \quad \textit{for all } T \in (0, \widetilde{T})$$

and solves (QP), *then $\widetilde{T} \le T_+(f_0)$ and $\tilde{f} = f$ on $[0, \widetilde{T})$.*

Criterion for global existence: *if $f : [0, T] \cap [0, T_+(f_0)) \to \mathcal{O}_\alpha$ is uniformly continuous for all $T > 0$, then*

$$T_+(f_0) = \infty \quad \textit{or} \quad T_+(f_0) < \infty \ \textit{ and } \ \operatorname{dist}(f(t), \partial\mathcal{O}_\alpha) \to 0 \ \textit{ for } t \to T_+(f_0).$$

Continuous dependence of initial data: *the mapping $[(t, f_0) \mapsto f(t; f_0)]$ defines a semiflow on $\mathcal{O}_\alpha$ and, if $\Phi \in C^\omega(\mathcal{O}_\beta, \mathcal{L}(\mathbb{E}_1, \mathbb{E}_0))$, then*

$$[(t, f_0) \mapsto f(t; f_0)] : \{(t, f_0) : f_0 \in \mathcal{O}_\alpha, t \in (0, T_+(f_0))\} \to \mathbb{E}_\alpha$$

is a real-analytic map too.

As usual, $[\cdot\,,\cdot]_\theta$ denotes the complex interpolation functor. We choose for our particular problem $\mathbb{E}_i \in \{H^s(\mathbb{R}) : 0 \le s \le 3\}$, $i = 1, 2$, and in this context we rely on the well-known interpolation property

$$[H^{s_0}(\mathbb{R}), H^{s_1}(\mathbb{R})]_\theta = H^{(1-\theta)s_0+\theta s_1}(\mathbb{R}), \quad \theta \in (0, 1), \ \ -\infty < s_0 \le s_1 < \infty; \tag{1-4}$$

see, e.g., [Triebel 1978, Remark 2, Section 2.4.2].

The proof of Theorem 1.5 uses to a large extent the linear theory developed in [Amann 1995, Chapter II]. The main ideas of the proof of Theorem 1.5 can be found in [Amann 1986; 1988]. The uniqueness claim in Theorem 1.5 is slightly stronger compared to the result in [Amann 1993, Section 12] and it turns out to be quite useful when establishing the uniqueness in Theorems 1.1–1.2. For this reason we present in Appendix B the proof of Theorem 1.5.

In order to use Theorem 1.5 in the study of the Muskat problem (1-1), we have to write this evolution problem in the form (QP) and to establish then the property (1-3). With respect to this goal, we use the estimate provided in (1-2) and many techniques of nonlinear analysis.

2. The equations of motion and the equivalent formulation

We present the equations governing the dynamic of the fluids system and we prove, for a certain class of solutions, that the latter are equivalent to the system (1-1). The Muskat problem was originally proposed as a model for the encroachment of water into an oil sand, and therefore it is natural to assume that both fluids are incompressible, of Newtonian type, and immiscible. Since for flows in porous media the conservation of momentum equation can be replaced by Darcy's law, see, e.g., [Bear 1972], the equations governing the dynamic of the fluids are

$$\begin{cases} \operatorname{div} v_\pm(t) = 0 & \text{in } \Omega_\pm(t), \\ v_\pm(t) = -(k/\mu)(\nabla p_\pm(t) + (0, \rho_\pm g)) & \text{in } \Omega_\pm(t) \end{cases} \tag{2-1a}$$

for $t > 0$, where, using the subscript $\pm$ for the fluid located at $\Omega_\pm(t)$, we denote by $v_\pm(t) := (v_\pm^1(t), v_\pm^2(t))$ the velocity vector and $p_\pm(t)$ the pressure of the fluid $\pm$. These equations are supplemented by the natural boundary conditions at the free surface

$$\begin{cases} p_+(t) - p_-(t) = \sigma\kappa(f(t)) & \text{on } [y=f(t,x)], \\ \langle v_+(t) \mid \nu(t)\rangle = \langle v_-(t) \mid \nu(t)\rangle & \text{on } [y=f(t,x)], \end{cases} \tag{2-1b}$$

where $\nu(t)$ is the unit normal at $[y=f(t,x)]$ pointing into $\Omega_+(t)$ and $\langle \cdot \mid \cdot \rangle$ is the inner product on $\mathbb{R}^2$. Furthermore, the far-field boundary conditions

$$\begin{cases} f(t,x) \to 0 & \text{for } |x| \to \infty, \\ v_\pm(t,x,y) \to 0 & \text{for } |(x,y)| \to \infty \end{cases} \tag{2-1c}$$

state that the fluid motion is localized, the fluids being close to the rest state far away from the origin. The motion of the interface $[y=f(t,x)]$ is coupled to that of the fluids through the kinematic boundary condition

$$\partial_t f(t) = \langle v_\pm(t) \mid (-f'(t), 1)\rangle \quad \text{on } [y=f(t,x)]. \tag{2-1d}$$

Finally, the interface at time $t = 0$ is assumed to be known,

$$f(0) = f_0. \tag{2-1e}$$

The equations (2-1) are known as the Muskat problem and they determine completely the dynamic of the system. We now show that the Muskat problem (2-1) is equivalent to the system (1-1) presented in the Introduction. The proof uses classical results on Cauchy-type integrals defined on regular curves; see, e.g., [Lu 1993]. More precisely, we establish the following equivalence result.

Proposition 2.1 (equivalence of the two formulations). *Let* $\sigma \geq 0$ *and* $T \in (0, \infty]$. *The following are equivalent*:

(i) *The Muskat problem* (2-1) *for*[2]

$$f \in C^1((0,T), L_2(\mathbb{R})) \cap C([0,T), L_2(\mathbb{R})), \qquad f(t) \in H^5(\mathbb{R}) \quad \textit{for all } t \in (0,T),$$
$$v_\pm(t) \in C(\overline{\Omega_\pm(t)}) \cap C^1(\Omega_\pm(t)), \quad p_\pm(t) \in C^1(\overline{\Omega_\pm(t)}) \cap C^2(\Omega_\pm(t)) \quad \textit{for all } t \in (0,T).$$

(ii) *The evolution problem* (1-1) *for*

$$f \in C^1((0,T), L_2(\mathbb{R})) \cap C([0,T), L_2(\mathbb{R})), \qquad f(t) \in H^5(\mathbb{R}) \quad \textit{for all } t \in (0,T).$$

Proof. We first establish the implication (i) $\Rightarrow$ (ii). Assuming that we are given a solution to (2-1) as in (i), we have to show that the first equation of (1-1) holds for each $t \in (0,T)$. Therefore, we fix $t \in (0,T)$ and we do not write in the arguments that follow the dependence of the physical variables of time t explicitly. In the following, $\mathbf{1}_E$ is the characteristic function of the set E. Introducing the global velocity field $v := (v^1, v^2) := v_- \mathbf{1}_{[y \leq f(x)]} + v_+ \mathbf{1}_{[y > f(x)]}$, Stokes' theorem together with (2-1a) and (2-1b) yields that the vorticity, which for two-dimensional flows corresponds to the scalar function $\omega := \partial_x v^2 - \partial_y v^1$, is supported on the free boundary, that is,

$$\langle \omega, \varphi \rangle = \int_{\mathbb{R}} \bar{\omega}(x) \varphi(x, f(x))\, dx \quad \text{for all } \varphi \in C_0^\infty(\mathbb{R}^2),$$

where

$$\bar{\omega} := \frac{k}{\mu} [\sigma \kappa(f) - \Delta_\rho f]'.$$

We next prove that the velocity is defined by the Biot–Savart law, that is, $v = \tilde{v}$ in $\mathbb{R}^2 \setminus [y{=}f(x)]$, where

$$\tilde{v}(x,y) := \frac{1}{2\pi} \int_{\mathbb{R}} \frac{(-(y - f(s)), x - s)}{(x-s)^2 + (y - f(s))^2} \bar{\omega}(s)\, ds \quad \text{in } \mathbb{R}^2 \setminus [y{=}f(x)]. \tag{2-2}$$

To this end we compute the limits $\tilde{v}_-(x, f(x))$ and $\tilde{v}_+(x, f(x))$ of $\tilde{v}$ at $(x, f(x))$ when we approach this point from below the interface $[y{=}f(x)]$ or from above, respectively. Using the well-known Plemelj formula, see, e.g., [Lu 1993], due to the fact that $f \in H^4(\mathbb{R})$ and after changing variables, we find the expressions

$$\tilde{v}_\pm(x, f(x)) = \frac{1}{2\pi} \mathrm{PV}\!\int_{\mathbb{R}} \frac{(-(f(x) - f(x-s)), s)}{s^2 + (f(x) - f(x-s))^2} \bar{\omega}(x-s)\, ds \mp \frac{1}{2} \frac{(1, f'(x))\bar{\omega}(x)}{1 + f'^2(x)}, \quad x \in \mathbb{R}, \tag{2-3}$$

where the principal value needs to be taken only at 0. In view of Lemma A.2 and of $f \in H^5(\mathbb{R})$, the restrictions $\tilde{v}_\pm$ of $\tilde{v}$ to $\Omega_\pm$ satisfy $\tilde{v}_\pm \in C(\overline{\Omega}_\pm) \cap C^1(\Omega_\pm)$ and moreover $\tilde{v}_\pm$ vanish at infinity. Next, we define the pressures $\tilde{p}_\pm \in C^1(\overline{\Omega}_\pm) \cap C^2(\Omega_\pm)$ by the formula

$$\tilde{p}_\pm(x,y) := c_\pm - \frac{\mu}{k} \int_0^x \tilde{v}^1_\pm(s, \pm d)\, ds - \frac{\mu}{k} \int_{\pm d}^y \tilde{v}^2_\pm(x, s)\, ds - \rho_\pm g y, \quad (x,y) \in \overline{\Omega}_\pm, \tag{2-4}$$

[2]The regularity $f(t) \in H^5(\mathbb{R})$, $t \in (0,T)$, is not optimal; that is, the two formulations are still equivalent if $f(t) \in H^r(\mathbb{R})$, $t \in (0,T)$, for $r < 5$ suitably chosen. In fact, if $\sigma = 0$, we may take $r = 3$. However, as stated in Theorem 1.3, $f(t) \in H^\infty(\mathbb{R})$ for all $t \in (0,T)$, and there is no reason for seeking the optimal range for r.

where d is a positive constant satisfying $d > \|f\|_\infty$ and $c_\pm \in \mathbb{R}$. For a proper choice of the constants $c_\pm$, it is not difficult to see that the pair $(\tilde p_\pm, \tilde v_\pm)$ satisfies (2-1a)–(2-1c). Let $V_\pm := v_\pm - \tilde v_\pm$, $V := (V^1, V^2) := V_- \mathbf{1}_{[y \le f(x)]} + V_+ \mathbf{1}_{[y > f(x)]} \in C(\mathbb{R}^2)$, and

$$\psi_\pm(x, y) := \int_{f(x)}^{y} V^1_\pm(x, s)\, ds - \int_0^x \langle V_\pm(s, f(s)) | (-f'(s), 1)\rangle\, ds \quad \text{for } (x, y) \in \overline{\Omega}_\pm,$$

be the stream function associated to $V_\pm$. Recalling (2-1a)–(2-1c), we deduce that the function $\psi := \psi_- \mathbf{1}_{[y \le f]} + \psi_+ \mathbf{1}_{[y > f]}$ satisfies $\Delta\psi = 0$ in $\mathcal{D}'(\mathbb{R}^2)$. Hence, ψ is the real part of a holomorphic function $u : \mathbb{C} \to \mathbb{C}$. Since u' is also holomorphic and $u' = \partial_x \psi - i\partial_y \psi = -(V^2, V^1)$ is bounded and vanishes for $|(x, y)| \to \infty$ it follows that $u' = 0$; hence $V = 0$. This proves that $v_\pm = \tilde v_\pm$.

We now infer from (2-1d) and (2-3) that the dynamic of the free boundary separating the fluids is described by the evolution equation

$$\begin{aligned}\partial_t f(t, x) &= \frac{k}{2\pi\mu} f'(t, x)\, \mathrm{PV}\!\int_{\mathbb{R}} \frac{f(t, x) - f(t, x - s)}{s^2 + (f(t, x) - f(t, x - s))^2} [\sigma\kappa(f) - \Delta_\rho f]'(t, x - s)\, ds \\ &\quad + \frac{k}{2\pi\mu}\, \mathrm{PV}\!\int_{\mathbb{R}} \frac{s}{s^2 + (f(t, x) - f(t, x - s))^2} [\sigma\kappa(f) - \Delta_\rho f]'(t, x - s)\, ds\end{aligned}$$

for $t > 0$ and $x \in \mathbb{R}$. This equation can be further simplified by using the formula

$$\int_{\delta<|x|<1/\delta} \frac{\partial}{\partial s}\big(\ln(s^2 + (f(x) - f(x-s))^2)\big)\, ds = \ln \frac{1+\delta^2(f(x)-f(x-1/\delta))^2}{1+\delta^2(f(x)-f(x+1/\delta))^2} \frac{1+(f(x)-f(x+\delta))^2/\delta^2}{1+(f(x)-f(x-\delta))^2/\delta^2}$$

for $\delta \in (0, 1)$ and $x \in \mathbb{R}$. Letting $\delta \to 0$, we get

$$\begin{aligned}0 &= \tfrac{1}{2}\, \mathrm{PV}\!\int_{\mathbb{R}} \frac{\partial}{\partial s}\big(\ln(s^2 + (f(x) - f(x - s))^2)\big)\, ds \\ &= \mathrm{PV}\!\int_{\mathbb{R}} \frac{s}{s^2 + (f(t, x) - f(t, x - s))^2}\, ds + \mathrm{PV}\!\int_{\mathbb{R}} \frac{(f(t, x) - f(t, x - s) f'(t, x - s)}{s^2 + (f(t, x) - f(t, x - s))^2}\, ds,\end{aligned}$$

and now the principal value needs to be taken in the first integral at zero and at infinity. Using this identity, we have shown that the mapping $[t \to f(t)]$ satisfies the evolution problem (1-1).

The implication (ii) $\Rightarrow$ (i) is now obvious. $\square$

3. The Muskat problem without surface tension: mapping properties

In Sections 3 and 4 we consider the stable case (a) mentioned on page 282. In this regime, after rescaling time, we may rewrite (1-1) in the abstract form

$$\dot f = \Phi(f)[f], \quad t > 0, \quad f(0) = f_0, \tag{3-1}$$

where $\Phi(f)$ is the linear operator formally defined by

$$\Phi(f)[h](x) := \mathrm{PV}\!\int_{\mathbb{R}} \frac{y(h'(x) - h'(x - y))}{y^2 + (f(x) - f(x - y))^2}\, dy. \tag{3-2}$$

We show in the next two sections that the mapping Φ satisfies all the assumptions of Theorem 1.5 if we make the following choices: $\mathbb{E}_0 := H^1(\mathbb{R})$, $\mathbb{E}_1 := H^2(\mathbb{R})$, $\mathbb{E}_\alpha = H^s(\mathbb{R})$ with $s \in \left(\frac{3}{2}, 2\right)$, and $\mathcal{O}_\beta := H^{\bar{s}}(\mathbb{R})$ with $\bar{s} \in \left(\frac{3}{2}, s\right)$. The first goal is to prove that

$$\Phi \in C^{1-}(H^s(\mathbb{R}), \mathcal{L}(H^2(\mathbb{R}), H^1(\mathbb{R}))) \tag{3-3}$$

for each $s \in \left(\frac{3}{2}, 2\right)$. Because the property (3-3) holds for all $s \in \left(\frac{3}{2}, 2\right)$, the parameter $\bar{s}$ will appear only in the proof of Theorem 1.1, which we present at the end of Section 4.

For the sake of brevity we set

$$\delta_{[x,y]} f := f(x) - f(x-y) \quad \text{for } x, y \in \mathbb{R},$$

and therewith

$$\Phi(f)[h](x) = \mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[x,y]} h'/y}{1 + (\delta_{[x,y]} f/y)^2}\, dy.$$

Boundedness of some multilinear singular integral operators. We first consider some multilinear operators which are related to Φ.[3] The estimates in Lemmas 3.1 and 3.4 enable us in particular to establish the regularity property (3-3). Lemma 3.1 is reconsidered later on, see Lemma 5.3, in a particular context when showing that Φ is in fact real-analytic.

Lemma 3.1. *Given $n, m \in \mathbb{N}$, $r \in \left(\frac{3}{2}, 2\right)$, $a_1, \ldots, a_{n+1}, b_1, \ldots, b_m \in H^r(\mathbb{R})$, and a function $c \in L_2(\mathbb{R})$ we define*

$$A_{m,n}(a_1, \ldots, a_{n+1})[b_1, \ldots, b_m, c](x) := \mathrm{PV}\int_{\mathbb{R}} \frac{\prod_{i=1}^{m} (\delta_{[x,y]} b_i/y)}{\prod_{i=1}^{n+1} [1 + (\delta_{[x,y]} a_i/y)^2]} \frac{\delta_{[x,y]} c}{y}\, dy.$$

Then:

(i) *There exists a constant C, depending only on r, n, m, and $\max_{i=1,\ldots,n+1} \|a_i\|_{H^r}$, such that*

$$\|A_{m,n}(a_1, \ldots, a_{n+1})[b_1, \ldots, b_m, c]\|_2 \le C\|c\|_2 \prod_{i=1}^{m} \|b_i\|_{H^r} \tag{3-4}$$

for all $b_1, \ldots, b_m \in H^r(\mathbb{R})$ and $c \in L_2(\mathbb{R})$.

(ii) $A_{m,n} \in C^{1-}\big((H^r(\mathbb{R}))^{n+1}, \mathcal{L}_{m+1}((H^r(\mathbb{R}))^m \times L_2(\mathbb{R}), L_2(\mathbb{R}))\big)$.

Remark 3.2. We note that

$$\Phi(f)[h] = A_{0,0}(f)[h'] \tag{3-5}$$

for all $f \in H^s(\mathbb{R})$, $s \in \left(\frac{3}{2}, 2\right)$, and $h \in H^2(\mathbb{R})$, and

$$A_{0,0}(0)[c](x) = \mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[x,y]} c}{y}\, dy = -\,\mathrm{PV}\int_{\mathbb{R}} \frac{c(x-y)}{y}\, dy = -\pi H c(x),$$

where H denotes the Hilbert transform [Stein 1993].

[3] As usual, the empty product is set to be equal to 1.

Remark 3.3. In the proof of Lemma 3.1 we split the operator $A_{m,n} := A_{m,n}(a_1, \dots, a_{n+1})$ into two operators

$$A_{m,n} = A^1_{m,n} - A^2_{m,n}.$$

If we keep $b_1, \dots, b_m$ fixed, then $A^1_{m,n}$ is a multiplication-type operator

$$A^1_{m,n}[b_1, \dots, b_m, c](x) := c(x)\,\mathrm{PV}\!\int_{\mathbb{R}} \frac{1}{y} \frac{\prod_{i=1}^m (\delta_{[x,y]} b_i / y)}{\prod_{i=1}^{n+1} [1 + (\delta_{[x,y]} a_i / y)^2]}\, dy,$$

while $A^2_{m,n}$ is the singular integral operator

$$A^2_{m,n}[b_1, \dots, b_m, c](x) := \mathrm{PV}\!\int_{\mathbb{R}} K(x, y) c(x - y)\, dy,$$

with the kernel K defined by

$$K(x, y) := \frac{1}{y} \frac{\prod_{i=1}^m (\delta_{[x,y]} b_i / y)}{\prod_{i=1}^{n+1} [1 + (\delta_{[x,y]} a_i / y)^2]} \quad \text{for } x \in \mathbb{R}, \quad y \neq 0.$$

Our proof shows that both operators $A^i_{m,n}$, $1 \le i \le 2$, satisfy (3-4). While the boundedness of $A^1_{m,n}$ follows by direct computation, the boundedness of $A^2_{m,n}$ follows from the estimate on the norm of operator defined in (1-2) and an argument due to Calderón as it appears in the proof of [Meyer and Coifman 1997, Theorem 9.7.11]. In fact, the arguments in the proof of Lemma 3.1 show that given Lipschitz functions $a_1, \dots, a_{n+m} : \mathbb{R} \to \mathbb{R}$, the singular integral operator

$$B_{n,m}(a_1, \dots, a_{n+m})[h](x) := \mathrm{PV}\!\int_{\mathbb{R}} \frac{h(x-y)}{y} \frac{\prod_{i=1}^n (\delta_{[x,y]} a_i / y)}{\prod_{i=n+1}^{n+m} [1 + (\delta_{[x,y]} a_i / y)^2]}\, dy$$

belongs to $\mathcal{L}(L_2(\mathbb{R}))$ and $\|B_{n,m}(a_1, \dots, a_{n+m})\|_{\mathcal{L}(L_2(\mathbb{R}))} \le C \prod_{i=1}^n \|a_i'\|_\infty$, where C is a constant depending only on n, m and $\max_{i=n+1,\dots,n+m} \|a_i'\|_\infty$.

It is worth pointing out that $B_{0,0} = B_{0,1}(0) = \pi H$.

Proof of Lemma 3.1. The multilinear operator $A^1_{m,n}$ is bounded provided that the mapping

$$\left[x \mapsto \mathrm{PV}\!\int_{\mathbb{R}} \frac{1}{y} \frac{\prod_{i=1}^m (\delta_{[x,y]} b_i / y)}{\prod_{i=1}^{n+1} [1 + (\delta_{[x,y]} a_i / y)^2]}\, dy \right]$$

belongs to $L_\infty(\mathbb{R})$. To establish this boundedness property we note that

$$\begin{aligned}\int_{\delta<|y|<1/\delta} \frac{1}{y} \frac{\prod_{i=1}^m (\delta_{[x,y]} b_i / y)}{\prod_{i=1}^{n+1} [1+(\delta_{[x,y]} a_i / y)^2]}\, dy &= \int_\delta^{1/\delta} \frac{1}{y} \frac{\prod_{i=1}^m (\delta_{[x,y]} b_i / y)}{\prod_{i=1}^{n+1} [1+(\delta_{[x,y]} a_i / y)^2]} - \frac{1}{y} \frac{\prod_{i=1}^m (-\delta_{[x,-y]} b_i / y)}{\prod_{i=1}^{n+1} [1+(\delta_{[x,-y]} a_i / y)^2]}\, dy \\ &=: \int_\delta^{1/\delta} I(x, y)\, dy\end{aligned}$$

for $\delta \in (0,1)$ and $x \in \mathbb{R}$, where

$$I(x,y) := \frac{1}{y}\frac{1}{\prod_{i=1}^{n+1}[1+(\delta_{[x,y]}a_i/y)^2]}\left(\prod_{i=1}^{m}(\delta_{[x,y]}b_i/y) - \prod_{i=1}^{m}(-\delta_{[x,-y]}b_i/y)\right) + \frac{1}{y}\left(\prod_{i=1}^{m}(-\delta_{[x,-y]}b_i/y)\right)\frac{\prod_{i=1}^{n+1}[1+(\delta_{[x,-y]}a_i/y)^2] - \prod_{i=1}^{n+1}[1+(\delta_{[x,y]}a_i/y)^2]}{\prod_{i=1}^{n+1}[1+(\delta_{[x,y]}a_i/y)^2][1+(\delta_{[x,-y]}a_i/y)^2]}.$$

We further have

$$\begin{aligned}\frac{1}{y}&\left(\prod_{i=1}^{m}(\delta_{[x,y]}b_i/y) - \prod_{i=1}^{m}(-\delta_{[x,-y]}b_i/y)\right)\\ &= \frac{1}{y}\prod_{i=1}^{m}\frac{b_i(x)-b_i(x-y)}{y} - \frac{1}{y}\prod_{i=1}^{m}\frac{b_i(x+y)-b_i(x)}{y}\\ &= -\sum_{i=1}^{m}\frac{b_i(x+y)-2b_i(x)+b_i(x-y)}{y^2}\left[\prod_{j=1}^{i-1}(\delta_{[x,y]}b_j/y)\right]\left[\prod_{j=i+1}^{m}(-\delta_{[x,-y]}b_j/y)\right],\end{aligned}$$

and similarly

$$\begin{aligned}\frac{1}{y}&\left(\prod_{i=1}^{n+1}[1+(\delta_{[x,-y]}a_i/y)^2] - \prod_{i=1}^{n+1}[1+(\delta_{[x,y]}a_i/y)^2]\right)\\ &= \sum_{i=1}^{n+1}\left[\prod_{j=1}^{i-1}[1+(\delta_{[x,-y]}a_j/y)^2]\right]\left[\prod_{j=i+1}^{n+1}[1+(\delta_{[x,y]}a_j/y)^2]\right]\\ &\qquad\times\frac{a_i(x+y)-a_i(x-y)}{y}\,\frac{a_i(x+y)-2a_i(x)+a_i(x-y)}{y^2}.\end{aligned}$$

Let us now observe that

$$|I(x,y)| \le \frac{2^{m+1}[1+4(n+1)\max_{i=1,\dots,n+1}\|a_i\|_\infty^2]}{y^2}\prod_{i=1}^{m}\|b_i\|_\infty \quad \text{for } x\in\mathbb{R},\quad y\ge 1. \tag{3-6}$$

Furthermore, since $r-\frac{1}{2}\in(1,2)$, we find, by taking advantage of $H^r(\mathbb{R}) \hookrightarrow \mathrm{BC}^{r-1/2}(\mathbb{R})$, that

$$\frac{|f(x+y)-2f(x)+f(x-y)|}{y^{r-1/2}} \le 4[f']_{r-3/2} \le C\|f\|_{H^r} \quad \text{for all } f\in H^r(\mathbb{R}),\quad x\in\mathbb{R},\quad y>0; \tag{3-7}$$

see [Lunardi 1995, Relation (0.2.2)]. Here $[\,\cdot\,]_{r-3/2}$ denotes the usual Hölder seminorm. Using (3-7), it follows that

$$|I(x,y)| \le Cy^{r-5/2}\left[\sum_{i=1}^{m}\left(\|b_i\|_{H^r}\prod_{j=1,\,j\neq i}^{m}\|b_j'\|_\infty\right) + \left(\prod_{i=1}^{m}\|b_i'\|_\infty\right)\sum_{i=1}^{n+1}\|a_i'\|_\infty\|a_i\|_{H^r}\right] \tag{3-8}$$

for $x \in \mathbb{R}$, $y \in (0, 1)$. Combining (3-6) and (3-8) yields

$$\sup_{x\in\mathbb{R}} \int_0^\infty |I(x, y)|\, dy \le C \prod_{i=1}^m \|b_i\|_{H^r},$$

where C depends only on r, n, m, and $\max_{i=1,\dots,n+1} \|a_i\|_{H^r}$. The latter estimate shows that (3-4) is satisfied when $A_{m,n}$ is replaced by $A^1_{m,n}$.

To deal with $A^2_{m,n}$, we define the functions $F : \mathbb{R}^{n+m+1} \to \mathbb{R}$ and $A : \mathbb{R} \to \mathbb{R}^{n+m+1}$ by

$$F(u_1, \dots, u_{n+1}, v_1, \dots, v_m) = \frac{\prod_{i=1}^m v_i}{\prod_{i=1}^{n+1}(1+u_i^2)} \quad \text{and} \quad A := (a_1, \dots, a_{n+1}, b_1, \dots, b_m),$$

where $b_i \in H^r(\mathbb{R})$ satisfy $\|b_i'\|_\infty \le 1,\ 1 \le i \le m$. The function F is smooth and A is Lipschitz continuous with a Lipschitz constant $L := \sqrt{m + (n+1)\max_{i=1,n+1} \|a_i'\|^2_\infty} \ge \|A'\|_\infty$. We further observe that

$$K(x, y) = \frac{1}{y} F\left(\frac{\delta_{[x,y]}A}{y}\right),$$

with $|\delta_{[x,y]}A/y| \le L$. Let $\widetilde{F}$ be a smooth function on $\mathbb{R}^{n+m+1}$ which is $4L$-periodic in each variable and which matches F on $[-L, L]^{n+m+1}$. Expanding $\widetilde{F}$ by its Fourier series

$$\widetilde{F} = \sum_{p\in\mathbb{Z}^{n+m+1}} \alpha_p e^{i(\pi/2L)\langle p|\cdot\rangle},$$

the associated sequence $(\alpha_p)_p$ is rapidly decreasing. Furthermore, we can write the kernel K as

$$K(x, y) = \sum_{p\in\mathbb{Z}^{n+m+1}} \alpha_p K_p(x, y), \quad x \in \mathbb{R},\ \ y \ne 0,$$

with

$$K_p(x, y) := \frac{1}{y} \exp\left(i \frac{\pi}{2L} \frac{\delta_{[x,y]}\langle p \mid A\rangle}{y}\right), \quad x \in \mathbb{R},\ \ y \ne 0,\ \ p \in \mathbb{Z}^{n+m+1}.$$

The kernels $K_p,\ p \in \mathbb{Z}^{n+m+1}$, define operators in $\mathcal{L}(L_2(\mathbb{R}))$ of the type (1-2) and with norms bounded by

$$C\left(1 + \frac{\pi}{2L}|p|\|A'\|_\infty\right) \le C(1 + |p|), \quad p \in \mathbb{Z}^{n+m+1},$$

with a universal constant C independent of p. Hence, the associated series is absolutely convergent in $\mathcal{L}(L_2(\mathbb{R}))$, meaning that the operator $A^2_{m,n}(a_1, \dots, a_{n+1})[b_1, \dots, b_m, \cdot\,]$ belongs to $\mathcal{L}(L_2(\mathbb{R}))$ and

$$\|A^2_{m,n}(a_1, \dots, a_{n+1})[b_1, \dots, b_m, c]\|_2 \le C\big(n, m, \max_{i=1,\dots,n+1} \|a_i'\|_\infty\big)\|c\|_2$$

for all $c \in L_2(\mathbb{R})$ and for all $b_i \in H^r(\mathbb{R})$ that satisfy $\|b_i'\|_\infty \le 1$. The desired estimate (3-4) follows now by using the linearity of $A^2_{m,n}$ in each argument. The claim (i) is now obvious.

Concerning (ii), we note that

$$A_{m,n}(\tilde a_1,\dots,\tilde a_{n+1})[b_1,\dots,b_m,c]-A_{m,n}(a_1,\dots,a_{n+1})[b_1,\dots,b_m,c] \\ =\sum_{i=1}^{n+1}A_{m+2,n+1}(\tilde a_1,\dots,\tilde a_i,a_i,\dots,a_{n+1})[a_i+\tilde a_i,a_i-\tilde a_i,b_1,\dots,b_m,c],$$

and the desired assertion follows now from (i). □

We consider once more the operators $A_{m,n}$ defined in Lemma 3.5 in the case when $m\ge 1$, but defined on a different Hilbert-space product where a weaker regularity of b_m is balanced by a higher regularity of the variable c. The estimates in Lemma 3.4 are slightly related to the ones announced in [Calderon et al. 1978, Theorem 4] and, except for that reference, we did not find similar results.

Lemma 3.4. *Let $n\in\mathbb{N}$, $1\le m\in\mathbb{N}$, $r\in\big(\tfrac32,2\big)$, $\tau\in\big(\tfrac52-r,1\big)$, and $a_1,\dots,a_{n+1}\in H^r(\mathbb{R})$ be given. Then*:

(i) *There exists a constant C, depending only on r and τ, such that*

$$\|A_{m,n}(a_1,\dots,a_{n+1})[b_1,\dots,b_m,c]\|_2\le C\|c\|_{H^\tau}\|b_m\|_{H^{r-1}}\prod_{i=1}^{m-1}\|b_i'\|_\infty$$

for all $b_1,\dots,b_m\in H^r(\mathbb{R})$ and all $c\in H^1(\mathbb{R})$. In particular, $A_{m,n}(a_1,\dots,a_{n+1})$ extends to a bounded operator

$$A_{m,n}(a_1,\dots,a_{n+1})\in\mathcal{L}_{m+1}((H^r(\mathbb{R}))^{m-1}\times H^{r-1}(\mathbb{R})\times H^\tau(\mathbb{R}),L_2(\mathbb{R})).$$

(ii) $A_{m,n}\in C^{1-}\big((H^r(\mathbb{R}))^{n+1},\mathcal{L}_{m+1}((H^r(\mathbb{R}))^{m-1}\times H^{r-1}(\mathbb{R})\times H^\tau(\mathbb{R}),L_2(\mathbb{R}))\big)$.

Proof. The claim (ii) is again a direct consequence of (i), so that we are left to prove the first claim. To this end we write

$$A_{m,n}(a_1,\dots,a_{n+1})[b_1,\dots,b_m,c](x)=\int_{\mathbb{R}}K(x,y)\,dy,$$

where

$$K(x,y):=\frac{\prod_{i=1}^{m-1}(\delta_{[x,y]}b_i/y)}{\prod_{i=1}^{n+1}[1+(\delta_{[x,y]}a_i/y)^2]}\frac{\delta_{[x,y]}b_m}{y}\frac{\delta_{[x,y]}c}{y}\quad\text{for } x\in\mathbb{R},\ \ y\ne0.$$

Using Minkowski's integral inequality, we compute

$$\bigg(\int_{\mathbb{R}}\bigg|\int_{\mathbb{R}}K(x,y)\,dy\bigg|^2dx\bigg)^{1/2}\le\int_{\mathbb{R}}\bigg(\int_{\mathbb{R}}|K(x,y)|^2\,dx\bigg)^{1/2}dy,$$

and exploiting the fact $H^{r-1}(\mathbb{R})\hookrightarrow \mathrm{BC}^{r-3/2}(\mathbb{R})$, we get

$$\int_{\mathbb{R}}|K(x,y)|^2\,dx\le\frac{C}{y^{7-2r}}\|b_m\|^2_{H^{r-1}}\bigg(\prod_{i=1}^{m-1}\|b_i'\|_\infty^2\bigg)\int_{\mathbb{R}}|c-\tau_yc|^2\,dx \\ =\frac{C}{y^{7-2r}}\|b_m\|^2_{H^{r-1}}\bigg(\prod_{i=1}^{m-1}\|b_i'\|_\infty^2\bigg)\int_{\mathbb{R}}|\mathcal{F}c(\xi)|^2|e^{iy\xi}-1|^2\,d\xi.$$

Since

$$|e^{iy\xi}-1|^2 \le C[(1+|\xi|^2)^\tau y^{2\tau}\mathbf{1}_{(-1,1)}(y)+\mathbf{1}_{[|y|\ge 1]}(y)], \quad y,\xi\in\mathbb{R},$$

it follows that

$$\int_{\mathbb{R}}|K(x,y)|^2\,dx \le C\|c\|_{H^\tau}^2\|b_m\|_{H^{r-1}}^2\biggl(\prod_{i=1}^{m-1}\|b_i'\|_\infty^2\biggr)\biggl[y^{2(r+\tau)-7}\mathbf{1}_{(-1,1)}(y)+\frac{1}{y^{7-2r}}\mathbf{1}_{[|y|\ge 1]}(y)\biggr],$$

and we conclude that

$$\int_{\mathbb{R}}\biggl(\int_{\mathbb{R}}|K(x,y)|^2\,dx\biggr)^{1/2}dy \le C\|c\|_{H^\tau}\|b_m\|_{H^{r-1}}\prod_{i=1}^{m-1}\|b_i'\|_\infty.$$

The claim (i) follows at once. ☐

Mapping properties. We now use Lemmas 3.1 and 3.4 to prove that the mapping Φ defined by (3-2) is well-defined and locally Lipschitz continuous as an operator from $H^s(\mathbb{R})$ into the Banach space $\mathcal{L}(H^2(\mathbb{R}),H^1(\mathbb{R}))$ for each $s\in\bigl(\frac{3}{2},2\bigr)$.

Lemma 3.5. *Given $s\in\bigl(\frac{3}{2},2\bigr)$, it holds that*

$$\Phi\in C^{1-}(H^s(\mathbb{R}),\mathcal{L}(H^2(\mathbb{R}),H^1(\mathbb{R}))).$$

Proof. We first prove that $\Phi(f)\in\mathcal{L}(H^2(\mathbb{R}),H^1(\mathbb{R}))$ for each $f\in H^s(\mathbb{R})$. Remark 3.2 and Lemma 3.1 (with $r=s$) yield that $\Phi(f)\in\mathcal{L}(H^2(\mathbb{R}),L_2(\mathbb{R}))$. In order to establish that $\Phi(f)[h]\in H^1(\mathbb{R})$, we let $\{\tau_\varepsilon\}_{\varepsilon\in\mathbb{R}}$ denote the C_0-group of right translations on $L_2(\mathbb{R})$, that is, $\tau_\varepsilon f(x):=f(x-\varepsilon)$ for $f\in L_2(\mathbb{R})$ and $x,\varepsilon\in\mathbb{R}$. Given $\varepsilon\in(0,1)$, it holds that

$$\begin{aligned}\frac{\tau_\varepsilon(\Phi(f)[h])-\Phi(f)[h]}{\varepsilon}&=\frac{\tau_\varepsilon(A_{0,0}[f][h'])-A_{0,0}(f)[h']}{\varepsilon}=\frac{A_{0,0}(\tau_\varepsilon f)[\tau_\varepsilon h']-A_{0,0}(f)[h']}{\varepsilon}\\&=A_{0,0}(\tau_\varepsilon f)\biggl[\frac{\tau_\varepsilon h'-h'}{\varepsilon}\biggr]-A_{2,1}(\tau_\varepsilon f,f)\biggl[\tau_\varepsilon f+f,\frac{\tau_\varepsilon f-f}{\varepsilon},h'\biggr]\end{aligned}$$

and the convergences

$$\tau_\varepsilon f\xrightarrow[\varepsilon\to 0]{} f \quad\text{in } H^s(\mathbb{R}),\qquad \frac{\tau_\varepsilon f-f}{\varepsilon}\xrightarrow[\varepsilon\to 0]{}-f'\quad\text{in } H^{s-1}(\mathbb{R}),\qquad \frac{\tau_\varepsilon h-h}{\varepsilon}\xrightarrow[\varepsilon\to 0]{}-h'\quad\text{in } H^1(\mathbb{R}),$$

together with Lemma 3.1 (with $r=s$) and Lemma 3.4 $\bigl($with $r=s$, $\tau\in\bigl(\frac{5}{2}-s,1\bigr)\bigr)$ imply that $\Phi(f)[h]\in H^1(\mathbb{R})$ and

$$(\Phi(f)[h])'=A_{0,0}(f)[h'']-2A_{2,1}(f,f)[f,f',h']. \tag{3-9}$$

This proves that $\Phi(f)\in\mathcal{L}(H^2(\mathbb{R}),H^1(\mathbb{R}))$. Finally, the local Lipschitz continuity of Φ follows from the local Lipschitz continuity properties established in Lemmas 3.1 and 3.4. ☐

4. The Muskat problem without surface tension: the generator property

We now fix $f \in H^s(\mathbb{R}),\ s \in \left(\frac{3}{2}, 2\right)$. The goal of this section is to prove that $\Phi(f)$, regarded as an unbounded operator in $H^1(\mathbb{R})$ with definition domain $H^2(\mathbb{R})$, is the generator of a strongly continuous and analytic semigroup in $\mathcal{L}(H^1(\mathbb{R}))$, that is,

$$-\Phi(f) \in \mathcal{H}(H^2(\mathbb{R}), H^1(\mathbb{R})).$$

In order to establish this property we first approximate locally the operator $\Phi(f)$, in a sense to be made precise in Theorem 4.2, by Fourier multipliers and carry then the desired generator property, which we establish for the Fourier multipliers, back to the original operator, see Theorem 4.4. A similar approach was followed in [Escher 1994; Escher et al. 2018; Escher and Simonett 1995; 1997] in the context of spaces of continuous functions. The situation here is different as we consider Sobolev spaces on the line. The method though can be adapted to this setting after exploiting the structure of the operator $\Phi(f)$, especially the fact that the functions f and f' both vanish at infinity. As a result of this decay property we can use localization families with a finite number of elements, and this fact enables us to introduce for each localization family an equivalent norm on the Sobolev spaces $H^k(\mathbb{R}),\ k \in \mathbb{N}$, which is suitable for the further analysis, see Lemma 4.1. We start by choosing for each $\varepsilon \in (0, 1)$, a finite ε-localization family, that is, a family

$$\{\pi_j^\varepsilon : -N+1 \le j \le N\} \subset C^\infty(\mathbb{R}, [0, 1]),$$

with $N = N(\varepsilon) \in \mathbb{N}$ sufficiently large, such that

- $\operatorname{supp} \pi_j^\varepsilon$ is an interval of length less or equal to ε for all $|j| \le N-1$; (4-1)
- $\operatorname{supp} \pi_N^\varepsilon \subset (-\infty, -x_N] \cup [x_N, \infty)$ and $x_N \ge \varepsilon^{-1}$; (4-2)
- $\operatorname{supp} \pi_j^\varepsilon \cap \operatorname{supp} \pi_l^\varepsilon = \varnothing$ if $[|j-l| \ge 2, \max\{|j|, |l|\} \le N-1]$ or $[|l| \le N-2,\ j = N]$; (4-3)
- $\sum_{j=-N+1}^{N} (\pi_j^\varepsilon)^2 = 1$; (4-4)
- $\|(\pi_j^\varepsilon)^{(k)}\|_\infty \le C\varepsilon^{-k}$ for all $k \in \mathbb{N},\ -N+1 \le j \le N$. (4-5)

Such ε-localization families can be easily constructed. Additionally, we choose for each $\varepsilon \in (0, 1)$ a second family

$$\{\chi_j^\varepsilon : -N+1 \le j \le N\} \subset C^\infty(\mathbb{R}, [0, 1])$$

with the properties

- $\chi_j^\varepsilon = 1$ on $\operatorname{supp} \pi_j^\varepsilon$; (4-6)
- $\operatorname{supp} \chi_j^\varepsilon$ is an interval of length less or equal to 3ε for $|j| \le N-1$; (4-7)
- $\operatorname{supp} \chi_N^\varepsilon \subset [|x| \ge x_N - \varepsilon]$. (4-8)

Each ε-localization family $\{\pi_j^\varepsilon : -N+1 \le j \le N\}$ defines a norm on $H^k(\mathbb{R}),\ k \in \mathbb{N}$, which is equivalent to the standard H^k-norm.

Lemma 4.1. *Given $\varepsilon \in (0,1)$, let $\{\pi_j^\varepsilon : -N+1 \le j \le N\} \subset C^\infty(\mathbb{R},[0,1])$ be a family with the properties* (4-1)–(4-5). *Then, for each $k \in \mathbb{N}$, the mapping*

$$\Big[h \mapsto \sum_{j=-N+1}^{N} \|\pi_j^\varepsilon h\|_{H^k}\Big] : H^k(\mathbb{R}) \to [0,\infty)$$

defines a norm on $H^k(\mathbb{R})$ which is equivalent to the standard H^k-norm.

Proof. The proof is a simple exercise. □

We now consider the mapping

$$[\tau \mapsto \Phi(\tau f)] : [0,1] \to \mathcal{L}(H^2(\mathbb{R}), H^1(\mathbb{R})).$$

As a consequence of Lemma 3.5, this mapping continuously transforms the operator $\Phi(f)$, for which we want to establish the generator property, into the operator $\Phi(0) = -\pi(-\partial_x^2)^{1/2}$. Indeed, since the Hilbert transform is a Fourier multiplier with symbol $[\xi \mapsto -i \operatorname{sign}(\xi)]$, we obtain together with Remark 3.2 that

$$\mathcal{F}(\Phi(0)[h])(\xi) = -\pi\mathcal{F}(Hh')(\xi) = i\pi \operatorname{sign}(\xi)\mathcal{F}(h')(\xi) = -\pi|\xi|(\mathcal{F}h)(\xi) = -\pi\mathcal{F}((-\partial_x^2)^{1/2}h)(\xi)$$

for $\xi \in \mathbb{R}$. The parameter τ will allow us to use a continuity argument when showing that the resolvent set of $\Phi(f)$ contains a positive real number; see the proof Theorem 4.4.

Our next goal is to prove that the operator $\Phi(\tau f)$ can be locally approximated for each $\tau \in [0,1]$ by Fourier multipliers, as stated below. The estimate (4-9) with $j = N$ uses to a large extent the fact that f and f' vanish at infinity.

Theorem 4.2. *Let $f \in H^s(\mathbb{R})$, $s \in \big(\frac{3}{2}, 2\big)$, and $\mu > 0$ be given.*

Then, there exist $\varepsilon \in (0,1)$, a finite ε-localization family $\{\pi_j^\varepsilon : -N+1 \le j \le N\}$ satisfying (4-1)–(4-5), *a constant $K = K(\varepsilon)$, and for each $j \in \{-N+1,\dots,N\}$ and $\tau \in [0,1]$ there exist operators*

$$\mathbb{A}_{j,\tau} \in \mathcal{L}(H^2(\mathbb{R}), H^1(\mathbb{R}))$$

such that

$$\|\pi_j^\varepsilon \Phi(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_{H^1} \le \mu\|\pi_j^\varepsilon h\|_{H^2} + K\|h\|_{H^{(11-2s)/4}} \tag{4-9}$$

for all $j \in \{-N+1,\dots,N\}$, $\tau \in [0,1]$, and $h \in H^2(\mathbb{R})$. The operators $\mathbb{A}_{j,\tau}$ are defined by

$$\mathbb{A}_{j,\tau} := \bigg[\mathrm{PV}\!\int_{\mathbb{R}} \frac{1}{y}\frac{1}{1+\tau^2(\delta_{[x_j^\varepsilon,y]}f/y)^2}\,dy\bigg]\partial_x - \frac{\pi}{1+(\tau f'(x_j^\varepsilon))^2}(-\partial_x^2)^{1/2}, \quad |j| \le N-1, \tag{4-10}$$

where x_j^ε is a point belonging to $\operatorname{supp}\pi_j^\varepsilon$, *and*

$$\mathbb{A}_{N,\tau} := -\pi(-\partial_x^2)^{1/2}. \tag{4-11}$$

Proof. Let $\{\pi_j^\varepsilon : -N+1 \le j \le N\}$ be an ε-localization family satisfying the properties (4-1)–(4-5) and $\{\chi_j^\varepsilon : -N+1 \le j \le N\}$ be an associated family satisfying (4-6)–(4-8), with $\varepsilon \in (0,1)$ which will be fixed below. We first infer from Lemma A.1 that for each $\tau \in [0, 1]$ the function

$$a_\tau(x) := \mathrm{PV}\int_{\mathbb{R}} \frac{1}{y}\frac{1}{1+\tau^2(\delta_{[x,y]}f/y)^2}\,dy, \quad x \in \mathbb{R},$$

belongs to $\mathrm{BC}^\alpha(\mathbb{R}) \cap C_0(\mathbb{R})$, with $\alpha := \frac{1}{2}s - \frac{3}{4}$. We now write

$$\mathbb{A}_{j,\tau} := \mathbb{A}^1_{j,\tau} - \mathbb{A}^2_{j,\tau},$$

where

$$\mathbb{A}^1_{j,\tau} := \alpha_{j,\tau}\partial_x, \quad \mathbb{A}^2_{j,\tau} := \beta_{j,\tau}(-\partial_x^2)^{1/2},$$

and

$$\alpha_{j,\tau} := \begin{cases} a_\tau(x_j^\varepsilon), & |j| \le N-1, \\ 0, & j = N, \end{cases} \qquad \beta_{j,\tau} := \begin{cases} \dfrac{\pi}{1+(\tau f'(x_j^\varepsilon))^2}, & |j| \le N-1, \\ \pi, & j = N. \end{cases} \tag{4-12}$$

Let now $h \in H^2(\mathbb{R})$ be arbitrary. In the following we shall denote by C constants which are independent of ε (and, of course, of $h \in H^2(\mathbb{R})$, $\tau \in [0, 1]$, and $j \in \{-N+1, \dots, N\}$), while the constants that we denote by K may depend only upon ε.

Step 1: We first infer from Lemma 3.5 that

$$\begin{aligned}\|\pi_j^\varepsilon\Phi(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_{H^1} &\le \|\pi_j^\varepsilon\Phi(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_2 + \|(\pi_j^\varepsilon\Phi(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h])'\|_2 \\ &\le (1+\|(\pi_j^\varepsilon)'\|_\infty)\|A_{0,0}(\tau f)[h']\|_2 + \|\mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_2 \\ &\quad + 2\|A_{2,1}(\tau f, \tau f)[f, f', h']\|_2 + \|\pi_j^\varepsilon A_{0,0}(\tau f)[h''] - \mathbb{A}_{j,\tau}[(\pi_j^\varepsilon h)']\|_2.\end{aligned}$$

Using Lemma 3.4 $\big($with $r = s$ and $\tau = \frac{7}{4} - \frac{1}{2}s\big)$ and Lemma 3.1 (with $r = s$), it follows that

$$\|\pi_j^\varepsilon\Phi(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_{H^1} \le K\|h\|_{H^{(11-2s)/4}} + \|\pi_j^\varepsilon A_{0,0}(\tau f)[h''] - \mathbb{A}_{j,\tau}[(\pi_j^\varepsilon h)']\|_2. \tag{4-13}$$

We are left to estimate the L_2-norm of the highest-order term $\pi_j^\varepsilon A_{0,0}(\tau f)[h''] - \mathbb{A}_{j,\tau}[(\pi_j^\varepsilon h)']$, and for this we need several steps.

Step 2: With the notation introduced in Remark 3.3 we have

$$A_{0,0}(\tau f)[h''] = a_\tau h'' - B_{0,1}(\tau f)[h''],$$

and therewith

$$\begin{aligned}&\|\pi_j^\varepsilon A_{0,0}(\tau f)[h''] - \mathbb{A}_{j,\tau}[(\pi_j^\varepsilon h)']\|_2 \\ &\qquad \le \|\pi_j^\varepsilon a_\tau h'' - \mathbb{A}^1_{j,\tau}[(\pi_j^\varepsilon h)']\|_2 + \|\pi_j^\varepsilon B_{0,1}(\tau f)[h''] - \mathbb{A}^2_{j,\tau}[(\pi_j^\varepsilon h)']\|_2.\end{aligned} \tag{4-14}$$

By virtue of Lemma A.1, in particular of the estimate (A-1), and of $\chi_j^\varepsilon = 1$ on $\operatorname{supp}\pi_j^\varepsilon$, we get for $|j| \le N-1$

$$\begin{aligned}\|\pi_j^\varepsilon a_\tau h'' - \mathbb{A}^1_{j,\tau}[(\pi_j^\varepsilon h)'\|_2 &= \|a_\tau\pi_j^\varepsilon h'' - a_\tau(x_j^\varepsilon)(\pi_j^\varepsilon h)''\|_2 \\ &\le \|(a_\tau - a_\tau(x_j^\varepsilon))(\pi_j^\varepsilon h)''\|_2 + K\|h\|_{H^1} \\ &= \|(a_\tau - a_\tau(x_j^\varepsilon))\chi_j^\varepsilon(\pi_j^\varepsilon h)''\|_2 + K\|h\|_{H^1} \\ &= \|(a_\tau - a_\tau(x_j^\varepsilon))\chi_j^\varepsilon\|_\infty\|(\pi_j^\varepsilon h)''\|_2 + K\|h\|_{H^1} \\ &\le \tfrac{1}{2}\mu\|\pi_j^\varepsilon h\|_{H^2} + K\|h\|_{H^1},\end{aligned} \tag{4-15}$$

provided that ε is sufficiently small. We have used here (and also later on without explicit mentioning) the fact that $|\operatorname{supp}\chi_j^\varepsilon|\le 3\varepsilon$. Since $\mathbb{A}^1_{N,\tau}=0$, we obtain from (A-2) for ε sufficiently small that

$$\begin{aligned}\|\pi_N^\varepsilon a_\tau h''-\mathbb{A}^1_{N,\tau}[(\pi_N^\varepsilon h)']\|_2=\|\pi_N^\varepsilon a_\tau h''\|_2&\le\|a_\tau\chi_N^\varepsilon\|_\infty\|(\pi_N^\varepsilon h)''\|_2+K\|h\|_{H^1}\\&\le\tfrac12\mu\|\pi_N^\varepsilon h\|_{H^2}+K\|h\|_{H^1}.\end{aligned}\tag{4-16}$$

Step 3: We are left with the term $\|\pi_j^\varepsilon B_{0,1}(\tau f)[h'']-\mathbb{A}^2_{j,\tau}[(\pi_j^\varepsilon h)']\|_2$, and we consider first the case $|j|\le N-1$ (see Step 4 for $j=N$). Observing that $\pi(-\partial_x^2)^{1/2}=B_{0,1}(0)\circ\partial_x$, it follows that

$$\pi_j^\varepsilon B_{0,1}(\tau f)[h'']-\mathbb{A}^2_{j,\tau}[(\pi_j^\varepsilon h)']=T_1[h]-T_2[h],$$

where

$$\begin{aligned}T_1[h]&:=\pi_j^\varepsilon B_{0,1}(\tau f)[h'']-\frac{1}{1+(\tau f'(x_j^\varepsilon))^2}B_{0,1}(0)[\pi_j^\varepsilon h''],\\T_2[h]&:=\frac{1}{1+(\tau f'(x_j^\varepsilon))^2}B_{0,1}(0)[(\pi_j^\varepsilon)''h+2(\pi_j^\varepsilon)'h'].\end{aligned}$$

Since by Remark 3.3

$$\|T_2[h]\|_2\le K\|h\|_{H^1},\tag{4-17}$$

we are left to estimate $T_1[h]$, which is further decomposed as

$$T_1[h]=T_{11}[h]-T_{12}[h],$$

with

$$\begin{aligned}T_{11}[h](x)&:=\mathrm{PV}\int_{\mathbb{R}}\bigg[\frac{1}{1+\tau^2(\delta_{[x,y]}f/y)^2}-\frac{1}{1+(\tau f'(x_j^\varepsilon))^2}\bigg]\frac{(\chi_j^\varepsilon\pi_j^\varepsilon h'')(x-y)}{y}\,dy,\\T_{12}[h](x)&:=\mathrm{PV}\int_{\mathbb{R}}\frac{\delta_{[x,y]}\pi_j^\varepsilon/y}{1+\tau^2(\delta_{[x,y]}f/y)^2}h''(x-y)\,dy.\end{aligned}$$

Integrating by parts, we obtain the relation

$$\begin{aligned}T_{12}[h]=B_{0,1}(\tau f)[(\pi_j^\varepsilon)'h']-B_{1,1}(\pi_j^\varepsilon,\tau f)[h']-2\tau^2B_{2,2}(\pi_j^\varepsilon,f,\tau f,\tau f)[f'h']\\+2\tau^2B_{3,2}(\pi_j^\varepsilon,f,f,\tau f,\tau f)[h'],\end{aligned}$$

and Remark 3.3 leads us to

$$\|T_{12}[h]\|_2\le K\|h\|_{H^1}.\tag{4-18}$$

In order to deal with the term $T_{11}[h]$ we let $F_j\in C(\mathbb{R})$ denote the Lipschitz function that satisfies

$$F_j=f\quad\text{on }\operatorname{supp}\chi_j^\varepsilon,\qquad F_j'=f'(x_j^\varepsilon)\quad\text{on }\mathbb{R}\setminus\operatorname{supp}\chi_j^\varepsilon,\tag{4-19}$$

and we observe that

$$\begin{aligned}T_{11}[h](x)&:=\tau^2\,\mathrm{PV}\int_{\mathbb{R}}\frac{[\delta_{[x,y]}(f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}}-f)/y][\delta_{[x,y]}(f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}}+f)/y]}{[1+\tau^2(\delta_{[x,y]}f/y)^2][1+(\tau f'(x_j^\varepsilon))^2]}\frac{(\chi_j^\varepsilon\pi_j^\varepsilon h'')(x-y)}{y}\,dy\\&=\frac{\tau^2}{1+(\tau f'(x_j^\varepsilon))^2}(T_{111}[h]-T_{112}[h])(x),\end{aligned}$$

where

$$T_{111}[h] := \chi_j^\varepsilon B_{2,1}(f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} - f,\, f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} + f,\, \tau f)[\pi_j^\varepsilon h''],$$

$$T_{112}[h](x) := \mathrm{PV}\int_{\mathbb{R}} \frac{[\delta_{[x,y]}(f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} - f)/y][\delta_{[x,y]}(f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} + f)/y](\delta_{[x,y]}\chi_j^\varepsilon/y)}{1+\tau^2(\delta_{[x,y]}f/y)^2}(\pi_j^\varepsilon h'')(x-y)\,dy.$$

Integrating by parts as in the case of $T_{12}[h]$, it follows from Remark 3.3 that

$$\|T_{112}[h]\|_2 \le K\|h\|_{H^1}. \tag{4-20}$$

On the other hand, (4-19), Remark 3.3 and the Hölder continuity of f' yield

$$\begin{aligned}
\|T_{111}[h]\|_2 &= \|\chi_j^\varepsilon B_{2,1}(f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} - f,\, f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} + f,\, \tau f)[\pi_j^\varepsilon h'']\|_2 \\
&= \|\chi_j^\varepsilon B_{2,1}(f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} - F_j,\, f'(x_j^\varepsilon)\mathrm{id}_{\mathbb{R}} + F_j,\, \tau f)[\pi_j^\varepsilon h'']\|_2 \\
&\le C\|f'(x_j^\varepsilon) - F_j'\|_\infty \|\pi_j^\varepsilon h''\|_2 \\
&= C\|f'(x_j^\varepsilon) - f'\|_{L_\infty(\operatorname{supp}\chi_j^\varepsilon)} \|\pi_j^\varepsilon h''\|_2 \\
&\le \tfrac{1}{2}\mu\|\pi_j^\varepsilon h\|_{H^2} + K\|h\|_{H^1}.
\end{aligned} \tag{4-21}$$

The desired estimate (4-9) follows for $|j| \le N-1$ from (4-13)–(4-15) and (4-17), (4-18), (4-20), and (4-21).

Step 4: We are left with the term $\|\pi_N^\varepsilon B_{0,1}(\tau f)[h''] - \mathbb{A}_{N,\tau}^2[(\pi_N^\varepsilon h)']\|_2$, which we decompose as

$$\begin{aligned}
&(\pi_N^\varepsilon B_{0,1}(\tau f)[h''] - \mathbb{A}_{N,\tau}^2[(\pi_N^\varepsilon h)'])(x) \\
&\qquad = \pi_N^\varepsilon(x)\,\mathrm{PV}\int_{\mathbb{R}} \frac{h''(x-y)}{y}\frac{1}{1+\tau^2(\delta_{[x,y]}f/y)^2}\,dy - \mathrm{PV}\int_{\mathbb{R}} \frac{(\pi_N^\varepsilon h)''(x-y)}{y}\,dy \\
&\qquad =: T_1[h](x) + T_2[h](x) - T_3[h](x),
\end{aligned}$$

where

$$\begin{aligned}
T_1[h] &:= -\tau^2 B_{2,1}(f, f, \tau f)[\pi_N^\varepsilon h''], \\
T_2[h](x) &:= \mathrm{PV}\int_{\mathbb{R}} h''(x-y)\frac{\delta_{[x,y]}\pi_N^\varepsilon}{y}\frac{1}{1+\tau^2(\delta_{[x,y]}f/y)^2}\,dy, \\
T_3[h] &:= B_{0,1}(0)[(\pi_N^\varepsilon)''h + 2(\pi_N^\varepsilon)'h'].
\end{aligned}$$

For the difference $T_2[h] - T_3[h]$ we find, as in the previous step (see (4-17) and (4-18)), that

$$\|T_2[h] - T_3[h]\|_2 \le K\|h\|_{H^1}. \tag{4-22}$$

When dealing with $T_1[h]$, we introduce the function $F_N \in W_\infty^1(\mathbb{R})$ by the formula

$$F_N(x) := \begin{cases} f(x), & |x| \ge x_N - \varepsilon, \\ \dfrac{x + x_N - \varepsilon}{2(x_N-\varepsilon)} f(x_N - \varepsilon) + \dfrac{x_N - \varepsilon - x}{2(x_N - \varepsilon)} f(-x_N + \varepsilon), & |x| \le x_N - \varepsilon. \end{cases}$$

The relation (4-2) implies $\|F_N\|_\infty + \|F_N'\|_\infty \to 0$ for $\varepsilon \to 0$. Moreover, it holds that

$$T_1[h](x) = -\tau^2 \,\mathrm{PV}\!\int_{\mathbb{R}} \frac{(\chi_N^\varepsilon \pi_N^\varepsilon h'')(x-y)}{y} \frac{(\delta_{[x,y]} f/y)^2}{1+\tau^2(\delta_{[x,y]} f/y)^2} =: T_{11}[h](x) - T_{12}[h](x),$$

where

$$T_{11}[h](x) := \tau^2 \,\mathrm{PV}\!\int_{\mathbb{R}} (\pi_N^\varepsilon h'')(x-y) \frac{(\delta_{[x,y]} f/y)^2 (\delta_{[x,y]} \chi_N^\varepsilon/y)}{1+\tau^2(\delta_{[x,y]} f/y)^2}\, dy,$$
$$T_{12}[h] := \tau^2 \chi_N^\varepsilon B_{2,1}(f, f, \tau f)[\pi_N^\varepsilon h''].$$

Recalling that $\operatorname{supp} \pi_N^\varepsilon \subset \operatorname{supp} \chi_N^\varepsilon \subset [|x| \geq x_N - \varepsilon]$ and that $f = F_N$ on $\operatorname{supp} \chi_N^\varepsilon$, it follows by Remark 3.3 that

$$\begin{aligned}\|T_{12}[h]\|_2 &= \|\tau^2 \chi_N^\varepsilon B_{2,1}(F_N, F_N, \tau f)[\pi_N^\varepsilon h'']\|_2 \leq \|B_{2,1}(F_N, F_N, \tau f)[\pi_N^\varepsilon h'']\|_2 \leq C\|F_N'\|_\infty^2 \|\pi_N^\varepsilon h''\|_2 \\ &\leq \tfrac{1}{2}\mu \|\pi_j^\varepsilon h\|_{H^2} + K\|h\|_{H^1} \end{aligned} \tag{4-23}$$

for small ε. As $T_{11}[h]$ can be estimated in the same manner as the term $T_{112}[h]$ in the previous step, we obtain together with (4-22) and (4-23) that

$$\|\pi_N^\varepsilon B_{0,1}(\tau f)[h''] - \mathbb{A}_{N,\tau}^2[(\pi_N^\varepsilon h)']\|_2 \leq \tfrac{1}{2}\mu \|\pi_j^\varepsilon h\|_{H^2} + K\|h\|_{H^1} \tag{4-24}$$

if ε is sufficiently small. The claim (4-9) follows for $j = N$ from (4-13)–(4-14), (4-16), and (4-24). □

The operators $\mathbb{A}_{\tau,j}$ found in Theorem 4.2 are generators of strongly continuous analytic semigroups in $\mathcal{L}(H^1(\mathbb{R}))$ and they satisfy resolvent estimates which are uniform with respect to $x_j^\varepsilon \in \mathbb{R}$ and $\tau \in [0, 1]$; see Proposition 4.3 below. To be more precise, in Proposition 4.3 and in the proof of Theorem 4.4, the Sobolev spaces $H^k(\mathbb{R})$, $k \in \{1, 2\}$, consist of complex-valued functions and $\mathbb{A}_{j,\tau}$ are the natural extensions (complexifications) of the operators introduced in Theorem 4.2.

Proposition 4.3. *Let $f \in H^s(\mathbb{R})$, $s \in (\tfrac{3}{2}, 2)$, be fixed. Given $x_0 \in \mathbb{R}$ and $\tau \in [0, 1]$, let*

$$\mathbb{A}_{x_0,\tau} := \alpha_\tau \partial_x - \beta_\tau (-\partial_x^2)^{1/2},$$

where

$$\alpha_\tau \in \{0, a_\tau(x_0)\} \quad \text{and} \quad \beta_\tau \in \left\{\pi, \frac{\pi}{1+(\tau f'(x_0))^2}\right\},$$

with a_τ denoting the function defined in Lemma A.1. Then, there exists a constant $\kappa_0 \geq 1$ such that

$$\lambda - \mathbb{A}_{x_0,\tau} \in \operatorname{Isom}(H^2(\mathbb{R}), H^1(\mathbb{R})), \tag{4-25}$$
$$\kappa_0 \|(\lambda - \mathbb{A}_{x_0,\tau})[h]\|_{H^1} \geq |\lambda| \cdot \|h\|_{H^1} + \|h\|_{H^2} \tag{4-26}$$

for all $x_0 \in \mathbb{R}$, $\tau \in [0, 1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 1$, and $h \in H^2(\mathbb{R})$.

Proof. The constants α_τ, β_τ defined above satisfy, in view of (A-3),

$$|\alpha_\tau| \leq 4\left(\|f\|_\infty^2 + \frac{2\|f'\|_\infty [f']_{s-3/2}}{s-\tfrac{3}{2}}\right) \quad \text{and} \quad \beta_\tau \in \left[\frac{\pi}{1+\max|f'|^2}, \pi\right]. \tag{4-27}$$

Furthermore, the operator $\mathbb{A}_{x_0,\tau}$ is a Fourier multiplier with symbol

$$m_\tau(\xi) := -\beta_\tau|\xi| + i\alpha_\tau\xi, \quad \xi \in \mathbb{R}.$$

Given $\operatorname{Re}\lambda \geq 1$, it is easy to see that the operator $R(\lambda, \mathbb{A}_{x_0,\tau})$ defined by

$$\mathcal{F}(R(\lambda, \mathbb{A}_{x_0,\tau})[h]) = \frac{1}{\lambda - m_\tau}\mathcal{F}h, \quad h \in H^1(\mathbb{R}),$$

belongs to $\mathcal{L}(H^1(\mathbb{R}), H^2(\mathbb{R}))$ and that it is the inverse of $\lambda - \mathbb{A}_{x_0,\tau}$. Moreover, for each $\operatorname{Re}\lambda \geq 1$ and $h \in H^2(\mathbb{R})$, we have

$$\begin{aligned}\|(\lambda - \mathbb{A}_{x_0,\tau})[h]\|_{H^1}^2 &= \int_{\mathbb{R}} (1+|\xi|^2)|\mathcal{F}((\lambda - \mathbb{A}_{x_0,\tau})[h])|^2(\xi)\,d\xi = \int_{\mathbb{R}} (1+|\xi|^2)|\lambda - m_\tau(\xi)|^2|\mathcal{F}h|^2(\xi)\,d\xi \\ &\geq \min\{1, \beta_\tau^2\}\int_{\mathbb{R}} (1+|\xi|^2)^2|\mathcal{F}h|^2(\xi)\,d\xi = \min\{1, \beta_\tau^2\}\|h\|_{H^2}^2. \end{aligned} \tag{4-28}$$

Appealing to the inequality

$$\begin{aligned}\frac{|\lambda|^2}{|\lambda - m_\tau(\xi)|^2} &= \frac{(\operatorname{Re}\lambda)^2}{(\operatorname{Re}\lambda + \beta_\tau|\xi|)^2 + (\operatorname{Im}\lambda - \alpha_\tau\xi)^2} + \frac{(\operatorname{Im}\lambda)^2}{(\operatorname{Re}\lambda + \beta_\tau|\xi|)^2 + (\operatorname{Im}\lambda - \alpha_\tau\xi)^2} \\ &\leq 1 + \frac{2(\operatorname{Im}\lambda - \alpha_\tau\xi)^2 + 2\alpha_\tau^2\xi^2}{(\operatorname{Re}\lambda + \beta_\tau|\xi|)^2 + (\operatorname{Im}\lambda - \alpha_\tau\xi)^2} \leq 1 + 2\left[1 + \left(\frac{\alpha_\tau}{\beta_\tau}\right)^2\right] \leq 3\left[1 + \left(\frac{\alpha_\tau}{\beta_\tau}\right)^2\right]\end{aligned}$$

for $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 1$, the estimate (4-26) follows from the relations (4-27) and (4-28). □

We now establish the desired generation result.

Theorem 4.4. *Let $f \in H^s(\mathbb{R})$, $s \in \left(\frac{3}{2}, 2\right)$, be given. Then*

$$-\Phi(f) \in \mathcal{H}(H^2(\mathbb{R}), H^1(\mathbb{R})). \tag{4-29}$$

Proof. Let $\kappa_0 \geq 1$ be the constant determined in Proposition 4.3. Setting $\mu := \frac{1}{2}\kappa_0$, we deduce from Theorem 4.2 that there exists a constant $\varepsilon \in (0, 1)$, an ε-localization family $\{\pi_j^\varepsilon : -N+1 \leq j \leq N\}$ that satisfies (4-1)–(4-5), a constant $K = K(\varepsilon)$, and for each $-N+1 \leq j \leq N$ and $\tau \in [0, 1]$ operators $\mathbb{A}_{j,\tau} \in \mathcal{L}(H^2(\mathbb{R}), H^1(\mathbb{R}))$ such that

$$\|\pi_j^\varepsilon \Phi(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_{H^1} \leq \frac{1}{2\kappa_0}\|\pi_j^\varepsilon h\|_{H^2} + K\|h\|_{H^{(11-2s)/4}} \tag{4-30}$$

for all $-N+1 \leq j \leq N$, $\tau \in [0, 1]$, and $h \in H^2(\mathbb{R})$. In view of Proposition 4.3, it holds that

$$\kappa_0\|(\lambda - \mathbb{A}_{j,\tau})[\pi_j^\varepsilon h]\|_{H^1} \geq |\lambda| \cdot \|\pi_j^\varepsilon h\|_{H^1} + \|\pi_j^\varepsilon h\|_{H^2} \tag{4-31}$$

for all $-N+1 \leq j \leq N$, $\tau \in [0, 1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \geq 1$, and $h \in H^2(\mathbb{R})$. The relations (4-30)–(4-31) lead us to

$$\begin{aligned}\kappa_0\|\pi_j^\varepsilon(\lambda - \Phi(\tau f))[h]\|_{H^1} &\geq \kappa_0\|(\lambda - \mathbb{A}_{j,\tau})[\pi_j^\varepsilon h]\|_{H^1} - \kappa_0\|\pi_j^\varepsilon \Phi(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_{H^1} \\ &\geq |\lambda| \cdot \|\pi_j^\varepsilon h\|_{H^1} + \tfrac{1}{2}\|\pi_j^\varepsilon h\|_{H^2} - \kappa_0 K\|h\|_{H^{(11-2s)/4}}\end{aligned}$$

for all $-N+1 \le j \le N$, $\tau \in [0,1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge 1$, and $h \in H^2(\mathbb{R})$. Summing up over $j \in \{-N+1, \dots, N\}$, we infer from Lemma 4.1 that there exists a constant $C \ge 1$ with the property that

$$C\|h\|_{H^{(11-2s)/4}} + C\|(\lambda - \Phi(\tau f))[h]\|_{H^1} \ge |\lambda| \cdot \|h\|_{H^1} + \|h\|_{H^2}$$

for all $\tau \in [0,1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge 1$, and $h \in H^2(\mathbb{R})$. Using (1-4) together with Young's inequality, we may find constants $\kappa \ge 1$ and $\omega > 0$ such that

$$\kappa\|(\lambda - \Phi(\tau f))[h]\|_{H^1} \ge |\lambda| \cdot \|h\|_{H^1} + \|h\|_{H^2} \tag{4-32}$$

for all $\tau \in [0,1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge \omega$, and $h \in H^2(\mathbb{R})$. Furthermore, combining the property

$$(\omega - \Phi(\tau f))|_{\tau=0} = \omega - \Phi(0) = \omega + \pi(-\partial_x^2)^{1/2} \in \operatorname{Isom}(H^2(\mathbb{R}), H^1(\mathbb{R}))$$

with (4-32), the method of continuity, see, e.g., [Gilbarg and Trudinger 1998, Theorem 5.2], yields that

$$\omega - \Phi(f) \in \operatorname{Isom}(H^2(\mathbb{R}), H^1(\mathbb{R})). \tag{4-33}$$

The relations (4-32) (with $\tau = 1$), (4-33), and [Lunardi 1995, Corollary 2.1.3] lead us to the desired claim (4-29). □

We are now in a position to prove the well-posedness result Theorem 1.1.

Proof of Theorem 1.1. Let $s \in \left(\frac{3}{2}, 2\right)$ and $\bar{s} \in \left(\frac{3}{2}, s\right)$ be given. Combining Lemma 3.5 and Theorem 4.4 yields

$$-\Phi \in C^{1-}(H^{\bar{s}}(\mathbb{R}), \mathcal{H}(H^2(\mathbb{R}), H^1(\mathbb{R}))).$$

Setting $\alpha := s-1$ and $\beta := \bar{s}-1$, we have $0 < \beta < \alpha < 1$ and (1-4) yields

$$H^{\bar{s}}(\mathbb{R}) = [H^1(\mathbb{R}), H^2(\mathbb{R})]_\beta \quad \text{and} \quad H^s(\mathbb{R}) = [H^1(\mathbb{R}), H^2(\mathbb{R})]_\alpha.$$

It follows now from Theorem 1.5 that (1-1), or equivalently (3-1), possesses a maximally defined solution

$$f := f(\cdot\,; f_0) \in C([0, T_+(f_0)), H^s(\mathbb{R})) \cap C((0, T_+(f_0)), H^2(\mathbb{R})) \cap C^1((0, T_+(f_0)), H^1(\mathbb{R}))$$

with

$$f \in C^{s-\bar{s}}([0,T], H^{\bar{s}}(\mathbb{R})) \quad \text{for all } T < T_+(f_0).$$

Concerning uniqueness, we now show that any classical solution

$$\tilde{f} \in C([0,\widetilde{T}), H^s(\mathbb{R})) \cap C((0,\widetilde{T}), H^2(\mathbb{R})) \cap C^1((0,\widetilde{T})), H^1(\mathbb{R})), \quad \widetilde{T} \in (0,\infty],$$

satisfies

$$\tilde{f} \in C^\eta([0,T], H^{\bar{s}}(\mathbb{R})) \quad \text{for all } T \in (0,\widetilde{T}), \tag{4-34}$$

where $\eta := (s-\bar{s})/s \in (0, s-\bar{s})$. This proves then the uniqueness claim of Theorem 1.1. We pick thus $T \in (0,\widetilde{T})$ arbitrarily. Then it follows directly from Lemma 3.1(i) that

$$\sup_{(0,T]} \|\partial_t \tilde{f}\|_2 \le C;$$

hence $\tilde{f} \in \mathrm{BC}^1((0, T], L_2(\mathbb{R}))$. Since $\tilde{f} \in C([0, T], H^s(\mathbb{R}))$, we conclude form (1-4), the previous bound, and the mean value theorem, that

$$\|\tilde{f}(t) - \tilde{f}(s)\|_{H^{\bar{s}}} \le \|\tilde{f}(t) - \tilde{f}(s)\|_2^{1-\bar{s}/s} \|\tilde{f}(t) - \tilde{f}(s)\|_{H^s}^{\bar{s}/s} \le C|t-s|^{\eta}, \quad t, s \in [0, T],$$

which proves (4-34).

Assume now that $T_+(f_0) < \infty$ and

$$\sup_{[0,T_+(f_0))} \|f(t)\|_{H^s} < \infty.$$

Arguing as above, we find that

$$\|f(t) - f(s)\|_{H^{(s+\bar{s})/2}} \le C|t-s|^{(s-\bar{s})/2s}, \quad t, s \in [0, T_+(f_0)).$$

The criterion for global existence in Theorem 1.5 applied for $\alpha := (s + \bar{s} - 2)/2$ and $\beta := \bar{s} - 1$ implies that the solution can be continued on an interval $[0, \tau)$ with $\tau > T_+(f_0)$. Moreover, it holds that

$$f \in C^{(s-\bar{s})/2}([0, T], H^{\bar{s}}(\mathbb{R})) \quad \text{for all } T \in (0, \tau).$$

The uniqueness claim in Theorem 1.5 leads us to a contradiction. Hence our assumption was false and $T_+(f_0) = \infty$. □

5. Instantaneous real-analyticity

We now improve the regularity of the solutions found in Theorems 1.1 and 1.2. To this end we first show that the mapping Φ defined by (3-2) is actually real-analytic; see Proposition 5.1. As $[f \mapsto \Phi(f)]$ is not a Nemytskij-type operator, we cannot use classical results for such operators, as presented, e.g., in [Runst and Sickel 1996]. Instead, we directly estimate the rest of the associated Taylor series. We conclude the section with the proof of Theorem 1.3, which is obtained, via Proposition 5.1, from the real-analyticity property of the semiflow as stated in Theorem 1.5, applied in the context of a nonlinear evolution problem related to (1-1).

Proposition 5.1. *Given $s \in \left(\tfrac{3}{2}, 2\right)$, it holds that*

$$\Phi \in C^{\omega}(H^s(\mathbb{R}), \mathcal{L}(H^2(\mathbb{R}), H^1(\mathbb{R}))). \tag{5-1}$$

Proof. Let $\phi : \mathbb{R} \to \mathbb{R}$ be the map defined by $\phi(x) := (1 + x^2)^{-1}$, $x \in \mathbb{R}$. Then, given $f_0 \in H^s(\mathbb{R})$, it holds that

$$\Phi(f_0)[h](x) = \mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[x,y]}h'}{y} \phi\left(\frac{\delta_{[x,y]}f_0}{y}\right) dy, \quad h \in H^2(\mathbb{R}).$$

Given $n \in \mathbb{N}$, we let

$$\begin{aligned}\partial^n \Phi(f_0)[f_1, \dots, f_n][h](x) &:= \mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[x,y]}h'}{y} \left(\prod_{i=1}^n \frac{\delta_{[x,y]}f_i}{y}\right) \phi^{(n)}\left(\frac{\delta_{[x,y]}f_0}{y}\right) dy \\ &= \sum_{k=0,\, n+k\in 2\mathbb{N}}^{n} a_k^n A_{n+k,n}(f_0, \dots, f_0)[\underbrace{f_0, \dots, f_0}_{k \text{ times}}, f_1, \dots, f_n, h'](x)\end{aligned}$$

for $f_i \in H^s(\mathbb{R})$, $1 \le i \le n$, $h \in H^2(\mathbb{R})$, and $x \in \mathbb{R}$, where $a_k^n, n \in \mathbb{N}, 0 \le k \le n$, are defined in Lemma 5.2. Arguing as in the proof of Lemma 3.5, it follows from Lemmas 3.1 and 3.4 that $\partial^n\Phi(f_0) \in \mathcal{L}^n_{\mathrm{sym}}(H^s(\mathbb{R}), \mathcal{L}(H^2(\mathbb{R}), H^1(\mathbb{R})))$; that is, $\partial^n\Phi(f_0)$ is a bounded n-linear and symmetric operator.

Moreover, given $f_0, f \in H^s(\mathbb{R})$, $n \in \mathbb{N}^*$, and $h \in H^2(\mathbb{R})$, Fubini's theorem combined with Lebesgue's dominated convergence theorem and the continuity of the mapping

$$\left[\tau \mapsto \mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[\,\cdot\,,y]}h'}{y}\left(\frac{\delta_{[\,\cdot\,,y]}f}{y}\right)^{n+1}\phi^{(n+1)}\left(\frac{\delta_{[\,\cdot\,,y]}(f_0+\tau f)}{y}\right)dy\right] : [0,1] \to H^1(\mathbb{R}),$$

yield that

$$\begin{aligned}
\Phi(f_0+f)[h](x) &- \sum_{k=0}^{n}\frac{\partial^k\Phi(f_0)[f]^k[h](x)}{k!}\\
&= \mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[x,y]}h'}{y}\left(\frac{\delta_{[x,y]}f}{y}\right)^{n+1}\int_0^1 \frac{(1-\tau)^n}{n!}\phi^{(n+1)}\left(\frac{\delta_{[x,y]}(f_0+\tau f)}{y}\right)d\tau\,dy\\
&= \int_0^1 \frac{(1-\tau)^n}{n!}\,\mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[x,y]}h'}{y}\left(\frac{\delta_{[x,y]}f}{y}\right)^{n+1}\phi^{(n+1)}\left(\frac{\delta_{[x,y]}(f_0+\tau f)}{y}\right)dy\,d\tau,
\end{aligned}$$

and

$$\begin{aligned}
&\left\|\Phi(f_0+f)[h] - \sum_{k=0}^{n}\frac{\partial^k\Phi(f_0)[f]^k[h]}{k!}\right\|_{H^1}\\
&\qquad\le \frac{1}{n!}\max_{\tau\in[0,1]}\left\|\mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[\,\cdot\,,y]}h'}{y}\left(\frac{\delta_{[\,\cdot\,,y]}f}{y}\right)^{n+1}\phi^{(n+1)}(\delta_{[\,\cdot\,,y]}f_\tau/y)\,dy\right\|_{H^1},
\end{aligned}\tag{5-2}$$

where $f_\tau := f_0+\tau f$, $0 \le \tau \le 1$. In order to estimate the right-hand side of (5-2) we note that

$$\begin{aligned}
&\left\|\mathrm{PV}\int_{\mathbb{R}} \frac{\delta_{[\,\cdot\,,y]}h'}{y}\left(\frac{\delta_{[\,\cdot\,,y]}f}{y}\right)^{n+1}\phi^{(n+1)}(\delta_{[\,\cdot\,,y]}f_\tau/y)\,dy\right\|_{H^1}\\
&\quad\le \sum_{k=0,\,n+k+1\in2\mathbb{N}}^{n+1}|a_k^{n+1}|\,\|A_{n+k+1,n+1}(f_\tau,\dots,f_\tau)[\underbrace{f_\tau,\dots,f_\tau}_{k\text{ times}},\underbrace{f,\dots,f}_{n+1\text{ times}},h']\|_{H^1}\\
&\quad\le \sum_{k=0,\,n+k+1\in2\mathbb{N}}^{n+1}|a_k^{n+1}|\,\|A_{n+k+1,n+1}(f_\tau,\dots,f_\tau)[\underbrace{f_\tau,\dots,f_\tau}_{k\text{ times}},\underbrace{f,\dots,f}_{n+1\text{ times}},h']\|_2\\
&\qquad+ \sum_{k=0,\,n+k+1\in2\mathbb{N}}^{n+1}|a_k^{n+1}|\,\|A_{n+k+1,n+1}(f_\tau,\dots,f_\tau)[\underbrace{f_\tau,\dots,f_\tau}_{k\text{ times}},\underbrace{f,\dots,f}_{n+1\text{ times}},h'']\|_2\\
&\qquad+k \sum_{k=0,\,n+k+1\in2\mathbb{N}}^{n+1}|a_k^{n+1}|\,\|A_{n+k+1,n+1}(f_\tau,\dots,f_\tau)[\underbrace{f_\tau,\dots,f_\tau}_{k-1\text{ times}},\underbrace{f,\dots,f}_{n+1\text{ times}},f_\tau',h']\|_2\\
&\qquad-2(n+2) \sum_{k=0,\,n+k+1\in2\mathbb{N}}^{n+1}|a_k^{n+1}|\,\|A_{n+k+3,n+2}(f_\tau,\dots,f_\tau)[\underbrace{f_\tau,\dots,f_\tau}_{k+1\text{ times}},\underbrace{f,\dots,f}_{n+1\text{ times}},f_\tau',h']\|_2.
\end{aligned}\tag{5-3}$$

Combining the results of Lemmas 5.2–5.4, we conclude that there exists an integer $p > 0$ and a positive constant C (depending only on $\|f_0\|_{H^s}$) such that for all $f \in H^s(\mathbb{R})$ with $\|f\|_{H^s} \le 1$ and all $n \ge 3$ we have

$$\left\| \Phi(f_0+f) - \sum_{k=0}^{n} \frac{\partial^k \Phi(f_0)[f]^k}{k!} \right\|_{\mathcal{L}(H^2(\mathbb{R}),H^1(\mathbb{R}))} \le C^{n+1} n^p \|f\|_{H^s}^{n+1}.$$

The claim follows. □

The following technical results are used in the proof of Proposition 5.1.

Lemma 5.2. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be defined by $\phi(x) := (1+x^2)^{-1}$, $x \in \mathbb{R}$. Given $n \in \mathbb{N}$, it holds that*

$$\phi^{(n)}(x) = \frac{1}{(1+x^2)^{n+1}} \sum_{k=0}^{n} a_k^n x^k,$$

where the coefficients $a_k^n \in \mathbb{R}$ satisfy $|a_k^n| \le 4^n (n+2)!$ for all $0 \le k \le n$. Moreover, $a_k^n = 0$ if $n+k \notin 2\mathbb{N}$.

Proof. The claim for $n \in \{0, 1, 2, 3\}$ is obvious. Assume that the claim holds for some integer $n \ge 3$. Since

$$(1+x^2)^{n+2} \phi^{(n+1)}(x) = (1+x^2) \sum_{k=1}^{n} k a_k^n x^{k-1} - 2(n+1)x \sum_{k=0}^{n} a_k^n x^k,$$

the coefficient a_k^{n+1}, $0 \le k \le n+1$, of x^k satisfies

$$\begin{aligned} |a_{n+1}^{n+1}| &\le n|a_n^n| + 2(n+1)|a_n^n| \le 4(n+1)|a_n^n| \le 4^{n+1}(n+3)!\,, \\ |a_n^{n+1}| &\le (n-1)|a_{n-1}^n| + 2(n+1)|a_{n-1}^n| = 0, \end{aligned}$$

and for $n-1 \ge k \ge 2$ we have

$$|a_k^{n+1}| \le (k+1)|a_{k+1}^n| + (k-1)|a_{k-1}^n| + 2(n+1)|a_{k-1}^n| \le 4^{n+1}(n+3)!\,,$$

while

$$\begin{aligned} |a_1^{n+1}| &\le 2|a_2^n| + 2(n+1)|a_0^n| \le 4^{n+1}(n+3)!\,, \\ |a_0^{n+1}| &\le |a_1^n| \le 4^{n+1}(n+3)!\,. \end{aligned}$$

The conclusion is now obvious. □

In the next lemma we estimate the first two terms that appear on the right-hand side of (5-3).

Lemma 5.3. *Let $n, k \in \mathbb{N}$ satisfy $n \ge 3$ and $0 \le k \le n+1$, and let $s \in \left(\frac{3}{2}, 2\right)$. Given $f, f_\tau \in H^s(\mathbb{R})$, it holds that*

$$\big\| A_{n+k+1,n+1}(f_\tau, \dots, f_\tau)[\underbrace{f_\tau, \dots, f_\tau}_{k \text{ times}}, \underbrace{f, \dots, f}_{n+1 \text{ times}}, \cdot\,] \big\|_{\mathcal{L}(L_2(\mathbb{R}))} \le C^n n^4 \max\{1, \|f_\tau\|_{H^s}^4\} \|f\|_{H^s}^{n+1}, \tag{5-4}$$

with a constant $C \ge 1$ independent of n, k, f, and f_τ.

Proof. Much as in the proof of Lemma 3.1 we write

$$A_{n+k+1,n+1}(f_\tau, \dots, f_\tau)[f_\tau, \dots, f_\tau, f, \dots, f, \cdot\,] = M - S,$$

where M is the multiplication operator

$$M[h](x) := h(x)\,\mathrm{PV}\int_{\mathbb{R}} \frac{1}{y}\left(\frac{\delta_{[x,y]}f}{y}\right)^{n+1} \frac{(\delta_{[x,y]}f_\tau/y)^k}{[1+(\delta_{[x,y]}f_\tau/y)^2]^{n+2}}\,dy$$

and S is the singular integral operator

$$S[h](x) := \mathrm{PV}\int_{\mathbb{R}} \left(\frac{\delta_{[x,y]}f}{y}\right)^{n+1} \frac{(\delta_{[x,y]}f_\tau/y)^k}{[1+(\delta_{[x,y]}f_\tau/y)^2]^{n+2}} \frac{h(x-y)}{y}\,dy$$

for $h \in L_2(\mathbb{R})$. Arguing as in proof of Lemma 3.1, it follows that

$$\|M\|_{\mathcal{L}(L_2(\mathbb{R}))} \le nC^n \max\{1, \|f_\tau\|_{H^s}\}\|f\|_{H^s}^{n+1} \tag{5-5}$$

with a constant $C \ge 1$ independent of n, k, f, and f_τ.

In order to deal with the operator S we consider the functions $F:\mathbb{R}^2 \to \mathbb{R}$ and $A:\mathbb{R}\to\mathbb{R}^2$ defined by

$$F(x_1,x_2) := \frac{x_1^{n+1}x_2^k}{(1+x_2^2)^{n+2}}, \quad A := (A_1, A_2) := (f, f_\tau).$$

The function F is smooth, A is Lipschitz continuous, and we set

$$a_j := \|A_j'\|_\infty, \quad 1 \le j \le 2.$$

Since S is the singular integral operator with kernel

$$K(x,y) := \frac{1}{y}F\left(\frac{\delta_{[x,y]}A}{y}\right), \quad x \in \mathbb{R},\ \ y \ne 0,$$

and $|\delta_{[x,y]}A_j/y| \le a_j$ for $1 \le j \le 2$, it is natural to introduce a smooth periodic function $\widetilde{F}$ on $\mathbb{R}^2$, which is $4a_j$-periodic in the variable x_j, $1 \le j \le 2$, and which matches F on $\prod_{j=1}^2[-a_j, a_j]$. More precisely, we choose $\varphi \in C_0^\infty(\mathbb{R}, [0,1])$ with $\varphi = 1$ on $[|x| \le 1]$ and $\varphi = 0$ on $[|x| \ge 2]$ and we define $\widetilde{F}$ to be the periodic extension of

$$\left[(x_1,x_2) \mapsto F(x_1,x_2)\prod_{j=1}^2 \varphi\left(\frac{x_j}{a_j}\right)\right] : Q \to \mathbb{R},$$

where $Q := \prod_{j=1}^2[-2a_j, 2a_j]$. We now expand $\widetilde{F}$ by its Fourier series

$$\widetilde{F}(x_1,x_2) = \sum_{p\in\mathbb{Z}^2} \alpha_p \exp\left(i\sum_{j=1}^2 \frac{p_jx_j}{T_j}\right),$$

where

$$T_j := \frac{2a_j}{\pi}, \qquad \alpha_p := \frac{1}{4^2a_1a_2}\int_Q \widetilde{F}(x_1,x_2)\exp\left(-i\sum_{j=1}^2\frac{p_jx_j}{T_j}\right)d(x_1,x_2), \quad p \in \mathbb{Z}^2,$$

and observe that

$$K(x,y) = \frac{1}{y}\widetilde{F}\left(\frac{\delta_{[x,y]}A}{y}\right) = \sum_{p\in\mathbb{Z}^2}\alpha_pK_p(x,y), \quad x\in\mathbb{R},\ \ y\ne 0,$$

with

$$K_p(x, y) := \frac{1}{y} \exp\left(i \frac{\delta_{[x,y]}\left(\sum_{j=1}^{2}(p_j/T_j)A_j\right)}{y}\right), \quad x \in \mathbb{R}, \ \ y \neq 0, \ \ p \in \mathbb{Z}^2.$$

The kernels K_p define operators in $\mathcal{L}(L_2(\mathbb{R}))$ of type (1-2) and the norms of these operators can be estimated from above by

$$C\left(1 + \left\| \sum_{j=1}^{2} \frac{p_j}{T_j} A_j' \right\|_\infty\right) \leq C(1 + |p|), \quad p \in \mathbb{Z}^2.$$

Since $\sum_{p\in\mathbb{Z}^2}(1+|p|^3)^{-1} < \infty$ we get

$$\|S\|_{\mathcal{L}(L_2(\mathbb{R}))} \leq C \sum_{p\in\mathbb{Z}^2} |\alpha_p|(1+|p|) \leq C \sup_{p\in\mathbb{Z}^2} [(1+|p|^4)|\alpha_p|].$$

We estimate next the quantity $\sup_{p\in\mathbb{Z}^2}(1+|p|^4)|\alpha_p|$. To this end we write

$$\alpha_p = \frac{1}{4^2} \prod_{j=1}^{2} \frac{I_j}{a_j},$$

where

$$I_1 := \int_{-2a_1}^{2a_1} x_1^{n+1} \varphi\left(\frac{x_1}{a_1}\right) e^{-ip_1x_1/T_1}\, dx_1, \quad I_2 := \int_{-2a_2}^{2a_2} \frac{x_2^k}{(1+x_2^2)^{n+2}} \varphi\left(\frac{x_2}{a_2}\right) e^{-ip_2x_2/T_2}\, dx_2.$$

Since $\varphi = 0$ in $[|x| \geq 2]$ and $n \geq 3$, integration by parts leads us, in the case when $p_1 \neq 0$, to

$$|I_1| \leq \left(\frac{T_1}{|p_1|}\right)^4 \int_{-2a_1}^{2a_1} \left| \left(x_1^{n+1} \varphi\left(\frac{x_1}{a_1}\right)\right)^{(4)} \right| dx_1 \leq C \frac{2^n n^4 a_1^{n+2}}{p_1^4}, \tag{5-6}$$

and similarly, since $x_2 \leq 1 + x_2^2$, we find for $p_2 \neq 0$ that

$$|I_2| \leq C \frac{n^4 \max\{a_2, a_2^5\}}{p_2^4}. \tag{5-7}$$

The estimates

$$|I_1| \leq C 2^n a_1^{n+2}, \quad |I_2| \leq C a_2, \tag{5-8}$$

are valid for all $p \in \mathbb{Z}^2$. Combining (5-6)–(5-8), we arrive at

$$\sup_{p\in\mathbb{Z}^2} (1+|p|^4)|\alpha_p| \leq C 2^n n^4 \max\{1, a_2^4\} a_1^{n+1},$$

which leads us to

$$\|S\|_{\mathcal{L}(L_2(\mathbb{R}))} \leq C 2^n n^4 \max\{1, \|f_\tau'\|_\infty^4\} \|f'\|_\infty^{n+1} \leq n^4 C^n \max\{1, \|f_\tau\|_{H^s}^4\} \|f\|_{H^s}^{n+1}.$$

This inequality together with (5-5) proves the desired claim. □

In the next lemma we estimate the last two terms on the right-hand side of (5-3) in the proof of Proposition 5.1.

Lemma 5.4. *Let $n, k \in \mathbb{N}$ satisfy $n \geq 1$ and $0 \leq k \leq n+1$. Let further $l \in \{0, 1\}$ and $s \in \left(\frac{3}{2}, 2\right)$. Given $f, f_\tau \in H^s(\mathbb{R})$, it holds that*

$$\bigl\| A_{n+k+1+2l,n+1+l}(f_\tau, \dots, f_\tau)[\underbrace{f_\tau, \dots, f_\tau}_{k-1+2l \text{ times}}, \underbrace{f, \dots, f}_{n+1 \text{ times}}, f'_\tau, \cdot\,] \bigr\|_{\mathcal{L}(H^1(\mathbb{R}), L_2(\mathbb{R}))} \leq C^n \|f_\tau\|_{H^s} \|f\|_{H^s}^{n+1}, \quad (5\text{-}9)$$

with a constant $C \geq 1$ independent of n, k, f, and f_τ.

Proof. The proof is similar to that of Lemma 3.4. □

We are now in a position to prove Theorem 1.3 when $\sigma = 0$, where we use a parameter trick which appears, in other forms, also in [Angenent 1990; Escher and Simonett 1996; Prüss et al. 2015]. We present a new idea which uses only the abstract result Theorem 1.5 in the context of an evolution problem related to (1-1), and not explicitly the maximal regularity property as in the above-mentioned papers. The proof when $\sigma > 0$ is almost identical and is also discussed below, but it relies on some properties established in Section 6.

Proof of Theorem 1.3. Assume first that $\sigma = 0$. We then pick $f_0 \in H^s(\mathbb{R})$, $s \in \left(\frac{3}{2}, 2\right)$, and we let $f = f(\,\cdot\,; f_0) : [0, T_+(f_0)) \to H^s(\mathbb{R})$ denote the unique maximal solution to (1-1), whose existence is guaranteed by Theorem 1.1. We further choose $\lambda_1, \lambda_2 \in (0, \infty)$ and we define

$$f_{\lambda_1,\lambda_2}(t, x) := f(\lambda_1 t, x + \lambda_2 t), \qquad x \in \mathbb{R}, \ \ 0 \leq t < T_+ := T_+(f_0)/\lambda_1.$$

Classical arguments show that

$$f_{\lambda_1,\lambda_2} \in C([0, T_+), H^s(\mathbb{R})) \cap C((0, T_+), H^2(\mathbb{R})) \cap C^1((0, T_+), H^1(\mathbb{R})).$$

We next introduce the function $u := (u_1, u_2, u_3) : [0, T_+) \to \mathbb{R}^2 \times H^s(\mathbb{R})$, where

$$(u_1, u_2)(t) := (\lambda_1, \lambda_2), \qquad u_3(t) := f_{\lambda_1,\lambda_2}(t), \qquad 0 \leq t < T_+,$$

and we note that u solves the quasilinear evolution problem

$$\dot{u} = \Psi(u)[u], \qquad t > 0, \qquad u(0) = (\lambda_1, \lambda_2, f_0), \tag{5-10}$$

with $\Psi : (0, \infty)^2 \times H^s(\mathbb{R}) \to \mathcal{L}(\mathbb{R}^2 \times H^2(\mathbb{R}), \mathbb{R}^2 \times H^1(\mathbb{R}))$ denoting the operator defined by

$$\Psi((v_1, v_2, v_3))[(u_1, u_2, u_3)] := (0, 0, v_1 \Phi(v_3)[u_3] + v_2 \partial_x u_3). \tag{5-11}$$

Proposition 5.1 immediately yields

$$\Psi \in C^\omega\bigl((0, \infty)^2 \times H^s(\mathbb{R}), \mathcal{L}(\mathbb{R}^2 \times H^2(\mathbb{R}), \mathbb{R}^2 \times H^1(\mathbb{R}))\bigr) \quad \text{for all } s \in \left(\tfrac{3}{2}, 2\right).$$

Given $v := (v_1, v_2, v_3) \in (0, \infty)^2 \times H^s(\mathbb{R})$, the operator $\Psi(v)$ can be represented as a matrix

$$\Psi(v) = \begin{pmatrix} 0 & 0 \\ 0 & v_1 \Phi(v_3) + v_2 \partial_x \end{pmatrix} : \mathbb{R}^2 \times H^2(\mathbb{R}) \to \mathbb{R}^2 \times H^1(\mathbb{R}),$$

and we infer from [Amann 1995, Corollary I.1.6.3] that $-\Psi(v) \in \mathcal{H}(\mathbb{R}^2 \times H^2(\mathbb{R}), \mathbb{R}^2 \times H^1(\mathbb{R}))$ if and only if

$$-(v_1\Phi(v_3) + v_2\partial_x) \in \mathcal{H}(H^2(\mathbb{R}), H^1(\mathbb{R})). \tag{5-12}$$

We note that $v_2\partial_x$ is a first-order Fourier multiplier and its symbol is purely imaginary. Therefore, obvious modifications of the arguments presented in the proofs of Theorem 4.2 and Proposition 4.3 enable us to conclude that the property (5-12) is satisfied for each $(v_1, v_2, v_3) \in (0,\infty)^2 \times H^s(\mathbb{R})$ and $s \in \left(\frac{3}{2}, 2\right)$. Setting $\mathbb{F}_0 := \mathbb{R}^2 \times H^1(\mathbb{R})$ and $\mathbb{F}_1 := \mathbb{R}^2 \times H^2(\mathbb{R})$, it holds that

$$[\mathbb{F}_0, \mathbb{F}_1]_\theta = \mathbb{R}^2 \times H^{1+\theta}(\mathbb{R}), \quad \theta \in (0,1),$$

and we may now apply Theorem 1.5 in the context of the quasilinear parabolic problem (5-10) to conclude (much as in the proof of Theorem 1.1), for each $u_0 = (\lambda_1, \lambda_2, f_0) \in (0,\infty)^2 \times H^s(\mathbb{R})$, $s \in \left(\frac{3}{2}, 2\right)$, the existence of a unique maximal solution

$$u := u(\cdot\,; u_0) \in C([0, T_+(u_0)), (0,\infty)^2 \times H^s(\mathbb{R})) \cap C((0, T_+(u_0)), \mathbb{F}_1) \cap C^1((0, T_+(u_0)), \mathbb{F}_0).$$

Additionally, the set

$$\Omega := \{(\lambda_1, \lambda_2, f_0, t) : t \in (0, T_+((\lambda_1, \lambda_2, f_0)))\}$$

is open in $(0,\infty)^2 \times H^s(\mathbb{R}) \times (0,\infty)$ and

$$[(\lambda_1, \lambda_2, f_0, t) \mapsto u(t; (\lambda_1, \lambda_2, f_0))] : \Omega \to \mathbb{R}^2 \times H^s(\mathbb{R})$$

is a real-analytic map.

So, if we fix $f_0 \in H^s(\mathbb{R})$, then we may conclude that

$$\frac{T_+(f_0)}{\lambda_1} = T_+((\lambda_1, \lambda_2, f_0)) \quad \text{for all } (\lambda_1, \lambda_2) \in (0,\infty)^2.$$

As we want to prove that $f = f(\cdot\,; f_0)$ is real-analytic in $(0, T_+(f_0)) \times \mathbb{R}$, it suffices to establish the real-analyticity property in a small ball around (t_0, x_0) for each $x_0 \in \mathbb{R}$ and $t_0 \in (0, T_+(f_0))$. Let thus $(t_0, x_0) \in (0, T_+(f_0)) \times \mathbb{R}$ be arbitrary. For $(\lambda_1, \lambda_2) \in \mathbb{B}((1,1), \varepsilon) \subset (0,\infty)^2$, with ε chosen suitably small, we have that

$$t_0 < T_+((\lambda_1, \lambda_2, f_0)) \quad \text{for all } (\lambda_1, \lambda_2) \in \mathbb{B}((1,1), \varepsilon),$$

and therewith

$$\mathbb{B}((1,1), \varepsilon) \times \{f_0\} \times \{t_0\} \subset \Omega.$$

Moreover, since $u_3(\cdot\,; u_0) = f_{\lambda_1,\lambda_2}$, the restriction

$$[(\lambda_1, \lambda_2) \mapsto f_{\lambda_1,\lambda_2}(t_0)] : \mathbb{B}((1,1), \varepsilon) \to H^s(\mathbb{R}) \tag{5-13}$$

is a real-analytic map. Since $[h \mapsto h(x_0 - t_0)] : H^s(\mathbb{R}) \to \mathbb{R}$ is a linear operator, the composition

$$[(\lambda_1, \lambda_2) \mapsto f(\lambda_1 t_0, x_0 - t_0 + \lambda_2 t_0)] : \mathbb{B}((1,1), \varepsilon) \to \mathbb{R} \tag{5-14}$$

is real-analytic too. Furthermore, for $\delta > 0$ small, the mapping $\varphi : \mathbb{B}((t_0, x_0), \delta) \to \mathbb{B}((1, 1), \varepsilon)$ with

$$\varphi(t, x) := \left(\frac{t}{t_0}, \frac{x - x_0 + t_0}{t_0}\right) \tag{5-15}$$

is well-defined and real-analytic, and therefore the composition of the functions defined by (5-14) and (5-15), that is, the mapping

$$[(t, x) \mapsto f(t, x)] : \mathbb{B}((t_0, x_0), \delta) \to \mathbb{R},$$

is also real-analytic. This proves the first claim.

Finally, the property $f \in C^\omega((0, T_+(f_0)), H^k(\mathbb{R}))$ for arbitrary $k \in \mathbb{N}$, is an immediate consequence of (5-13).

The arguments presented above carry over to the case when $\sigma > 0$ (with the obvious modifications). If $\sigma > 0$, the operator $v_2 \partial_x$ appearing in (5-12) can be regarded as being a lower-order perturbation and therefore the generator property of $\Psi(v)$ follows in this case directly from the corresponding property of the original operator; see Theorem 6.3. □

6. The Muskat problem with surface tension and gravity effects

We now consider surface-tension forces acting at the interface between the fluids; that is, we take $\sigma > 0$. The motion of the fluids may also be influenced by gravity, but we make no restrictions on Δ_ρ, which is now an arbitrary real number. If we model flows in a vertical Hele-Shaw cell, this means in particular that the lower fluid may be less dense than the fluid above. Since Δ_ρ can be zero, (1-1) is also a model for fluid motions in a horizontal Hele-Shaw cell as for these flows the effects due to gravity are usually neglected, that is, $g = 0$. Again, we rescale the time appropriately and rewrite (1-1) as the system

$$\begin{cases} \partial_t f(t,x) = f'(t,x)\,\mathrm{PV}\displaystyle\int_{\mathbb{R}} \frac{f(t,x) - f(t,x-y)}{y^2 + (f(t,x) - f(t,x-y))^2} (\kappa(f))'(t, x-y)\,dy \\ \qquad\qquad + \mathrm{PV}\displaystyle\int_{\mathbb{R}} \frac{y}{y^2 + (f(t,x) - f(t,x-y))^2} (\kappa(f))'(t, x-y)\,dy \\ \qquad\qquad + \Theta\,\mathrm{PV}\displaystyle\int_{\mathbb{R}} \frac{y(f'(t,x) - f'(t,x-y))}{y^2 + (f(t,x) - f(t,x-y))^2}\,dy \qquad \text{for } t > 0,\ x \in \mathbb{R}, \\ f(0, \cdot) = f_0, \end{cases} \tag{6-1}$$

with

$$\Theta := \frac{\Delta\rho}{\sigma} \in \mathbb{R}.$$

Since

$$(\kappa(f))' = \frac{f'''}{(1 + f'^2)^{3/2}} - 3\frac{f' f''^2}{(1 + f'^2)^{5/2}},$$

we observe that the first equation of (6-1) is again quasilinear, but now this property is due to the fact that $(\kappa(f))'$ is an affine function in the variable f'''.

To be more precise we set

$$\begin{aligned}\Phi_\sigma(f)[h](x) := f'(x)\,\mathrm{PV}\!\int_{\mathbb{R}} \frac{\delta_{[x,y]}f}{y^2+(\delta_{[x,y]}f)^2}\Big(\frac{h'''}{(1+f'^2)^{3/2}} - 3\frac{f'f''h''}{(1+f'^2)^{5/2}}\Big)(x-y)\,dy \\ + \mathrm{PV}\!\int_{\mathbb{R}} \frac{y}{y^2+(\delta_{[x,y]}f)^2}\Big(\frac{h'''}{(1+f'^2)^{3/2}} - 3\frac{f'f''h''}{(1+f'^2)^{5/2}}\Big)(x-y)\,dy \\ + \Theta\,\mathrm{PV}\!\int_{\mathbb{R}} \frac{y(\delta_{[x,y]}h')}{y^2+(\delta_{[x,y]}f)^2}\,dy,\end{aligned} \tag{6-2}$$

and we recast the problem (6-1) in the compact form

$$\dot f = \Phi_\sigma(f)[f], \quad t>0, \quad f(0)=f_0. \tag{6-3}$$

We emphasize that there are also other ways to write (6-1) as a quasilinear problem. For example the terms containing only $f^{(l)}$, $0 \le l \le 2$, can be viewed as a nonlinear function $[f \mapsto F(f)]$ which would appear as an additive term to the right-hand side of (6-3), with $\Phi_\sigma(f)[h]$ modified accordingly. However, the formulation (6-2)–(6-3) appears to us as being optimal as it allows us to consider the largest set of initial data among all formulations. To be more precise, the operator introduced by (6-2) satisfies, with the notation in Lemma 3.1 and Remark 3.3, the relation

$$\begin{aligned}&\Phi_\sigma(f)[h]\\ &= f'B_{1,1}(f,f)\Big[\frac{h'''}{(1+f'^2)^{3/2}} - 3\frac{f'f''h''}{(1+f'^2)^{5/2}}\Big] + B_{0,1}(f)\Big[\frac{h'''}{(1+f'^2)^{3/2}} - 3\frac{f'f''h''}{(1+f'^2)^{5/2}}\Big] + \Theta A_{0,0}(f)[h'],\end{aligned}$$

and we now claim, based on the results in Section 5, that

$$\Phi_\sigma \in C^\omega(H^2(\mathbb{R}), \mathcal{L}(H^3(\mathbb{R}), L_2(\mathbb{R}))). \tag{6-4}$$

Indeed, arguing as in Section 5, it follows that

$$\begin{aligned}\Big[f \mapsto \Big[h \mapsto \mathrm{PV}\!\int_{\mathbb{R}} \frac{\delta_{[\,\cdot\,,y]}f}{y^2+(\delta_{[\,\cdot\,,y]}f)^2} h(\cdot - y)\,dy\Big]\Big] &\in C^\omega(H^2(\mathbb{R}), \mathcal{L}(L_2(\mathbb{R}))),\\ \Big[f \mapsto \Big[h \mapsto \mathrm{PV}\!\int_{\mathbb{R}} \frac{y}{y^2+(\delta_{[\,\cdot\,,y]}f)^2} h(\cdot - y)\,dy\Big]\Big] &\in C^\omega(H^2(\mathbb{R}), \mathcal{L}(L_2(\mathbb{R}))),\\ \Big[f \mapsto \Big[h \mapsto \mathrm{PV}\!\int_{\mathbb{R}} \frac{y(\delta_{[\,\cdot\,,y]}h')}{y^2+(\delta_{[\,\cdot\,,y]}f)^2}\,dy\Big]\Big] &\in C^\omega(H^2(\mathbb{R}), \mathcal{L}(H^3(\mathbb{R}), L_2(\mathbb{R}))).\end{aligned} \tag{6-5}$$

Moreover, classical arguments, see, e.g., [Runst and Sickel 1996, Theorem 5.5.3/4], yield that

$$\Big[f \mapsto \Big[h \mapsto \frac{h'''}{(1+f'^2)^{3/2}} - 3\frac{f'f''h''}{(1+f'^2)^{5/2}}\Big]\Big] \in C^\omega(H^2(\mathbb{R}), \mathcal{L}(H^3(\mathbb{R}), L_2(\mathbb{R}))). \tag{6-6}$$

The relations (6-5)–(6-6) immediately imply (6-4).

In the following, we prove that $\Phi_\sigma(f)$ is, for each $f \in H^2(\mathbb{R})$, the generator of a strongly continuous and analytic semigroup in $\mathcal{L}(L_2(\mathbb{R}))$, that is,

$$-\Phi_\sigma(f) \in \mathcal{H}(H^3(\mathbb{R}), L_2(\mathbb{R})).$$

To this end we write

$$\Phi_\sigma = \Phi_{\sigma,1} + \Phi_{\sigma,2}, \tag{6-7}$$

where

$$\Phi_{\sigma,1}(f)[h] := f' B_{1,1}(f, f)\Big[\frac{h'''}{(1+f'^2)^{3/2}}\Big] + B_{0,1}(f)\Big[\frac{h'''}{(1+f'^2)^{3/2}}\Big], \tag{6-8}$$

$$\Phi_{\sigma,2}(f)[h] := -3f' B_{1,1}(f, f)\Big[\frac{f'f''h''}{(1+f'^2)^{5/2}}\Big] - 3B_{0,1}(f)\Big[\frac{f'f''h''}{(1+f'^2)^{5/2}}\Big] + \Theta A_{0,0}(f)[h'] \tag{6-9}$$

for $f \in H^2(\mathbb{R})$, $h \in H^3(\mathbb{R})$. Since $\Phi_{\sigma,2}(f) \in \mathcal{L}(H^{8/3}(\mathbb{R}), L_2(\mathbb{R}))$ and $[L_2(\mathbb{R}), H^3(\mathbb{R})]_{8/9} = H^{8/3}(\mathbb{R})$, we can view $\Phi_{\sigma,2}(f)$ as being a lower-order perturbation, see [Lunardi 1995, Proposition 2.4.1], and we only need to establish the generator property for the leading-order term $\Phi_{\sigma,1}(f)$. Much as in Section 4, we consider a continuous mapping

$$[\tau \mapsto \Phi_{\sigma,1}(\tau f)] : [0, 1] \to \mathcal{L}(H^3(\mathbb{R}), L_2(\mathbb{R})),$$

which transforms the operator $\Phi_{\sigma,1}(f)$ into the operator

$$\Phi_{\sigma,1}(0) = B_{0,1}(0) \circ \partial_x^3 = -\pi(\partial_x^4)^{3/4},$$

where $(\partial_x^4)^{3/4}$ is the Fourier multiplier with symbol $[\xi \mapsto |\xi|^3]$. We now establish the following result.

Theorem 6.1. *Let $f \in H^2(\mathbb{R})$ and $\mu > 0$ be given.*

Then, there exist $\varepsilon \in (0, 1)$, a finite ε-localization family $\{\pi_j^\varepsilon : -N+1 \le j \le N\}$ satisfying (4-1)–(4-5), *a constant $K = K(\varepsilon)$, and for each $j \in \{-N+1, \dots, N\}$ and $\tau \in [0, 1]$ there exist operators*

$$\mathbb{A}_{j,\tau} \in \mathcal{L}(H^3(\mathbb{R}), L_2(\mathbb{R}))$$

such that

$$\|\pi_j^\varepsilon \Phi_{\sigma,1}(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_2 \le \mu \|\pi_j^\varepsilon h\|_{H^3} + K\|h\|_{H^2} \tag{6-10}$$

for all $j \in \{-N+1, \dots, N\}$, $\tau \in [0, 1]$, and $h \in H^3(\mathbb{R})$. The operators $\mathbb{A}_{j,\tau}$ are defined by

$$\mathbb{A}_{j,\tau} := -\frac{\pi}{(1+\tau^2 f'^2(x_j^\varepsilon))^{3/2}} (\partial_x^4)^{3/4}, \quad |j| \le N-1, \tag{6-11}$$

where x_j^ε is a point belonging to $\operatorname{supp} \pi_j^\varepsilon$, *and*

$$\mathbb{A}_{N,\tau} := -\pi(\partial_x^4)^{3/4}. \tag{6-12}$$

Proof. Let $\{\pi_j^\varepsilon : -N+1 \le j \le N\}$ be an ε-localization family satisfying the properties (4-1)–(4-5) and $\{\chi_j^\varepsilon : -N+1 \le j \le N\}$ be an associated family satisfying (4-6)–(4-8), with $\varepsilon \in (0, 1)$ which will be fixed below.

To deal with both terms of $\Phi_{\sigma,1}(\tau f)$, see (6-8), at once, we consider the operator

$$K_a(\tau f)[h] := f'_{a,\tau} B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{h'''}{(1+\tau^2 f'^2)^{3/2}}\Big],$$

where, for $a \in \{0, 1\}$, we set

$$f_{a,\tau} := (1-a)\tau f + a\, \mathrm{id}_{\mathbb{R}}.$$

For $a = 0$ we recover the first term in the definition of $\Phi_{\sigma,1}(\tau f)[h]$, while for $a = 1$ the expression matches the second one.

In the following, $h \in H^3(\mathbb{R})$ is arbitrary. Again, constants which are independent of ε (and, of course, of $h \in H^3(\mathbb{R})$, $\tau \in [0, 1]$, $a \in \{0, 1\}$, and $j \in \{-N+1, \dots, N\}$) are denoted by C, while the constants that we denote by K may depend only upon ε. We further let

$$\mathbb{A}^a_{j,\tau} := -\pi \frac{f'^2_{a,\tau}(x^\varepsilon_j)}{(1+\tau^2 f'^2(x^\varepsilon_j))^{5/2}}(\partial_x^4)^{3/4} \quad \text{for } |j| \le N-1,$$

and

$$\mathbb{A}^a_{N,\tau} := -\pi a^2 (\partial_x^4)^{3/4}.$$

We analyze the cases $j = N$ and $|j| \le N-1$ separately.

The case $|j| \le N-1$. For $|j| \le N-1$ we write

$$\pi^\varepsilon_j K_a(\tau f)[h] - \mathbb{A}^a_{j,\tau}[\pi^\varepsilon_j h] := T_1[h] + T_2[h] + T_3[h], \tag{6-13}$$

where

$$\begin{aligned}
T_1[h] &:= \pi^\varepsilon_j K_a(\tau f)[h] - f'_{a,\tau}(x^\varepsilon_j) B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi^\varepsilon_j h'''}{(1+\tau^2 f'^2)^{3/2}}\Big],\\
T_2[h] &:= f'_{a,\tau}(x^\varepsilon_j) B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi^\varepsilon_j h'''}{(1+\tau^2 f'^2)^{3/2}}\Big] - \frac{f'_{a,\tau}(x^\varepsilon_j)}{(1+\tau^2 f'^2(x^\varepsilon_j))^{3/2}} B_{1,1}(f_{a,\tau}, \tau f)[\pi^\varepsilon_j h'''],\\
T_3[h] &:= \frac{f'_{a,\tau}(x^\varepsilon_j)}{(1+\tau^2 f'^2(x^\varepsilon_j))^{3/2}} B_{1,1}(f_{a,\tau}, \tau f)[\pi^\varepsilon_j h'''] - \mathbb{A}^a_{j,\tau}[\pi^\varepsilon_j h].
\end{aligned}$$

We consider first the term $T_1[h]$. The identity $\chi^\varepsilon_j \pi^\varepsilon_j = 1$ on $\operatorname{supp} \pi^\varepsilon_j$ and integration by parts lead us to the relation

$$\begin{aligned}
T_1[h] = \chi^\varepsilon_j (f'_{a,\tau} - f'_{a,\tau}(x^\varepsilon_j)) B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi^\varepsilon_j h'''}{(1+\tau^2 f'^2)^{3/2}}\Big]&\\
+(1-\chi^\varepsilon_j)(f'_{a,\tau}(x^\varepsilon_j) - f'_{a,\tau}) B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi^{\varepsilon\prime}_j h''}{(1+\tau^2 f'^2)^{3/2}} - 3\tau^2 \frac{\pi^\varepsilon_j f' f'' h''}{(1+\tau^2 f'^2)^{5/2}}\Big]&\\
+(f'_{a,\tau}(x^\varepsilon_j) - f'_{a,\tau})\Big\{B_{1,1}(\chi^\varepsilon_j, \tau f)\Big[\frac{\pi^\varepsilon_j f'_{a,\tau} h''}{(1+\tau^2 f'^2)^{3/2}}\Big] - 2B_{2,1}(f_{a,\tau}, \chi^\varepsilon_j, \tau f)\Big[\frac{\pi^\varepsilon_j h''}{(1+\tau^2 f'^2)^{3/2}}\Big]\Big\}&\\
-2\tau^2 (f'_{a,\tau}(x^\varepsilon_j) - f'_{a,\tau}) B_{3,2}(f_{a,\tau}, \chi^\varepsilon_j, f, \tau f, \tau f)\Big[\frac{\pi^\varepsilon_j f' h''}{(1+\tau^2 f'^2)^{3/2}}\Big]&\\
+2\tau^2 (f'_{a,\tau}(x^\varepsilon_j) - f'_{a,\tau}) B_{4,2}(f_{a,\tau}, \chi^\varepsilon_j, f, f, \tau f, \tau f)\Big[\frac{\pi^\varepsilon_j h''}{(1+\tau^2 f'^2)^{3/2}}\Big]&
\end{aligned}$$

$$
\begin{aligned}
&+3\tau^2 f'_{a,\tau}\left\{\pi_j^\varepsilon B_{1,1}(f_{a,\tau},\tau f)\left[\frac{f'f''h''}{(1+\tau^2 f'^2)^{5/2}}\right]-B_{1,1}(f_{a,\tau},\tau f)\left[\frac{\pi_j^\varepsilon f'f''h''}{(1+\tau^2 f'^2)^{5/2}}\right]\right\}\\
&+f'_{a,\tau}\left\{B_{1,1}(\pi_j^\varepsilon,\tau f)\left[\frac{f'_{a,\tau}h''}{(1+\tau^2 f'^2)^{3/2}}\right]+B_{1,1}(f_{a,\tau},\tau f)\left[\frac{\pi_j^{\varepsilon\prime}h''}{(1+\tau^2 f'^2)^{3/2}}\right]\right\}\\
&-2f'_{a,\tau}B_{2,1}(\pi_j^\varepsilon,f_{a,\tau},\tau f)\left[\frac{h''}{(1+\tau^2 f'^2)^{3/2}}\right]-2\tau^2 f'_{a,\tau}B_{3,2}(\pi_j^\varepsilon,f_{a,\tau},f,\tau f,\tau f)\left[\frac{f'h''}{(1+\tau^2 f'^2)^{3/2}}\right]\\
&+2\tau^2 f'_{a,\tau}B_{4,2}(\pi_j^\varepsilon,f_{a,\tau},f,f,\tau f,\tau f)\left[\frac{h''}{(1+\tau^2 f'^2)^{3/2}}\right].
\end{aligned}
$$

Using Remark 3.3, the interpolation property (1-4), Young's inequality, and the Hölder continuity of $f'_{a,\tau}$, it follows that

$$
\begin{aligned}
\|T_1[h]\|_2 &\le C[\|\chi_j^\varepsilon(f'_{a,\tau}-f'_{a,\tau}(x_j^\varepsilon))\|_\infty\|\pi_j^\varepsilon h'''\|_2+\|\pi_j^\varepsilon h''\|_\infty]+K\|h\|_{H^2}\\
&\le \tfrac{1}{3}\mu\|\pi_j^\varepsilon h\|_{H^3}+K\|h\|_{H^2}
\end{aligned}
\tag{6-14}
$$

provided that ε is sufficiently small.

Furthermore, we have

$$
T_2[h]=\frac{\tau^2 f'_{a,\tau}(x_j^\varepsilon)}{(1+\tau^2 f'^2(x_j^\varepsilon))^{3/2}}B_{1,1}(f_{a,\tau},\tau f)[Q(f'(x_j^\varepsilon)-f')\pi_j^\varepsilon h'''],
$$

where

$$
Q:=\frac{(f'(x_j^\varepsilon)+f')[(1+\tau^2 f'^2)^2+(1+\tau^2 f'^2)(1+\tau^2 f'^2(x_j^\varepsilon))+(1+\tau^2 f'^2(x_j^\varepsilon))^2]}{(1+\tau^2 f'^2)^{3/2}[(1+\tau^2 f'^2)^{3/2}+(1+\tau^2 f'^2(x_j^\varepsilon))^{3/2}]},
$$

and therewith

$$
\|T_2[h]\|_2\le C\|\chi_j^\varepsilon(f'_{a,\tau}-f'_{a,\tau}(x_j^\varepsilon))\|_\infty\|\pi_j^\varepsilon h'''\|_2\le \tfrac{1}{3}\mu\|\pi_j^\varepsilon h\|_{H^3}+K\|h\|_{H^2}
\tag{6-15}
$$

if ε is sufficiently small.

Finally, arguing as in Step 3 of the proof of Theorem 4.4, we deduce that for ε sufficiently small we have

$$
\|T_3[h]\|_2\le \tfrac{1}{3}\mu\|\pi_j^\varepsilon h\|_{H^3}+K\|h\|_{H^2}.
\tag{6-16}
$$

Gathering (6-13)–(6-16), we have established the desired estimate (6-10) for $|j|\le N-1$.

The case $j=N$. For $j=N$ we write

$$
\pi_N^\varepsilon K_a(\tau f)[h]-\mathbb{A}_{N,\tau}^a[\pi_N^\varepsilon h]=:S_1[h]+S_2[h]+S_3[h]+S_4[h],
\tag{6-17}
$$

where

$$
\begin{aligned}
S_1[h]&:=\pi_N^\varepsilon K_a(\tau f)[h]-aB_{1,1}(f_{a,\tau},\tau f)\left[\frac{\pi_N^\varepsilon h'''}{(1+\tau^2 f'^2)^{3/2}}\right],\\
S_2[h]&:=aB_{1,1}(f_{a,\tau},\tau f)\left[\frac{\pi_N^\varepsilon h'''}{(1+\tau^2 f'^2)^{3/2}}\right]-aB_{1,1}(f_{a,\tau},\tau f)[\pi_N^\varepsilon h'''],
\end{aligned}
$$

$$S_3[h] := aB_{1,1}(f_{a,\tau}, \tau f)[\pi_N^\varepsilon h'''] - a^2 B_{0,1}(0)[\pi_N^\varepsilon h'''],$$

$$S_4[h] := a^2 B_{0,1}(0)[\pi_N^\varepsilon h'''] - \mathbb{A}_{N,\tau}^a[\pi_N^\varepsilon h].$$

Much as for $T_1[h]$, we derive the identity

$$\begin{aligned}
S_1[h] = {} & \tau(1-a)\chi_N^\varepsilon f' B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi_N^\varepsilon h'''}{(1+\tau^2 f'^2)^{3/2}}\Big] \\
& -\tau(1-a) f'(1-\chi_N^\varepsilon) B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi_N^{\varepsilon\prime} h''}{(1+\tau^2 f'^2)^{3/2}} - 3\tau^2 \frac{\pi_N^\varepsilon f' f'' h''}{(1+\tau^2 f'^2)^{5/2}}\Big] \\
& -\tau(1-a) f'\Big\{B_{1,1}(\chi_N^\varepsilon, \tau f)\Big[\frac{\pi_N^\varepsilon f'_{a,\tau} h''}{(1+\tau^2 f'^2)^{3/2}}\Big] - 2B_{2,1}(f_{a,\tau}, \chi_N^\varepsilon, \tau f)\Big[\frac{\pi_N^\varepsilon h''}{(1+\tau^2 f'^2)^{3/2}}\Big]\Big\} \\
& +2\tau^3(1-a) f' B_{3,2}(f_{a,\tau}, \chi_N^\varepsilon, f, \tau f, \tau f)\Big[\frac{\pi_N^\varepsilon f' h''}{(1+\tau^2 f'^2)^{3/2}}\Big] \\
& -2\tau^3(1-a) f' B_{4,2}(f_{a,\tau}, \chi_N^\varepsilon, f, f, \tau f, \tau f)\Big[\frac{\pi_N^\varepsilon h''}{(1+\tau^2 f'^2)^{3/2}}\Big] \\
& +3\tau^2 f'_{a,\tau}\Big\{\pi_N^\varepsilon B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{f' f'' h''}{(1+\tau^2 f'^2)^{5/2}}\Big] - B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi_N^\varepsilon f' f'' h''}{(1+\tau^2 f'^2)^{5/2}}\Big]\Big\} \\
& + f'_{a,\tau}\Big\{B_{1,1}(\pi_N^\varepsilon, \tau f)\Big[\frac{f'_{a,\tau} h''}{(1+\tau^2 f'^2)^{3/2}}\Big] + B_{1,1}(f_{a,\tau}, \tau f)\Big[\frac{\pi_N^{\varepsilon\prime} h''}{(1+\tau^2 f'^2)^{3/2}}\Big]\Big\} \\
& -2 f'_{a,\tau} B_{2,1}(\pi_N^\varepsilon, f_{a,\tau}, \tau f)\Big[\frac{h''}{(1+\tau^2 f'^2)^{3/2}}\Big] - 2\tau^2 f'_{a,\tau} B_{3,2}(\pi_N^\varepsilon, f_{a,\tau}, f, \tau f, \tau f)\Big[\frac{f' h''}{(1+\tau^2 f'^2)^{3/2}}\Big] \\
& +2\tau^2 f'_{a,\tau} B_{4,2}(\pi_N^\varepsilon, f_{a,\tau}, f, f, \tau f, \tau f)\Big[\frac{h''}{(1+\tau^2 f'^2)^{3/2}}\Big].
\end{aligned}$$

Recalling that f' vanishes at infinity, we obtain by virtue of Remark 3.3, the interpolation property (1-4), and Young's inequality that

$$\|S_1[h]\|_2 \le \tfrac{1}{3}\mu\|\pi_j^\varepsilon h\|_{H^3} + K\|h\|_{H^2} \tag{6-18}$$

provided that ε is sufficiently small. Furthermore, Remark 3.3 implies that for ε sufficiently small

$$\begin{aligned}
\|S_2[h]\|_2 &= a\Big\|B_{11}(f_{a,\tau}, \tau f)\Big[\frac{\pi_N^\varepsilon h'''}{(1+\tau^2 f'^2)^{3/2}}[1-(1+\tau^2 f'^2)^{3/2}]\Big]\Big\|_2 \\
&\le C\|\pi_N^\varepsilon h'''\|_2 \|\chi_N^\varepsilon[1-(1+\tau^2 f'^2)^{3/2}]\|_\infty \le \tfrac{1}{3}\mu\|\pi_j^\varepsilon h\|_{H^3} + K\|h\|_{H^2}.
\end{aligned} \tag{6-19}$$

Since $a(1-a)=0$, we compute that

$$S_3[h] = -a^2 B_{2,1}(f, f, \tau f)[\pi_\varepsilon^N h'''],$$

and the arguments presented in Step 4 of the proof of Theorem 4.2 yield

$$\|S_3[h]\|_2 \le \tfrac{1}{3}\mu\|\pi_j^\varepsilon h\|_{H^3} + K\|h\|_{H^2} \tag{6-20}$$

for ε sufficiently small. Finally,

$$\|S_4[h]\|_2 = a^2\|B_{0,1}(0)[3(\pi_N^\varepsilon)'h'' + 3(\pi_N^\varepsilon)''h' + (\pi_N^\varepsilon)'''h]\|_2 \le K\|h\|_{H^2}, \tag{6-21}$$

and combining (6-17)–(6-21) we obtain the estimate (6-10) for $j = N$. $\square$

The Fourier multipliers defined by (6-11)–(6-12) are generators of strongly continuous analytic semigroups in $\mathcal{L}(L_2(\mathbb{R}))$ and they satisfy resolvent estimates which are uniform with respect to $x_j^\varepsilon \in \mathbb{R}$ and $\tau \in [0, 1]$. More precisely, we have the following result.

Proposition 6.2. *Let $f \in H^2(\mathbb{R})$ be fixed. Given $x_0 \in \mathbb{R}$ and $\tau \in [0, 1]$, let*

$$\mathbb{A}_{x_0,\tau} := -\frac{\pi}{(1+\tau^2 f'^2(x_0))^{3/2}}(\partial_x^4)^{3/4}.$$

Then, there exists a constant $\kappa_0 \ge 1$ such that

$$\lambda - \mathbb{A}_{x_0,\tau} \in \mathrm{Isom}(H^3(\mathbb{R}), L_2(\mathbb{R})), \tag{6-22}$$

$$\kappa_0\|(\lambda - \mathbb{A}_{x_0,\tau})[h]\|_2 \ge |\lambda| \cdot \|h\|_2 + \|h\|_{H^3} \tag{6-23}$$

for all $x_0 \in \mathbb{R}$, $\tau \in [0, 1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge 1$, and $h \in H^3(\mathbb{R})$.

Proof. The proof is similar to that of Proposition 4.3 and therefore we omit it. $\square$

We now conclude with the following general result.

Theorem 6.3. *Let $f \in H^2(\mathbb{R})$ be given. Then*

$$-\Phi_\sigma(f) \in \mathcal{H}(H^3(\mathbb{R}), L_2(\mathbb{R})).$$

Proof. As mentioned in the discussion preceding Theorem 6.1, we only need to prove the claim for the leading-order term $\Phi_{\sigma,1}(f)$. Let $\kappa_0 \ge 1$ be the constant determined in Proposition 6.2 and let $\mu := \frac{1}{2}\kappa_0$. By virtue of Theorem 6.1 there exist constants $\varepsilon \in (0, 1)$ and $K = K(\varepsilon) > 0$, an ε-localization family $\{\pi_j^\varepsilon : -N+1 \le j \le N\}$ that satisfies (4-1)–(4-5), and for each $-N+1 \le j \le N$ and $\tau \in [0, 1]$ operators $\mathbb{A}_{j,\tau} \in \mathcal{L}(H^3(\mathbb{R}), L_2(\mathbb{R}))$ such that

$$\|\pi_j^\varepsilon\Phi_{\sigma,1}(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_2 \le \frac{1}{2\kappa_0}\|\pi_j^\varepsilon h\|_{H^3} + K\|h\|_{H^2} \tag{6-24}$$

for all $-N+1 \le j \le N$, $\tau \in [0, 1]$, and $h \in H^3(\mathbb{R})$. Furthermore, Proposition 6.2 implies

$$\kappa_0\|(\lambda - \mathbb{A}_{j,\tau})[\pi_j^\varepsilon h]\|_2 \ge |\lambda| \cdot \|\pi_j^\varepsilon h\|_2 + \|\pi_j^\varepsilon h\|_{H^3} \tag{6-25}$$

for all $-N+1 \le j \le N$, $\tau \in [0, 1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge 1$, and $h \in H^3(\mathbb{R})$. Combining (6-24)–(6-25), we find

$$\begin{aligned}\kappa_0\|\pi_j^\varepsilon(\lambda - \Phi_{\sigma,1}(\tau f))[h]\|_2 &\ge \kappa_0\|(\lambda - \mathbb{A}_{j,\tau})[\pi_j^\varepsilon h]\|_2 - \kappa_0\|\pi_j^\varepsilon\Phi_{\sigma,1}(\tau f)[h] - \mathbb{A}_{j,\tau}[\pi_j^\varepsilon h]\|_2\\ &\ge |\lambda| \cdot \|\pi_j^\varepsilon h\|_2 + \tfrac{1}{2}\|\pi_j^\varepsilon h\|_{H^3} - \kappa_0 K\|h\|_{H^2}\end{aligned}$$

for all $-N+1 \le j \le N$, $\tau \in [0,1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge 1$, and $h \in H^3(\mathbb{R})$. Summing up over $j \in \{-N+1, \dots, N\}$, we infer from Lemma 4.1 that there exists a constant $C \ge 1$ with the property that

$$C\|h\|_{H^2} + C\|(\lambda - \Phi(\tau f))[h]\|_2 \ge |\lambda| \cdot \|h\|_2 + \|h\|_{H^3}$$

for all $\tau \in [0,1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge 1$, and $h \in H^3(\mathbb{R})$. Using (1-4) and Young's inequality, it follows that there exist constants $\kappa \ge 1$ and $\omega > 0$ with the property that

$$\kappa\|(\lambda - \Phi_{\sigma,1}(\tau f))[h]\|_2 \ge |\lambda| \cdot \|h\|_2 + \|h\|_{H^3} \tag{6-26}$$

for all $\tau \in [0,1]$, $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda \ge \omega$, $h \in H^3(\mathbb{R})$. Since $(\omega - \Phi_{\sigma,1}(\tau f))|_{\tau=0} \in \operatorname{Isom}(H^3(\mathbb{R}), L_2(\mathbb{R}))$, the method of continuity together with (6-26) yields that

$$\omega - \Phi_{\sigma,1}(f) \in \operatorname{Isom}(H^3(\mathbb{R}), L_2(\mathbb{R})). \tag{6-27}$$

The claim follows from (6-26) (with $\tau = 1$), (6-27), and [Lunardi 1995, Proposition 2.4.1 and Corollary 2.1.3]. □

We are now in a position to prove the well-posedness for the Muskat problem with surface tension.

Proof of Theorem 1.2. Let $s \in (2,3)$, $\bar{s} = 2$, $1 > \alpha := \frac{1}{3}s > \beta := \frac{2}{3} > 0$. Combining (6-4) and Theorem 6.3, it follows that

$$-\Phi_\sigma \in C^\omega(H^2(\mathbb{R}), \mathcal{H}(H^3(\mathbb{R}), L_2(\mathbb{R}))).$$

Since

$$H^2(\mathbb{R}) = [L_2(\mathbb{R}), H^3(\mathbb{R})]_\beta \quad \text{and} \quad H^s(\mathbb{R}) = [L_2(\mathbb{R}), H^3(\mathbb{R})]_\alpha,$$

we now infer from Theorem 1.5 that (1-1), or equivalently (6-3), possesses a maximally defined solution

$$f := f(\cdot\,; f_0) \in C([0, T_+(f_0)), H^s(\mathbb{R})) \cap C((0, T_+(f_0)), H^3(\mathbb{R})) \cap C^1((0, T_+(f_0)), L_2(\mathbb{R}))$$

with

$$f \in C^{(s-2)/3}([0,T], H^2(\mathbb{R})) \quad \text{for all } T < T_+(f_0).$$

Concerning the uniqueness of solutions, we next show that any classical solution

$$\tilde{f} \in C([0,\widetilde{T}), H^s(\mathbb{R})) \cap C((0,\widetilde{T}), H^3(\mathbb{R})) \cap C^1((0,\widetilde{T})), L_2(\mathbb{R})), \quad \widetilde{T} \in (0,\infty],$$

to (6-3) satisfies

$$\tilde{f} \in C^\eta([0,T], H^2(\mathbb{R})) \quad \text{for all } T \in (0,\widetilde{T}), \tag{6-28}$$

where $\eta := (s-2)/(s+1)$. To this end, we recall that

$$\Phi_\sigma(f)[f] = f' B_{1,1}(f,f)[(\kappa(f))'] + B_{0,1}(f)[(\kappa(f))'] + \Theta A_{0,0}(f)[f'] \quad \text{for } f \in H^3(\mathbb{R}). \tag{6-29}$$

Let $T \in (0,\widetilde{T})$ be fixed. Lemma 3.1(i) implies that

$$\sup_{[0,T]} \|A_{0,0}(\tilde{f})[\tilde{f}']\|_2 \le C. \tag{6-30}$$

We now consider the highest-order terms in (6-29). Arguing as in Lemma 3.5, it follows from Remark 3.3 that $B_{0,1}(f)[\kappa(f)],\ B_{1,1}(f,f)[\kappa(f)]\in H^1(\mathbb{R})$ for all $f\in H^3(\mathbb{R})$, with

$$(B_{0,1}(f)[\kappa(f)])' = B_{0,1}(f)[(\kappa(f))'] - 2B_{2,2}(f',f,f,f)[\kappa(f)],$$
$$(B_{1,1}(f,f)[\kappa(f)])' = B_{1,1}(f,f)[(\kappa(f))'] + B_{1,1}(f',f)[\kappa(f)] - 2B_{3,2}(f',f,f,f,f)[\kappa(f)].$$

Furthermore, given $t\in(0,T]$ and $\varphi\in H^1(\mathbb{R})$, integration by parts together with $\tilde f\in C([0,T],H^s(\mathbb{R}))$ leads us to

$$\begin{aligned}&\left|\int_{\mathbb{R}} \tilde f'(t)(B_{1,1}(\tilde f(t),\tilde f(t))[\kappa(\tilde f(t))])'\varphi\,dx\right|\\ &\qquad=\left|\int_{\mathbb{R}} \tilde f''(t)B_{1,1}(\tilde f(t),\tilde f(t))[\kappa(\tilde f(t))]\varphi\,dx\right|+\left|\int_{\mathbb{R}} \tilde f'(t)B_{1,1}(\tilde f(t),\tilde f(t))[\kappa(\tilde f(t))]\varphi'\,dx\right|\le C\|\varphi\|_{H^1},\end{aligned}$$

so that

$$\sup_{(0,T]}\|\tilde f'(B_{1,1}(\tilde f,\tilde f)[\kappa(\tilde f)])'\|_{H^{-1}}\le C, \tag{6-31}$$

and similarly

$$\sup_{(0,T]}\|(B_{0,1}(\tilde f)[\kappa(\tilde f)])'\|_{H^{-1}}\le C. \tag{6-32}$$

We now estimate the term $f'B_{1,1}(f',f)[\kappa(f)]$ with $f\in H^3(\mathbb{R})$ in the $H^{-1}(\mathbb{R})$-norm. To this end, we rely on the formula

$$B_{1,1}(f',f)[\kappa(f)] = T_1(f) - T_2(f) - T_3(f),$$

where

$$\begin{aligned}T_1(f)(x) &:= \int_0^\infty \frac{\kappa(f)(x-y)-\kappa(f)(x+y)}{y}\,\frac{f'(x)-f'(x-y)}{y}\,\frac{1}{1+(\delta_{[x,y]}f/y)^2}\,dy,\\ T_2(f)(x) &:= \int_0^\infty \frac{\kappa(f)(x+y)}{y}\,\frac{f'(x+y)-2f'(x)+f'(x-y)}{y}\,\frac{1}{1+(\delta_{[x,y]}f/y)^2}\,dy,\\ T_3(f)(x) &:= \int_0^\infty \frac{\kappa(f)(x+y)}{y}\,\frac{f'(x)-f'(x+y)}{y}\,\frac{1}{[1+(\delta_{[x,y]}f/y)^2][1+(\delta_{[x,-y]}f/y)^2]}\\ &\qquad\qquad\times\frac{f(x+y)-f(x-y)}{y}\,\frac{f(x+y)-2f(x)+f(x-y)}{y}\,dy.\end{aligned}$$

We estimate the terms $\tilde f'T_i(\tilde f)$, $1\le i\le 3$, separately. Given $t\in(0,T]$ and $\varphi\in H^1(\mathbb{R})$, we compute

$$\begin{aligned}&\left|\int_{\mathbb{R}} \tilde f'(t)T_1(\tilde f(t))\varphi\,dx\right|\\ &\quad\le C\|\varphi\|_\infty\int_0^\infty\!\!\int_{\mathbb{R}} \frac{|\kappa(\tilde f(t))(x-y)-\kappa(\tilde f(t))(x+y)|}{y}\,\frac{|\tilde f'(t,x)-\tilde f'(t,x-y)|}{y}\,dx\,dy\\ &\quad\le C\|\varphi\|_\infty\int_0^\infty \frac{1}{y^2}\Big(\int_{\mathbb{R}}|\kappa(\tilde f(t))(x-y)-\kappa(\tilde f(t))(x+y)|^2\,dx\Big)^{1/2}\Big(\int_{\mathbb{R}}|\tilde f'(t,x)-\tilde f'(t,x-y)|^2\,dx\Big)^{1/2}dy\\ &\quad= C\|\varphi\|_\infty\int_0^\infty \frac{1}{y^2}\Big(\int_{\mathbb{R}}|\mathcal{F}(\kappa(\tilde f(t)))|^2(\xi)|e^{i2\xi y}-1|^2\,d\xi\Big)^{1/2}\Big(\int_{\mathbb{R}}|\mathcal{F}(\tilde f'(t))|^2(\xi)|e^{iy\xi}-1|^2\,d\xi\Big)^{1/2}dy,\end{aligned}$$

and since

$$\begin{aligned}|e^{iy\xi}-1|^2 &\le C(1+|\xi|^2)[y^2\mathbf{1}_{(0,1)}(y)+\mathbf{1}_{[y\ge 1]}(y)],\\ |e^{i2y\xi}-1|^2 &\le C(1+|\xi|^2)^{s-2}[y^{2(s-2)}\mathbf{1}_{(0,1)}(y)+\mathbf{1}_{[y\ge 1]}(y)],\end{aligned}\qquad y>0,\ \ \xi\in\mathbb{R},$$

it follows that

$$\begin{aligned}\left|\int_{\mathbb{R}}\tilde f'(t)T_1(\tilde f(t))\varphi\,dx\right| &\le C\|\varphi\|_\infty\|\kappa(\tilde f(t))|_{H^{s-2}}\|\tilde f(t)\|_{H^1}\int_0^\infty y^{s-3}\mathbf{1}_{(0,1)}(y)+y^{-2}\mathbf{1}_{[y\ge 1]}(y)\,dy\\ &\le C\|\varphi\|_{H^1}.\end{aligned}\tag{6-33}$$

To bound the curvature term in the $H^{s-2}(\mathbb{R})$-norm we have use the inequality

$$\|\kappa(f)\|_{H^{s-2}}\le C\|(1+f'^2)^{-3/2}\|_{BC^{s-3/2}}\|f\|_{H^s}\quad\text{for all } f\in H^s(\mathbb{R}).$$

Similarly we have

$$\begin{aligned}\left|\int_{\mathbb{R}}\tilde f'(t)T_2(\tilde f(t))\varphi\,dx\right| &\le C\|\varphi\|_\infty\int_0^\infty\frac{1}{y^2}\left(\int_{\mathbb{R}}|\kappa(\tilde f(t))(x+y)|^2\,dx\right)^{1/2}\\ &\qquad\times\left(\int_{\mathbb{R}}|\tilde f'(t,x+y)-2\tilde f'(t,x)+\tilde f'(t,x-y)|^2\,dx\right)^{1/2}dy\\ &\le C\|\varphi\|_\infty\int_0^\infty\frac{1}{y^2}\left(\int_{\mathbb{R}}|\mathcal{F}(\tilde f'(t))|^2(\xi)|e^{iy\xi}-2+e^{-iy\xi}|^2\,d\xi\right)^{1/2}dy\\ &\le C\|\varphi\|_\infty\int_0^\infty y^{s-3}\mathbf{1}_{(0,1)}(y)+y^{-2}\mathbf{1}_{[y\ge 1]}(y)\,dy\\ &\le C\|\varphi\|_{H^1}\end{aligned}\tag{6-34}$$

by virtue of

$$|e^{iy\xi}-2+e^{-iy\xi}|^2\le C(1+|\xi|^2)^{s-1}[y^{2(s-1)}\mathbf{1}_{(0,1)}(y)+\mathbf{1}_{[y\ge 1]}(y)],\qquad y>0,\ \ \xi\in\mathbb{R}.$$

Lastly, since $H^{s-1}(\mathbb{R})\hookrightarrow \mathrm{BC}^{s-3/2}(\mathbb{R})$ for $s\neq\frac{5}{2}$ $\big($the estimate (6-35) holds though also for $s=\frac{5}{2}\big)$ and $H^s(\mathbb{R})\hookrightarrow \mathrm{BC}^1(\mathbb{R})$, the inequality

$$|e^{iy\xi}-2+e^{-iy\xi}|^2\le C(1+|\xi|^2)^2[y^4\mathbf{1}_{(0,1)}(y)+\mathbf{1}_{[y\ge 1]}(y)],\qquad y>0,\ \ \xi\in\mathbb{R},$$

leads us to

$$\begin{aligned}\left|\int_{\mathbb{R}}\tilde f'(t)T_3(\tilde f(t))\varphi\,dx\right| &\le C\|\varphi\|_\infty\int_0^\infty\frac{y^{\min\{1,s-3/2\}}}{y^3}\left(\int_{\mathbb{R}}|\mathcal{F}(\tilde f(t))|^2(\xi)|e^{iy\xi}-2+e^{-iy\xi}|^2\,d\xi\right)^{1/2}dy\\ &\le C\|\varphi\|_\infty\int_0^\infty y^{\min\{0,s-5/2\}}\mathbf{1}_{(0,1)}(y)+y^{-2}\mathbf{1}_{[y\ge 1]}(y)\,dy\\ &\le C\|\varphi\|_{H^1}.\end{aligned}\tag{6-35}$$

Gathering (6-33)–(6-35), we conclude that

$$\sup_{(0,T]}\|\tilde f'B_{1,1}(\tilde f',\tilde f)[\kappa(\tilde f)]\|_{H^{-1}}\le C,\tag{6-36}$$

and similarly we obtain

$$\sup_{(0,T]}\big[\|\tilde f' B_{3,2}(\tilde f', \tilde f, \tilde f, \tilde f, \tilde f)[\kappa(\tilde f)]\|_{H^{-1}} + \|B_{2,2}(\tilde f', \tilde f, \tilde f, \tilde f)[\kappa(\tilde f)]\|_{H^{-1}}\big] \le C. \tag{6-37}$$

Combining (6-30)–(6-32), (6-36), and (6-37), it follows that $\tilde f \in \mathrm{BC}^1((0,T], H^{-1}(\mathbb{R}))$. Recalling that $\eta = (s-2)/(s+1)$, (1-4) together with the mean value theorem yields

$$\|\tilde f(t) - \tilde f(s)\|_{H^2} \le \|\tilde f(t) - \tilde f(s)\|^{\eta}_{H^{-1}} \|\tilde f(t) - \tilde f(s)\|^{1-\eta}_{H^s} \le C|t-s|^{\eta}, \quad t, s \in [0, T],$$

which proves (6-28) and the uniqueness claim in Theorem 1.2.

Finally, let us assume that $T_+(f_0) < \infty$ and that

$$\sup_{[0,T_+(f_0))} \|f(t)\|_{H^s} < \infty.$$

Arguing as above, we find that

$$\|f(t) - f(s)\|_{H^{(s+2)/2}} \le C|t-s|^{(s-2)/(2s+2)}, \quad t, s \in [0, T_+(f_0)).$$

The criterion for global existence in Theorem 1.5 applied for $\alpha := \frac{1}{6}(s+2)$ and $\beta := \frac{2}{3}$ implies that the solution can be continued on an interval $[0,\tau)$ with $\tau > T_+(f_0)$ and that

$$f \in C^{(s-2)/6}([0,T], H^2(\mathbb{R})) \quad \text{for all } T \in (0,\tau).$$

The uniqueness claim in Theorem 1.5 leads us to a contradiction. Hence our assumption was false and $T_+(f_0) = \infty$. □

Appendix A: Some technical results

The following lemma is used in the proof of Theorem 4.2.

Lemma A.1. *Given $f \in H^s(\mathbb{R})$, $s \in \big(\frac{3}{2}, 2\big)$, and $\tau \in [0,1]$, let $a_\tau : \mathbb{R} \to \mathbb{R}$ be defined by*

$$a_\tau(x) := \mathrm{PV}\int_{\mathbb{R}} \frac{y}{y^2 + \tau^2(\delta_{[x,y]} f)^2}\, dy, \quad x \in \mathbb{R}.$$

Let further $\alpha := \frac{1}{2}s - \frac{3}{4} \in (0,1)$. Then, $a_\tau \in \mathrm{BC}^\alpha(\mathbb{R}) \cap C_0(\mathbb{R})$,

$$\sup_{\tau\in[0,1]} \|a_\tau\|_{\mathrm{BC}^\alpha} < \infty, \tag{A-1}$$

and, given $\varepsilon_0 > 0$, there exists $\eta > 0$ such that

$$\sup_{\tau\in[0,1]} \sup_{|x|\ge\eta} |a_\tau(x)| \le \varepsilon_0. \tag{A-2}$$

Proof. It holds that

$$a_\tau(x) = \tau^2 \lim_{\delta\to 0} \int_\delta^{1/\delta} \frac{f(x+y) - 2f(x) + f(x-y)}{y^2} \frac{f(x+y) - f(x-y)}{y} \frac{y^4}{\Pi(x,y)}\, dy, \quad x \in \mathbb{R},$$

with

$$\Pi(x,y) := [y^2+\tau^2(\delta_{[x,-y]}f)^2][y^2+\tau^2(\delta_{[x,y]}f)^2].$$

Letting

$$I(x,y) := \tau^2 \frac{f(x+y)-2f(x)+f(x-y)}{y^2}\frac{f(x+y)-f(x-y)}{y}\frac{y^4}{\Pi(x,y)}, \quad (x,y)\in\mathbb{R}\times(0,\infty),$$

it follows that

$$|I(x,y)| \le 8\biggl(\|f\|_\infty^2 \frac{1}{y^3}\mathbf{1}_{[1,\infty)}(y) + \|f'\|_\infty [f']_{s-3/2}\frac{1}{y^{5/2-s}}\mathbf{1}_{(0,1)}(y)\biggr), \quad (x,y)\in\mathbb{R}\times(0,\infty). \tag{A-3}$$

The latter estimate was obtained by using the fact that $f\in \mathrm{BC}^{s-1/2}(\mathbb{R})$, $s-\frac{1}{2}\in(1,2)$, together with (3-7). Hence,

$$a_\tau(x) = \int_0^\infty I(x,y)\,dy, \quad x\in\mathbb{R},$$

and $\sup_{\tau\in[0,1]}\|a_\tau\|_\infty<\infty$. To estimate the Hölder seminorm of a_τ, we compute for $x,x'\in\mathbb{R}$ that

$$|a_\tau(x)-a_\tau(x')| \le \int_0^\infty |I(x,y)-I(x',y)|\,dy \le T_1+T_2+T_3, \tag{A-4}$$

where

$$\begin{aligned}
T_1 &:= \int_0^\infty \frac{|f(x+y)-2f(x)+f(x-y)|}{y^2}\frac{|[f(x+y)-f(x-y)]-[f(x'+y)-f(x'-y)]|}{y}\frac{y^4}{\Pi(x,y)}\,dy,\\
T_2 &:= \int_0^\infty \frac{|[f(x+y)-2f(x)+f(x-y)]-[f(x'+y)-2f(x')+f(x'-y)]|}{y^2}\\
&\qquad\qquad \times\frac{|f(x'+y)-f(x'-y)|}{y}\frac{y^4}{\Pi(x,y)}\,dy,\\
T_3 &:= \int_0^\infty \frac{|f(x'+y)-2f(x')+f(x'-y)|}{y^2}\frac{|f(x'+y)-f(x'-y)|}{y}\frac{|\Pi(x,y)-\Pi(x',y)|}{y^4}\,dy.
\end{aligned}$$

Using the mean value theorem, we have

$$\begin{aligned}
\frac{|[f(x+y)-f(x-y)]-[f(x'+y)-f(x'-y)]|}{y} &\le 2\int_0^1 |f'(x+(2\tau-1)y)-f'(x'+(2\tau-1)y)|\,d\tau\\
&\le 2[f']_{s-3/2}|x-x'|^{s-3/2}, \quad y>0,
\end{aligned}$$

and, much as above, we find that

$$|T_1| \le C\|f\|_{H^s}^2|x-x'|^{2\alpha}. \tag{A-5}$$

To deal with the second term we appeal to the formula

$$\begin{aligned}
f(x+y)-2f(x)+f(x-y) = y[f'(x+y)-f'(x-y)] &+ y\int_0^1 f'(x+\tau y)-f'(x+y)\,d\tau\\
&- y\int_0^1 f'(x-\tau y)-f'(x-y)\,d\tau \quad \text{for } x,y\in\mathbb{R},
\end{aligned}$$

and we get

$$\frac{|[f(x+y)-2f(x)+f(x-y)]-[f(x'+y)-2f(x')+f(x'-y)]|}{y^2} \le T_{2a}+T_{2b}+T_{2c},$$

where

$$\begin{aligned}
T_{2a} &:= \frac{|[f'(x+y)-f'(x-y)]-[f'(x'+y)-f'(x'-y)]|}{y} \\
&\le 2[f']_{2\alpha}\Big(\frac{1}{y}\mathbf{1}_{[1,\infty)}(y)|x-x'|^{2\alpha}+2\frac{1}{y^{1-\alpha}}\mathbf{1}_{(0,1)}(y)|x-x'|^{\alpha}\Big), \\
T_{2b} &:= \frac{1}{y}\int_0^1 \big|[f'(x+\tau y)-f'(x+y)]-[f'(x'+\tau y)-f'(x'+y)]\big|\,d\tau \\
&\le 2[f']_{2\alpha}\Big(\frac{1}{y}\mathbf{1}_{[1,\infty)}(y)|x-x'|^{2\alpha}+\frac{1}{y^{1-\alpha}}\mathbf{1}_{(0,1)}(y)|x-x'|^{\alpha}\Big), \\
T_{2c} &:= \frac{1}{y}\int_0^1 \big|[f'(x-\tau y)-f'(x-y)]-[f'(x'-\tau y)-f'(x'-y)]\big|\,d\tau \\
&\le 2[f']_{2\alpha}\Big(\frac{1}{y}\mathbf{1}_{[1,\infty)}(y)|x-x'|^{2\alpha}+\frac{1}{y^{1-\alpha}}\mathbf{1}_{(0,1)}(y)|x-x'|^{\alpha}\Big),
\end{aligned}$$

and therewith

$$|T_2| \le C\|f\|_{H^s}^2(|x-x'|^{\alpha}+|x-x'|^{2\alpha}). \tag{A-6}$$

Finally, since

$$\frac{|\Pi(x,y)-\Pi(x',y)|}{y^4} \le 4\|f'\|_\infty(1+\|f'\|_\infty^2)[f']_{2\alpha}|x-x'|^{2\alpha},$$

we infer from (A-3) that

$$|T_3| \le C\|f\|_{H^s}^4(1+\|f\|_{H^s}^2)(|x-x'|^{\alpha}+|x-x'|^{2\alpha}). \tag{A-7}$$

The relation (A-1) is a simple consequence of (A-4)–(A-7) and of $\sup_{\tau\in[0,1]}\|a_\tau\|_\infty<\infty$.

To prove that a_τ vanishes at infinity, let $\varepsilon_0>0$ be arbitrary. We write

$$a_\tau(x)=\int_0^M I(x,y)\,dy+\int_M^\infty I(x,y)\,dy, \quad x\in\mathbb{R},$$

for some $M>1$ with

$$\frac{4\|f\|_\infty^2}{M^2}\le\frac{\varepsilon_0}{2}.$$

Recalling (A-3), it follows that for all $x\in\mathbb{R}$ we have

$$\int_M^\infty |I(x,y)|\,dy \le 8\|f\|_\infty^2\int_M^\infty \frac{1}{y^3}\,dy=\frac{4\|f\|_\infty^2}{M^2}\le\frac{\varepsilon_0}{2}.$$

Let $\beta\in(0,1)$ be chosen such that $\frac{3}{2}+\beta<s$. Since $f\in C_0(\mathbb{R})$, there exists $\eta>M$ with

$$|f(y)| \le \left[\frac{\big(s-\frac{3}{2}-\beta\big)\varepsilon_0 M^{3/2+\beta-s}}{32([f']_{s-3/2}\|f'\|_\infty^{1-\beta}+1)}\right]^{1/\beta} \quad \text{for all } |y|\ge\eta-M.$$

Using this inequality, it follows that for all $|x| \geq \eta$ we have

$$|a_\tau(x)| \leq \int_0^M |I(x,y)|\,dy + \tfrac{1}{2}\varepsilon_0 \leq \tfrac{1}{2}\varepsilon_0 + 8[f']_{s-3/2}\|f'\|_\infty^{1-\beta} \int_0^M \frac{|f(x+y)-f(x-y)|^\beta}{y^{5/2+\beta-s}}\,dy \leq \varepsilon_0;$$

hence $a_\tau \in C_0(\mathbb{R})$ and (A-2) holds true. □

The next result is used in Proposition 2.1.

Lemma A.2. *Given $f \in H^5(\mathbb{R})$ and $\bar\omega \in H^2(\mathbb{R})$, set*

$$\tilde v(x,y) := \frac{1}{2\pi}\int_{\mathbb{R}} \frac{(-(y-f(s)),\, x-s)}{(x-s)^2+(y-f(s))^2}\,\bar\omega(s)\,ds \quad \text{in } \mathbb{R}^2 \setminus [y=f(x)]. \tag{A-8}$$

Let further $\Omega_- := [y<f(x)]$, $\Omega_+ := [y>f(x)]$, and $\tilde v_\pm := \tilde v|_{\Omega_\pm}$. Then, $\tilde v_\pm \in C(\overline{\Omega}_\pm) \cap C^1(\Omega_\pm)$ and

$$\tilde v_\pm(x,y) \to 0 \quad \text{for} \quad |(x,y)| \to \infty. \tag{A-9}$$

Proof. It is easy to see that $\tilde v_\pm \in C^1(\Omega_\pm)$. Plemelj's formula further shows that $\tilde v_\pm \in C(\overline{\Omega}_\pm)$ and

$$\tilde v_\pm(x, f(x)) = \frac{1}{2\pi}\,\mathrm{PV}\!\int_{\mathbb{R}} \frac{(-(f(x)-f(x-s)),\, s)}{s^2+(f(x)-f(x-s))^2}\,\bar\omega(x-s)\,ds \mp \frac{1}{2}\frac{(1, f'(x))\bar\omega(x)}{1+f'^2(x)}, \quad x \in \mathbb{R},$$

or equivalently, with the notation in Remark 3.3,

$$\tilde v_\pm|_{[y=f(x)]} = \pm\frac{1}{2}\frac{(1,f')\bar\omega}{1+f'^2} - \frac{1}{2\pi}B_{1,1}(f,f)[\bar\omega] + \frac{i}{2\pi}B_{0,1}(f)[\bar\omega].$$

Recalling that $f \in H^5(\mathbb{R})$ and $\bar\omega \in H^2(\mathbb{R})$, the arguments in the proof of Lemma 3.5 show that $B_{1,1}(f,f)[\bar\omega]$ and $B_{0,1}(f)[\bar\omega]$ belong to $H^1(\mathbb{R})$; thus

$$\tilde v_\pm(x, f(x)) \to 0 \quad \text{for } |x| \to \infty. \tag{A-10}$$

Furthermore, since f and $\bar\omega$ vanish at infinity, we find, much as in the proof of (A-2), that

$$\sup_{[y\geq n]} |\tilde v_+| + \sup_{[y\leq -n]} |\tilde v_-| \to 0 \quad \text{for } n \to \infty \tag{A-11}$$

and that, for arbitrary $0<a<b$,

$$\sup_{[a\leq y\leq b]\cap[|x|\geq n]} |\tilde v_+| + \sup_{[-b\leq y\leq -a]\cap[|x|\geq n]} |\tilde v_-| \to 0 \quad \text{for } n \to \infty. \tag{A-12}$$

Finally, arguing along the lines of the proof of Privalov's theorem, see [Lu 1993, Theorem 3.1.1] (the lengthy details, which for $\bar\omega \in W^1_\infty(\mathbb{R})$ are simpler than in that book, are left to the interested reader), it follows that there exists a constant C, which depends only on f and $\bar\omega$, such that for each $z=(x,y) \in \mathbb{R}^2 \setminus [y=f(x)]$ with $y \in [-\|f\|_\infty - 1, \|f\|_\infty + 1]$ the following inequalities hold:

$$\begin{aligned} |\tilde v_+(z) - \tilde v_+(x,f(x))| &\leq C|y-f(x)|^{1/2} \quad \text{if } y > f(x),\\ |\tilde v_-(z) - \tilde v_-(x,f(x))| &\leq C|y-f(x)|^{1/2} \quad \text{if } y < f(x). \end{aligned} \tag{A-13}$$

The relation (A-9) is an obvious consequence of (A-10)–(A-13). □

Appendix B: The proof of Theorem 1.5

This section is dedicated to the proof of Theorem 1.5. In the following $\mathbb{E}_0$ and $\mathbb{E}_1$ denote complex Banach spaces[4] and we assume that the embedding $\mathbb{E}_1 \hookrightarrow \mathbb{E}_0$ is dense. In view of [Amann 1995, Theorem I.1.2.2], we may represent the set $\mathcal{H}(\mathbb{E}_1, \mathbb{E}_0)$ of negative analytic generators as

$$\mathcal{H}(\mathbb{E}_1, \mathbb{E}_0) = \bigcup_{\substack{\kappa \geq 1 \\ \omega > 0}} \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa, \omega),$$

where, given $\kappa \geq 1$ and $\omega > 0$, the class $\mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa, \omega)$ consists of the operators $\mathbb{A} \in \mathcal{L}(\mathbb{E}_1, \mathbb{E}_0)$ having the properties

- $\omega + \mathbb{A} \in \mathrm{Isom}(\mathbb{E}_1, \mathbb{E}_0)$, and
- $\kappa^{-1} \leq \dfrac{\|(\lambda + \mathbb{A})x\|_0}{|\lambda| \cdot \|x\|_0 + \|x\|_1} \leq \kappa$ for all $0 \neq x \in \mathbb{E}_1$ and all $\operatorname{Re} \lambda \geq \omega$.

Given $\mathbb{A} \in \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa, \omega)$ and $r \in (0, \kappa^{-1})$, it follows from [Amann 1995, Theorem I.1.3.1(i)] that

$$\mathbb{A} + B \in \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa/(1-\kappa r), \omega) \quad \text{for all } \|B\|_{\mathcal{L}(\mathbb{E}_1, \mathbb{E}_0)} \leq r. \tag{B-1}$$

This property shows in particular that $\mathcal{H}(\mathbb{E}_1, \mathbb{E}_0)$ is an open subset of $\mathcal{L}(\mathbb{E}_1, \mathbb{E}_0)$.

The proof of Theorem 1.5 uses to a large extent the powerful theory of parabolic evolution operators developed in [Amann 1995]. The following result is a direct consequence of Theorem II.5.1.1, Lemma II.5.1.3 and Lemma II.5.1.4 in that paper.

Proposition B.1. *Let $T > 0$, $\rho \in (0, 1)$, $L \geq 0$, $\kappa \geq 1$, and $\omega > 0$ be given constants. Moreover, let $\mathcal{A} \subset C^\rho([0, T], \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0))$ be a family satisfying*

- $[\mathbb{A}]_{\rho,[0,T]} := \sup\limits_{t \neq s \in [0,T]} \dfrac{\|\mathbb{A}(t) - \mathbb{A}(s)\|}{|t-s|^\rho} \leq L$ *for all* $\mathbb{A} \in \mathcal{A}$, *and*
- $\mathbb{A}(t) \in \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa, \omega)$ *for all* $\mathbb{A} \in \mathcal{A}$ *and* $t \in [0, T]$.

Then, given $\mathbb{A} \in \mathcal{A}$, there exists a unique parabolic evolution operator[5] $U_{\mathbb{A}}$ for $\mathbb{A}$ possessing $\mathbb{E}_1$ as a regularity subspace. Moreover, the following hold:

(i) *There exists a constant $C > 0$ such that*

$$\|U_{\mathbb{A}}(t, s)\|_{\mathcal{L}(\mathbb{E}_j)} + (t-s)\|U_{\mathbb{A}}(t, s)\|_{\mathcal{L}(\mathbb{E}_0, \mathbb{E}_1)} \leq C \tag{B-2}$$

for all $(t, s) \in \Delta_T^* := \{(t, s) \in [0, T]^2 : 0 \leq s < t \leq T\}$, $j \in \{0, 1\}$, and all $\mathbb{A} \in \mathcal{A}$.

(ii) *Let $\Delta_T := \{(t, s) \in [0, T]^2 : 0 \leq s \leq t \leq T\}$ and $0 \leq \beta \leq \alpha \leq 1$. Then, given $x \in \mathbb{E}_\alpha$, it holds that $U_{\mathbb{A}}(\cdot, \cdot)x \in C(\Delta_T, \mathbb{E}_\alpha)$. Moreover, $U_{\mathbb{A}} \in C(\Delta_T^*, \mathcal{L}(\mathbb{E}_\beta, \mathbb{E}_\alpha))$, and there exists a constant $C > 0$ such that*

$$(t-s)^{\alpha-\beta}\|U_{\mathbb{A}}(t, s)\|_{\mathcal{L}(\mathbb{E}_\beta, \mathbb{E}_\alpha)} \leq C \tag{B-3}$$

for all $(t, s) \in \Delta_T^*$ and all $\mathbb{A} \in \mathcal{A}$.

[4] The proof of Theorem 1.5 in the context of real Banach spaces is identical.

[5] In the sense of [Amann 1995, Section II].

(iii) *Given $0 \le \beta < 1$ and $0 < \alpha \le 1$, there exists a constant $C > 0$ such that*

$$(t-s)^{\beta-\alpha}\|(U_{\mathbb{A}} - U_{\mathbb{B}})(t,s)\|_{\mathcal{L}(\mathbb{E}_\alpha,\mathbb{E}_\beta)} \le C \max_{\tau\in[s,t]} \|\mathbb{A}(\tau) - \mathbb{B}(\tau)\|_{\mathcal{L}(\mathbb{E}_1,\mathbb{E}_0)} \tag{B-4}$$

for all $(t,s) \in \Delta_T^*$ and all $\mathbb{A}$, $\mathbb{B} \in \mathcal{A}$.

Let now $\mathcal{A}$ be a family as in Proposition B.1. Given $\mathbb{A} \in \mathcal{A}$ and $x \in \mathbb{E}_0$, we consider the linear problem

$$\dot{u} + \mathbb{A}(t)u = 0, \quad t \in (0,T], \quad u(0) = x. \tag{B-5}$$

Using the fundamental properties of the parabolic evolution operator $U_{\mathbb{A}}$ associated to $\mathbb{A}$, it follows from [Amann 1995, Remark II.2.1.2] that (B-5) has a unique classical solution $u := u(\,\cdot\,; x, \mathbb{A})$, that is,

$$u := u(\,\cdot\,; x, \mathbb{A}) \in C^1((0,T],\mathbb{E}_0) \cap C((0,T],\mathbb{E}_1) \cap C([0,T],\mathbb{E}_0)$$

and u satisfies the equation of (B-5) pointwise. This solution is given by the expression

$$u(t) = U_{\mathbb{A}}(t,0)x, \quad t \in [0,T].$$

If $x \in \mathbb{E}_\alpha$ for some $\alpha \in (0,1)$, we may use the relations (B-2)–(B-4) to derive additional regularity properties for the solution, as stated below.

Proposition B.2. *Let $\mathcal{A}$ be a family as in Proposition B.1. The following hold true*:

(i) *Let $0 \le \beta \le \alpha < 1$ and $x \in \mathbb{E}_\alpha$. Then $u \in C^{\alpha-\beta}([0,T],\mathbb{E}_\beta)$ and there exists $C > 0$ such that*

$$\|u(t) - u(s)\|_\beta \le C(t-s)^{\alpha-\beta}\|x\|_\alpha \tag{B-6}$$

for all $(t,s) \in \Delta_T$, $x \in \mathbb{E}_\alpha$, and $\mathbb{A} \in \mathcal{A}$.

(ii) *Let $0 \le \beta < \alpha \le 1$. Then, there exists $C > 0$ such that*

$$\|u(t;x,\mathbb{A}) - u(t;x,\mathbb{B})\|_\beta \le Ct^{\alpha-\beta} \max_{\tau\in[0,t]} \|\mathbb{A}(\tau) - \mathbb{B}(\tau)\|_{\mathcal{L}(\mathbb{E}_1,\mathbb{E}_0)}\|x\|_\alpha \tag{B-7}$$

for all $t \in [0,T]$, $x \in \mathbb{E}_\alpha$, and $\mathbb{A}$, $\mathbb{B} \in \mathcal{A}$.

Proof. The claim (i) follows from [Amann 1995, Theorem II.5.3.1], while (ii) is a consequence of Theorem II.5.2.1 of the same book. □

By means of a contraction argument we now obtain as a preliminary result the following (uniform) local existence theorem, which stays at the basis of Theorem 1.5.

Proposition B.3. *Let the assumptions of Theorem 1.5 be satisfied and let $\bar{f} \in \mathcal{O}_\alpha := \mathcal{O}_\beta \cap \mathbb{E}_\alpha$. Then, there exist constants $\delta = \delta(\bar{f}) > 0$ and $r = r(\bar{f}) > 0$ with the property that for all $f_0 \in \mathcal{O}_\alpha$ with $\|f_0 - \bar{f}\|_\alpha \le r$ the problem*

$$\dot{f} = \Phi(f)[f], \quad t > 0, \quad f(0) = f_0, \tag{QP}$$

possesses a classical solution

$$f \in C([0,\delta],\mathcal{O}_\alpha) \cap C((0,\delta],\mathbb{E}_1) \cap C^1((0,\delta],\mathbb{E}_0) \cap C^{\alpha-\beta}([0,\delta],\mathbb{E}_\beta).$$

Moreover, *if* h *is a further solution to* (QP) *with*

$$h \in C((0,\delta], \mathbb{E}_1) \cap C^1((0,\delta], \mathbb{E}_0) \cap C^\eta([0,\delta], \mathcal{O}_\beta) \quad \textit{for some } \eta \in (0, \alpha-\beta],$$

then $f \equiv h$.

Proof. Existence: We first note that $\mathcal{O}_\alpha$ is an open subset of $\mathbb{E}_\alpha$; see, e.g., [Amann 1995, Section I.2.11]. Since by assumption $-\Phi \in C^{1-}(\mathcal{O}_\beta, \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0))$, it follows from (B-1) there exist constant $R > 0$, $L > 0$, $\kappa \geq 1$, and $\omega > 0$ such that

$$\|\Phi(f) - \Phi(g)\|_{\mathcal{L}(\mathbb{E}_1,\mathbb{E}_0)} \leq L\|f - g\|_\beta \quad \text{for all } f,\ g \in \overline{\mathbb{B}}_{\mathbb{E}_\beta}(\bar{f}, R) \subset \mathcal{O}_\beta, \tag{B-8}$$

$$-\Phi(f) \in \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa, \omega) \quad \text{for all } f \in \overline{\mathbb{B}}_{\mathbb{E}_\beta}(\bar{f}, R). \tag{B-9}$$

Let $\rho \in (0, \alpha - \beta)$ be fixed. If $r > 0$ is sufficiently small, it holds that

$$\overline{\mathbb{B}}_{\mathbb{E}_\alpha}(\bar{f}, r) \subset \overline{\mathbb{B}}_{\mathbb{E}_\beta}(\bar{f}, R) \cap \mathcal{O}_\alpha. \tag{B-10}$$

Given $\delta > 0$, $r > 0$ such that (B-10) holds (r and δ will be fixed later on) and $f_0 \in \overline{\mathbb{B}}_{\mathbb{E}_\alpha}(\bar{f}, r)$, we define the set

$$\mathbb{M} := \big\{ f \in C([0,\delta], \overline{\mathbb{B}}_{\mathbb{E}_\beta}(\bar{f}, R)) : f(0) = f_0 \text{ and } \|f(t) - f(s)\|_\beta \leq |t-s|^\rho \text{ for all } t, s \in [0,\delta] \big\}.$$

Since $\mathbb{M}$ is a closed subset of $C([0,\delta], \mathbb{E}_\beta)$, it is also a (nonempty) complete metric space. Given $f \in \mathbb{M}$, we define

$$\mathbb{A}_f(t) := -\Phi(f(t)), \quad t \in [0,\delta].$$

As a direct consequence of (B-8) and of the definition of $\mathbb{M}$, it follows that

$$\|\mathbb{A}_f(t) - \mathbb{A}_f(s)\|_{\mathcal{L}(\mathbb{E}_1,\mathbb{E}_0)} \leq L\|f(t) - f(s)\|_\beta \leq L|t-s|^\rho, \quad t, s \in [0,\delta],$$

and (B-9) yields that $\mathbb{A}_f(t) \in \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa, \omega)$ for all $f \in \mathbb{M}$ and all $t \in [0,\delta]$. Proposition B.1 ensures the existence of a parabolic evolution operator $U_{\mathbb{A}_f}$ for $\mathbb{A}_f$. Given $f \in \mathbb{M}$, it is natural to consider the linear evolution problem

$$\dot{g} + \mathbb{A}_f(t) g = 0, \quad t \in (0,\delta], \quad g(0) = f_0, \tag{B-11}$$

which has, in view of Proposition B.2, a unique classical solution

$$g := \Gamma(f) := U_{\mathbb{A}_f}(\,\cdot\,, 0) f_0 \in C^{\alpha-\beta}([0,\delta], \mathbb{E}_\beta) \cap C([0,\delta], \mathbb{E}_\alpha).$$

The existence part of Proposition B.1 reduces to proving that $\Gamma : \mathbb{M} \to \mathbb{M}$ is a strict contraction for suitable r and δ. Clearly $\Gamma(f)(0) = f_0$. Moreover, (B-6) yields

$$\|\Gamma(f)(t) - \Gamma(f)(s)\|_\beta \leq C|t-s|^{\alpha-\beta}\|f_0\|_\alpha \leq C\delta^{\alpha-\beta-\rho}(\|\bar{f}\|_\alpha + r)|t-s|^\rho \leq |t-s|^\rho \quad \text{for all } t, s \in [0,\delta],$$

provided that

$$C\delta^{\alpha-\beta-\rho}(\|\bar{f}\|_\alpha + r) \leq 1. \tag{B-12}$$

The latter estimate (with $s = 0$) yields

$$\|\Gamma(f)(t) - \bar f\|_\beta \le \|\Gamma(f)(t) - \Gamma(f)(0)\|_\beta + \|f_0 - \bar f\|_\beta \le \delta^\rho + r\|i_{\mathbb{E}_\alpha \hookrightarrow \mathbb{E}_\beta}\|_{\mathcal{L}(\mathbb{E}_\alpha, \mathbb{E}_\beta)} \le R \quad \text{for all } t \in [0, \delta],$$

if we additionally require that

$$\delta^\rho + r\|i_{\mathbb{E}_\alpha \hookrightarrow \mathbb{E}_\beta}\|_{\mathcal{L}(\mathbb{E}_\alpha, \mathbb{E}_\beta)} \le R. \tag{B-13}$$

We now assume that r and δ are chosen such that (B-12)–(B-13) hold true. It then follows that $\Gamma : \mathbb{M} \to \mathbb{M}$ is a well-defined map. Furthermore, given $f,\ h \in \mathbb{M}$, the estimate (B-7) together with (B-8) yields

$$\begin{aligned}
\|\Gamma(f)(t) - \Gamma(h)(t)\|_\beta &= \|U_{\mathbb{A}_f}(t, 0) f_0 - U_{\mathbb{A}_h}(t, 0) f_0\|_\beta \\
&\le C t^{\alpha-\beta} \max_{\tau \in [0,t]} \|\mathbb{A}_f(\tau) - \mathbb{A}_h(\tau)\|_{\mathcal{L}(\mathbb{E}_1, \mathbb{E}_0)} \|f_0\|_\alpha \\
&\le C\delta^{\alpha-\beta} L(\|\bar f\|_\alpha + r) \max_{t \in [0,\delta]} \|f(t) - h(t)\|_\beta \\
&\le \tfrac{1}{2} \max_{t \in [0,\delta]} \|f(t) - h(t)\|_\beta \quad \text{for all } t \in [0, \delta],
\end{aligned}$$

provided that

$$C\delta^{\alpha-\beta} L(\|\bar f\|_\alpha + r) \le \tfrac{1}{2}. \tag{B-14}$$

Hence, if r and δ are chosen such that also (B-14) is satisfied, then Γ is a strict contraction and Banach's fixed-point theorem ensures that Γ has a fixed point. This proves the existence part.

Uniqueness: Let f be a solution to (QP) as found above and let $h \not\equiv f$ be a further classical solution such that $h \in C^\eta([0, \delta], \mathbb{E}_\beta)$ for some $\eta \in (0, \alpha - \beta]$. The real number

$$t_0 := \max\{t \in [0, \delta] : f|_{[0,t]} = h|_{[0,t]}\}$$

satisfies $0 \le t_0 < \delta$ and $f = h$ on $[0, t_0]$. Since $f(t_0) \in \mathcal{O}_\alpha$, there exist $R > 0,\ L > 0,\ \kappa \ge 1$, and $\omega > 0$ such that

$$\begin{aligned}
\|\Phi(u) - \Phi(v)\|_{\mathcal{L}(\mathbb{E}_1, \mathbb{E}_0)} \le L\|u - v\|_\beta \quad &\text{for all } u,\ v \in \overline{\mathbb{B}}_{\mathbb{E}_\beta}(f(t_0), R) \subset \mathcal{O}_\beta, \\
-\Phi(u) \in \mathcal{H}(\mathbb{E}_1, \mathbb{E}_0, \kappa, \omega) \quad &\text{for all } u \in \overline{\mathbb{B}}_{\mathbb{E}_\beta}(f(t_0), R).
\end{aligned}$$

Given $\delta_0 \in (t_0, \delta]$, the set

$$\mathbb{M}_0 := \left\{ h \in C([0, \delta_0 - t_0], \overline{\mathbb{B}}_{\mathbb{E}_\beta}(f(t_0), R)) : h(0) = f(t_0),\ \frac{\|h(t) - h(s)\|_\beta}{|t - s|^{\eta/2}} \le 1 \text{ for all } t \ne s \in [0, \delta_0 - t_0] \right\}$$

is a (nonempty) complete metric space. Letting $\mathbb{A}_h(t) := -\Phi(h(t))$ for $h \in \mathbb{M}_0$ and $t \in [0, \delta_0 - t_0]$, we may argue as in the existence part of this proof to conclude that the linear problem

$$\dot u + \mathbb{A}_h(t) u = 0, \quad t \in (0, \delta_0 - t_0], \quad h(0) = f(t_0)$$

has a unique classical solution $\Gamma_0(h) \in C^{\alpha-\beta}([0, \delta_0 - t_0], \mathbb{E}_\beta) \cap C([0, \delta_0 - t_0], \mathbb{E}_\alpha)$. Furthermore, $\Gamma_0 : \mathbb{M}_0 \to \mathbb{M}_0$ is a $\frac{1}{2}$-contraction provided that δ_0 is sufficiently close to t_0; hence Γ_0 has a unique fixed point. But, if $\delta_0 - t_0$ is sufficiently small, then it can be easily seen that $f(\cdot + t_0)|_{[0,\delta_0 - t_0]}$ and $h(\cdot + t_0)|_{[0,\delta_0 - t_0]}$ both belong to $\mathbb{M}_0$ and these functions are therefore fixed points of Γ_0. This implies $f = h$ on $[0, \delta_0]$ for some $\delta_0 > t_0$, in contradiction with the definition of t_0. This proves the uniqueness claim. □

We are now in a position to prove Theorem 1.5.

Proof of Theorem 1.5. Let $f_0 \in \mathcal{O}_\alpha$ be given. According to Proposition B.3 (with $\bar{f} := f_0$), there exists $\delta > 0$ and a classical solution

$$f \in C([0, \delta], \mathcal{O}_\alpha) \cap C((0, \delta], \mathbb{E}_1) \cap C^1((0, \delta], \mathbb{E}_0) \cap C^{\alpha-\beta}([0, \delta], \mathbb{E}_\beta)$$

to (QP). This solution can be continued as follows. Applying Proposition B.3 (with $\bar{f} := f(\delta)$), we find $r > 0$ and $\delta_1 > 0$ such that

$$\dot{h} = \Phi(h)[h], \quad t \in (0, \delta_1], \quad h(0) = f_1 \tag{B-15}$$

has a classical solution $h \in C([0, \delta_1], \mathcal{O}_\alpha) \cap C((0, \delta_1], \mathbb{E}_1) \cap C^1((0, \delta_1], \mathbb{E}_0) \cap C^{\alpha-\beta}([0, \delta_1], \mathbb{E}_\beta)$ for each $f_1 \in \mathcal{O}_\alpha$ with $\|f_1 - f(t_0)\|_\alpha \le r$. Let $t_0 \in (0, \delta)$ be such that

$$t_0 + \delta_1 > \delta \quad \text{and} \quad \|f(t_0) - \bar{f}\|_\alpha \le r.$$

Hence it is possible to choose $f_1 := f(t_0)$ as an initial value in (B-15). Since $f(\,\cdot\, + t_0) : [0, \delta - t_0] \to \mathcal{O}_\alpha$ and $h : [0, \delta - t_0] \to \mathcal{O}_\alpha$ are both classical solutions to

$$\dot{h} = \Phi(h)[h], \quad t \in (0, \delta - t_0], \quad h(0) = f(t_0),$$

by Proposition B.3 they must coincide. Consequently, the function $F : [0, t_0 + \delta_1] \to \mathcal{O}_\alpha$ defined by

$$F(t) := \begin{cases} f(t), & t \in [0, \delta], \\ h(t - t_0), & t \in [\delta, t_0 + \delta_1] \end{cases}$$

is a classical solution to (QP) which extends f. The maximal solution $f = f(\,\cdot\,; f_0) : I(f_0) \to \mathcal{O}_\alpha$ in Theorem 1.5 is defined by setting

$$\begin{aligned} I(f_0) &:= \bigcup \bigl\{[0, \delta] : \text{(QP) has a classical solution } f_\delta \text{ on } [0, \delta] \text{ with } f_\delta \in C^{\alpha-\beta}([0, \delta], \mathbb{E}_\beta)\bigr\}, \\ f(t) &:= f_\delta(t) \quad \text{for } t \in [0, \delta]. \end{aligned}$$

The construction above shows that f is well-defined and that $I(f_0) = [0, T_+(f_0))$ with $T_+(f_0) \le \infty$. This proves the existence claim in Theorem 1.5. The uniqueness assertion is an immediate consequence of Proposition B.3.

We now prove the criterion for global existence. Hence, let us assume that the unique classical maximal solution $f = f(\,\cdot\,; f_0) : [0, T_+(f_0)) \to \mathcal{O}_\alpha$ to (QP) is uniformly continuous when restricted to each interval $[0, T] \cap [0, T_+(f_0))$, with $T > 0$ arbitrary. We further assume that $\tau := T_+(f_0) < \infty$; otherwise we are done. Then, since f is uniformly continuous on $[0, \tau)$, it is straightforward to see that the limit

$$f(\tau) := \lim_{t \nearrow \tau} f(t)$$

exists in $\overline{\mathcal{O}}_\alpha$. If $\operatorname{dist}(f(t), \partial\mathcal{O}_\alpha) \not\to 0$ for $t \to \tau$, it must hold that $f(\tau) \in \mathcal{O}_\alpha$. Proceeding as above, we may extend in view of Proposition B.3 this maximal solution to an interval $[0, \tau + \delta_1)$ for some $\delta_1 > 0$, which is a contradiction and we are done.

Finally, the semiflow property of the solution map $[(t, f_0) \mapsto f(t; f_0)]$ stated at the end of Theorem 1.5 is proven in detail in [Amann 1988, Theorem 8.1]. Furthermore, if Φ is additionally smooth, then proceeding as in Theorem 11.3 of the same paper one may show that the semiflow map is also smooth. For real-analytic Φ, the real-analyticity of $[(t, f_0) \mapsto f(t; f_0)]$ follows by estimating the Fréchet derivatives of the flow map, which is a rather tedious and lengthy procedure which we refrain from presenting here. $\square$

Acknowledgement

The author would like to thank Christoph Thiele for the discussion on an issue related to the analysis in Section 3.

References

[Amann 1986] H. Amann, "Quasilinear parabolic systems under nonlinear boundary conditions", *Arch. Rational Mech. Anal.* **92**:2 (1986), 153–192. MR Zbl

[Amann 1988] H. Amann, "Dynamic theory of quasilinear parabolic equations, I: Abstract evolution equations", *Nonlinear Anal.* **12**:9 (1988), 895–919. MR Zbl

[Amann 1993] H. Amann, "Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems", pp. 9–126 in *Function spaces, differential operators and nonlinear analysis* (Friedrichroda, 1992), edited by H.-J. Schmeisser and H. Triebel, Teubner-Texte Math. **133**, Teubner, Stuttgart, 1993. MR Zbl

[Amann 1995] H. Amann, *Linear and quasilinear parabolic problems, I: Abstract linear theory*, Monographs in Mathematics **89**, Birkhäuser, Boston, 1995. MR Zbl

[Ambrose 2004] D. M. Ambrose, "Well-posedness of two-phase Hele-Shaw flow without surface tension", *European J. Appl. Math.* **15**:5 (2004), 597–607. MR Zbl

[Ambrose 2014] D. M. Ambrose, "The zero surface tension limit of two-dimensional interfacial Darcy flow", *J. Math. Fluid Mech.* **16**:1 (2014), 105–143. MR Zbl

[Angenent 1990] S. B. Angenent, "Nonlinear analytic semiflows", *Proc. Roy. Soc. Edinburgh Sect. A* **115**:1-2 (1990), 91–107. MR Zbl

[Bazaliy and Vasylyeva 2014] B. V. Bazaliy and N. Vasylyeva, "The two-phase Hele-Shaw problem with a nonregular initial interface and without surface tension", *Zh. Mat. Fiz. Anal. Geom.* **10**:1 (2014), 3–43. MR Zbl

[Bear 1972] J. Bear, *Dynamics of fluids in porous media*, Elsevier, New York, 1972. Zbl

[Berselli et al. 2014] L. C. Berselli, D. Córdoba, and R. Granero-Belinchón, "Local solvability and turning for the inhomogeneous Muskat problem", *Interfaces Free Bound.* **16**:2 (2014), 175–213. MR Zbl

[Calderon et al. 1978] A. P. Calderon, C. P. Calderon, E. Fabes, M. Jodeit, and N. M. Rivière, "Applications of the Cauchy integral on Lipschitz curves", *Bull. Amer. Math. Soc.* **84**:2 (1978), 287–290. MR Zbl

[Castro et al. 2011] A. Castro, D. Córdoba, C. L. Fefferman, F. Gancedo, and M. López-Fernández, "Turning waves and breakdown for incompressible flows", *Proc. Natl. Acad. Sci. USA* **108**:12 (2011), 4754–4759. MR Zbl

[Castro et al. 2012] A. Castro, D. Córdoba, C. Fefferman, F. Gancedo, and M. López-Fernández, "Rayleigh–Taylor breakdown for the Muskat problem with applications to water waves", *Ann. of Math.* (2) **175**:2 (2012), 909–948. MR Zbl

[Castro et al. 2013] A. Castro, D. Córdoba, C. Fefferman, and F. Gancedo, "Breakdown of smoothness for the Muskat problem", *Arch. Ration. Mech. Anal.* **208**:3 (2013), 805–909. MR Zbl

[Cheng et al. 2016] C. H. A. Cheng, R. Granero-Belinchón, and S. Shkoller, "Well-posedness of the Muskat problem with H^2 initial data", *Adv. Math.* **286** (2016), 32–104. MR Zbl

[Constantin et al. 2013] P. Constantin, D. Córdoba, F. Gancedo, and R. M. Strain, "On the global existence for the Muskat problem", *J. Eur. Math. Soc.* (*JEMS*) **15**:1 (2013), 201–227. MR Zbl

[Constantin et al. 2016] P. Constantin, D. Córdoba, F. Gancedo, L. Rodríguez-Piazza, and R. M. Strain, "On the Muskat problem: global in time results in 2D and 3D", *Amer. J. Math.* **138**:6 (2016), 1455–1494. MR Zbl

[Constantin et al. 2017] P. Constantin, F. Gancedo, R. Shvydkoy, and V. Vicol, "Global regularity for 2D Muskat equations with finite slope", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **34**:4 (2017), 1041–1074. MR Zbl

[Córdoba and Gancedo 2007] D. Córdoba and F. Gancedo, "Contour dynamics of incompressible 3-D fluids in a porous medium with different densities", *Comm. Math. Phys.* **273**:2 (2007), 445–471. MR Zbl

[Córdoba and Gancedo 2010] D. Córdoba and F. Gancedo, "Absence of squirt singularities for the multi-phase Muskat problem", *Comm. Math. Phys.* **299**:2 (2010), 561–575. MR Zbl

[Córdoba et al. 2011] A. Córdoba, D. Córdoba, and F. Gancedo, "Interface evolution: the Hele-Shaw and Muskat problems", *Ann. of Math.* (2) **173**:1 (2011), 477–542. MR Zbl

[Córdoba et al. 2013] A. Córdoba, D. Córdoba, and F. Gancedo, "Porous media: the Muskat problem in three dimensions", *Anal. PDE* **6**:2 (2013), 447–497. MR Zbl

[Córdoba et al. 2014] D. Córdoba Gazolaz, R. Granero-Belinchón, and R. Orive-Illera, "The confined Muskat problem: differences with the deep water regime", *Commun. Math. Sci.* **12**:3 (2014), 423–455. MR Zbl

[Da Prato and Grisvard 1979] G. Da Prato and P. Grisvard, "Equations d'évolution abstraites non linéaires de type parabolique", *Ann. Mat. Pura Appl.* (4) **120** (1979), 329–396. MR Zbl

[Ehrnström et al. 2013] M. Ehrnström, J. Escher, and B.-V. Matioc, "Steady-state fingering patterns for a periodic Muskat problem", *Methods Appl. Anal.* **20**:1 (2013), 33–46. MR Zbl

[Escher 1994] J. Escher, "The Dirichlet–Neumann operator on continuous functions", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **21**:2 (1994), 235–266. MR Zbl

[Escher and Matioc 2011] J. Escher and B.-V. Matioc, "On the parabolicity of the Muskat problem: well-posedness, fingering, and stability results", *Z. Anal. Anwend.* **30**:2 (2011), 193–218. MR Zbl

[Escher and Simonett 1995] J. Escher and G. Simonett, "Maximal regularity for a free boundary problem", *NoDEA Nonlinear Differential Equations Appl.* **2**:4 (1995), 463–510. MR Zbl

[Escher and Simonett 1996] J. Escher and G. Simonett, "Analyticity of the interface in a free boundary problem", *Math. Ann.* **305**:3 (1996), 439–459. MR Zbl

[Escher and Simonett 1997] J. Escher and G. Simonett, "Classical solutions of multidimensional Hele-Shaw models", *SIAM J. Math. Anal.* **28**:5 (1997), 1028–1047. MR Zbl

[Escher et al. 2012] J. Escher, A.-V. Matioc, and B.-V. Matioc, "A generalized Rayleigh–Taylor condition for the Muskat problem", *Nonlinearity* **25**:1 (2012), 73–92. MR Zbl

[Escher et al. 2018] J. Escher, B.-V. Matioc, and C. Walker, "The domain of parabolicity for the Muskat problem", *Indiana Univ. Math. J.* **67**:2 (2018), 679–737.

[Friedman and Tao 2003] A. Friedman and Y. Tao, "Nonlinear stability of the Muskat problem with capillary pressure at the free boundary", *Nonlinear Anal.* **53**:1 (2003), 45–80. MR Zbl

[Gancedo and Strain 2014] F. Gancedo and R. M. Strain, "Absence of splash singularities for surface quasi-geostrophic sharp fronts and the Muskat problem", *Proc. Natl. Acad. Sci. USA* **111**:2 (2014), 635–639. MR Zbl

[Gilbarg and Trudinger 1998] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Grundlehren der Mathematischen Wissenschaften **224**, Springer, 1998.

[Gómez-Serrano and Granero-Belinchón 2014] J. Gómez-Serrano and R. Granero-Belinchón, "On turning waves for the inhomogeneous Muskat problem: a computer-assisted proof", *Nonlinearity* **27**:6 (2014), 1471–1498. MR Zbl

[Granero-Belinchón 2014] R. Granero-Belinchón, "Global existence for the confined Muskat problem", *SIAM J. Math. Anal.* **46**:2 (2014), 1651–1680. MR Zbl

[Hong et al. 1997] J. Hong, Y. Tao, and F. Yi, "Muskat problem with surface tension", *J. Partial Differential Equations* **10**:3 (1997), 213–231. MR Zbl

[Lu 1993] J. K. Lu, *Boundary value problems for analytic functions*, Series in Pure Mathematics **16**, World Scientific, River Edge, NJ, 1993. MR Zbl

[Lunardi 1995] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Progress in Nonlinear Differential Equations and their Applications **16**, Birkhäuser, Basel, 1995. MR Zbl

[Meyer and Coifman 1997] Y. Meyer and R. Coifman, *Wavelets: Calderón–Zygmund and multilinear operators*, Cambridge Studies in Advanced Mathematics **48**, Cambridge University Press, 1997. MR Zbl

[Murai 1986] T. Murai, "Boundedness of singular integral operators of Calderón type, VI", *Nagoya Math. J.* **102** (1986), 127–133. MR Zbl

[Muskat 1934] M. Muskat, "Two fluid systems in porous media: the encroachment of water into an oil sand", *Physics* **5** (1934), 250–264. Zbl

[Prüss and Simonett 2016a] J. Prüss and G. Simonett, *Moving interfaces and quasilinear parabolic evolution equations*, Monographs in Mathematics **105**, Springer, 2016. MR Zbl

[Prüss and Simonett 2016b] J. Prüss and G. Simonett, "On the Muskat problem", *Evol. Equ. Control Theory* **5**:4 (2016), 631–645. MR Zbl

[Prüss et al. 2015] J. Prüss, Y. Shao, and G. Simonett, "On the regularity of the interface of a thermodynamically consistent two-phase Stefan problem with surface tension", *Interfaces Free Bound.* **17**:4 (2015), 555–600. MR Zbl

[Runst and Sickel 1996] T. Runst and W. Sickel, *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, De Gruyter Series in Nonlinear Analysis and Applications **3**, de Gruyter, Berlin, 1996. MR Zbl

[Saffman and Taylor 1958] P. G. Saffman and G. Taylor, "The penetration of a fluid into a porous medium or Hele-Shaw cell containing a more viscous liquid", *Proc. Roy. Soc. London. Ser. A* **245** (1958), 312–329. MR Zbl

[Siegel et al. 2004] M. Siegel, R. E. Caflisch, and S. Howison, "Global existence, singular solutions, and ill-posedness for the Muskat problem", *Comm. Pure Appl. Math.* **57**:10 (2004), 1374–1411. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR Zbl

[Tofts 2017] S. Tofts, "On the existence of solutions to the Muskat problem with surface tension", *J. Math. Fluid Mech.* **19**:4 (2017), 581–611. MR Zbl

[Triebel 1978] H. Triebel, *Interpolation theory, function spaces, differential operators*, North-Holland Mathematical Library **18**, North-Holland Publishing Co., Amsterdam-New York, 1978. MR Zbl

[Yi 1996] F. Yi, "Local classical solution of Muskat free boundary problem", *J. Partial Differential Equations* **9**:1 (1996), 84–96. MR Zbl

Received 18 Oct 2016. Revised 17 Jan 2018. Accepted 7 May 2018.

BOGDAN-VASILE MATIOC: bogdan.matioc@ur.de
Fakultät für Mathematik, Universität Regensburg, Regensburg, Germany

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.333

MAXIMAL GAIN OF REGULARITY IN VELOCITY AVERAGING LEMMAS

DIOGO ARSÉNIO AND NADER MASMOUDI

We investigate new settings of velocity averaging lemmas in kinetic theory where a maximal gain of half a derivative is obtained. Specifically, we show that if the densities f and g in the transport equation $v\cdot\nabla_x f = g$ belong to $L^r_x L^{r'}_v$, where $2n/(n+1) < r \leq 2$ and $n \geq 1$ is the dimension, then the velocity averages belong to $H^{1/2}_x$.

We further explore the setting where the densities belong to $L^{4/3}_x L^2_v$ and show, by completing the work initiated by Pierre-Emmanuel Jabin and Luis Vega on the subject, that velocity averages almost belong to $W^{n/(4(n-1)),4/3}_x$ in this case, in any dimension $n \geq 2$, which strongly indicates that velocity averages should almost belong to $W^{1/2,2n/(n+1)}_x$ whenever the densities belong to $L^{2n/(n+1)}_x L^2_v$.

These results and their proofs bear a strong resemblance to the famous and notoriously difficult problems of boundedness of Bochner–Riesz multipliers and Fourier restriction operators, and to smoothing conjectures for Schrödinger and wave equations, which suggests interesting links between kinetic theory, dispersive equations and harmonic analysis.

1. Introduction and main results

Velocity averaging lemmas are a category of regularity results concerning the kinetic transport equation

$$(\partial_t + v\cdot\nabla_x) f(t,x,v) = g(t,x,v), \tag{1-1}$$

where $(t,x,v) \in [0,\infty)\times\mathbb{R}^n\times\mathbb{R}^n$, or its stationary counterpart

$$v\cdot\nabla_x f(x,v) = g(x,v), \tag{1-2}$$

where $(x,v)\in\mathbb{R}^n\times\mathbb{R}^n$, with $n \geq 1$.

We would like to thank Pierre-Emmanuel Jabin for sharing his insight on the subject with us. Masmoudi was partially supported by NSF grant no. DMS-1716466.
MSC2010: primary 35B65; secondary 42B37, 82C40.
Keywords: velocity averaging lemmas, kinetic theory, kinetic transport equation.

Variants of the above equations are also relevant. Indeed, different spatial and velocity domains, as well as nonlinear velocity fields (consider the relativistic case), are sometimes studied. Nevertheless, for the sake of simplicity, we will focus exclusively on the Euclidean stationary setting (1-2), which, we believe, captures the essential features of kinetic transport (at least as far as velocity averaging is concerned). We refer the interested reader to Appendix C, where we establish an equivalence of velocity averaging lemmas for velocities in $\mathbb{R}^n$ and in $\mathbb{S}^{n-1}$. In particular, this provides a rather general method to adapt the results contained in the present work to settings where velocities belong to a manifold of codimension 1, which includes the nonstationary transport equation (1-1).

The classical velocity averaging lemma was established first in [Golse et al. 1988]. It asserts that if $f, g \in L^2_{x,v}$ satisfy the transport relation (1-2), then the velocity averages of f enjoy the regularization

$$\int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\,dv \in H_x^{\frac{1}{2}}$$

for any given $\varphi \in L^\infty_c(\mathbb{R}^n)$ (that is, any measurable function bounded almost everywhere with compact support). Note that such regularity results had already been suggested in weaker forms in [Agoshkov 1984; Golse et al. 1985].

An extension of this fundamental result to the $L^p_{x,v}$ setting, with $1 < p < \infty$, was also obtained in [Golse et al. 1988] and later substantially improved in [Bézard 1994; DeVore and Petrova 2001; DiPerna et al. 1991]. Generally speaking, such generalizations are deduced by interpolating the preceding $L^2_{x,v}$ case with the degenerate $L^1_{x,v}$ and $L^\infty_{x,v}$ cases. In this setting, it is established that, for any $\varphi \in L^\infty_c(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\,dv \in W_x^{s,p}$$

whenever $f, g \in L^p_{x,v}$, with $s = 1 - \frac{1}{p}$ if $p \le 2$ and for any $0 \le s < \frac{1}{p}$ if $p > 2$.

When $p \le 2$, the optimality of the regularity index $s = 1 - \frac{1}{p}$ in the preceding result was shown in [Lions 1995] through a straightforward dimensional analysis. As for the case $p > 2$, it was also argued in that paper that the regularity of velocity averages cannot be improved beyond the value $s = \frac{1}{p}$, but this optimality argument remains incomplete, for it requires the use of a larger class of velocity weights $\varphi(v)$ with unbounded support. In fact, it turns out that, in general, the value $s = \frac{1}{p}$ is not optimal in the range $2 < p < \infty$, for it is possible to largely improve this regularity index beyond the value $s = \frac{1}{p}$ in dimension $n = 1$, as stated in the following one-dimensional theorem.

Theorem 1.1. *In dimension $n = 1$, let $f, g \in L^p_{x,v}$, with $2 < p < \infty$, be such that* (1-2) *holds true. Then,*

$$\int_{\mathbb{R}} f(x,v)\,\varphi(v)\,dv \in W_x^{s,p}$$

for all $0 \le s < 1 - \frac{1}{p}$ and any $\varphi \in L^\infty_c(\mathbb{R})$.

This result clearly follows from the more general Theorem 4.3, by setting $p = r$ therein, which is established later on in Section 4.

The question of the optimality of the value $s = \frac{1}{p}$, when $p > 2$, in higher dimensions $n \geq 2$ was finally definitely settled in [DeVore and Petrova 2001, Theorem 1.3], where a remarkable construction of a convoluted counterexample shows the necessity of the constraint $s \leq \frac{1}{p}$ whenever $p > 2$ and $n \geq 2$. Note however that it remained so far unknown whether the endpoint value $s = \frac{1}{p}$ is admissible or not. It turns out that, as a byproduct of our methods, we are able to settle this question here by showing that the endpoint value $s = \frac{1}{p}$ is indeed admissible when $p > 2$ (see Theorem 3.6).

On the whole, the maximal gain of regularity, when $n \geq 2$, clearly happens for the value $p = 2$, where half a derivative is gained by averaging in velocity.

It is to be emphasized that the refined interpolation methods used in [DeVore and Petrova 2001] yield more precise results. Indeed, it is established therein that, for each $1 < p < \infty$, the velocity averages actually belong to the Besov space $B^s_{p,p}(\mathbb{R}^n)$ with $s = \min\{\frac{1}{p}, 1 - \frac{1}{p}\}$, which is smaller than $W^{s,p}(\mathbb{R}^n)$ for values $1 < p \leq 2$, and that this is optimal. Nevertheless, for the sake of simplicity, we will only focus here on standard Sobolev spaces and we will omit the more precise formulations of velocity averaging lemmas in Besov spaces, which can be easily deduced from our proofs if needed (we refer to the proof of Proposition 3.2 in Section 3 for some more details on this matter).

Numerous generalizations of velocity averaging lemmas are available. For instance, several settings where f and g belong to distinct spaces (possibly with different homogeneity) of different kinds (Besov, Sobolev, etc.) with mixed integrability and regularity in space and velocity have been considered in [Arsénio and Masmoudi 2014; Bézard 1994; DiPerna et al. 1991; Jabin and Vega 2004; Westdickenberg 2002]. Naturally, the ensuing gain of regularity on the velocity averages depends then on the different parameters used to characterize these function spaces. In these more general settings, the phenomena of dispersion (as discovered in [Castella and Perthame 1996]) and hypoellipticity (as discovered in [Bouchut 2002]; see also [Arsénio and Saint-Raymond 2011]) in kinetic transport equations come into play and, loosely speaking, interact with the regularization due to velocity averaging to produce new interesting results. We refer to [Arsénio and Masmoudi 2014; Westdickenberg 2002] and [Arsénio and Masmoudi 2014; Arsénio and Saint-Raymond 2011; Jabin and Vega 2004] for such results combining velocity averaging with the dispersive and hypoelliptic effects, respectively. Note that none of these phenomena is fully distinct from the others.

It was argued in [Arsénio and Masmoudi 2014; Westdickenberg 2002] that the influence of dispersion on velocity averaging produces a gain of integrability which can be interpreted, through Sobolev embeddings, as a regularity gain which is sometimes larger than half a derivative and even possibly close or equal to a whole derivative (note that the gain of regularity can never be larger than a whole derivative, for the transport operator is a differential operator of order 1). Furthermore, the hypoelliptic phenomenon may also produce a regularity gain close or equal to a whole derivative on the velocity average, see [Arsénio and Masmoudi 2014; Jabin and Vega 2004], but this requires assuming some regularity on f and g a priori.

In this article, we will exclusively focus on the gain of regularity due to velocity averaging, possibly combined with dispersion (without interpreting the gain of regularity through Sobolev embeddings as was done in [Arsénio and Masmoudi 2014; Westdickenberg 2002], though), and will mostly ignore the aforementioned effects produced by hypoellipticity that were analyzed in [Arsénio and Masmoudi

2014; Jabin and Vega 2004]. To this end, we will only consider settings where f and g belong to mixed Lebesgue spaces and no a priori regularity is assumed. In this case, it is largely agreed that the gain of regularity cannot exceed half a derivative in dimension $n \geq 2$ (but there is no proof of this general assertion, yet).

Thus, so far, the maximal gain of half a derivative is only known to be attained when f and g both belong to $L^2_{x,v}$. In the present work, we explore new settings of velocity averaging lemmas where a maximal gain of half a derivative is obtained. Our first main result shows that it is possible to gain exactly half a derivative even if f and g do not belong to $L^2_{x,v}$.

Theorem 1.2. *In any dimension $n \geq 1$, let $f, g \in L^r_x L^{r'}_v$, with $\frac{2n}{n+1} < r \leq 2$, be such that* (1-2) *holds true. Then,*

$$\int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\,dv \in H^{\frac{1}{2}}_x$$

for any $\varphi \in L^\infty_c(\mathbb{R}^n)$.

This result clearly follows from the more general Theorem 3.6 by setting $a = 2$ therein, which is established later on in Section 3. It is based on a TT^*-argument combined with the dispersion due to kinetic transport studied in Section 2 and velocity averaging.

Such a result had already been hinted at in [Jabin and Vega 2003; 2004], where it was established that, in two dimensions only ($n = 2$), velocity averages of f belong to H^s_x, for any $0 \leq s < \frac{1}{2}$, provided f and g belong to $L^{\frac{4}{3}}_x L^\infty_v$; see [Jabin and Vega 2004, Theorem 1.3].

In fact, in [Jabin and Vega 2003; 2004], the authors further identified another case which could potentially lead to a gain of almost half a derivative on the velocity averages. More precisely, they showed that, in two dimensions only ($n = 2$), velocity averages of f belong to $W^{s,\frac{4}{3}}_x$ for any $0 \leq s < \frac{1}{2}$, provided f and g belong to $L^{\frac{4}{3}}_x L^2_v$ and under the peculiar assumption that $g(x,v)\,\varphi(v)$ is an even function in v; see [Jabin and Vega 2004, Theorem 1.2]. The latter assumption is rather unnatural and it remained unclear whether this evenness condition could be removed or not.

By building upon the work from [Jabin and Vega 2004], combining our methods with the remarkable proof of Theorem 1.2 therein, we are able to bring a definitive answer to this two-dimensional problem, which is precisely the content of the following result.

Theorem 1.3. *In dimension $n = 2$, let $f, g \in L^{\frac{4}{3}}_x L^2_v$ be such that* (1-2) *holds true. Then,*

$$\int_{\mathbb{R}^2} f(x,v)\,\varphi(v)\,dv \in W^{s,\frac{4}{3}}_x$$

for all $0 \leq s < \frac{1}{2}$ and any $\varphi \in L^\infty_c(\mathbb{R}^2)$.

This result clearly follows from the more general Theorem 5.4, by setting $r = \frac{4}{3}$ therein, which is proved later on in Section 5. Its proof follows from the analysis of the boundedness of some adjoint transport operator on the dual space $L^4_x = (L^{\frac{4}{3}}_x)'$ and uses crucially the trivial fact that the exponent 4 is an even integer to control the square of this adjoint transport operator in L^2_x rather than the operator itself in L^4_x. This fact, among other characteristics of the proof, is strikingly reminiscent of the proofs of boundedness of Bochner–Riesz multipliers and Fourier restriction operators in two dimensions. We refer to [Grafakos 2009] for more on these subjects from harmonic analysis.

In higher dimensions, we extend the preceding result into the following theorem.

Theorem 1.4. *In any dimension $n \geq 3$, let $f, g \in L^{\frac{4}{3}}_x L^2_v$ be such that* (1-2) *holds true.*

Then,

$$\int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\,dv \in W^{s,\frac{4}{3}}_x$$

for all $0 \leq s < \frac{n}{4(n-1)}$ and any $\varphi \in L^\infty_c(\mathbb{R}^n)$.

This result clearly follows from the more general Theorem 6.8 by setting $r = \frac{4}{3}$ therein, which is proved later on in Section 6.

Observe that, employing rather general interpolation methods, it is possible to deduce a large variety of velocity averaging results, similar to those asserted in the above theorems, combining spaces for f and g with distinct integrabilities. We refer to [Arsénio and Masmoudi 2014] (see in particular the very general Theorem 4.7 therein) for such interpolation techniques.

It is likely that Theorem 1.4 may be largely improved. Indeed, note that a formal interpolation would yield

$$(L^1_x L^2_v, L^{\frac{2n}{n+1}}_x L^2_v)_{\frac{n}{2(n-1)}} = L^{\frac{4}{3}}_x L^2_v \quad \text{and} \quad (L^1_x, W^{\frac{1}{2},\frac{2n}{n+1}}_x)_{\frac{n}{2(n-1)}} = W^{\frac{n}{4(n-1)},\frac{4}{3}}_x, \tag{1-3}$$

whence formally extrapolating the above regularity result has us believe that, for any $\varphi \in L^\infty_c(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\,dv \in W^{s,\frac{2n}{n+1}}_x \tag{1-4}$$

for all $0 \leq s < \frac{1}{2}$, whenever $f, g \in L^{\frac{2n}{n+1}}_x L^2_v$ (see [Arsénio 2015] for a survey of velocity averaging lemmas and more on such conjectures; see also Figure 2 and the related comments following the proof of Theorem 6.8, below). In other words, Theorem 1.4 would follow from a formal interpolation of (1-4) with the degenerate L^1 case.

However, we do not know how to prove this estimate…

Finally, we would like to emphasize that, in this work, we investigate velocity averaging for its own sake, as a functional analytic study. Indeed, the search for maximal regularity in velocity averaging lemmas has already proved a challenging and interesting endeavor requiring diverse and original methods (extending beyond the classical settings of velocity averaging), producing interesting new results and leading to exciting research perspectives.

However, it should not be overlooked that velocity averaging lemmas also enjoy concrete applications to a wide variety of fundamental problems from kinetic theory. Such applications include, for instance, the existence of renormalized solutions to the Boltzmann equation [DiPerna and Lions 1989] and the convergence of such solutions to Leray solutions of the Navier–Stokes equations in a viscous incompressible hydrodynamic regime [Golse and Saint-Raymond 2004].

The investigation of sharp versions of averaging lemmas, such as the ones presented in this work, may lead to fundamental applications, as well. Indeed, we believe that such results may be very useful in establishing optimal regularity estimates in nonlinear conservation laws, through the study of their corresponding kinetic formulations. In particular, averaging lemmas with mixed integrability in x and v may be crucial in such attempts, for kinetic formulations are often based on densities which display distinct integrability or regularity properties in each variable.

We refer to [Lions et al. 1994a] (see Theorem 4 therein; see also [Lions et al. 1994b, Proposition 7] in the context of isentropic gas dynamics) for an early application of velocity averaging lemmas to kinetic formulations of scalar conservation laws, showing the existence of a regularizing phenomena as a truly nonlinear effect in hyperbolic equations. Nevertheless, the smoothness properties obtained through such applications have so far fallen short of the expected optimal regularity. In fact, other methods have already succeeded in establishing better results, see [Golse and Perthame 2013], which are sharp. However, one should keep in mind that the versions of velocity averaging lemmas used in these works were not sharp in the first place (for the kinetic formulation under consideration). In fact, it is likely that sharp versions of velocity averaging lemmas would yield sharp regularity properties in conservation laws, when kinetic formulations are available, which would largely expand the possibilities of reaching optimal regularity results in nonlinear conservation laws.

However, such research would require significant efforts and we will therefore not delve any further into this realm of applications, leaving it for subsequent works.

2. The transport operator and dispersive estimates

Let $f(x,v), g(x,v) \in \mathcal{S}(\mathbb{R}^n_x \times \mathbb{R}^n_v)$ ($\mathcal{S}$ denotes the Schwartz space of rapidly decaying functions) be a solution of the transport equation (1-2). Then, introducing some cutoff function $\rho \in \mathcal{S}(\mathbb{R})$ such that $\rho(0) = 1$ and recalling the Fourier inversion formula

$$\rho(r) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{irs} \hat{\rho}(s)\, ds = \int_{\mathbb{R}} e^{-irs} \tilde{\rho}(s)\, ds,$$

where

$$\hat{\rho}(r) = \int_{\mathbb{R}} e^{-irs} \rho(s)\, ds,$$
$$\tilde{\rho}(r) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{irs} \rho(s)\, ds,$$

one can show that

$$f(x,v) = A_t f(x,v) + t B_t g(x,v), \tag{2-1}$$

with

$$A_t f(x,v) = \int_{\mathbb{R}} f(x - stv, v)\, \tilde{\rho}(s)\, ds,$$
$$B_t g(x,v) = \int_{\mathbb{R}} g(x - stv, v) \tilde{\tau}(s)\, ds,$$

where $\tau(s) = (1 - \rho(s))/(is)$ and $t \in \mathbb{R}$ is an interpolation parameter. We refer the reader to [Arsénio and Masmoudi 2014, Section 3] for full details on the derivation of this decomposition formula.

Further considering the Fourier transform in the space variable only,

$$\hat{f}(\eta,v)=\mathcal{F}_x f(\eta,v)=\int_{\mathbb{R}^n} e^{-i\eta\cdot x} f(x,v)\,dx,$$

it holds that

$$\begin{aligned}\mathcal{F}_x A_t f(\eta,v)&=\rho(t\eta\cdot v)\mathcal{F}_x f(\eta,v),\\ \mathcal{F}_x B_t g(\eta,v)&=\tau(t\eta\cdot v)\mathcal{F}_x g(\eta,v).\end{aligned}\tag{2-2}$$

Notice that τ is smooth near the origin, for $\rho(0)=1$, and that

$$\tilde{\tau}(s)=\frac{1}{2\pi}\int_{\mathbb{R}} e^{isr}\frac{1-\rho(r)}{ir}\,dr=\mathbb{1}_{\{s\geq 0\}}-\int_{-\infty}^{s}\tilde{\rho}(\sigma)\,d\sigma.\tag{2-3}$$

In particular, $\tilde{\tau}$ is bounded pointwise and, if $\tilde{\rho}$ is compactly supported, so is $\tilde{\tau}$. Observe, however, that it is not possible to isolate the origin from the support of $\tilde{\tau}$.

Generally speaking, the estimates established in this work clearly apply to both operators A_t and B_t. Nevertheless, for the sake of simplicity, we only formulate our results in terms of the operator A_t. The corresponding results for B_t are easily deduced by replacing ρ by τ.

In this section, we study the dispersive properties of the operators A_t and B_t, which will serve in the proof of Theorem 3.6 below. To this end, we will use the following basic dispersive estimate established in [Castella and Perthame 1996]:

$$\|f(x-tv,v)\|_{L^p_xL^r_v}\leq\frac{1}{|t|^{n\left(\frac{1}{r}-\frac{1}{p}\right)}}\|f(x,v)\|_{L^r_xL^p_v}\quad\text{for all } 1\leq r\leq p\leq\infty.\tag{2-4}$$

Our first dispersive estimate on A_t below is an elementary application of (2-4) to the operator A_t. We will not make any direct use of this simple result later on. It does, however, provide some insight into the dispersive properties of A_t and, therefore, we list it here for completeness.

Proposition 2.1. *For any given* $1\leq r\leq p\leq\infty$, *the operator* A_t *satisfies the estimate*

$$\|A_tf\|_{L^p_xL^r_v}\leq\frac{1}{|t|^{n\left(\frac{1}{r}-\frac{1}{p}\right)}}\left\|\frac{\tilde{\rho}(s)}{s^{n\left(\frac{1}{r}-\frac{1}{p}\right)}}\right\|_{L^1}\|f\|_{L^r_xL^p_v}$$

for all $t\neq 0$.

Proof. This result is a simple extension of the standard dispersive estimate (2-4). Indeed, we have

$$\|A_tf\|_{L^p_xL^r_v}\leq\int_{\mathbb{R}}\|f(x-stv,v)\|_{L^p_xL^r_v}|\tilde{\rho}(s)|\,ds\leq\int_{\mathbb{R}}\frac{1}{|st|^{n\left(\frac{1}{r}-\frac{1}{p}\right)}}\|f(x,v)\|_{L^r_xL^p_v}|\tilde{\rho}(s)|\,ds.\qquad\square$$

Next, combining the dispersion of the free transport flow (2-4) with a TT^*-argument yields the following proposition.

Proposition 2.2. *Let* $2\leq a\leq\infty$, $2<q\leq\infty$ *and* $1\leq r\leq p\leq\infty$ *be such that*

$$\frac{2}{a}=\frac{1}{p}+\frac{1}{r}\quad\textit{and}\quad n\left(\frac{1}{r}-\frac{1}{p}\right)=\frac{2}{q}.$$

Then, the operator A_t *satisfies the estimate*

$$\|A_t f\|_{L^a_{x,v}} \le \frac{C}{|t|^{\frac{1}{q}}} \|\tilde{\rho}\|_{L^{q'}} \|f\|_{L^r_x L^p_v}$$

for all $t \neq 0$, *where* $C > 0$ *only depends on* q.

Proof. First of all, notice that the case $q = \infty$, so that $a = p = r$, is obvious, with a constant $C = 1$. We may therefore assume, without any loss of generality, that $q < \infty$.

Thus, we estimate, using the dispersion (2-4),

$$\begin{aligned}
\|A_t f\|^2_{L^a_{x,v}} &= \||A_t f|^2\|_{L^{a/2}_{x,v}} \\
&= \left\| \int_{\mathbb{R}\times\mathbb{R}} f(x - stv, v)\, \tilde{\rho}(s)\, f(x - \sigma tv, v)\, \tilde{\rho}(\sigma)\, ds\, d\sigma \right\|_{L^{a/2}_{x,v}} \\
&\le \int_{\mathbb{R}\times\mathbb{R}} \|f(x, v) f(x - (\sigma - s)tv, v)\|_{L^{a/2}_{x,v}} |\tilde{\rho}(s)\tilde{\rho}(\sigma)|\, ds\, d\sigma \\
&\le \int_{\mathbb{R}\times\mathbb{R}} \|f(x, v)\|_{L^r_x L^p_v} \|f(x - (\sigma - s)tv, v)\|_{L^p_x L^r_v} |\tilde{\rho}(s)\, \tilde{\rho}(\sigma)|\, ds\, d\sigma \\
&\le \frac{1}{|t|^{\frac{2}{q}}} \|f\|^2_{L^r_x L^p_v} \int_{\mathbb{R}\times\mathbb{R}} \frac{1}{|\sigma - s|^{\frac{2}{q}}} |\tilde{\rho}(s)\, \tilde{\rho}(\sigma)|\, ds\, d\sigma.
\end{aligned}$$

Hence, by virtue of the Hardy–Littlewood–Sobolev inequality,

$$\|A_t f\|_{L^a_{x,v}} \le \frac{C}{|t|^{\frac{1}{q}}} \|\tilde{\rho}\|_{L^{q'}} \|f\|_{L^r_x L^p_v},$$

where $C > 0$ only depends on q. □

The preceding proposition only accepts parameters in the range $2 \le a \le \infty$. The next proposition handles the range $1 \le a \le 2$. It is obtained by interpolating the estimate from the preceding proposition with the degenerate L^1 case. Figure 1 represents the range of validity of the parameters $\frac{1}{p}$ and $\frac{1}{r}$ for both Propositions 2.2 and 2.3. More precisely, the shaded region therein delimited by the points $(0, 0)$, $\left(\frac{1}{n}, 0\right)$, $\left(\frac{n+1}{2n}, \frac{n-1}{2n}\right)$ and $\left(\frac{1}{2}, \frac{1}{2}\right)$ is handled by Proposition 2.2, whereas the shaded region bounded by the points $\left(\frac{1}{2}, \frac{1}{2}\right)$, $\left(\frac{n+1}{2n}, \frac{n-1}{2n}\right)$ and $(1, 1)$ concerns Proposition 2.3.

Proposition 2.3. *Let* $1 \le a \le 2$, $a' < q \le \infty$ *and* $1 \le r \le p \le \infty$ *be such that*

$$\frac{2}{a} = \frac{1}{p} + \frac{1}{r} \quad \text{and} \quad n\Big(\frac{1}{r} - \frac{1}{p}\Big) = \frac{2}{q}.$$

Then, the operator A_t *satisfies the estimate*

$$\|A_t f\|_{L^a_{x,v}} \le \frac{C}{|t|^{\frac{1}{q}}} \|\tilde{\rho}\|_{L^{q'}} \|f\|_{L^r_x L^p_v}$$

for all $t \neq 0$, *where* $C > 0$ *only depends on* q *and* a.

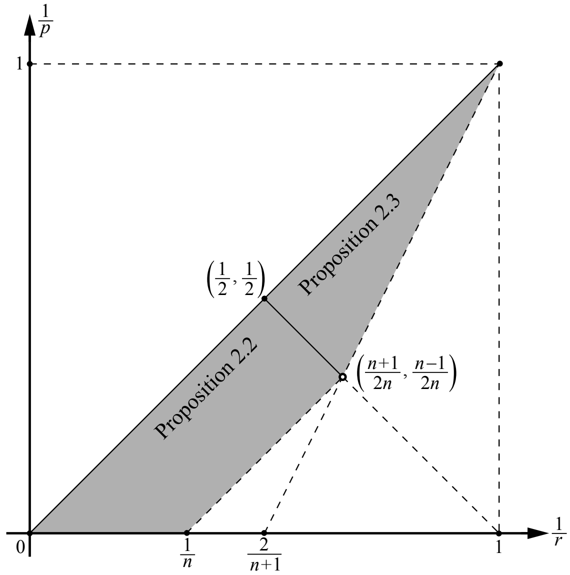

Figure 1. Range of validity of the parameters $\frac{1}{r}$ and $\frac{1}{p}$ in Propositions 2.2 and 2.3.

Proof. This result will follow from the interpolation of the case $a = 2$ from Proposition 2.2 and the trivial estimate

$$\|A_t f\|_{L^1_{x,v}} \le \|\tilde\rho\|_{L^1} \|f\|_{L^1_{x,v}}. \tag{2-5}$$

Thus, without any loss of generality, we assume that $1 < a < 2$ and we define $0 < \theta < 1,\ 2 < q_1 \le \infty$ and $1 \le r_1 \le p_1 \le \infty$ by

$$\theta = \frac{2}{a'} = \frac{1}{p'} + \frac{1}{r'}, \qquad p_1 = \theta r', \qquad r_1 = \theta p', \qquad q_1 = \theta q,$$

so that

$$\frac{1}{a} = \frac{1-\theta}{1} + \frac{\theta}{2}, \quad \frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{\infty} + \frac{\theta}{q_1}, \quad \frac{1}{r} = \frac{1-\theta}{1} + \frac{\theta}{r_1},$$

and

$$1 = \frac{1}{p_1} + \frac{1}{r_1} \quad \text{and} \quad n\Big(\frac{1}{r_1} - \frac{1}{p_1}\Big) = \frac{2}{q_1}.$$

In particular, notice that, since $\frac{2}{q_1} < 1$, necessarily $1 < r_1 \le p_1 < \infty$. On Figure 1, the point $\left(\frac{1}{r_1}, \frac{1}{p_1}\right)$ lies somewhere on the half open segment $\left[\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{n+1}{2n}, \frac{n-1}{2n}\right)\right)$.

It follows then from Proposition 2.2 that

$$\|A_t f\|_{L^2_{x,v}} \le \frac{C}{|t|^{\frac{1}{q_1}}} \|\tilde\rho\|_{L^{q_1'}} \|f\|_{L^{r_1}_x L^{p_1}_v}, \tag{2-6}$$

where $C > 0$ only depends on q_1.

Now, standard results from complex interpolation theory of Lebesgue spaces, see [Bergh and Löfström 1976, Section 5.1], establish that

$$(L^1_{x,v}, L^2_{x,v})_{[\theta]} = L^a_{x,v}, \quad (L^1_{x,v}, L^{r_1}_x L^{p_1}_v)_{[\theta]} = L^r_x L^p_v \quad \text{and} \quad (L^1, L^{q_1'})_{[\theta]} = L^{q'}.$$

Therefore, interpolating estimates (2-5) and (2-6) (these estimates remain valid for complex-valued functions), which are multilinear in $\tilde{\rho}$ and f (use the multilinear complex interpolation theorem [Bergh and Löfström 1976, Theorem 4.4.1]), we arrive at

$$\|A_t f\|_{L^a_{x,v}} \le \frac{C}{|t|^{\frac{1}{q}}} \|\tilde{\rho}\|_{L^{q'}} \|f\|_{L^r_x L^p_v},$$

where $C > 0$ only depends on q and a. □

Note that the adjoint operator of A_t satisfies $A_t^* = A_{-t}$. Combining Propositions 2.2 and 2.3 with a duality argument yields the following result.

Proposition 2.4. *Let* $1 \le a \le \infty$, $\max\{2, a\} < q \le \infty$ *and* $1 \le r \le p \le \infty$ *be such that*

$$\frac{2}{a} = \frac{1}{p} + \frac{1}{r} \quad \text{and} \quad n\Big(\frac{1}{r} - \frac{1}{p}\Big) = \frac{2}{q}.$$

Then, the operator A_t *satisfies the estimate*

$$\|A_t f\|_{L^p_x L^r_v} \le \frac{C}{|t|^{\frac{1}{q}}} \|\tilde{\rho}\|_{L^{q'}} \|f\|_{L^a_{x,v}}$$

for all $t \neq 0$, *where* $C > 0$ *only depends on* q *and* a.

Proof. This result easily follows from a duality argument. Indeed, by Proposition 2.2 (if $1 \le a \le 2$) or Proposition 2.3 (if $2 \le a \le \infty$), we have

$$\begin{aligned}\left|\int_{\mathbb{R}^n\times\mathbb{R}^n} A_t f(x,v) g(x,v)\,dx\,dv\right| &= \left|\int_{\mathbb{R}^n\times\mathbb{R}^n} f(x,v) A_{-t} g(x,v)\,dx\,dv\right| \\ &\le \|f\|_{L^a_{x,v}} \|A_{-t} g\|_{L^{a'}_{x,v}} \\ &\le \|f\|_{L^a_{x,v}} \frac{C}{|t|^{\frac{1}{q}}} \|\tilde{\rho}\|_{L^{q'}} \|g\|_{L^{p'}_x L^{r'}_v},\end{aligned}$$

where $C > 0$ only depends on q and a. Then, taking the supremum over all $g \in L^{p'}_x L^{r'}_v$ easily concludes the proof of the proposition. □

3. Dispersion and velocity averaging

We proceed now to combining the dispersive estimates from the previous section with the classical regularizing effects due to velocity averaging. This will eventually lead to our first main result Theorem 3.6.

To this end, we consider, for any $t \neq 0$ and $\varphi(v) \in L^\infty_c(\mathbb{R}^n)$, the velocity averaging operator T_t defined, for all $f(x,v) \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$, by

$$T_t f(x) = \int_{\mathbb{R}^n} A_t f(x,v)\,\varphi(v)\,dv.$$

In particular, for all $g(x) \in \mathcal{S}(\mathbb{R}^n)$, one has, by duality,

$$\int_{\mathbb{R}^n} T_t f(x) g(x)\,dx = \int_{\mathbb{R}^n\times\mathbb{R}^n} f(x,v) T_t^* g(x,v)\,dx\,dv,$$

where the adjoint operator T_t^* is defined by

$$T_t^* g(x,v) = A_{-t}(g(x)\,\varphi(v)) = \int_{\mathbb{R}} g(x+stv)\,\tilde{\rho}(s)\,ds\,\varphi(v).$$

We will also consider the operators T_t and T_t^* defined with B_t instead of A_t.

For clarity, throughout this section, we will always consider the same given velocity weight $\varphi(v) \in L_c^\infty(\mathbb{R}^n)$ and we will assume that its support is contained inside a closed ball of radius $R > 0$ centered at the origin.

We begin by applying the classical Hilbertian methods of velocity averaging from [Golse et al. 1988] to the operator T_t and its adjoint T_t^*. The resulting estimates are recorded in the following proposition. For the sake of completeness and convenience of the reader, we provide a complete justification of these results.

Proposition 3.1. *The operator T_t and its adjoint T_t^* satisfy the estimates*

$$\|(1-\Delta_x)^{\frac14} T_t f\|_{L^2_x} \le C\left(\|\rho\|_{L^\infty}\,\|\varphi\|_{L^2_v} + \frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho\|_{L^2}\,\|\varphi\|_{L^\infty_v}\right)\|f\|_{L^2_{x,v}},$$

$$\|(1-\Delta_x)^{\frac14} T_t^* g\|_{L^2_{x,v}} \le C\left(\|\rho\|_{L^\infty}\,\|\varphi\|_{L^2_v} + \frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho\|_{L^2}\,\|\varphi\|_{L^\infty_v}\right)\|g\|_{L^2_x}$$

for all $t \neq 0$, where $C > 0$ only depends on the dimension.

Proof. We deal with the estimate on the adjoint operator T_t^* first. Thus, it is readily seen, by Plancherel's theorem and using the Fourier representation (2-2) of A_t, that

$$\|T_t^* g\|_{L^2_{x,v}} = \frac{1}{(2\pi)^{\frac n2}}\|\rho(-t\eta\cdot v)\hat{g}(\eta)\varphi(v)\|_{L^2_{\eta,v}} \le \|\rho\|_{L^\infty}\,\|g\|_{L^2_x}\,\|\varphi\|_{L^2_v}. \tag{3-1}$$

Furthermore, using again Plancherel's theorem, we find that

$$\begin{aligned}\|(-\Delta_x)^{\frac14} T_t^* g\|_{L^2_{x,v}} &= \frac{1}{(2\pi)^{\frac n2}}\big\||\eta|^{\frac12}\rho(-t\eta\cdot v)\hat{g}(\eta)\varphi(v)\big\|_{L^2_{\eta,v}} \\ &= \frac{1}{(2\pi)^{\frac n2}}\left\|\rho\left(-t\frac{\eta}{|\eta|}\cdot v\right)\hat{g}(\eta)\varphi\left(\frac{v}{|\eta|}\cdot\frac{\eta}{|\eta|}\frac{\eta}{|\eta|} + \left(v - v\cdot\frac{\eta}{|\eta|}\frac{\eta}{|\eta|}\right)\right)\right\|_{L^2_{\eta,v}},\end{aligned}$$

where we have rescaled the variable v by a factor $|\eta|$ in the direction $\frac{\eta}{|\eta|}$ only.

Then, writing $v' = (v_2, \dots, v_n)$ and recalling that the support of φ is contained in a closed ball of radius $R > 0$ centered at the origin, we deduce that

$$\begin{aligned}\|(-\Delta_x)^{\frac14} T_t^* g\|_{L^2_{x,v}} &\le \|\rho(-tv_1)\mathbb{1}_{\{|v'|\le R\}}\|_{L^2_v}\,\|g\|_{L^2_x}\,\|\varphi\|_{L^\infty_v} \\ &= \left(\frac{|\mathbb{S}^{n-2}|}{n-1}\right)^{\frac12}\frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho(s)\|_{L^2_s}\,\|g\|_{L^2_x}\,\|\varphi\|_{L^\infty_v}.\end{aligned} \tag{3-2}$$

Finally, combining estimates (3-1) and (3-2) establishes the estimate on T_t^*.

The estimate on T_t is then easily deduced from the estimate on T_t^* by a duality argument. ☐

Interpolating the preceding result with the degenerate L^1 case yields the following proposition.

Proposition 3.2. *For any given $1 \le a \le 2$, the operator T_t and its adjoint T_t^* satisfy the estimates*

$$\|(1-\Delta_x)^{\frac{s}{2}} T_t f\|_{L^a_x} \le C(\|\tilde\rho\|_{L^1}\|\varphi\|_{L^\infty_v})^{1-2s}\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v} + \frac{R^{\frac{n-1}{2}}}{|t|^{\frac{1}{2}}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)^{2s}\|f\|_{L^a_{x,v}},$$

$$\|(1-\Delta_x)^{\frac{s}{2}} T_t^* g\|_{L^a_{x,v}} \le C(\|\tilde\rho\|_{L^1}\|\varphi\|_{L^1_v})^{1-2s}\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v} + \frac{R^{\frac{n-1}{2}}}{|t|^{\frac{1}{2}}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)^{2s}\|g\|_{L^a_x}$$

for all $t \neq 0$, where $s = 1 - \frac{1}{a}$ and $C > 0$ only depends on a and the dimension.

Proof. We deal with the estimate on T_t first.

It was established in [DeVore and Petrova 2001, Theorem 3.2] that the real interpolation space $(L^1, H^{\frac{1}{2}})_{2s,a}$, where $s = 1 - \frac{1}{a}$ with $1 < a < 2$, is precisely the Besov space $B^s_{a,a}$, which is continuously embedded into the classical fractional Sobolev space $W^{s,a}$; that is,

$$(L^1_x, H^{\frac{1}{2}}_x)_{2s,a} \subset W^{s,a}_x.$$

Note that it would be possible to formulate a better result by using below the smaller Besov space $B^s_{a,a}$, as in [DeVore and Petrova 2001]. However, for the sake of simplicity, we choose not to do so and stick to Sobolev spaces. We refer to [DeVore and Petrova 2001] for more details on this.

Next, it is well known from the real interpolation theory of Lebesgue spaces, see [Bergh and Löfström 1976, Theorem 5.2.1], that

$$(L^1_{x,v}, L^2_{x,v})_{2s,a} = L^a_{x,v}.$$

Therefore, the first part of this result easily follows from the real interpolation of the classical estimate on T_t from Proposition 3.1 with the case $p = 1$ of the simple estimate

$$\|T_t f\|_{L^p_x} \le \|\tilde\rho\|_{L^1}\|\varphi\|_{L^{p'}_v}\|f\|_{L^p_{x,v}},$$

valid for any $1 \le p \le \infty$.

There only remains to establish the estimate on the adjoint operator T_t^*. To this end, note that T_t^* commutes with the differentiation in x so that the estimate on T_t^* from Proposition 3.1 can be recast as

$$\|T_t^* g\|_{L^2_{x,v}} \le C\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v} + \frac{R^{\frac{n-1}{2}}}{|t|^{\frac{1}{2}}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)\|(1-\Delta_x)^{-\frac{1}{4}} g\|_{L^2_x},$$

where $C > 0$ only depends on the dimension.

We wish now to complex interpolate the preceding estimate with the elementary control

$$\|T_t^* g\|_{L^1_{x,v}} \le \|\tilde\rho\|_{L^1}\|\varphi\|_{L^1_v}\|g\|_{L^1_x} \le C\|\tilde\rho\|_{L^1}\|\varphi\|_{L^1_v}\|g\|_{h^1_x},$$

where h^1 denotes the local Hardy space; see [Runst and Sickel 1996, Section 2.1.2] for a definition.

To this end, we use the results from complex interpolation theory, see [Bergh and Löfström 1976, Theorem 5.1.1; Runst and Sickel 1996, Section 2.5.2],

$$(L^1_{x,v}, L^2_{x,v})_{[2s]} = L^a_{x,v} \quad \text{and} \quad (h^1_x, H^{-\frac{1}{2}}_x)_{[2s]} = W^{-s,a}_x,$$

to deduce that

$$\|T_t^* g\|_{L^a_{x,v}} \leq C(\|\tilde\rho\|_{L^1}\|\varphi\|_{L^1_v})^{1-2s}\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v}+\frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)^{2s}\|(1-\Delta_x)^{-\frac{s}{2}}g\|_{L^a_x}.$$

Finally, we easily conclude the proof of the proposition by using again that T_t^* commutes with the differentiation in x and replacing g by $(1-\Delta_x)^{\frac{s}{2}}g$ in the above estimate. □

Combining now the preceding proposition with a duality argument yields the following result.

Proposition 3.3. *For any given $2 \leq a \leq \infty$, the operator T_t and its adjoint T_t^* satisfy the estimates*

$$\|(1-\Delta_x)^{\frac{s}{2}}T_t f\|_{L^a_x} \leq C(\|\tilde\rho\|_{L^1}\|\varphi\|_{L^1_v})^{1-2s}\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v}+\frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)^{2s}\|f\|_{L^a_{x,v}},$$

$$\|(1-\Delta_x)^{\frac{s}{2}}T_t^* g\|_{L^a_{x,v}} \leq C(\|\tilde\rho\|_{L^1}\|\varphi\|_{L^\infty_v})^{1-2s}\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v}+\frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)^{2s}\|g\|_{L^a_x}$$

for all $t \neq 0$, where $s = \frac{1}{a}$ and $C > 0$ only depends on a and the dimension.

Proof. These estimates follow straightforwardly from Proposition 3.2 through a duality argument.

Indeed, by Proposition 3.2, noticing that both T_t and T_t^* commute with differentiation in x, it holds that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n}(1-\Delta_x)^{\frac{s}{2}}T_t f(x)g(x)\,dx\right| \\
&\qquad = \left|\int_{\mathbb{R}^n\times\mathbb{R}^n} f(x,v)(1-\Delta_x)^{\frac{s}{2}}T_t^* g(x,v)\,dx\,dv\right| \\
&\qquad \leq \|f\|_{L^a_{x,v}}\|(1-\Delta_x)^{\frac{s}{2}}T_t^* g\|_{L^{a'}_{x,v}} \\
&\qquad \leq C(\|\tilde\rho\|_{L^1}\|\varphi\|_{L^1_v})^{1-2s}\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v}+\frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)^{2s}\|f\|_{L^a_{x,v}}\|g\|_{L^{a'}_x},
\end{aligned}$$

where $C > 0$ only depends on a and the dimension. Then, taking the supremum over all $g \in L^{a'}_x$ easily concludes the proof of the first estimate on T_t.

Similarly, using Proposition 3.2 again, it holds that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n\times\mathbb{R}^n}(1-\Delta_x)^{\frac{s}{2}}T_t^* g(x,v)f(x,v)\,dx\,dv\right| \\
&\qquad = \left|\int_{\mathbb{R}^n} g(x)(1-\Delta_x)^{\frac{s}{2}}T_t f(x)\,dx\right| \\
&\qquad \leq \|g\|_{L^a_x}\|(1-\Delta_x)^{\frac{s}{2}}T_t f\|_{L^{a'}_x} \\
&\qquad \leq C(\|\tilde\rho\|_{L^1}\|\varphi\|_{L^\infty_v})^{1-2s}\left(\|\rho\|_{L^\infty}\|\varphi\|_{L^2_v}+\frac{R^{\frac{n-1}{2}}}{|t|^{\frac12}}\|\rho\|_{L^2}\|\varphi\|_{L^\infty_v}\right)^{2s}\|g\|_{L^a_x}\|f\|_{L^{a'}_{x,v}},
\end{aligned}$$

where $C > 0$ only depends on a and the dimension. Finally, taking the supremum over all $f \in L^{a'}_{x,v}$ easily concludes the proof of the proposition. □

From now on, in this section, we assume that the cutoff function $\rho(r)$ may be decomposed as a product $\rho(r)=\rho_1(r)\rho_2(r)$, so that $\tilde\rho(s)=\tilde\rho_1*\tilde\rho_2(s)$. Naturally, we will denote by A_t^i, T_t^i and T_t^{i*}, where $i=1,2$, the respective operators A_t, T_t and T_t^* where we replace the cutoff ρ by ρ_i. It is then readily seen that

$$A_t=A_t^1A_t^2=A_t^2A_t^1,\quad T_t=T_t^1A_t^2=T_t^2A_t^1,\quad T_t^*=A_{-t}^2T_t^{1*}=A_{-t}^1T_t^{2*}. \tag{3-3}$$

As shown below, this useful trick allows us to combine the previous regularity results from this section with the dispersive estimates from Section 2 to obtain new estimates on the operators T_t and T_t^*.

Proposition 3.4. *Let* $1\le a\le\infty$, $\max\{2,a\}<q\le\infty$ *and* $1\le r\le p\le\infty$ *be such that*

$$\frac{2}{a}=\frac{1}{p}+\frac{1}{r}\quad\textit{and}\quad n\Big(\frac{1}{r}-\frac{1}{p}\Big)=\frac{2}{q}.$$

Then, the operator T_t^* *satisfies the estimate*

$$\|(1-\Delta_x)^{\frac{s}{2}}T_t^*g\|_{L_x^pL_v^r}\le\frac{C}{|t|^{\frac{1}{q}+s}}\|\tilde\rho_1\|_{L^{q'}}\|\tilde\rho_2\|_{L^1}^{1-2s}(\|\rho_2\|_{L^\infty}+\|\rho_2\|_{L^2})^{2s}\|g\|_{L_x^a}$$

for all $t\neq0$ *such that* $|t|\le1$, *where* $s=\min\{1-\frac{1}{a},\frac{1}{a}\}$ *and* $C>0$ *only depends on* q, a, φ *and the dimension.*

Proof. We treat the case $1\le a\le2$ first, so that $q>2$ and $s=1-\frac{1}{a}$. Writing $T_t^*=A_{-t}^1T_t^{2*}$ and then successively employing Propositions 2.4 and 3.2, we find, noticing A_{-t}^1 commutes with differentiation in x, that

$$\begin{aligned}
&\|(1-\Delta_x)^{\frac{s}{2}}T_t^*g\|_{L_x^pL_v^r}\\
&\qquad=\|A_{-t}^1(1-\Delta_x)^{\frac{s}{2}}T_t^{2*}g\|_{L_x^pL_v^r}\\
&\qquad\le\frac{C}{|t|^{\frac{1}{q}}}\|\tilde\rho_1\|_{L^{q'}}\|(1-\Delta_x)^{\frac{s}{2}}T_t^{2*}g\|_{L_{x,v}^a}\\
&\qquad\le\frac{C}{|t|^{\frac{1}{q}}}\|\tilde\rho_1\|_{L^{q'}}(\|\tilde\rho_2\|_{L^1}\|\varphi\|_{L_v^1})^{1-2s}\left(\|\rho_2\|_{L^\infty}\|\varphi\|_{L_v^2}+\frac{R^{\frac{n-1}{2}}}{|t|^{\frac{1}{2}}}\|\rho_2\|_{L^2}\|\varphi\|_{L_v^\infty}\right)^{2s}\|g\|_{L_x^a},
\end{aligned}$$

where $C>0$ only depends on q, a and the dimension. Since $|t|\le1$, this concludes the proof of the proposition when $a\le2$.

The case $a\ge2$ is handled similarly. One now has that $q>a$ and $s=\frac{1}{a}$. Therefore, applying successively Propositions 2.4 and 3.3, we find that

$$\begin{aligned}
&\|(1-\Delta_x)^{\frac{s}{2}}T_t^*g\|_{L_x^pL_v^r}\\
&\qquad=\|A_{-t}^1(1-\Delta_x)^{\frac{s}{2}}T_t^{2*}g\|_{L_x^pL_v^r}\\
&\qquad\le\frac{C}{|t|^{\frac{1}{q}}}\|\tilde\rho_1\|_{L^{q'}}\|(1-\Delta_x)^{\frac{s}{2}}T_t^{2*}g\|_{L_{x,v}^a}\\
&\qquad\le\frac{C}{|t|^{\frac{1}{q}}}\|\tilde\rho_1\|_{L^{q'}}(\|\tilde\rho_2\|_{L^1}\|\varphi\|_{L_v^\infty})^{1-2s}\left(\|\rho_2\|_{L^\infty}\|\varphi\|_{L_v^2}+\frac{R^{\frac{n-1}{2}}}{|t|^{\frac{1}{2}}}\|\rho_2\|_{L^2}\|\varphi\|_{L_v^\infty}\right)^{2s}\|g\|_{L_x^a},
\end{aligned}$$

where $C > 0$ only depend on q, a and the dimension. Since $|t| \leq 1$, this concludes the proof of the proposition. ☐

Combining the previous result with a duality argument yields estimates on the operator T_t, which are contained in the following proposition.

Proposition 3.5. *Let* $1 \leq a \leq \infty$, $\max\{2, a'\} < q \leq \infty$ *and* $1 \leq r \leq p \leq \infty$ *be such that*

$$\frac{2}{a} = \frac{1}{p} + \frac{1}{r} \quad \text{and} \quad n\Big(\frac{1}{r} - \frac{1}{p}\Big) = \frac{2}{q}.$$

Then, the operator T_t satisfies the estimate

$$\|(1-\Delta_x)^{\frac{s}{2}} T_t f\|_{L^a_x} \leq \frac{C}{|t|^{\frac{1}{q}+s}} \|\tilde{\rho}_1\|_{L^{q'}} \|\tilde{\rho}_2\|_{L^1}^{1-2s} (\|\rho_2\|_{L^\infty} + \|\rho_2\|_{L^2})^{2s} \|f\|_{L^r_x L^p_v} \tag{3-4}$$

for all $t \neq 0$ such that $|t| \leq 1$, where $s = \min\{1 - \frac{1}{a}, \frac{1}{a}\}$ and $C > 0$ only depends on q, a, φ and the dimension.

Proof. This estimate follows straightforwardly from Proposition 3.4 through a duality argument.

Indeed, using Proposition 3.4, we find, since T_t^* commutes with differentiation in x, that

$$\begin{aligned}\left|\int_{\mathbb{R}^n} (1-\Delta_x)^{\frac{s}{2}} T_t f(x) g(x)\, dx\right| &= \left|\int_{\mathbb{R}^n\times\mathbb{R}^n} f(x,v)(1-\Delta_x)^{\frac{s}{2}} T_t^* g(x,v)\, dx\, dv\right| \\ &\leq \|f\|_{L^r_x L^p_v} \|(1-\Delta_x)^{\frac{s}{2}} T_t^* g\|_{L^{r'}_x L^{p'}_v} \\ &\leq \frac{C}{|t|^{\frac{1}{q}+s}} \|\tilde{\rho}_1\|_{L^{q'}} \|\tilde{\rho}_2\|_{L^1}^{1-2s} (\|\rho_2\|_{L^\infty} + \|\rho_2\|_{L^2})^{2s} \|f\|_{L^r_x L^p_v} \|g\|_{L^{a'}_x},\end{aligned}$$

where $C > 0$ only depends on q, a, φ and the dimension. Finally, taking the supremum over all $g \in L^{a'}_x$ easily concludes the proof of the proposition.

Note that, in order to deduce this result, we could just as well have applied here a combination of Propositions 2.2 and 2.3 with Propositions 3.3 and 3.2, respectively. ☐

We proceed now to the main theorem of this section. It contains Theorem 1.2 presented in the Introduction as special case (corresponding to the case $a = 2$ below) and provides a considerable extension of the classical velocity averaging lemma in $L^2_{x,v}$ (corresponding to the case $a = p = r = 2$ below). Indeed, observe that the case $a = 2$ therein yields a maximal gain of regularity of half a derivative on velocity averages for a variety of parameters, which was previously known to occur only in the classical $L^2_{x,v}$ setting.

Theorem 3.6. *In any dimension $n \geq 1$, let $1 \leq a \leq \infty$, $\max\{2, a'\} < q \leq \infty$ and $1 \leq r \leq p \leq \infty$ be such that*

$$\frac{2}{a} = \frac{1}{p} + \frac{1}{r} \quad \text{and} \quad n\Big(\frac{1}{r} - \frac{1}{p}\Big) = \frac{2}{q}.$$

Then, for any $f, g \in L^r_x(\mathbb{R}^n; L^p_v(\mathbb{R}^n))$ such that (1-2) *holds true, one has*

$$\int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\, dv \in W^{s,a}_x(\mathbb{R}^n)$$

for any $\varphi \in L^\infty_c(\mathbb{R}^n)$, *where* $s = \min\left\{1-\frac{1}{a}, \frac{1}{a}\right\}$. *Furthermore, one has the estimate*

$$\left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^n} f\varphi\, dv \right\|_{L^a_x} \le C(\|f\|_{L^r_x L^p_v} + \|g\|_{L^r_x L^p_v}),$$

where $C > 0$ *only depends on* φ *and constant parameters.*

Proof. We consider first the refined interpolation formula (2-1), which is valid for solutions of the transport equation (1-2), for some given cutoff $\rho \in \mathcal{S}(\mathbb{R})$. Clearly, further differentiating (2-1) in x and then averaging in v yields

$$\left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^n} f\varphi\, dv \right\|_{L^a_x} \le \left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^n} A_t f\varphi\, dv \right\|_{L^a_x} + t \left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^n} B_t g\varphi\, dv \right\|_{L^a_x}. \tag{3-5}$$

We wish now to apply Proposition 3.5 to the preceding estimate. To this end, according to (3-3), we take the decompositions

$$\rho(r) = \rho_1(r)\rho_2(r) \quad \text{and} \quad \tau(r) = \frac{1-\rho(r)}{ir} = \tau_1(r)\tau_2(r),$$

where

$$\rho_1(r) = \frac{1}{(1+r^2)^{\frac{\beta}{2}}}, \qquad \tau_1(r) = \frac{1}{(1+r^2)^{\frac{\beta}{2}}},$$
$$\rho_2(r) = (1+r^2)^{\frac{\beta}{2}}\rho(r), \quad \tau_2(r) = (1+r^2)^{\frac{\beta}{2}}\tau(r)$$

for some fixed $\frac{1}{q} < \beta < \frac{1}{2}$. In view of the technical Lemma B.1 from Appendix B, it then holds that

$$\tilde\rho_1, \tilde\tau_1 \in L^{q'}, \quad \tilde\rho_2, \tilde\tau_2 \in L^1 \quad \text{and} \quad \rho_2, \tau_2 \in L^\infty \cap L^2.$$

All constants involving norms of the cutoff functions ρ_1, ρ_2, τ_1 and τ_2 in the right-hand side of (3-4) are therefore finite.

Thus, applying Proposition 3.5 to estimate (3-5), we conclude, for any $0 < t < 1$, that

$$\left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^n} f\varphi\, dv \right\|_{L^a_x} \le C\left(\frac{1}{t^{\frac{1}{q}+s}} \|f\|_{L^r_x L^p_v} + t^{1-(\frac{1}{q}+s)} \|g\|_{L^r_x L^p_v} \right),$$

where $C > 0$ only depends on constant parameters. □

4. The one-dimensional case

In the previous section, by combining kinetic dispersion with velocity averaging, we have established, in Theorem 3.6, a whole new range of regularity results on the solutions of the kinetic transport equation (1-2). The results from Theorem 3.6 are valid in any dimension $n \ge 1$. In one dimension ($n = 1$), it turns out that it is possible to obtain more results for a wide range of parameters which are not covered by Theorem 3.6. This is due to the fact that, in one dimension, spatial frequencies are always parallel to velocities.

In the present section, we explore this one-dimensional setting, which provides a good test case for velocity averaging lemmas in mixed Lebesgue spaces and allows one to get familiar with the decompositions used in this work in a much simpler setting. It does not, however, set a road map for the remaining more involved sections concerning higher dimensions, for it heavily relies on the elementary structure of the transport equation in one dimension.

We use the same notation as in the previous sections.

Proposition 4.1. *In dimension $n = 1$, let $1 < p < \infty$, $1 \le r < \infty$, $0 \le s < \frac{1}{r}$ and $\rho \in \mathcal{S}(\mathbb{R})$ be such that $\tilde\rho$ has its support contained inside a ball of radius $r_0 > 0$ centered at the origin.*

Then, the operator T_t^* satisfies the estimate

$$\|(1-\Delta_x)^{\frac{s}{2}} T_t^* g\|_{L^p_x L^r_v} \le \frac{C}{|t|^s} (r_0 \|\tilde\rho\|_{L^\infty})^{1-\frac{1}{r}} (\|\tilde\rho\|_{L^1} + \|(-\Delta)^{\frac{sr}{2}} \tilde\rho\|_{L^1})^{\frac{1}{r}} \|g\|_{L^p_x}$$

for all $t \neq 0$ such that $|t| \le 1$, where $C > 0$ only depends on p, s and φ.

Proof. First, it is readily seen, for any $1 \le r \le p \le \infty$, that

$$\|T_t^* g\|_{L^p_x L^r_v} \le \|T_t^* g\|_{L^r_v L^p_x} \le \|\tilde\rho\|_{L^1} \|\varphi\|_{L^r_v} \|g\|_{L^p_x}. \tag{4-1}$$

When the restriction $r \le p$ is not satisfied, the above estimate fails and we need a more convoluted estimate to handle this case. To this end, we write that

$$\begin{aligned} |T_t^* g(x,v)| &= \left| \int_{\mathbb{R}} g(x+stv)\tilde\rho(s)\, ds \varphi(v) \right| = \left| \frac{1}{tv} \int_{\mathbb{R}} g(x+r)\tilde\rho\left(\frac{r}{tv}\right) dr\, \varphi(v) \right| \\ &\le \|\tilde\rho\|_{L^\infty} \left| \frac{1}{tv} \int_{\{|r| \le |tv| r_0\}} g(x+r)\, dr\, \varphi(v) \right| \le 2r_0 \|\tilde\rho\|_{L^\infty} Mg(x) |\varphi(v)|, \end{aligned}$$

where Mg denotes the Hardy–Littlewood maximal function of g defined by

$$Mg(x) = \sup_{\delta > 0} \frac{1}{2\delta} \int_{\{|y| \le \delta\}} |g(x-y)|\, dy.$$

Recall that the Hardy–Littlewood maximal operator $g \mapsto Mg$ is bounded over $L^p(\mathbb{R})$ for any $1 < p < \infty$, and maps $L^1(\mathbb{R})$ into the standard weak-L^1 space $L^{1,\infty}(\mathbb{R})$; see [Grafakos 2008, Theorem 2.1.6]. It therefore easily follows from the previous estimate that

$$\|T_t^* g\|_{L^{1,\infty}_x L^r_v} \le C r_0 \|\tilde\rho\|_{L^\infty} \|\varphi\|_{L^r_v} \|g\|_{L^1_x}$$

and

$$\|T_t^* g\|_{L^p_x L^r_v} \le C r_0 \|\tilde\rho\|_{L^\infty} \|\varphi\|_{L^r_v} \|g\|_{L^p_x} \tag{4-2}$$

for any $1 < p < \infty$ and $1 \le r \le \infty$, where $C > 0$ only depends on p.

Next, we further compute, exploiting the one-dimensional structure of the operators, for any $0 < \alpha < 1$,

$$\begin{aligned} (-\Delta_x)^{\frac{\alpha}{2}} T_t^* g(x,v) &= \int_{\mathbb{R}} (-\Delta_x)^{\frac{\alpha}{2}} g(x+stv)\tilde\rho(s)\, ds\, \varphi(v) \\ &= \frac{1}{|tv|^\alpha} \int_{\mathbb{R}} (-\Delta_s)^{\frac{\alpha}{2}} (g(x+stv))\tilde\rho(s)\, ds\, \varphi(v) \\ &= \frac{1}{|tv|^\alpha} \int_{\mathbb{R}} g(x+stv)(-\Delta)^{\frac{\alpha}{2}} \tilde\rho(s)\, ds\, \varphi(v), \end{aligned}$$

whence, for any $1 \le p \le \infty$,

$$\|(-\Delta_x)^{\frac{\alpha}{2}} T_t^* g\|_{L^p_x L^1_v} \le \|(-\Delta_x)^{\frac{\alpha}{2}} T_t^* g\|_{L^1_v L^p_x} \le \frac{1}{|t|^\alpha} \|(-\Delta)^{\frac{\alpha}{2}} \tilde\rho\|_{L^1} \left\| \frac{1}{|v|^\alpha} \varphi \right\|_{L^1_v} \|g\|_{L^p_x},$$

which, when combined with (4-1), yields

$$\|(1-\Delta_x)^{\frac{\alpha}{2}} T_t^* g\|_{L^p_x L^1_v} \le C\left(\|\tilde\rho\|_{L^1} \|\varphi\|_{L^1_v} + \frac{1}{|t|^\alpha} \|(-\Delta)^{\frac{\alpha}{2}} \tilde\rho\|_{L^1} \left\| \frac{1}{|v|^\alpha} \varphi \right\|_{L^1_v} \right) \|g\|_{L^p_x}, \tag{4-3}$$

where $C > 0$ only depends on p and α.

We wish now to interpolate the bound (4-2), where we set $r = \infty$, with (4-3). To this end, recalling that T_t^* commutes with differentiation in x, we first recast (4-3) as

$$\|T_t^* g\|_{L^p_x L^1_v} \le C\left(\|\tilde\rho\|_{L^1} \|\varphi\|_{L^1_v} + \frac{1}{|t|^\alpha} \|(-\Delta)^{\frac{\alpha}{2}} \tilde\rho\|_{L^1} \left\| \frac{1}{|v|^\alpha} \varphi \right\|_{L^1_v} \right) \|g\|_{W_x^{-\alpha,p}}, \tag{4-4}$$

and then we use the standard results from complex interpolation theory, see [Bergh and Löfström 1976, Sections 5.1 and 6.4], valid for any $1 < p, r < \infty$,

$$(L^p_x L^\infty_v, L^p_x L^1_v)_{[\frac{1}{r}]} = L^p_x L^r_v \quad \text{and} \quad (L^p_x, W_x^{-\alpha,p})_{[\frac{1}{r}]} = W_x^{-\frac{\alpha}{r},p},$$

to deduce from the interpolation of (4-2) and (4-4) that

$$\|T_t^* g\|_{L^p_x L^r_v} \le C(r_0 \|\tilde\rho\|_{L^\infty} \|\varphi\|_{L^\infty_v})^{1-\frac{1}{r}} \left(\|\tilde\rho\|_{L^1} \|\varphi\|_{L^1_v} + \frac{1}{|t|^\alpha} \|(-\Delta)^{\frac{\alpha}{2}} \tilde\rho\|_{L^1} \left\| \frac{1}{|v|^\alpha} \varphi \right\|_{L^1_v} \right)^{\frac{1}{r}} \|g\|_{W_x^{-\alpha/r,p}},$$

where $C > 0$ only depends on p and α. □

Note that it would be possible to improve the gain of regularity in the preceding proposition by assuming that the support of the velocity weight $\varphi(v)$ does not contain the origin. However, this is a rather unnatural setting which we prefer to avoid here.

Combining the previous result with a duality argument yields estimates on the operator T_t, which are contained in the following proposition.

Proposition 4.2. *In dimension $n = 1$, let $1 < p < \infty$, $1 < r \le \infty$, $0 \le s < 1 - \frac{1}{r}$ and $\rho \in \mathcal{S}(\mathbb{R})$ be such that $\tilde\rho$ has its support contained inside a ball of radius $r_0 > 0$ centered at the origin.*

Then, the operator T_t satisfies the estimate

$$\|(1-\Delta_x)^{\frac{s}{2}} T_t f\|_{L^p_x} \le \frac{C}{|t|^s} (r_0 \|\tilde\rho\|_{L^\infty})^{\frac{1}{r}} (\|\tilde\rho\|_{L^1} + \|(-\Delta)^{\frac{sr'}{2}} \tilde\rho\|_{L^1})^{1-\frac{1}{r}} \|f\|_{L^p_x L^r_v} \tag{4-5}$$

for all $t \neq 0$ such that $|t| \le 1$, where $C > 0$ only depends on p, s and φ.

Proof. This estimate follows straightforwardly from Proposition 4.1 through a duality argument.

Indeed, using Proposition 4.1, we find, since T_t^* commutes with differentiation in x, that

$$\begin{aligned}\left|\int_{\mathbb{R}}(1-\Delta_x)^{\frac{s}{2}}T_t f(x)g(x)\,dx\right| &= \left|\int_{\mathbb{R}\times\mathbb{R}} f(x,v)(1-\Delta_x)^{\frac{s}{2}}T_t^* g(x,v)\,dx\,dv\right| \\ &\le \|f\|_{L_x^p L_v^r}\|(1-\Delta_x)^{\frac{s}{2}}T_t^* g\|_{L_x^{p'}L_v^{r'}} \\ &\le \frac{C}{|t|^s}(r_0\|\tilde\rho\|_{L^\infty})^{\frac{1}{r}}(\|\tilde\rho\|_{L^1}+\|(-\Delta)^{\frac{sr'}{2}}\tilde\rho\|_{L^1})^{1-\frac{1}{r}}\|f\|_{L_x^p L_v^r}\|g\|_{L_x^{p'}},\end{aligned}$$

where $C>0$ only depends on p, s and φ. Finally, taking the supremum over all $g\in L_x^{p'}$ easily concludes the proof of the proposition. □

We proceed now to the main theorem of this section.

Theorem 4.3. *In dimension $n=1$, let $1<p<\infty$ and $1<r\le\infty$.*

Then, for any $f,g\in L_x^p(\mathbb{R};L_v^r(\mathbb{R}))$ such that (1-2) *holds true, one has*

$$\int_{\mathbb{R}} f(x,v)\,\varphi(v)\,dv\in W_x^{s,p}(\mathbb{R})$$

for any $\varphi\in L_c^\infty(\mathbb{R})$ and any $0\le s<1-\frac{1}{r}$. Furthermore, one has the estimate

$$\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}} f\varphi\,dv\right\|_{L_x^p}\le C(\|f\|_{L_x^p L_v^r}+\|g\|_{L_x^p L_v^r}),$$

where $C>0$ only depends on φ and constant parameters.

Proof. We consider first the refined interpolation formula (2-1), which is valid for solutions of the transport equation (1-2) for some given cutoff $\rho\in\mathcal{S}(\mathbb{R})$. Clearly, further differentiating (2-1) in x and then averaging in v yields

$$\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}} f\varphi\,dv\right\|_{L_x^p}\le\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}} A_t f\varphi\,dv\right\|_{L_x^p}+t\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}} B_t g\varphi\,dv\right\|_{L_x^p}. \tag{4-6}$$

We wish now to apply Proposition 4.2 to the preceding estimate. To this end, note that $\tilde\rho$ and all of its derivatives clearly are bounded pointwise and integrable. In order to apply that result, we also further need to ask that $\tilde\rho$ be compactly supported, which is always possible.

Next, in view of (2-3), notice that $\tilde\tau$ also is bounded pointwise, integrable and compactly supported. Therefore, there only remains to check that $(-\Delta)^{\frac{\alpha}{2}}\tilde\tau$ is integrable for any $0<\alpha<1$. This, in fact, easily follows from a direct application of the technical Lemma B.2 from Appendix B to

$$\mathcal{F}[(-\Delta)^{\frac{\alpha}{2}}\tilde\tau]=|r|^\alpha\frac{1-\rho(r)}{ir}.$$

All constants involving norms of the cutoff functions ρ and τ in the right-hand side of (4-5) are therefore finite.

Thus, applying Proposition 4.2 to estimate (4-6), we conclude, for any $0<t<1$, that

$$\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}} f\varphi\,dv\right\|_{L_x^p}\le C\left(\frac{1}{t^s}\|f\|_{L_x^p L_v^r}+t^{1-s}\|g\|_{L_x^p L_v^r}\right),$$

where $C>0$ only depends on constant parameters. □

5. The two-dimensional case

Our study of the one-dimensional case in the previous section showed that it is possible to largely improve the classical velocity averaging results in that setting. In particular, we showed therein that the gain of regularity of velocity averages is, in some cases, substantially improved beyond the value $\frac{1}{2}$.

While such a general improvement is not achievable in higher dimensions ($n \geq 2$), in view of the counterexamples from [DeVore and Petrova 2001, Theorem 1.3] discussed in our Introduction, it is nevertheless possible, as shown below, to obtain new cases displaying a gain of regularity of velocity averages of almost half a derivative.

In two dimensions ($n = 2$), this was already strongly suggested in [Jabin and Vega 2004, Theorem 1.2]. Here, we build upon the work from that paper to obtain refined two-dimensional velocity averaging results displaying an almost maximal gain of regularity of half a derivative. In the next section, we will generalize these methods to higher dimensions ($n \geq 3$), without achieving a gain of half a derivative, though.

We define now, in any dimension $n \geq 1$, the velocity averaging operator on the sphere

$$S_t f(x) = \int_{\mathbb{S}^{n-1}} A_t f(x,v)\,dv = \int_{\mathbb{S}^{n-1}} \int_{\mathbb{R}} f(x-stv,v)\,\tilde{\rho}(s)\,ds\,dv,$$

and its adjoint operator

$$S_t^* g(x,v) = A_{-t}(g(x))(x,v) = \int_{\mathbb{R}} g(x+stv)\,\tilde{\rho}(s)\,ds,$$

so that

$$\int_{\mathbb{R}^n} S_t f(x) g(x)\,dx = \int_{\mathbb{R}^n \times \mathbb{S}^{n-1}} f(x,v) S_t^* g(x,v)\,dx\,dv.$$

We will also consider the operators S_t and S_t^* defined with B_t instead of A_t. These operators correspond to the kinetic transport equation (1-2) with velocities restricted to the sphere $v \in \mathbb{S}^{n-1}$ and are introduced here for mere convenience and simplicity of analysis later on.

This reduction to the sphere is possible here because the regularization phenomenon in the transport equation (1-2) comes from averaging in velocity directions rather than integration along velocity magnitudes. In fact, any bound established on S_t and S_t^* will yield a corresponding bound on T_t and T_t^*, respectively, as shown below in Proposition 5.3 (we also refer the reader to Appendix C for a discussion of the equivalence of velocity averaging lemmas with velocities in the full Euclidean space $\mathbb{R}^n$ and on the sphere $\mathbb{S}^{n-1}$).

Note that this is not true in general. For instance, in the two-dimensional time-dependent setting (1-1) with velocities restricted to the sphere $\mathbb{S}^1$ and $f, g \in L^2_{x,v}$, it was identified in [Bournaveas and Perthame 2001] that only a quarter of a derivative could be gained on the velocity average of f, whereas half a derivative is gained when velocities range in an open subset of $\mathbb{R}^2$ (which corresponds to a three-dimensional setting $(t,x) \in \mathbb{R}^{1+2}$ with velocities restricted to a manifold of dimension 2, much like the stationary case $x \in \mathbb{R}^3$ with $v \in \mathbb{S}^2$).

For completeness, we begin our analysis of the operators S_t and S_t^* by establishing their smoothing effect in L^2 employing the classical Hilbertian methods of velocity averaging from [Golse et al. 1988]. This result is valid in any dimension $n \geq 2$ and will also be used in the next section on higher-dimensional results.

Proposition 5.1. *In any dimension $n \geq 2$, the operator S_t and its adjoint S_t^* satisfy the estimates, for any $0 \leq s \leq \frac{1}{2}$,*

$$\|(1-\Delta_x)^{\frac{s}{2}} S_t f\|_{L^2_x} \leq C\left(\|\rho\|_{L^\infty} + \frac{1}{|t|^s}\left(\left\|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2} + \big\||r|^s\rho(r)\big\|_{L^\infty}\right)\right)\|f\|_{L^2_{x,v}},$$

$$\|(1-\Delta_x)^{\frac{s}{2}} S_t^* g\|_{L^2_{x,v}} \leq C\left(\|\rho\|_{L^\infty} + \frac{1}{|t|^s}\left(\left\|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2} + \big\||r|^s\rho(r)\big\|_{L^\infty}\right)\right)\|g\|_{L^2_x}$$

for all $t \neq 0$, where $C > 0$ only depends on the dimension.

Proof. This proof is almost identical to the general case of Proposition 3.1. Nevertheless, for later applications of this result, it is important to carefully keep track of the dependence of the constants on t and ρ.

We deal with the estimate on the adjoint operator S_t^* first. Thus, it is readily seen, by Plancherel's theorem, that

$$\|S_t^* g\|_{L^2_{x,v}} = \frac{1}{(2\pi)^{\frac{n}{2}}}\|\rho(-t\eta\cdot v)\hat{g}(\eta)\|_{L^2_{\eta,v}} \leq \|\rho\|_{L^\infty}\|g\|_{L^2_x}. \tag{5-1}$$

Furthermore, using again Plancherel's theorem, we find that

$$\begin{aligned}\|(-\Delta_x)^{\frac{s}{2}} T_t^* g\|_{L^2_{x,v}} &= \frac{1}{(2\pi)^{\frac{n}{2}}}\big\||\eta|^s\rho(-t\eta\cdot v)\hat{g}(\eta)\big\|_{L^2_{\eta,v}} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}}\frac{|\mathbb{S}^{n-2}|^{\frac{1}{2}}}{|t|^{\frac{1}{2}}}\left\|\left(|\eta|^{2s-1}\int_{-|t\eta|}^{|t\eta|}|\rho(r)|^2\left(1-\left(\frac{r}{|t\eta|}\right)^2\right)^{\frac{n-3}{2}}dr\right)^{\frac{1}{2}}\hat{g}(\eta)\right\|_{L^2_\eta} \\ &\leq \frac{1}{(2\pi)^{\frac{n}{2}}}\frac{|\mathbb{S}^{n-2}|^{\frac{1}{2}}}{|t|^s}\left\|\left(\int_{-|t\eta|}^{|t\eta|}\left|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right|^2\left(1-\left(\frac{r}{|t\eta|}\right)^2\right)^{\frac{n-3}{2}}dr\right)^{\frac{1}{2}}\hat{g}(\eta)\right\|_{L^2_\eta},\end{aligned}$$

with the convention that $|\mathbb{S}^0| = 2$ when $n = 2$.

Hence, if $n \geq 3$, we easily deduce that

$$\|(-\Delta_x)^{\frac{s}{2}} T_t^* g\|_{L^2_{x,v}} \leq \frac{|\mathbb{S}^{n-2}|^{\frac{1}{2}}}{|t|^s}\left\|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2}\|g(x)\|_{L^2_x}.$$

In the two-dimensional case, the bound on the cutoff ρ is only slightly more involved. We estimate, in this case, for any $N > 0$, that

$$\begin{aligned}\int_{-N}^{N}\left|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right|^2\frac{1}{\left(1-\left(\frac{r}{N}\right)^2\right)^{\frac{1}{2}}}dr &\leq \frac{2}{\sqrt{3}}\left\|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2}^2 + \int_{\frac{N}{2}}^{N}\left(\left|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right|^2 + \left|\frac{\rho(-r)}{|r|^{\frac{1}{2}-s}}\right|^2\right)\frac{1}{\left(1-\left(\frac{r}{N}\right)^2\right)^{\frac{1}{2}}}dr \\ &\leq \frac{2}{\sqrt{3}}\left\|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2}^2 + \frac{2}{3}\pi N\sup_{\frac{N}{2}\leq|r|\leq N}\left|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right|^2.\end{aligned}$$

It therefore follows that, in any dimension $n \geq 2$,

$$\|(-\Delta_x)^{\frac{s}{2}} S_t^* g\|_{L^2_{x,v}} \leq \frac{C}{|t|^s}\left(\left\|\frac{\rho(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2} + \left\||r|^s\rho(r)\right\|_{L^\infty}\right)\|g\|_{L^2_x}. \tag{5-2}$$

Finally, combining estimates (5-1) and (5-2) establishes the estimate on S_t^*.

The estimate on S_t is then easily deduced from the estimate on S_t^* by a duality argument. □

At this stage, we need to further introduce a classical Littlewood–Paley decomposition, which will be used in our proofs. To this end, let $\psi_0(\eta), \psi(\eta) \in C_c^\infty(\mathbb{R}^n)$ be compactly supported smooth cutoff functions, whose supports satisfy

$$\operatorname{supp}\psi_0 \subset \{|\eta| \leq 1\} \quad \text{and} \quad \operatorname{supp}\psi \subset \left\{\tfrac{1}{2} \leq |\eta| \leq 2\right\},$$

and such that

$$\psi_0(\eta) + \sum_{k=0}^{\infty}\psi\left(\frac{\eta}{2^k}\right) = 1 \quad \text{for all } \eta \in \mathbb{R}^n.$$

For any tempered distribution $f(x) \in \mathcal{S}'(\mathbb{R}^n)$, we define the dyadic blocks $\Delta_0 f(x), \Delta_{2^k} f(x) \in \mathcal{S}(\mathbb{R}^n)$, for each $k \in \mathbb{Z}$, by

$$\Delta_0 f = \mathcal{F}^{-1}\psi_0(\eta)\mathcal{F}f \quad \text{and} \quad \Delta_{2^k} f = \mathcal{F}^{-1}\psi\left(\frac{\eta}{2^k}\right)\mathcal{F}f.$$

so that

$$f = \Delta_0 f + \sum_{k=0}^{\infty}\Delta_{2^k} f \quad \text{in } \mathcal{S}'(\mathbb{R}^n). \tag{5-3}$$

As in Section 3, from now on, we assume that the cutoff function $\rho(r)$ may be decomposed as a product $\rho(r) = \rho_1(r)\rho_2(r)$, so that $\tilde{\rho}(s) = \tilde{\rho}_1 * \tilde{\rho}_2(s)$. We recall from (3-3) that

$$A_t = A_t^1 A_t^2 = A_t^2 A_t^1, \quad S_t = S_t^1 A_t^2 = S_t^2 A_t^1, \quad S_t^* = A_{-t}^2 S_t^{1*} = A_{-t}^1 S_t^{2*}, \tag{5-4}$$

where A_t^i, S_t^i and S_t^{i*}, with $i = 1, 2$, denote the respective operators A_t, S_t and S_t^* with the cutoff ρ replaced by ρ_i.

As shown in the results below, a key idea here is to use this trick to gain integration in one dimension along v through the straightforward estimate

$$\begin{aligned}
|S_t^* g(x,v)|^2 &= |A_{-t}^1 S_t^{2*} g(x,v)|^2 \\
&\leq \|\tilde{\rho}_1\|_{L^1}\int_{\mathbb{R}} |S_t^{2*} g(x+stv, v)|^2\,|\tilde{\rho}_1(s)|\,ds \\
&\leq \|\tilde{\rho}_1\|_{L^1}\|\tilde{\rho}_1\|_{L^\infty}\int_{[-r_1,r_1]} |S_t^{2*} g(x+stv, v)|^2\,ds,
\end{aligned} \tag{5-5}$$

where $\operatorname{supp}\tilde{\rho}_1 \subset [-r_1, r_1]$, for some $r_1 > 0$, and thus obtain new estimates on the adjoint operator S_t^*.

The next proposition contains an estimate which is central to the present two-dimensional setting. It strongly relies on the clever and elegant proof of Theorem 1.2 from [Jabin and Vega 2004], which it crucially improves by exploiting the structure of the operator A_t through the use of (5-5).

Proposition 5.2. *In dimension $n=2$, let $2\le p\le 4$, $0\le s<s_0<\frac{1}{2}$ and $\rho_1,\rho_2\in\mathcal{S}(\mathbb{R})$ be such that $\tilde\rho_1$ and $\tilde\rho_2$ have their supports contained inside balls of radii $r_1,r_2>0$, respectively, centered at the origin.*

Then, the operator S_t^* satisfies the estimate

$$\begin{aligned}
&\|(1-\Delta_x)^{\frac{s}{2}}S_t^*g\|_{L^p_xL^2_v}\\
&\le\frac{C}{|t|^{\frac{1}{2}}}\|\tilde\rho_1\|_{L^1}^{\frac{1}{2}}\|\tilde\rho_1\|_{L^\infty}^{\frac{1}{2}}(r_1+r_2)^{\frac{1}{4}}(1+r_1+r_2)^{\frac{1}{2}}\left(\left\|\frac{\rho_2(r)}{|r|^{\frac{1}{2}-s_0}}\right\|_{L^2}+\|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty}+\|\tilde\rho_2\|_{L^1}\right)\|g\|_{L^p_x}
\end{aligned}$$

for all $t\neq 0$ such that $|t|\le 1$, where $C>0$ only depends on fixed parameters.

Proof. First, notice that, for any $2\le p\le\infty$,

$$\|S_t^*g\|_{L^p_xL^2_v}\le\|S_t^*g\|_{L^2_vL^p_x}\le(2\pi)^{\frac{1}{2}}\|\tilde\rho\|_{L^1}\|g\|_{L^p_x}. \tag{5-6}$$

As for the regularity estimate, we employ the bound (5-5) and the standard Littlewood–Paley dyadic frequency decomposition previously introduced, to deduce, writing $g_k=\Delta_{2^k}g$ for convenience, for any $k\ge 0$, that

$$\begin{aligned}
\|S_t^*g_k\|_{L^4_xL^2_v}^4&=\int_{\mathbb{R}^2\times\mathbb{S}^1}|S_t^*g_k(x,v_2)|^2\left(\int_{\mathbb{S}^1}|S_t^*g_k(x,v_1)|^2\,dv_1\right)dx\,dv_2\\
&\le\|\tilde\rho_1\|_{L^1}\|\tilde\rho_1\|_{L^\infty}\int_{\mathbb{R}^2\times\mathbb{S}^1}\left(\int_{[-r_1,r_1]}|S_t^{2*}g_k(x+s_2tv_2,v_2)|^2\,ds_2\right)\\
&\qquad\qquad\times\left(\int_{\mathbb{S}^1}|S_t^*g_k(x,v_1)|^2\,dv_1\right)dx\,dv_2\\
&=\|\tilde\rho_1\|_{L^1}\|\tilde\rho_1\|_{L^\infty}\int_{\mathbb{R}^2\times\mathbb{S}^1}|S_t^{2*}g_k(x,v_2)|^2\\
&\qquad\qquad\times\left(\int_{[-r_1,r_1]}\int_{\mathbb{S}^1}|S_t^*g_k(x+s_2tv_2,v_1)|^2\,dv_1\,ds_2\right)dx\,dv_2\\
&\le\|\tilde\rho_1\|_{L^1}^2\|\tilde\rho_1\|_{L^\infty}^2\int_{\mathbb{R}^2\times\mathbb{S}^1}|S_t^{2*}g_k(x,v_2)|^2\\
&\qquad\qquad\times\left(\int_{[-r_1,r_1]^2}\int_{\mathbb{S}^1}|S_t^{2*}g_k(x+s_1tv_1+s_2tv_2,v_1)|^2\,dv_1\,ds_1\,ds_2\right)dx\,dv_2\\
&\le\|\tilde\rho_1\|_{L^1}^2\|\tilde\rho_1\|_{L^\infty}^2\|S_t^{2*}g_k\|_{L^2_{x,v}}^2\sup_{\substack{x\in\mathbb{R}^2\\ v_2\in\mathbb{S}^1}}I(x,v_2),
\end{aligned}$$

where

$$I(x,v_2)=\int_{[-r_1,r_1]^2}\int_{\mathbb{S}^1}|S_t^{2*}g_k(x+s_1tv_1+s_2tv_2,v_1)|^2\,dv_1\,ds_1\,ds_2.$$

Further using Proposition 5.1, we deduce that, for all $t>0$ and any given $0<s<\frac{1}{2}$,

$$\|S_t^*g_k\|_{L^4_xL^2_v}^4\le\frac{C}{|t|^{2s}2^{k2s}}\|\tilde\rho_1\|_{L^1}^2\|\tilde\rho_1\|_{L^\infty}^2\left(\left\|\frac{\rho_2(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2}^2+\left\||r|^s\rho_2(r)\right\|_{L^\infty}^2\right)\|g\|_{L^2_x}^2\sup_{\substack{x\in\mathbb{R}^2\\ v_2\in\mathbb{S}^1}}I(x,v_2). \tag{5-7}$$

We claim now that

$$\sup_{\substack{x\in\mathbb{R}^2\\ v_2\in\mathbb{S}^1}} I(x,v_2) \leq \frac{C(k+1)}{|t|2^{k2s}}(r_1+r_2)(1+r_1+r_2)^2\left(\left\|\frac{\rho_2(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2}^2+\|(1+|r|^s)\rho_2(r)\|_{L^\infty}^2+\|\tilde\rho_2\|_{L^1}^2\right)\|g\|_{L^\infty_x}^2, \tag{5-8}$$

where $C>0$ only depends on fixed parameters. In order to establish (5-8), we employ the change of variables $(s_1,s_2)\mapsto z=s_1tv_1+s_2tv_2$ whenever v_1 and v_2 form a basis, which holds almost everywhere. It is readily seen that the Jacobian determinant of this transformation is given by $t^2\sin\theta$, where $\theta\in[0,\pi]$ is the angle between v_1 and v_2 defined by $\cos\theta=v_1\cdot v_2$. Thus, noticing that $|z|=|s_1tv_1+s_2tv_2|\leq 2r_1|t|$, we infer

$$\begin{aligned} I(x,v_2) &= \int_{[-r_1,r_1]^2}\int_{\mathbb{S}^1}|S_t^{2*}g_k(x+s_1tv_1+s_2tv_2,v_1)|^2\,dv_1\,ds_1\,ds_2\\ &\leq \int_{\mathbb{R}^2\times\mathbb{S}^1}|S_t^{2*}g_k(x+z,v_1)|^2\frac{1}{t^2\sin\theta}\mathbb{1}_{\{|z|\leq 2r_1|t|,|z\cdot v_2^\perp|\leq r_1|t|\sin\theta\}}\,dz\,dv_1\\ &= \sum_{i=0}^\infty\int_{\mathbb{R}^2\times S_i}\left|\int_{\mathbb{R}}g_k(x+z+stv_1)\tilde\rho_2(s)\,ds\right|^2\frac{1}{t^2\sin\theta}\mathbb{1}_{\{|z|\leq 2r_1|t|,|z\cdot v_2^\perp|\leq r_1|t|\sin\theta\}}\,dz\,dv_1, \end{aligned} \tag{5-9}$$

where we have used the notation

$$v^\perp=\begin{pmatrix}v_1\\ v_2\end{pmatrix}^\perp=\begin{pmatrix}-v_2\\ v_1\end{pmatrix}$$

for any $v\in\mathbb{S}^1$, and have decomposed the domain of integration $v_1\in\mathbb{S}^1$ into

$$\mathbb{S}^1\setminus\{\sin\theta=0\}=\bigcup_{i=0}^\infty S_i,$$

with

$$S_i=\left\{v_1\in\mathbb{S}^1:\frac{1}{2^{i+1}}<|\sin\theta|\leq\frac{1}{2^i}\right\}.$$

Recall now that $\tilde\rho_2$ is supported inside a ball of radius $r_2>0$ centered at the origin. Therefore, we find that, in the last integrand above,

$$|(z+stv_1)\cdot v_2^\perp|\leq|z\cdot v_2^\perp|+|stv_1\cdot v_2^\perp|\leq(r_1+r_2)|t|\sin\theta\leq\frac{(r_1+r_2)|t|}{2^i},$$

and

$$|z+stv_1|\leq(2r_1+r_2)|t|.$$

Hence, considering a smooth cutoff function $\chi\in C_c^\infty(\mathbb{R})$ such that $\mathbb{1}_{\{|s|\leq 1\}}\leq\chi(s)\leq\mathbb{1}_{\{|s|\leq 2\}}$, setting for convenience $r_0=2(r_1+r_2)|t|$, and defining, for each given $x\in\mathbb{R}^2$, $v_2\in\mathbb{S}^1$ and $i\in\mathbb{N}$,

$$K^{i,k}_{x,v_2}(z)=g_k(x+z)\chi\left(\frac{|z\cdot v_2|}{r_0}\right)\chi\left(2^i\frac{|z\cdot v_2^\perp|}{r_0}\right),$$

we deduce from (5-9) that, using Proposition 5.1,

$$
\begin{aligned}
I(x,v_2) &\le \sum_{i=0}^{\infty} \int_{\mathbb{R}^2\times S_i} \left| \int_{\mathbb{R}} K^{i,k}_{x,v_2}(z+stv_1)\,\tilde\rho_2(s)\,ds \right|^2 \frac{1}{t^2\sin\theta}\,dz\,dv_1 \\
&\le \sum_{2^i\le r_0 2^{k2s}} \frac{2^{i+1}}{t^2} \int_{\mathbb{R}^2\times\mathbb{S}^1} |S_t^{2*} K^{i,k}_{x,v_2}(z,v_1)|^2\,dz\,dv_1 \\
&\qquad\qquad + \sum_{2^i> r_0 2^{k2s}} \frac{2^{i+1}}{t^2} \int_{\mathbb{R}^2\times S_i} |S_t^{2*} K^{i,k}_{x,v_2}(z,v_1)|^2\,dz\,dv_1 \\
&\le \sum_{2^i\le r_0 2^{k2s}} \frac{2^{i+1}}{t^2} \|S_t^{2*} K^{i,k}_{x,v_2}\|^2_{L^2_{z,v_1}} + \sum_{2^i> r_0 2^{k2s}} \frac{2^{i+1}}{t^2} |S_i|\,\|\tilde\rho_2\|^2_{L^1}\,\|K^{i,k}_{x,v_2}\|^2_{L^2_z} \\
&\le \frac{C}{|t|^{2s}} \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12-s}} \right\|^2_{L^2} + \|(1+|r|^s)\rho_2(r)\|^2_{L^\infty} \right) \sum_{2^i\le r_0 2^{k2s}} \frac{2^i}{t^2} \|(1-\Delta_z)^{-\frac{s}{2}} K^{i,k}_{x,v_2}\|^2_{L^2_z} \\
&\qquad\qquad + C \sum_{2^i> r_0 2^{k2s}} \frac{r_0^2}{t^2 2^i} \|\tilde\rho_2\|^2_{L^1}\,\|g_k\|^2_{L^\infty_x} \\
&\le \frac{C}{|t|^{2s}} \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12-s}} \right\|^2_{L^2} + \|(1+|r|^s)\rho_2(r)\|^2_{L^\infty} \right) \sum_{2^i\le r_0 2^{k2s}} \frac{2^i}{t^2} \|(1-\Delta_z)^{-\frac{s}{2}} K^{i,k}_{x,v_2}\|^2_{L^2_z} \\
&\qquad\qquad + C \frac{r_0}{t^2 2^{k2s}} \|\tilde\rho_2\|^2_{L^1}\,\|g\|^2_{L^\infty_x},
\end{aligned}
$$

where $C > 0$ is an independent constant.

Next, a direct application of the technical Lemma A.1 from Appendix A to the preceding estimate yields

$$
\begin{aligned}
I(x,v_2) &\le \frac{C}{|t|^{2s}} \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12-s}} \right\|^2_{L^2} + \|(1+|r|^s)\rho_2(r)\|^2_{L^\infty} \right) \sum_{2^i\le r_0 2^{k2s}} \frac{r_0^2}{t^2 2^{k2s}} \|g\|^2_{L^\infty_x} + C \frac{r_0}{t^2 2^{k2s}} \|\tilde\rho_2\|^2_{L^1}\,\|g\|^2_{L^\infty_x} \\
&\le C \frac{(k+1)(r_1+r_2)^2 \log(2+r_1+r_2)}{|t|^{2s} 2^{k2s}} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12-s}} \right\|^2_{L^2} + \|(1+|r|^s)\rho_2(r)\|^2_{L^\infty} \right) \|g\|^2_{L^\infty_x} + C \frac{r_1+r_2}{|t| 2^{k2s}} \|\tilde\rho_2\|^2_{L^1}\,\|g\|^2_{L^\infty_x},
\end{aligned}
$$

which establishes our claim (5-8).

Finally, combining estimates (5-7) and (5-8), we arrive at

$$
\begin{aligned}
\|S_t^* g_k\|^2_{L^4_x L^2_v} &\le \frac{C(k+1)^{\frac12}}{|t| 2^{k2s}} \|\tilde\rho_1\|_{L^1}\,\|\tilde\rho_1\|_{L^\infty} (r_1+r_2)^{\frac12} (1+r_1+r_2) \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12-s}} \right\|^2_{L^2} + \|(1+|r|^s)\rho_2(r)\|^2_{L^\infty} + \|\tilde\rho_2\|^2_{L^1} \right) \|g\|_{L^2_x}\,\|g\|_{L^\infty_x},
\end{aligned}
$$

where $C > 0$ is an independent constant.

In order to conclude, we write $|g(x)| = \int_0^\infty \mathbb{1}_{\{|g(x)|\geq s\}}\,ds$ to deduce from the preceding estimate, assuming g is nonnegative, that

$$\begin{aligned}
\|S_t^* g_k\|_{L^4_x L^2_v} &\leq \int_0^\infty \|S_t^* \Delta_{2^k} \mathbb{1}_{\{|g(x)|\geq s\}}\|_{L^4_x L^2_v}\,ds \\
&\leq \frac{C(k+1)^{\frac14}}{|t|^{\frac12} 2^{ks}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac14}(1+r_1+r_2)^{\frac12} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s}} \right\|_{L^2} + \|(1+|r|^s)\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \int_0^\infty |\{|g(x)|\geq s\}|^{\frac14}\,ds \\
&\leq \frac{C(k+1)^{\frac14}}{|t|^{\frac12} 2^{ks}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac14}(1+r_1+r_2)^{\frac12} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s}} \right\|_{L^2} + \|(1+|r|^s)\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^{4,1}_x},
\end{aligned}$$

where $L^{4,1}_x$ denotes a standard Lorentz space; see [Bergh and Löfström 1976, Section 1.3] or [Grafakos 2008, Section 1.4] for definitions and properties of Lorentz spaces. When, g is signed, we arrive at the same estimate simply by decomposing $g = g^+ - g^-$ into its positive and negative parts, treating each contribution separately, and then noticing that

$$\begin{aligned}
\|g^+\|_{L^{4,1}_x} + \|g^-\|_{L^{4,1}_x} &\leq C \int_0^\infty \left|\{|g^+(x)|\geq s\}\right|^{\frac14} + \left|\{|g^-(x)|\geq s\}\right|^{\frac14} ds \\
&\leq C \int_0^\infty \left(\left|\{|g^+(x)|\geq s\}\right| + \left|\{|g^-(x)|\geq s\}\right| \right)^{\frac14} ds \\
&\leq C \int_0^\infty \left|\{|g(x)|\geq s\}\right|^{\frac14} ds \leq C \|g\|_{L^{4,1}_x}.
\end{aligned}$$

Moreover, by allowing an arbitrarily small loss of regularity, that is, by replacing $0 < s < \frac12$ by a slightly smaller value, it is possible to replace the Lorentz space $L^{4,1}_x$ by the standard Lebesgue space L^4_x in the right-hand side of the above estimate.

Therefore, on the whole, for any $0 \leq s < s_0 < \frac12$, we have established the estimate

$$\begin{aligned}
\|S_t^* g_k\|_{L^4_x L^2_v} &\leq \frac{C}{|t|^{\frac12} 2^{ks}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac14}(1+r_1+r_2)^{\frac12} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^4_x},
\end{aligned}$$

where $C > 0$ only depends on constant parameters, which, when combined with the easy bound (5-6) for low frequencies, yields

$$\begin{aligned}
\|(1-\Delta_x)^{\frac{s}{2}} S_t^* g\|_{L^4_x L^2_v} &\leq \frac{C}{|t|^{\frac12}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac14}(1+r_1+r_2)^{\frac12} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^4_x}. \qquad \text{(5-10)}
\end{aligned}$$

Finally, recalling from complex interpolation theory, see [Bergh and Löfström 1976, Section 5.1], that, for any $2 < p < 4$,

$$(L^2_x L^2_v, L^4_x L^2_v)_{[2-\frac{4}{p}]} = L^p_x L^2_v \quad \text{and} \quad (L^2_x, L^4_x)_{[2-\frac{4}{p}]} = L^p_x,$$

we conclude the proof of the proposition by interpolating the estimate (5-10) with the classical estimate on S_t^* from Proposition 5.1. □

Next, we utilize the previous result on the adjoint operator S_t^* to deduce corresponding estimates on T_t and T_t^*.

Proposition 5.3. *In dimension $n = 2$, let $\frac{4}{3} \le r \le 2$, $2 \le p \le 4$, $0 \le s < s_0 < \frac{1}{2}$ and $\rho_1, \rho_2 \in \mathcal{S}(\mathbb{R})$ be such that $\tilde\rho_1$ and $\tilde\rho_2$ have their supports contained inside balls of radii $r_1, r_2 > 0$, respectively, centered at the origin.*

Then, the operators T_t and T_t^* satisfy the estimates

$$\begin{aligned}\|(1-\Delta_x)^{\frac{s}{2}} T_t f\|_{L^r_x} \le \frac{C}{|t|^{\frac12}} \|\varphi\|_{L^\infty} R^{\frac12} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac14}(1+r_1+r_2)^{\frac12} \\ \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|f\|_{L^r_x L^2_v} \end{aligned} \tag{5-11}$$

and

$$\begin{aligned}\|(1-\Delta_x)^{\frac{s}{2}} T_t^* g\|_{L^p_x L^2_v} \le \frac{C}{|t|^{\frac12}} \|\varphi\|_{L^\infty} R^{\frac12} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac14}(1+r_1+r_2)^{\frac12} \\ \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^p_x} \end{aligned}$$

for all $t \neq 0$ such that $|t| \le 1$, where $C > 0$ only depends on fixed parameters.

Proof. It is readily seen that

$$T_t^* g(x,v) = \int_{\mathbb{R}} g(x+stv)\tilde\rho(s)\,ds\,\varphi(v) = \varphi(v) S^*_{|v|t} g\left(x, \frac{v}{|v|}\right).$$

Therefore, for any $2 \le p \le \infty$, we compute in polar coordinates, recalling that the support of the velocity weight $\varphi \in L^\infty_c(\mathbb{R}^2)$ is contained inside a closed ball of radius $R > 0$ centered at the origin,

$$\begin{aligned}\|(1-\Delta_x)^{\frac{s}{2}} T_t^* g(x,v)\|_{L^p_x(\mathbb{R}^2; L^2_v(\mathbb{R}^2))} &= \left\| \left(\int_{\mathbb{R}^2} \left| (1-\Delta_x)^{\frac{s}{2}} T_t^* g(x,v) \right|^2 dv \right)^{\frac12} \right\|_{L^p_x} \\ &= \left\| \int_0^\infty r \int_{\mathbb{S}^1} \left| \varphi(rv)(1-\Delta_x)^{\frac{s}{2}} S^*_{rt} g(x,v) \right|^2 dv\,dr \right\|_{L^{p/2}_x}^{\frac12} \\ &\le \|\varphi\|_{L^\infty} \left(\int_0^R r \left\| \int_{\mathbb{S}^1} \left| (1-\Delta_x)^{\frac{s}{2}} S^*_{rt} g(x,v) \right|^2 dv \right\|_{L^{p/2}_x} dr \right)^{\frac12} \\ &\le \|\varphi\|_{L^\infty} \left(\int_0^R r \|(1-\Delta_x)^{\frac{s}{2}} S^*_{rt} g(x,v)\|^2_{L^p_x(\mathbb{R}^2; L^2_v(\mathbb{S}^1))} dr \right)^{\frac12}.\end{aligned}$$

Then, combining Proposition 5.2 with the above estimate, we find

$$\begin{aligned}\|(1-\Delta_x)^{\frac{s}{2}} T_t^* g\|_{L^p_x L^2_v} &\le \|\varphi\|_{L^\infty} \left(\int_0^R r \|(1-\Delta_x)^{\frac{s}{2}} S^*_{rt} g\|^2_{L^p_x L^2_v} \, dr \right)^{\frac{1}{2}} \\ &\le \frac{C}{|t|^{\frac{1}{2}}} \|\varphi\|_{L^\infty} R^{\frac{1}{2}} \|\tilde\rho_1\|^{\frac{1}{2}}_{L^1} \|\tilde\rho_1\|^{\frac{1}{2}}_{L^\infty} (r_1+r_2)^{\frac{1}{4}} (1+r_1+r_2)^{\frac{1}{2}} \\ &\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac{1}{2}-s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^p_x},\end{aligned}$$

where $C>0$ is an independent constant, which establishes the estimate on T_t^*.

The estimate on T_t is then easily deduced from the estimate on T_t^* by a duality argument. □

We proceed now to the main theorem of this section. Note that an equivalent version of this result with spherical averages and an identical regularity gain can be readily obtained by applying the methods from Appendix C.

Theorem 5.4. *In dimension $n=2$, let $\frac{4}{3} \le r \le 2$.*

Then, for any $f, g \in L^r_x(\mathbb{R}^2; L^2_v(\mathbb{R}^2))$ such that (1-2) *holds true, one has*

$$\int_{\mathbb{R}^2} f(x,v)\,\varphi(v)\,dv \in W^{s,r}_x(\mathbb{R}^2)$$

for any $\varphi \in L^\infty_c(\mathbb{R}^2)$ and any $0 \le s < \frac{1}{2}$. Furthermore, one has the estimate

$$\left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^2} f\varphi\,dv \right\|_{L^r_x} \le C(\|f\|_{L^r_x L^2_v} + \|g\|_{L^r_x L^2_v}),$$

where $C>0$ only depends on φ and constant parameters.

Proof. We consider first the refined interpolation formula (2-1), which is valid for solutions of the transport equation (1-2), for some given cutoff $\rho \in \mathcal{S}(\mathbb{R})$. Clearly, further differentiating (2-1) in x and then averaging in v yields

$$\left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^2} f\varphi\,dv \right\|_{L^r_x} \le \left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^2} A_t f\varphi\,dv \right\|_{L^r_x} + t \left\| (1-\Delta_x)^{\frac{s}{2}} \int_{\mathbb{R}^2} B_t g\varphi\,dv \right\|_{L^r_x}. \tag{5-12}$$

We wish now to apply Proposition 5.3 to the preceding estimate. To this end, according to (5-4), we take the decompositions

$$\rho(r) = \rho_1(r)\rho_2(r) \quad \text{and} \quad \tau(r) = \frac{1-\rho(r)}{ir} = \tau_1(r)\tau_2(r),$$

where

$$\tilde\rho_1(r) \in C^\infty_c(\mathbb{R}), \quad \tau_1(r) = \frac{1}{(1+r^2)^{\frac{1}{4}}},$$

$$\tilde\rho_2(r) \in C^\infty_c(\mathbb{R}), \quad \tau_2(r) = (1+r^2)^{\frac{1}{4}}\tau(r).$$

Clearly, all constants involving norms of the cutoff functions ρ_1 and ρ_2 in the right-hand side of (5-11) are finite and we may therefore straightforwardly apply Proposition 5.3 to control the first term in the right-hand side of (5-12). However, the same is not so obviously true concerning the cutoff functions τ_1 and τ_2. The application of Proposition 5.3 to the second term in the right-hand side of (5-12) will therefore require some substantial technical work, which we present now.

To this end, we employ a homogeneous Littlewood–Paley frequency decomposition, see (5-3), of τ_1 and τ_2 to write

$$\tau = \Big(\sum_{j\in\mathbb{Z}} \Delta_{2^j}\tau_1\Big)\Big(\sum_{k\in\mathbb{Z}} \Delta_{2^k}\tau_2\Big) = \sum_{j\in\mathbb{Z}} (\Delta_{2^j}\tau_1)\tau_3^j + \sum_{j\in\mathbb{Z}} (\Delta_{2^j}\tau_2)\tau_4^j,$$

where

$$\tau_3^j = \sum_{\substack{k\in\mathbb{Z}\\ k\le j}} \Delta_{2^k}\tau_2 = \mathcal{F}^{-1}\Big[\psi_0\Big(\frac{r}{2^{j+1}}\Big)\Big] * \tau_2, \quad \tau_4^j = \sum_{\substack{k\in\mathbb{Z}\\ k< j}} \Delta_{2^k}\tau_1 = \mathcal{F}^{-1}\Big[\psi_0\Big(\frac{r}{2^{j}}\Big)\Big] * \tau_1.$$

In view of the linearity of the operator T_t with respect to the cutoffs ρ or τ, we only need to verify the finiteness of the constants in (5-11) with $\Delta_{2^j}\tau_1$ and τ_3^j playing the roles of ρ_1 and ρ_2, respectively, and then with $\Delta_{2^j}\tau_2$ and τ_4^j instead of ρ_1 and ρ_2, respectively. It is to be emphasized here that the ensuing bounds on the cutoffs will then depend on $j\in\mathbb{Z}$. In order to guarantee the boundedness of T_t, we will therefore need to make sure that our method eventually yields constants that are summable in $j\in\mathbb{Z}$.

We evaluate now the norms involved in the right-hand side of (5-11) where we replace ρ_1 by $\Delta_{2^j}\tau_1$ (or $\Delta_{2^j}\tau_2$) and ρ_2 by τ_3^j (or τ_4^j). The bounds on $\Delta_{2^j}\tau_2$ and τ_4^j are handled in a strictly similar manner and so we omit the corresponding details.

First, note that a direct application of Lemma B.3 from Appendix B together with the fact that τ_1 and τ_2 are smooth so that their Fourier transforms decay faster than any inverse power at infinity, shows that

$$\begin{aligned} |\widetilde{\Delta_{2^j}\tau_1}(r)| &\le \frac{C}{|r|^{\frac12}(1+|r|)^N}\mathbb{1}_{\{2^{j-1}\le|r|\le 2^{j+1}\}}, \\ |\tilde{\tau}_3^j(r)| &\le \frac{C}{|r|^{\frac12}(1+|r|)^N}\mathbb{1}_{\{|r|\le 2^{j+1}\}} \end{aligned} \tag{5-13}$$

for any arbitrarily large $N>0$.

Furthermore, in view of Lemma B.4, it holds that each τ_3^j satisfies

$$|\tau_3^j(r)| \le \frac{C}{1+|r|^{\frac12}}$$

for some uniform $C>0$ independent of $j\in\mathbb{Z}$, whence, for any $0<s_0<\frac12$,

$$\frac{\tau_3^j(r)}{|r|^{\frac12-s_0}} \in L^2(\mathbb{R}) \quad\text{and}\quad (1+|r|^{s_0})\tau_3^j(r)\in L^\infty(\mathbb{R}), \tag{5-14}$$

uniformly in $j\in\mathbb{Z}$.

Therefore, using the bounds (5-13) and (5-14) to evaluate the terms involving $\rho_1=\Delta_{2^j}\tau_1$ and $\rho_2=\tau_3^j$ in the right-hand side of (5-11), we compute that the corresponding norm of the operator in (5-11) is no larger than a multiple of

$$\left(\frac{2^{\frac{j}{2}}}{1+2^{jN}}\right)^{\frac{1}{2}}\left(\frac{1}{2^{\frac{j}{2}}(1+2^{jN})}\right)^{\frac{1}{2}} 2^{\frac{j}{4}}(1+2^j)^{\frac{1}{2}} \le C\frac{2^{\frac{j}{4}}}{(1+2^j)^{N-\frac{1}{2}}},$$

which is summable over $j\in\mathbb{Z}$, provided $N>\frac{3}{4}$.

Thus, we conclude, according to Proposition 5.3, that the operators in the right-hand side of (5-12) are bounded.

It follows that, for any $0<t<1$,

$$\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}^2} f\varphi\,dv\right\|_{L^r_x} \le C\left(\frac{1}{t^{\frac{1}{2}}}\|f\|_{L^r_xL^2_v}+t^{\frac{1}{2}}\|g\|_{L^r_xL^2_v}\right),$$

where $C>0$ only depends on constant parameters. □

6. The higher-dimensional case

We move on now to the higher-dimensional case. More precisely, in the present section, we generalize the methods leading to Theorem 5.4 to establish an analog result valid in any dimension. Unfortunately, the ensuing result does not reach a maximal gain of regularity of half a derivative on the velocity averages, but only a gain of $\frac{n}{4(n-1)}$ derivatives, where $n\ge 3$ is the dimension. This drawback mainly stems from the fact that we work in the L^4_x setting, because our methods exploit the trivial fact that the exponent 4 is an even integer in order to control the square of some transport operator in L^2_x rather than the operator itself in L^4_x.

We begin with a few technical results. Loosely speaking, a key idea behind Proposition 5.2 consisted in noticing that $S_t^*g(x,v)$ is regular along the direction v and then using some duality argument in L^4_x to gain an integration variable in another nondegenerate direction. In higher dimensions, in order to carry out a similar strategy, we need to gain integration variables in $n-1$ nondegenerate directions. The next few lemmas will allow us to achieve such a dimensional build up of integration variables.

The following lemma generalizes estimate (5-8) from the proof of Proposition 5.2 and corresponds to a situation where we have already managed to build up the integration dimension all the way up to n (notice the n-dimensional integration in S in the estimate below).

Lemma 6.1. *In any dimension $n\ge 2$, let $0<s<\frac{1}{2}$, $r_1>0$ and $\rho_2\in\mathcal{S}(\mathbb{R})$ be such that $\tilde\rho_2$ has its support contained inside a ball of radius $r_2>0$ centered at the origin.*

Then, denoting $S=(s_1,\dots,s_n)\in\mathbb{R}^n$, it holds that, for any $k\in\mathbb{N}$,

$$\sup_{\substack{x\in\mathbb{R}^n\\ v_n\in\mathbb{S}^{n-1}}}\int_{[-r_1,r_1]^n}\int_{\mathbb{S}^{n-1}}\cdots\int_{\mathbb{S}^{n-1}}\left|S_t^{2*}\Delta_{2^k}g\left(x+\sum_{j=1}^n s_jtv_j,v_1\right)\right|^2 dv_1\cdots dv_{n-1}\,dS$$
$$\le \frac{C(k+1)}{|t|2^{k2s}}(r_1+r_2)^{n-1}(1+r_1+r_2)^2\left(\left\|\frac{\rho_2(r)}{|r|^{\frac{1}{2}-s}}\right\|_{L^2}^2+\|(1+|r|^s)\rho_2(r)\|_{L^\infty}^2+\|\tilde\rho_2\|_{L^1}^2\right)\|g\|_{L^\infty_x}^2$$

for all $t\neq 0$ such that $|t|\le 1$, where $C>0$ is independent of k, r_1, r_2 and ρ_2.

Proof. We assume that $\{v_2, \dots, v_n\}$ is a linearly independent set of vectors, which holds almost everywhere, and denote its span by H. Further let $u_1 \in \mathbb{R}^n$ be a unit vector orthogonal to H. The specific choice of unit vector is irrelevant, any such vector will do. Note that $\det(u_1, v_2, \dots, v_n) \neq 0$. Moreover, since $v_1 - (v_1 \cdot u_1)u_1$ belongs to H and recalling that the determinant is linear with respect to each of its column vectors, it holds that

$$\det(v_1, \dots, v_n) = v_1 \cdot u_1 \det(u_1, v_2, \dots, v_n).$$

We wish now to perform the change of variable $z = z(S) = \sum_{j=1}^n s_j t v_j$ in $\mathbb{R}^n$, whose Jacobian determinant is given by

$$\left| \frac{\partial z}{\partial S} \right| = t^n \det(v_1, \dots, v_n). \tag{6-1}$$

However, this operation becomes singular as v_1 approaches H, that is, as $v_1 \cdot u_1$ becomes small. Therefore, in order to deal with this degeneracy, we consider the following partition in v_1 of $\mathbb{S}^{n-1}$:

$$\mathbb{S}^{n-1} \setminus \{v_1 \cdot u_1 = 0\} = \bigcup_{i=0}^{\infty} S_i,$$

with

$$S_i = \left\{ v_1 \in \mathbb{S}^{n-1} : \frac{1}{2^{i+1}} |v_1 \cdot u_1| \leq \frac{1}{2^i} \right\}.$$

Then, defining $r_0 = n(r_1 + r_2)|t|$ and writing $g_k = \Delta_{2^k} g$, for convenience, one has the following straightforward estimate on $\bigcup_{2^i > r_0 2^{k2s}} S_i$:

$$\begin{aligned} \int_{\bigcup_{2^i > r_0 2^{k2s}} S_i} \left| S_t^{2*} g_k \left(x + \sum_{j=1}^n s_j t v_j, v_1 \right) \right|^2 dv_1 &\leq \sum_{2^i > r_0 2^{k2s}} |S_i| \, \|\tilde{\rho}_2\|_{L^1}^2 \, \|g\|_{L^\infty}^2 \\ &\leq \frac{C}{r_0 2^{k2s}} \|\tilde{\rho}_2\|_{L^1}^2 \, \|g\|_{L^\infty}^2, \end{aligned} \tag{6-2}$$

where $C > 0$ only depends on the dimension.

Now, on each domain S_i, with $2^i \leq r_0 2^{k2s}$, the Jacobian determinant (6-1) remains bounded away from zero. More precisely, for every $v_1 \in S_i$, it holds that

$$\begin{aligned} |t^n \det(u_1, v_2, \dots, v_n)| &\int_{[-r_1, r_1]^n} \left| S_t^{2*} g_k \left(x + \sum_{j=1}^n s_j t v_j, v_1 \right) \right|^2 dS \\ &= \frac{1}{|v_1 \cdot u_1|} \int_{z([-r_1, r_1]^n)} |S_t^{2*} g_k(x + z, v_1)|^2 \, dz \\ &\leq 2^{i+1} \int_{\left\{ |z| \leq n r_1 |t|, |z \cdot u_1| \leq \frac{r_1 |t|}{2^i} \right\}} |S_t^{2*} g_k(x + z, v_1)|^2 \, dz \\ &= 2^{i+1} \int_{\left\{ |z| \leq n r_1 |t|, |z \cdot u_1| \leq \frac{r_1 |t|}{2^i} \right\}} \left| \int_{\mathbb{R}} g_k(x + z + s t v_1) \tilde{\rho}_2(s) \, ds \right|^2 dz, \end{aligned} \tag{6-3}$$

where we have used that, for each $S\in[-r_1,r_1]^n$,

$$|z|=\left|\sum_{j=1}^n s_j t v_j\right|\le n r_1|t| \quad\text{and}\quad |z\cdot u_1|=\left|\left(\sum_{j=1}^n s_j t v_j\right)\cdot u_1\right|=|s_1 t v_1\cdot u_1|\le \frac{r_1|t|}{2^i}.$$

Next, further notice that, whenever $|z|\le n r_1|t|$, $|z\cdot u_1|\le r_1|t|/2^i$, $|s|\le r_2$ and $v_1\in S_i$, it holds that

$$|z+stv_1|\le (nr_1+r_2)|t| \quad\text{and}\quad |(z+stv_1)\cdot u_1|\le \frac{(r_1+r_2)|t|}{2^i}.$$

It therefore follows from (6-3) that

$$\begin{aligned}
&|t^n\det(u_1,v_2,\dots,v_n)|\int_{[-r_1,r_1]^n}\left|S_t^{2*}g_k\left(x+\sum_{j=1}^n s_j t v_j, v_1\right)\right|^2 dS\\
&\le 2^{i+1}\int_{\{|z|\le nr_1|t|,|z\cdot u_1|\le\frac{r_1|t|}{2^i}\}}\left|\int_{\mathbb{R}} g_k(x+z+stv_1)\mathbb{1}_{\{|z+stv_1|\le r_0,|(z+stv_1)\cdot u_1|\le\frac{r_0}{2^i}\}}\tilde\rho_2(s)\,ds\right|^2 dz\\
&\le 2^{i+1}\int_{\mathbb{R}^n}|S_t^{2*}K^{i,k}_{x,u_1}(z,v_1)|^2\,dz, \qquad (6\text{-}4)
\end{aligned}$$

where

$$K^{i,k}_{x,u_1}(z)=g_k(x+z)\chi\left(\frac{|z-(z\cdot u_1)u_1|}{r_0}\right)\chi\left(2^i\frac{|z\cdot u_1|}{r_0}\right),$$

and $\chi\in C_c^\infty(\mathbb{R})$ is a smooth cutoff function such that $\mathbb{1}_{\{|s|\le 1\}}\le\chi(s)\le\mathbb{1}_{\{|s|\le 2\}}$.

Further integrating (6-4) in $v_1\in S_i$ and then applying Proposition 5.1, we find that

$$\begin{aligned}
&|t^n\det(u_1,v_2,\dots,v_n)|\int_{S_i}\int_{[-r_1,r_1]^n}\left|S_t^{2*}g_k\left(x+\sum_{j=1}^n s_j t v_j, v_1\right)\right|^2 dS\,dv_1\\
&\qquad\le 2^{i+1}\int_{\mathbb{S}^{n-1}}\int_{\mathbb{R}^n}|S_t^{2*}K^{i,k}_{x,u_1}(z,v_1)|^2\,dz\,dv_1\\
&\qquad\le C\frac{2^i}{|t|^{2s}}\left(\left\|\frac{\rho_2(r)}{|r|^{\frac12-s}}\right\|_{L^2}^2+\|(1+|r|^s)\rho_2(r)\|_{L^\infty}^2\right)\|(1-\Delta_z)^{-\frac{s}{2}}K^{i,k}_{x,u_1}(z)\|^2_{L^2_z},
\end{aligned}$$

where $C>0$ only depends on the dimension. Moreover, a direct application of Lemma A.1 from Appendix A on paradifferential calculus yields

$$\|(1-\Delta_z)^{-\frac{s}{2}}K^{i,k}_{x,u_1}(z)\|^2_{L^2_z}\le\frac{Cr_0^n}{2^i 2^{k2s}}\|g\|^2_{L^\infty}$$

for every $i,k\in\mathbb{N}$ such that $2^i\le r_0 2^{k2s}$, where $C>0$ is independent of i, k and r_0, whence

$$\begin{aligned}
&|t^n\det(u_1,v_2,\dots,v_n)|\int_{S_i}\int_{[-r_1,r_1]^n}\left|S_t^{2*}g_k\left(x+\sum_{j=1}^n s_j t v_j, v_1\right)\right|^2 dS\,dv_1\\
&\qquad\le C\frac{r_0^n}{|t|^{2s}2^{k2s}}\left(\left\|\frac{\rho_2(r)}{|r|^{\frac12-s}}\right\|_{L^2}^2+\|(1+|r|^s)\rho_2(r)\|_{L^\infty}^2\right)\|g\|^2_{L^\infty}, \qquad (6\text{-}5)
\end{aligned}$$

where $C>0$ is independent of i, k, r_0 and ρ_2.

On the whole, combining (6-2), which is valid when $2^i > r_0 2^{k2s}$, with (6-5), which is valid when $2^i \le r_0 2^{k2s}$, we arrive at

$$\begin{aligned}
&\int_{[-r_1,r_1]^n}\int_{\mathbb{S}^{n-1}}\left|S_t^{2*}g_k\Big(x+\sum_{j=1}^n s_j t v_j, v_1\Big)\right|^2 dv_1\,dS\\
&\quad=\sum_{2^i\le r_0 2^{k2s}}\int_{[-r_1,r_1]^n}\int_{S_i}\left|S_t^{2*}g_k\Big(x+\sum_{j=1}^n s_j v_j, v_1\Big)\right|^2 dv_1\,dS\\
&\qquad\qquad+\int_{[-r_1,r_1]^n}\int_{\bigcup_{2^i>r_0 2^{k2s}}S_i}\left|S_t^{2*}g_k\Big(x+\sum_{j=1}^n s_j v_j, v_1\Big)\right|^2 dv_1\,dS\\
&\quad\le\frac{C}{|\det(u_1,v_2,\dots,v_n)|\,|t|^{2s}}\left(\left\|\frac{\rho_2(r)}{|r|^{\frac12-s}}\right\|_{L^2}^2+\|(1+|r|^s)\rho_2(r)\|_{L^\infty}^2\right)\\
&\qquad\qquad\times\sum_{2^i\le r_0 2^{k2s}}\frac{r_0^n}{|t|^n 2^{k2s}}\|g\|_{L^\infty}^2+\frac{Cr_1^n}{r_0 2^{k2s}}\|\tilde\rho_2\|_{L^1}^2\,\|g\|_{L^\infty}^2\\
&\quad\le C\frac{(k+1)(r_1+r_2)^n\log(2+r_1+r_2)}{|\det(u_1,v_2,\dots,v_n)|\,|t|^{2s}2^{k2s}}\\
&\qquad\qquad\times\left(\left\|\frac{\rho_2(r)}{|r|^{\frac12-s}}\right\|_{L^2}^2+\|(1+|r|^s)\rho_2(r)\|_{L^\infty}^2\right)\|g\|_{L^\infty}^2+\frac{Cr_1^{n-1}}{|t|2^{k2s}}\|\tilde\rho_2\|_{L^1}^2\,\|g\|_{L^\infty}^2.
\end{aligned}\tag{6-6}$$

Note that, when $n=2$, the proof is then finished for $|\det(u_1,v_2)|=1$. Therefore, when $n\ge 3$, there only remains to show that

$$\sup_{v_n\in\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}}\cdots\int_{\mathbb{S}^{n-1}}\frac{1}{|\det(u_1,v_2,\dots,v_n)|}\,dv_2\cdots dv_{n-1}<\infty,\tag{6-7}$$

which will clearly conclude the proof of the lemma upon integrating (6-6) in velocities $(v_2,\dots,v_{n-1})$ and combining the resulting estimate with (6-7).

In fact, the control (6-7) easily follows from a careful use of integration in spherical coordinates. Indeed, for each $2\le j\le n-1$ and any choice of orthonormal vectors $\{u_{j+1},\dots,u_n\}$, one has that (the unit vector u_1 is characterized here by the fact that it is orthogonal to the set $\{v_2,\dots,v_j,u_{j+1},\dots,u_n\}$)

$$\begin{aligned}
&\int_{\mathbb{S}^{n-1}}\cdots\int_{\mathbb{S}^{n-1}}\frac{1}{|\det(u_1,v_2,\dots,v_j,u_{j+1},\dots,u_n)|}\,dv_2\cdots dv_j\\
&\quad=\int_{\mathbb{S}^{j-1}\perp\{u_{j+1},\dots,u_n\}}\left(\int_{\mathbb{S}^{n-1}}\cdots\int_{\mathbb{S}^{n-1}}\frac{1}{|\det(u_1,v_2,\dots,v_{j-1},u_j,\dots,u_n)|}\,dv_2\cdots dv_{j-1}\right)du_j\\
&\qquad\qquad\times\int_0^\pi\cdots\int_0^\pi\sin^{n-3}\theta_n\cdots\sin^{j-2}\theta_{j+1}\,d\theta_{j+1}\cdots d\theta_n\\
&\quad\le C\sup_{\substack{u_j\in\mathbb{S}^{n-1}\\ u_j\cdot u_i=0\\ \text{for all } i=j+1,\dots,n}}\int_{\mathbb{S}^{n-1}}\cdots\int_{\mathbb{S}^{n-1}}\frac{1}{|\det(u_1,v_2,\dots,v_{j-1},u_j,\dots,u_n)|}\,dv_2\cdots dv_{j-1}.
\end{aligned}$$

Note that the unit vector u_1 above is also characterized by the fact that it is orthogonal to the set $\{v_2, \dots, v_{j-1}, u_j, \dots, u_n\}$. Hence, we deduce, for every $2 \le j \le n-1$, that

$$\begin{aligned}\sup_{\substack{u_{j+1},\dots,u_n\in\mathbb{S}^{n-1}\\ u_i\cdot u_k=0 \text{ if } i\neq k}} \int_{\mathbb{S}^{n-1}}\cdots\int_{\mathbb{S}^{n-1}} \frac{1}{|\det(u_1,v_2,\dots,v_j,u_{j+1},\dots,u_n)|}\,dv_2\cdots dv_j \\ \le C \sup_{\substack{u_j,\dots,u_n\in\mathbb{S}^{n-1}\\ u_i\cdot u_k=0 \text{ if } i\neq k}} \int_{\mathbb{S}^{n-1}\times\dots\times\mathbb{S}^{n-1}} \frac{1}{|\det(u_1,v_2,\dots,v_{j-1},u_j,\dots,u_n)|}\,dv_2\cdots dv_{j-1}.\end{aligned}$$

Applying now the preceding estimate $n-2$ times to reduce iteratively the number of integrations over spheres, we find that

$$\begin{aligned}\sup_{v_n\in\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}}\cdots\int_{\mathbb{S}^{n-1}} \frac{1}{|\det(u_1,v_2,\dots,v_n)|}\,dv_2\cdots dv_{n-1} \\ \le C \sup_{\substack{u_j,\dots,u_n\in\mathbb{S}^{n-1}\\ u_i\cdot u_k=0 \text{ if } i\neq k}} \int_{\mathbb{S}^{n-1}\times\dots\times\mathbb{S}^{n-1}} \frac{1}{|\det(u_1,v_2,\dots,v_{j-1},u_j,\dots,u_n)|}\,dv_2\cdots dv_{j-1} \\ \le C \sup_{\substack{u_2,\dots,u_n\in\mathbb{S}^{n-1}\\ u_i\cdot u_k=0 \text{ if } i\neq k}} \frac{1}{|\det(u_1,u_2,\dots,u_n)|},\end{aligned}$$

where the unit vector u_1 is orthogonal to $\{u_2, \dots, u_n\}$, which implies

$$|\det(u_1,u_2,\dots,u_n)| = 1,$$

and thus establishes (6-7). □

For convenience, we introduce now, for any integer $N \ge 2$, setting $S = (s_2, \dots, s_{N-1}) \in \mathbb{R}^{N-2}$ and $V = (v_1, \dots, v_{N-1}) \in (\mathbb{S}^{n-1})^{N-1}$, the following nonlinear operator:

$$I_N g = \int_{\mathbb{R}^n}\int_{\mathbb{S}^{n-1}} |S_t^* g(x,v_N)|^2 \left(\int_{[-r_1,r_1]^{N-2}}\int_{(\mathbb{S}^{n-1})^{N-1}} \left|S_t^* g\left(x+\sum_{j=2}^{N-1} s_j t v_j, v_1\right)\right|^2 dV\,dS\right) dv_N\,dx.$$

In particular, when $N = 2$, we have $I_2 g = \|S_t^* g\|^4_{L^4_x L^2_v}$.

Recall that, employing (5-5), it is possible to extract a one-dimensional integration from $S_t^* g(x, v_N)$ and $S_t^* g\big(x + \sum_{j=2}^{N-1} s_j t v_j, v_1\big)$ along v_N and v_1, respectively. Therefore, it is possible, at least formally, to gain an N-dimensional spatial integration in the above integrand by exploiting the integration along the variables s_j. Thus, loosely speaking, the number N represents the expected gain of spatial dimension on the domain of integration in I_N.

Prior to delving any further into our proofs, we take some time now to explain the general strategy behind the dimensional build up which will eventually allow us to apply Lemma 6.1 and establish the boundedness of $S_t^* : L^4_x \to W^{s,4}_x L^2_v$ for any $0 \le s < \frac{n}{4(n-1)}$, in Proposition 6.6, below.

More precisely, the aforementioned boundedness of S_t^* will be shown to follow from four properties of the nonlinear operator I_N:

• For $N=2$,

$$I_2 h = \|S_t^* h\|_{L_x^4 L_v^2}^4. \tag{6-8}$$

This property is a direct interpretation of the definition of I_2.

• For $N=n$ and any $0<s<\frac{1}{2}$, assuming for simplicity that $h(x)$ has frequencies localized inside an annulus of inner and outer radii comparable to 2^k, with $k\in\mathbb{N}$,

$$I_n h \le \frac{C}{2^{k4s}}\|h\|_{L_x^2}^2\|h\|_{L_x^\infty}^2, \tag{6-9}$$

where $C>0$ is independent of k. This estimate displays a gain of regularity, is a consequence of Lemma 6.1 and is established in Lemma 6.2, below.

• For any $N\ge 2$,

$$(I_N h)^2 \le \|S_t^* h\|_{L_x^4 L_v^2}^4 I_{2N-2} h, \tag{6-10}$$

which is a simple consequence of the Cauchy–Schwarz inequality (in x) followed by a careful change of variable. This estimate is established in Lemma 6.3, below.

• For any $N\ge 2$,

$$(I_N h)^2 \le C\|S^{2*} h\|_{L_{x,v}^4}^4 I_{2N-1} h, \tag{6-11}$$

where $C>0$ is an independent constant, which is a direct consequence of an application of (5-5) followed by a careful use of the Cauchy–Schwarz inequality with a change of variable. This estimate is established in Lemma 6.4, below.

The rule of the game of dimensional build up will then consist in employing estimates (6-10) and (6-11) to go from (6-8) to (6-9). In other words, by exploiting the mappings $N\mapsto 2N-2$ and $N\mapsto 2N-1$, for integers $N\ge 2$, we want to go from 2 to the dimension n. The fact that such a dimensional build up is actually possible is explained by the simple yet tricky Lemma 6.5, below.

Eventually, the appropriate combination of these estimates (and the handling of more technical difficulties) will give rise to the main result of this section, namely Theorem 6.8.

We proceed now with the actual preliminary results leading to Theorem 6.8.

For the sake of simplicity of notation, from now on, the variable S will denote the vector whose components are any number of integration variables $s_j\in[-r_1,r_1]$, whereas the variable V will denote the vector whose components are any number of integration variables $v_j\in\mathbb{S}^{n-1}$. At each step of our proofs, the exact meaning of S and V will be easily deduced from a careful inspection of the integrands and domains of integration.

Applying the preceding lemma combined with Proposition 5.1 to the above nonlinear operator I_N, when $N=n$ is the dimension, yields the following result.

Lemma 6.2. *In any dimension $n\ge 2$, let $0<s<\frac{1}{2}$ and $\rho_1,\rho_2\in\mathcal{S}(\mathbb{R})$ be such that $\tilde\rho_1$ and $\tilde\rho_2$ have their supports contained inside balls of radii $r_1,r_2>0$, respectively, centered at the origin.*

Then, it holds that, for any $k \in \mathbb{N}$,

$$I_n \Delta_{2^k} g \leq \frac{C(k+1)}{t^2 2^{k4s}} \|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 (r_1+r_2)^{n-1} (1+r_1+r_2)^2 \\ \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac{1}{2}-s}} \right\|_{L^2} + \|(1+|r|^s)\rho_2(r)\|_{L^\infty} + \|\tilde{\rho}_2\|_{L^1} \right)^4 \|g\|_{L_x^2}^2 \|g\|_{L_x^\infty}^2$$

for all $t \neq 0$ *such that* $|t| \leq 1$, *where* $C > 0$ *only depends on fixed parameters.*

Proof. In view of the simple estimate (5-5), it holds that

$$\begin{aligned} I_n \Delta_{2^k} g &\leq \|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} \int_{[-r_1,r_1]} |S_t^{2*} \Delta_{2^k} g(x+s_n t v_n, v_n)|^2 \, ds_n \\ &\quad \times \left(\int_{[-r_1,r_1]^{n-1}} \int_{(\mathbb{S}^{n-1})^{n-1}} \left| S_t^{2*} \Delta_{2^k} g\left(x + \sum_{j=1}^{n-1} s_j t v_j, v_1\right) \right|^2 dV\, dS \right) dv_n\, dx \\ &\leq \|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} |S_t^{2*} \Delta_{2^k} g(x, v_n)|^2 \\ &\quad \times \left(\int_{[-r_1,r_1]^{n}} \int_{(\mathbb{S}^{n-1})^{n-1}} \left| S_t^{2*} \Delta_{2^k} g\left(x + \sum_{j=1}^{n} s_j t v_j, v_1\right) \right|^2 dV\, dS \right) dv_n\, dx \\ &\leq \|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 \|S_t^{2*} \Delta_{2^k} g\|_{L_{x,v}^2}^2 \\ &\quad \times \sup_{\substack{x \in \mathbb{R}^n \\ v_n \in \mathbb{S}^{n-1}}} \int_{[-r_1,r_1]^{n}} \int_{(\mathbb{S}^{n-1})^{n-1}} \left| S_t^{2*} \Delta_{2^k} g\left(x + \sum_{j=1}^{n} s_j t v_j, v_1\right) \right|^2 dV\, dS. \end{aligned}$$

Therefore, applying Proposition 5.1 and Lemma 6.1 to the preceding estimate yields

$$I_n \Delta_{2^k} g \leq \frac{C(k+1)}{t^2 2^{k4s}} \|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 (r_1+r_2)^{n-1} (1+r_1+r_2)^2 \\ \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac{1}{2}-s}} \right\|_{L^2} + \|(1+|r|^s)\rho_2(r)\|_{L^\infty} + \|\tilde{\rho}_2\|_{L^1} \right)^4 \|g\|_{L_x^2}^2 \|g\|_{L_x^\infty}^2. \qquad \square$$

The next result explains how to increase the expected dimension of the domain of integration in the nonlinear operator I_N from N to $2N-2$.

Lemma 6.3. *In any dimension* $n \geq 2$, *it holds that, for any integer* $N \geq 2$,

$$(I_N g)^2 \leq \|S_t^* g\|_{L_x^4 L_v^2}^4 I_{2N-2} g.$$

Proof. First, by the Cauchy–Schwarz inequality, we find

$$\begin{aligned} I_N g &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{S}^{n-1}} |S_t^* g(x, v_N)|^2 \, dv_N \right) \left(\int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \left| S_t^* g\left(x + \sum_{j=2}^{N-1} s_j t v_j, v_1\right) \right|^2 dV\, dS \right) dx \\ &\leq \|S_t^* g\|_{L_x^4 L_v^2}^2 \left\| \int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \left| S_t^* g\left(x + \sum_{j=2}^{N-1} s_j t v_j, v_1\right) \right|^2 dV\, dS \right\|_{L_x^2}, \end{aligned}$$

whence

$$
\begin{aligned}
(I_N g)^2 \le \|S_t^* g\|^4_{L^4_x L^2_v} \int_{\mathbb{R}^n} \biggl(\int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \biggl| S_t^* g\biggl(x + \sum_{j=N}^{2N-3} s_j t v_j, v_{2N-2}\biggr)\biggr|^2 dV\, dS\biggr) \\
\times \biggl(\int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \biggl| S_t^* g\biggl(x + \sum_{j=2}^{N-1} s_j t v_j, v_1\biggr)\biggr|^2 dV\, dS\biggr) dx.
\end{aligned}
$$

Then, exploiting the integration in x to transfer the term $\sum_{j=N}^{2N-3} s_j t v_j$ in the above integrand, we deduce that

$$
\begin{aligned}
(I_N g)^2 \le \|S_t^* g\|^4_{L^4_x L^2_v} \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} |S_t^* g(x, v_{2N-2})|^2 \\
\times \biggl(\int_{[-r_1,r_1]^{2N-4}} \int_{(\mathbb{S}^{n-1})^{2N-3}} \biggl| S_t^* g\biggl(x + \sum_{j=2}^{2N-3} s_j t v_j, v_1\biggr)\biggr|^2 dV\, dS\biggr) dv_{2N-2}\, dx. \qquad \square
\end{aligned}
$$

The next result explains how to increase the expected dimension of the domain of integration in the nonlinear operator I_N from N to $2N-1$.

Lemma 6.4. *In any dimension $n \ge 2$, let $\rho_1 \in \mathcal{S}(\mathbb{R})$ be such that $\tilde\rho_1$ has its support contained inside a ball of radius $r_1 > 0$ centered at the origin.*

Then, it holds that, for any integer $N \ge 2$,

$$
(I_N g)^2 \le 4 \|\tilde\rho_1\|^2_{L^1} \, \|\tilde\rho_1\|^2_{L^\infty} \, r_1 \, \|S_t^{2*} g\|^4_{L^4_{x,v}} I_{2N-1} g.
$$

Proof. First, in view of the simple estimate (5-5), one has

$$
\begin{aligned}
I_N g &\le \|\tilde\rho_1\|_{L^1} \|\tilde\rho_1\|_{L^\infty} \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} \int_{[-r_1,r_1]} |S_t^{2*} g(x + s_N t v_N, v_N)|^2 \, ds_N \\
&\qquad \times \biggl(\int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \biggl| S_t^* g\biggl(x + \sum_{j=2}^{N-1} s_j t v_j, v_1\biggr)\biggr|^2 dV\, dS\biggr) dv_N\, dx \\
&= \|\tilde\rho_1\|_{L^1} \|\tilde\rho_1\|_{L^\infty} \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} |S_t^{2*} g(x, v_N)|^2 \\
&\qquad \times \biggl(\int_{[-r_1,r_1]^{N-1}} \int_{(\mathbb{S}^{n-1})^{N-1}} \biggl| S_t^* g\biggl(x + \sum_{j=2}^{N} s_j t v_j, v_1\biggr)\biggr|^2 dV\, dS\biggr) dv_N\, dx \\
&= \|\tilde\rho_1\|_{L^1} \|\tilde\rho_1\|_{L^\infty} \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} \bigl(|S_t^{2*} g(x, v_N)|^2 + |S_t^{2*} g(x, -v_N)|^2\bigr) \\
&\qquad \times \biggl(\int_{[0,r_1]} \int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \biggl| S_t^* g\biggl(x + \sum_{j=2}^{N} s_j t v_j, v_1\biggr)\biggr|^2 dV\, dS\, ds_N\biggr) dv_N\, dx.
\end{aligned}
$$

Hence, by the Cauchy–Schwarz inequality, we find

$$
\begin{aligned}
(I_N g)^2 &\leq 4\|\tilde{\rho}_1\|_{L^1}^2 \, \|\tilde{\rho}_1\|_{L^\infty}^2 \, \|S_t^{2*} g\|_{L^4_{x,v}}^4 \\
&\quad \times \left\| \int_{[0,r_1]} \int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \left| S_t^* g\Big(x + \sum_{j=2}^{N} s_j t v_j, v_1\Big)\right|^2 dV\, dS\, ds_N \right\|_{L^2_{x,v_N}}^2 \\
&= 4\|\tilde{\rho}_1\|_{L^1}^2 \, \|\tilde{\rho}_1\|_{L^\infty}^2 \, \|S_t^{2*} g\|_{L^4_{x,v}}^4 \\
&\quad \times \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} \left(\int_{[0,r_1]} \int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \left| S_t^* g\Big(x + \sum_{j=N}^{2N-2} s_j t v_j, v_{2N-1}\Big)\right|^2 dV\, dS\, ds_N \right) \\
&\qquad\quad \times \left(\int_{[0,r_1]} \int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-1}} \left| S_t^* g\Big(x + \sum_{j=2}^{N} s_j t v_j, v_1\Big)\right|^2 dV\, dS\, ds_N \right) dv_N\, dx.
\end{aligned}
$$

Finally, exploiting the integration in x to transfer first the term $s_N t v_N$ and then the term $\sum_{j=N+1}^{2N-2} s_j t v_j$ in the above integrand, we deduce that

$$
\begin{aligned}
(I_N g)^2 &\leq 4\|\tilde{\rho}_1\|_{L^1}^2 \, \|\tilde{\rho}_1\|_{L^\infty}^2 r_1 \|S_t^{2*} g\|_{L^4_{x,v}}^4 \\
&\quad \times \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} \left(\int_{[-r_1,r_1]^{N-2}} \int_{(\mathbb{S}^{n-1})^{N-2}} \left| S_t^* g\Big(x + \sum_{j=N+1}^{2N-2} s_j t v_j, v_{2N-1}\Big)\right|^2 dV\, dS \right) \\
&\qquad\quad \times \left(\int_{[-r_1,r_1]^{N-1}} \int_{(\mathbb{S}^{n-1})^{N}} \left| S_t^* g\Big(x + \sum_{j=2}^{N} s_j t v_j, v_1\Big)\right|^2 dV\, dS \right) dv_{2N-1}\, dx \\
&= 4\|\tilde{\rho}_1\|_{L^1}^2 \, \|\tilde{\rho}_1\|_{L^\infty}^2 r_1 \|S_t^{2*} g\|_{L^4_{x,v}}^4 \\
&\quad \times \int_{\mathbb{R}^n} \int_{\mathbb{S}^{n-1}} |S_t^* g(x, v_{2N-1})|^2 \\
&\qquad\quad \times \left(\int_{[-r_1,r_1]^{2N-3}} \int_{(\mathbb{S}^{n-1})^{2N-2}} \left| S_t^* g\Big(x + \sum_{j=2}^{2N-2} s_j t v_j, v_1\Big)\right|^2 dV\, dS \right) dv_{2N-1}\, dx. \qquad \square
\end{aligned}
$$

The following result is a simple technical lemma which, at first, may seem somewhat unrelated but will prove very useful later on for building up dimensions in the proof of Proposition 6.6.

Lemma 6.5. *Let the mappings $\Lambda_0, \Lambda_1 : \mathbb{N} \setminus \{0,1\} \to \mathbb{N} \setminus \{0,1\}$ be defined by*

$$\Lambda_0 k = 2k - 2 \quad \text{and} \quad \Lambda_1 k = 2k - 1.$$

Then, for any integer $n \geq 3$, there exists $L \in \mathbb{N}$ and $a_0, a_1, \ldots, a_L \in \{0,1\}$ such that

$$n = \Lambda_{a_0} \Lambda_{a_1} \cdots \Lambda_{a_L} 2,$$

and

$$n - 2 = \sum_{k=0}^{L} a_k 2^k.$$

Moreover, the above decomposition is unique provided $a_L = 1$.

Proof. We introduce first the auxiliary mappings $\widetilde{\Lambda}_0, \widetilde{\Lambda}_1 : \mathbb{N} \to \mathbb{N}$ defined by

$$\widetilde{\Lambda}_0 k = 2k \quad \text{and} \quad \widetilde{\Lambda}_1 k = 2k + 1.$$

In particular, for any $k \in \mathbb{N}$, it holds that

$$(\widetilde{\Lambda}_0 k) + 2 = \Lambda_0(k+2) \quad \text{and} \quad (\widetilde{\Lambda}_1 k) + 2 = \Lambda_1(k+2).$$

Next, let $L \in \mathbb{N}$ and $a_0, a_1, \dots, a_L \in \{0,1\}$ be the parameters appearing in the dyadic decomposition of the positive integer $n-2$:

$$n - 2 = \sum_{k=0}^{L} a_k 2^k.$$

Note that, assuming $a_L = 1$, the above choice of parameters is unique. Then, we have

$$n-2 = \widetilde{\Lambda}_{a_0}\left(\sum_{k=0}^{L-1} a_{k+1} 2^k\right) = \widetilde{\Lambda}_{a_0}\widetilde{\Lambda}_{a_1}\left(\sum_{k=0}^{L-2} a_{k+2} 2^k\right) = \cdots = \widetilde{\Lambda}_{a_0}\widetilde{\Lambda}_{a_1}\cdots\widetilde{\Lambda}_{a_L} 0.$$

It finally follows that

$$\begin{aligned} n &= 2 + \widetilde{\Lambda}_{a_0}\widetilde{\Lambda}_{a_1}\cdots\widetilde{\Lambda}_{a_L} 0 \\ &= \Lambda_{a_0}(2 + \widetilde{\Lambda}_{a_1}\cdots\widetilde{\Lambda}_{a_L} 0) \\ &= \Lambda_{a_0}\Lambda_{a_1}(2 + \widetilde{\Lambda}_{a_2}\cdots\widetilde{\Lambda}_{a_L} 0) = \cdots = \Lambda_{a_0}\Lambda_{a_1}\cdots\Lambda_{a_L} 2. \end{aligned}$$

□

Notice that, using the language of Lemma 6.5, it is possible to unify Lemmas 6.3 and 6.4 in the following estimate, for any $N \geq 2$:

$$(I_N g)^2 \leq (4\|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 r_1)^a \|S_t^* g\|_{L_x^4 L_v^2}^{4(1-a)} \|S_t^{2*} g\|_{L_{x,v}^4}^{4a} I_{\Lambda_a N} g, \tag{6-12}$$

where $a \in \{0,1\}$.

Now, appropriately combining Lemmas 6.2, 6.3 and 6.4, with the help of Lemma 6.5, we arrive at our main estimate on the operator S_t^*, which is recorded in the next proposition and generalizes Proposition 5.2 to higher dimensions.

Proposition 6.6. *In any dimension $n \geq 3$, let $2 \leq p \leq 4$, $0 \leq s < s_0 < \frac{1}{2}$ and $\rho_1, \rho_2 \in \mathcal{S}(\mathbb{R})$ be such that $\tilde{\rho}_1$ and $\tilde{\rho}_2$ have their supports contained inside balls of radii $r_1, r_2 > 0$, respectively, centered at the origin.*

Then, the operator S_t^* satisfies the estimate

$$\begin{aligned} \|(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}} S_t^* g\|_{L_x^p L_v^2} \leq \frac{C}{|t|^{\frac{n}{4(n-1)}}} \|\tilde{\rho}_1\|_{L^1}^{\frac{1}{2}} \|\tilde{\rho}_1\|_{L^\infty}^{\frac{1}{2}} (r_1+r_2)^{\frac{2n-3}{4(n-1)}} (1+r_1+r_2)^{\frac{1}{2(n-1)}} \\ \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac{1}{2}-s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde{\rho}_2\|_{L^1} \right) \|g\|_{L_x^p} \end{aligned}$$

for all $t \neq 0$ such that $|t| \leq 1$, where $C > 0$ only depends on fixed parameters.

Proof. First, notice that, for any $2 \le p \le \infty$,

$$\|S_t^* g\|_{L_x^p L_v^2} \le \|S_t^* g\|_{L_v^2 L_x^p} \le |\mathbb{S}^{n-1}|^{\frac{1}{2}} \|\tilde{\rho}\|_{L^1} \|g\|_{L_x^p}. \tag{6-13}$$

As for the regularity estimate, we employ the standard Littlewood–Paley dyadic frequency decomposition previously introduced to estimate $g_k = \Delta_{2^k} g$ for any $k \ge 0$.

To this end, we first decompose the dimension $n \ge 3$ according to Lemma 6.5,

$$n = \Lambda_{a_0}\Lambda_{a_1}\cdots\Lambda_{a_L} 2,$$

where $L \in \mathbb{N}$ and $a_0, a_1, \dots, a_L \in \{0,1\}$, and then apply successively estimate (6-12) to deduce that

$$\begin{aligned}
\|S_t^* g_k\|_{L_x^4 L_v^2}^{2^{L+1}} &= (I_2 g_k)^{2^{L-1}} \\
&\le (4\|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 r_1)^{a_L 2^{L-2}} \|S_t^* g_k\|_{L_x^4 L_v^2}^{(1-a_L)2^L} \|S_t^{2*} g_k\|_{L_{x,v}^4}^{a_L 2^L} (I_{\Lambda_{a_L} 2} g_k)^{2^{L-2}} \\
&\le (4\|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 r_1)^{a_{L-1} 2^{L-3} + a_L 2^{L-2}} \|S_t^* g_k\|_{L_x^4 L_v^2}^{(1-a_{L-1})2^{L-1} + (1-a_L)2^L} \\
&\qquad\qquad \times \|S_t^{2*} g_k\|_{L_{x,v}^4}^{a_{L-1}2^{L-1} + a_L 2^L} (I_{\Lambda_{a_{L-1}}\Lambda_{a_L} 2} g_k)^{2^{L-3}} \\
&\le \cdots \\
&\le (4\|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 r_1)^{\frac{1}{4}\sum_{j=0}^L a_j 2^j} \|S_t^* g_k\|_{L_x^4 L_v^2}^{\sum_{j=0}^L (1-a_j)2^j} \\
&\qquad\qquad \times \|S_t^{2*} g_k\|_{L_{x,v}^4}^{\sum_{j=0}^L a_j 2^j} (I_{\Lambda_{a_0}\Lambda_{a_1}\cdots\Lambda_{a_L} 2} g_k)^{\frac{1}{4}} \\
&= (4\|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 r_1)^{\frac{n-2}{4}} \|S_t^* g_k\|_{L_x^4 L_v^2}^{2^{L+1}-(n-1)} \|S_t^{2*} g_k\|_{L_{x,v}^4}^{n-2} (I_n g_k)^{\frac{1}{4}}.
\end{aligned}$$

Hence, it follows that

$$\begin{aligned}
\|S_t^* g_k\|_{L_x^4 L_v^2}^{n-1} &\le (4\|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 r_1)^{\frac{n-2}{4}} \|S_t^{2*} g_k\|_{L_{x,v}^4}^{n-2} (I_n g_k)^{\frac{1}{4}} \\
&\le (4\|\tilde{\rho}_1\|_{L^1}^2 \|\tilde{\rho}_1\|_{L^\infty}^2 r_1)^{\frac{n-2}{4}} \|S_t^{2*} g_k\|_{L_{x,v}^2}^{\frac{n-2}{2}} \|S_t^{2*} g_k\|_{L_{x,v}^\infty}^{\frac{n-2}{2}} (I_n g_k)^{\frac{1}{4}}.
\end{aligned}$$

Next, further applying Proposition 5.1 and Lemma 6.2 to the preceding bound yields

$$\begin{aligned}
\|S_t^* g_k\|_{L_x^4 L_v^2}^{n-1} \le \frac{C(k+1)^{\frac{1}{4}}}{|t|^{\frac{n}{4}} 2^{k\frac{n}{2}s}} \|\tilde{\rho}_1\|_{L^1}^{\frac{n-1}{2}} \|\tilde{\rho}_1\|_{L^\infty}^{\frac{n-1}{2}} (r_1+r_2)^{\frac{2n-3}{4}} (1+r_1+r_2)^{\frac{1}{2}} \\
\times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac{1}{2}-s}} \right\|_{L^2} + \|(1+|r|^s)\rho_2(r)\|_{L^\infty} + \|\tilde{\rho}_2\|_{L^1} \right)^{n-1} \|g\|_{L_x^2}^{\frac{n-1}{2}} \|g\|_{L_x^\infty}^{\frac{n-1}{2}},
\end{aligned}$$

where $C > 0$ is an independent constant.

The remainder of the demonstration follows the arguments from the end of the proof of Proposition 5.2, which we adapt to the present setting for completeness and convenience of the reader.

Thus, in order to conclude, we write $|g(x)| = \int_0^\infty \mathbb{1}_{\{|g(x)|\geq s\}}\,ds$ to deduce from the preceding estimate, assuming g is nonnegative, that

$$\begin{aligned}
\|S_t^* g_k\|_{L^4_x L^2_v} &\leq \int_0^\infty \|S_t^* \Delta_{2^k} \mathbb{1}_{\{|g(x)|\geq s\}}\|_{L^4_x L^2_v}\, ds \\
&\leq \frac{C(k+1)^{\frac{1}{4(n-1)}}}{|t|^{\frac{n}{4(n-1)}} 2^{k\frac{n}{2(n-1)}s}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac{2n-3}{4(n-1)}} (1+r_1+r_2)^{\frac{1}{2(n-1)}} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s}} \right\|_{L^2} + \|(1+|r|^s)\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \int_0^\infty |\{|g(x)|\geq s\}|^{\frac14}\, ds \\
&\leq \frac{C(k+1)^{\frac{1}{4(n-1)}}}{|t|^{\frac{n}{4(n-1)}} 2^{k\frac{n}{2(n-1)}s}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac{2n-3}{4(n-1)}} (1+r_1+r_2)^{\frac{1}{2(n-1)}} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s}} \right\|_{L^2} + \|(1+|r|^s)\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^{4,1}_x},
\end{aligned}$$

where $L^{4,1}_x$ denotes a standard Lorentz space; see [Bergh and Löfström 1976, Section 1.3] or [Grafakos 2008, Section 1.4] for definitions and properties of Lorentz spaces. When, g is signed, we arrive at the same estimate simply by decomposing $g = g^+ - g^-$ into its positive and negative parts, treating each contribution separately, and then noticing that

$$\begin{aligned}
\|g^+\|_{L^{4,1}_x} + \|g^-\|_{L^{4,1}_x} &\leq C \int_0^\infty \big|\{|g^+(x)|\geq s\}\big|^{\frac14} + \big|\{|g^-(x)|\geq s\}\big|^{\frac14}\, ds \\
&\leq C \int_0^\infty \big(\big|\{|g^+(x)|\geq s\}\big| + \big|\{|g^-(x)|\geq s\}\big|\big)^{\frac14}\, ds \\
&\leq C \int_0^\infty \big|\{|g(x)|\geq s\}\big|^{\frac14}\, ds \leq C\|g\|_{L^{4,1}_x}.
\end{aligned}$$

Moreover, by allowing an arbitrarily small loss of regularity, that is, by replacing $0 < s < \frac12$ by a slightly smaller value, it is possible to replace the Lorentz space $L^{4,1}_x$ by the standard Lebesgue space L^4_x in the right-hand side of the above estimate.

Therefore, on the whole, for any $0 \leq s < s_0 < \frac12$, we have established the estimate

$$\begin{aligned}
\|S_t^* g_k\|_{L^4_x L^2_v} &\leq \frac{C}{|t|^{\frac{n}{4(n-1)}} 2^{k\frac{n}{2(n-1)}s}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac{2n-3}{4(n-1)}} (1+r_1+r_2)^{\frac{1}{2(n-1)}} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^4_x},
\end{aligned}$$

where $C > 0$ only depends on constant parameters, which, when combined with the easy bound (6-13) for low frequencies, yields

$$\begin{aligned}
\|(1-\Delta_x)^{\frac{ns}{4(n-1)}} S_t^* g\|_{L^4_x L^2_v} &\leq \frac{C}{|t|^{\frac{n}{4(n-1)}}} \|\tilde\rho_1\|_{L^1}^{\frac12} \|\tilde\rho_1\|_{L^\infty}^{\frac12} (r_1+r_2)^{\frac{2n-3}{4(n-1)}} (1+r_1+r_2)^{\frac{1}{2(n-1)}} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac12 - s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde\rho_2\|_{L^1} \right) \|g\|_{L^4_x}. \qquad \text{(6-14)}
\end{aligned}$$

Finally, since S_t^* commutes with differentiation in x and recalling from complex interpolation theory, see [Bergh and Löfström 1976, Sections 5.1 and 6.4], that, for any $2 < p < 4$,

$$(L_x^2 L_v^2, L_x^4 L_v^2)_{[2-\frac{4}{p}]} = L_x^p L_v^2 \quad \text{and} \quad (W_x^{-s,2}, W_x^{-s\frac{n}{2(n-1)},4})_{[2-\frac{4}{p}]} = W_x^{-s\frac{2n+p-4}{p(n-1)},p},$$

we conclude the proof of the proposition by interpolating the estimate (6-14) with the classical estimate on S_t^* from Proposition 5.1. ☐

Next, we utilize the previous result on the adjoint operator S_t^* to deduce corresponding estimates on T_t and T_t^*.

Proposition 6.7. *In any dimension $n \geq 3$, let $\frac{4}{3} \leq r \leq 2$, $2 \leq p \leq 4$, $0 \leq s < s_0 < \frac{1}{2}$ and $\rho_1, \rho_2 \in \mathcal{S}(\mathbb{R})$ be such that $\tilde{\rho}_1$ and $\tilde{\rho}_2$ have their supports contained inside balls of radii $r_1, r_2 > 0$, respectively, centered at the origin.*

Then, the operators T_t and T_t^* satisfy the estimates

$$\begin{aligned}
&\|(1-\Delta_x)^{\frac{(2n+r'-4)s}{2r'(n-1)}} T_t f\|_{L_x^r} \\
&\quad \leq \frac{C}{|t|^{\frac{n}{4(n-1)}}} \|\varphi\|_{L^\infty} R^{\frac{n(2n-3)}{4(n-1)}} \|\tilde{\rho}_1\|_{L^1}^{\frac{1}{2}} \|\tilde{\rho}_1\|_{L^\infty}^{\frac{1}{2}} (r_1+r_2)^{\frac{2n-3}{4(n-1)}} (1+r_1+r_2)^{\frac{1}{2(n-1)}} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac{1}{2}-s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde{\rho}_2\|_{L^1} \right) \|f\|_{L_x^r L_v^2}
\end{aligned} \tag{6-15}$$

and

$$\begin{aligned}
&\|(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}} T_t^* g\|_{L_x^p L_v^2} \\
&\quad \leq \frac{C}{|t|^{\frac{n}{4(n-1)}}} \|\varphi\|_{L^\infty} R^{\frac{n(2n-3)}{4(n-1)}} \|\tilde{\rho}_1\|_{L^1}^{\frac{1}{2}} \|\tilde{\rho}_1\|_{L^\infty}^{\frac{1}{2}} (r_1+r_2)^{\frac{2n-3}{4(n-1)}} (1+r_1+r_2)^{\frac{1}{2(n-1)}} \\
&\qquad \times \left(\left\| \frac{\rho_2(r)}{|r|^{\frac{1}{2}-s_0}} \right\|_{L^2} + \|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty} + \|\tilde{\rho}_2\|_{L^1} \right) \|g\|_{L_x^p}
\end{aligned}$$

for all $t \neq 0$ such that $|t| \leq 1$, where $C > 0$ only depends on fixed parameters.

Proof. It is readily seen that

$$T_t^* g(x,v) = \int_{\mathbb{R}} g(x+stv)\tilde{\rho}(s)\,ds\,\varphi(v) = \varphi(v) S_{|v|t}^* g\left(x, \frac{v}{|v|}\right).$$

Therefore, for any $2 \leq p \leq \infty$, we compute in polar coordinates, recalling that the support of the velocity weight $\varphi \in L_c^\infty(\mathbb{R}^n)$ is contained inside a closed ball of radius $R > 0$ centered at the origin,

$$\begin{aligned}
&\|(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}} T_t^* g(x,v)\|_{L_x^p(\mathbb{R}^n; L_v^2(\mathbb{R}^n))} \\
&\qquad = \left\| \left(\int_{\mathbb{R}^n} |(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}} T_t^* g(x,v)|^2\,dv \right)^{\frac{1}{2}} \right\|_{L_x^p} \\
&\qquad = \left\| \int_0^\infty r^{n-1} \int_{\mathbb{S}^{n-1}} |\varphi(rv)(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}} S_{rt}^* g(x,v)|^2\,dv\,dr \right\|_{L_x^{p/2}}^{\frac{1}{2}}
\end{aligned}$$

$$\leq \|\varphi\|_{L^\infty}\left(\int_0^R r^{n-1}\left\|\int_{\mathbb{S}^{n-1}}|(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}}S_{rt}^* g(x,v)|^2\,dv\right\|_{L_x^{p/2}}dr\right)^{\frac{1}{2}}$$
$$\leq \|\varphi\|_{L^\infty}\left(\int_0^R r^{n-1}\|(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}}S_{rt}^* g(x,v)\|^2_{L_x^p(\mathbb{R}^n;L_v^2(\mathbb{S}^{n-1}))}\,dr\right)^{\frac{1}{2}}.$$

Then, combining Proposition 6.6 with the above estimate, we find

$$\begin{aligned}
&\|(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}}T_t^* g\|_{L_x^p L_v^2}\\
&\quad\leq \|\varphi\|_{L^\infty}\left(\int_0^R r^{n-1}\|(1-\Delta_x)^{\frac{(2n+p-4)s}{2p(n-1)}}S_{rt}^* g\|^2_{L_x^p L_v^2}\,dr\right)^{\frac{1}{2}}\\
&\quad\leq \frac{C}{|t|^{\frac{n}{4(n-1)}}}\|\varphi\|_{L^\infty}R^{\frac{n(2n-3)}{4(n-1)}}\|\tilde\rho_1\|_{L^1}^{\frac{1}{2}}\|\tilde\rho_1\|_{L^\infty}^{\frac{1}{2}}(r_1+r_2)^{\frac{2n-3}{4(n-1)}}(1+r_1+r_2)^{\frac{1}{2(n-1)}}\\
&\qquad\qquad\times\left(\left\|\frac{\rho_2(r)}{|r|^{\frac{1}{2}-s_0}}\right\|_{L^2}+\|(1+|r|^{s_0})\rho_2(r)\|_{L^\infty}+\|\tilde\rho_2\|_{L^1}\right)\|g\|_{L_x^p},
\end{aligned}$$

where $C>0$ is an independent constant, which establishes the estimate on T_t^*.

The estimate on T_t is then easily deduced from the estimate on T_t^* by a duality argument, which completes the proof of the proposition. □

We proceed now to the main theorem of this section.

Theorem 6.8. *In any dimension $n\geq 3$, let $\frac{4}{3}\leq r\leq 2$.*

Then, for any $f,g\in L_x^r(\mathbb{R}^n;L_v^2(\mathbb{R}^n))$ such that (1-2) *holds true, one has*

$$\int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\,dv\in W_x^{s,r}(\mathbb{R}^n)$$

for any $\varphi\in L_c^\infty(\mathbb{R}^n)$ and any

$$0\leq s<\frac{1}{2}\Big(3-\frac{4}{r}\Big)+\frac{n}{4(n-1)}\Big(\frac{4}{r}-2\Big).$$

Furthermore, one has the estimate

$$\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}^n} f\varphi\,dv\right\|_{L_x^r}\leq C(\|f\|_{L_x^r L_v^2}+\|g\|_{L_x^r L_v^2}),$$

where $C>0$ only depends on φ and constant parameters.

Proof. This demonstration follows the same ideas as the proof of Theorem 5.4. Nevertheless, for the sake of clarity and convenience of the reader, we provide a complete justification of this result.

We consider first the refined interpolation formula (2-1), which is valid for solutions of the transport equation (1-2), for some given cutoff $\rho\in\mathcal{S}(\mathbb{R})$. Clearly, further differentiating (2-1) in x and then averaging in v yields

$$\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}^n} f\varphi\,dv\right\|_{L_x^r}\leq\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}^n}A_t f\varphi\,dv\right\|_{L_x^r}+t\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}^n}B_t g\varphi\,dv\right\|_{L_x^r}. \tag{6-16}$$

We wish now to apply Proposition 6.7 to the preceding estimate. To this end, according to (5-4), we take the decompositions

$$\rho(r) = \rho_1(r)\rho_2(r) \quad \text{and} \quad \tau(r) = \frac{1-\rho(r)}{ir} = \tau_1(r)\tau_2(r),$$

where

$$\begin{aligned} \tilde{\rho}_1(r) &\in C_c^\infty(\mathbb{R}), \quad \tau_1(r) = \frac{1}{(1+r^2)^{\frac{1}{4}}}, \\ \tilde{\rho}_2(r) &\in C_c^\infty(\mathbb{R}), \quad \tau_2(r) = (1+r^2)^{\frac{1}{4}}\tau(r). \end{aligned}$$

Clearly, all constants involving norms of the cutoff functions ρ_1 and ρ_2 in the right-hand side of (5-11) are finite and we may therefore straightforwardly apply Proposition 6.7 to control the first term in the right-hand side of (6-16). However, the same is not so obviously true concerning the cutoff functions τ_1 and τ_2. The application of Proposition 6.7 to the second term in the right-hand side of (6-16) will therefore require some substantial technical work, which we present now.

To this end, we employ a homogeneous Littlewood–Paley frequency decomposition, see (5-3), of τ_1 and τ_2 to write that

$$\tau = \Big(\sum_{j\in\mathbb{Z}} \Delta_{2^j}\tau_1\Big)\Big(\sum_{k\in\mathbb{Z}} \Delta_{2^k}\tau_2\Big) = \sum_{j\in\mathbb{Z}} (\Delta_{2^j}\tau_1)\tau_3^j + \sum_{j\in\mathbb{Z}} (\Delta_{2^j}\tau_2)\tau_4^j,$$

where

$$\tau_3^j = \sum_{\substack{k\in\mathbb{Z}\\ k\le j}} \Delta_{2^k}\tau_2 = \mathcal{F}^{-1}\Big[\psi_0\Big(\frac{r}{2^{j+1}}\Big)\Big] * \tau_2, \quad \tau_4^j = \sum_{\substack{k\in\mathbb{Z}\\ k< j}} \Delta_{2^k}\tau_1 = \mathcal{F}^{-1}\Big[\psi_0\Big(\frac{r}{2^{j}}\Big)\Big] * \tau_1.$$

In view of the linearity of the operator T_t with respect to the cutoffs ρ or τ, we only need to verify the finiteness of the constants in (6-15) with $\Delta_{2^j}\tau_1$ and τ_3^j playing the roles of ρ_1 and ρ_2, respectively, and then with $\Delta_{2^j}\tau_2$ and τ_4^j instead of ρ_1 and ρ_2, respectively. It is to be emphasized here that the ensuing bounds on the cutoffs will then depend on $j \in \mathbb{Z}$. In order to guarantee the boundedness of T_t, we will therefore need to make sure that our method eventually yields constants that are summable in $j \in \mathbb{Z}$.

We evaluate now the norms involved in the right-hand side of (6-15) where we replace ρ_1 by $\Delta_{2^j}\tau_1$ (or $\Delta_{2^j}\tau_2$) and ρ_2 by τ_3^j (or τ_4^j). The bounds on $\Delta_{2^j}\tau_2$ and τ_4^j are handled in a strictly similar manner and so we omit the corresponding details.

First, note that a direct application of Lemma B.3 from Appendix B together with the fact that τ_1 and τ_2 are smooth so that their Fourier transforms decay faster than any inverse power at infinity, shows that

$$\begin{aligned} |\widetilde{\Delta_{2^j}\tau_1}(r)| &\le \frac{C}{|r|^{\frac{1}{2}}(1+|r|)^N}\mathbb{1}_{\{2^{j-1}\le|r|\le 2^{j+1}\}}, \\ |\tilde{\tau}_3^j(r)| &\le \frac{C}{|r|^{\frac{1}{2}}(1+|r|)^N}\mathbb{1}_{\{|r|\le 2^{j+1}\}}, \end{aligned} \tag{6-17}$$

for any arbitrarily large $N > 0$.

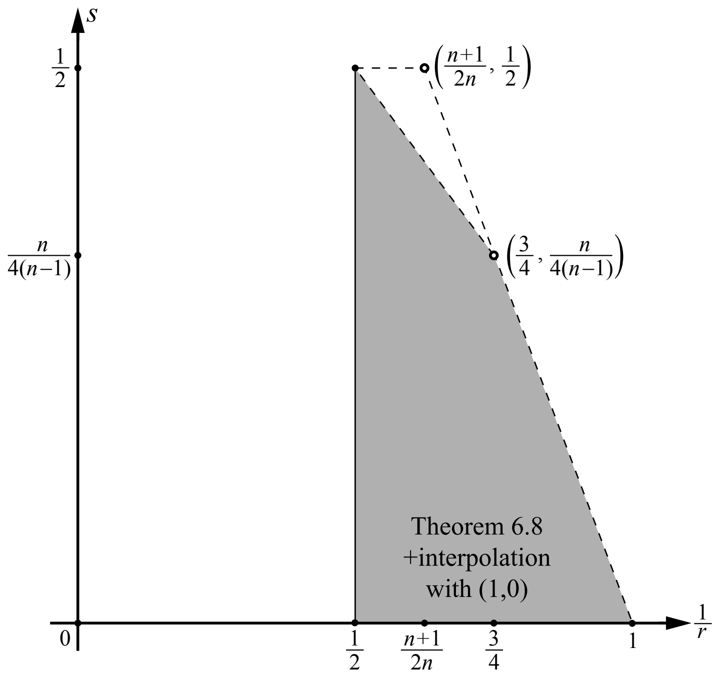

Figure 2. Range of validity of the parameters $\frac{1}{r}$ and s in Theorem 6.8 extended by interpolation with the degenerate L^1 case.

Furthermore, in view of Lemma B.4, it holds that each τ_3^j satisfies

$$|\tau_3^j(r)| \le \frac{C}{1+|r|^{\frac{1}{2}}}$$

for some uniform $C > 0$ independent of $j \in \mathbb{Z}$, whence, for any $0 < s_0 < \frac{1}{2}$,

$$\frac{\tau_3^j(r)}{|r|^{\frac{1}{2}-s_0}} \in L^2(\mathbb{R}) \quad \text{and} \quad (1+|r|^{s_0})\tau_3^j(r) \in L^\infty(\mathbb{R}), \tag{6-18}$$

uniformly in $j \in \mathbb{Z}$.

Therefore, using the bounds (6-17) and (6-18) to evaluate the terms involving $\rho_1 = \Delta_{2^j}\tau_1$ and $\rho_2 = \tau_3^j$ in the right-hand side of (6-15), we compute that the corresponding norm of the operator in (6-15) is no larger than a multiple of

$$\left(\frac{2^{\frac{j}{2}}}{1+2^{jN}}\right)^{\frac{1}{2}} \left(\frac{1}{2^{\frac{j}{2}}(1+2^{jN})}\right)^{\frac{1}{2}} 2^{j\frac{2n-3}{4(n-1)}}(1+2^j)^{\frac{1}{2(n-1)}} \le C\frac{2^{j\frac{2n-3}{4(n-1)}}}{(1+2^j)^{N-\frac{1}{2(n-1)}}},$$

which is summable over $j \in \mathbb{Z}$, provided $N > \frac{2n-1}{4(n-1)}$.

Thus, we conclude, according to Proposition 6.7, that the operators in the right-hand side of (6-16) are bounded.

It follows that, for any $0 < t < 1$,

$$\left\|(1-\Delta_x)^{\frac{s}{2}}\int_{\mathbb{R}^n} f\varphi\, dv\right\|_{L^r_x} \le C\left(\frac{1}{t^{\frac{n}{4(n-1)}}}\|f\|_{L^r_x L^2_v} + t^{1-\frac{n}{4(n-1)}}\|g\|_{L^r_x L^2_v}\right),$$

where $C > 0$ only depends on constant parameters. □

As already mentioned at the end of our Introduction, it is possible that Theorem 6.8 may be largely improved. In fact, the formal interpolation result (1-3) seems to indicate that Theorem 6.8 should hold for all parameters $\frac{2n}{n+1} \le r \le 2$ and $1 \le s < \frac{1}{2}$. The range of parameters defined by $\frac{3}{4} \le r \le 2$ and

$$0 \le s < \frac{1}{2}\left(3-\frac{4}{r}\right) + \frac{n}{4(n-1)}\left(\frac{4}{r}-2\right)$$

would then be recovered by interpolation with the degenerate L^1 case.

Indeed, Figure 2 represents the range of validity of the parameters $\frac{1}{r}$ and s in Theorem 6.8 extended by interpolation with the degenerate L^1 case. More precisely, Theorem 6.8 handles the region bounded by the points $\left(\frac{1}{2},0\right)$, $\left(\frac{3}{4},0\right)$, $\left(\frac{3}{4},\frac{n}{4(n-1)}\right)$ and $\left(\frac{1}{2},\frac{1}{2}\right)$, which yields the shaded region in Figure 2 when interpolated with the trivial L^1 case corresponding to the point $(1,0)$. Observe that the points $\left(\frac{n+1}{2n},\frac{1}{2}\right)$, $\left(\frac{3}{4},\frac{n}{4(n-1)}\right)$ and $(1,0)$ are all supported by the same line. It seems therefore natural to conjecture that a similar result should hold for all parameters encompassed by the area delimited by the points $\left(\frac{1}{2},0\right)$, $(1,0)$, $\left(\frac{n+1}{2n},\frac{1}{2}\right)$ and $\left(\frac{1}{2},\frac{1}{2}\right)$; see [Arsénio 2015] for more on such conjectures. This situation strongly resembles the corresponding existing conjectures for the boundedness of Bochner–Riesz multipliers and Fourier restriction operators.

Appendix A. Some paradifferential calculus

In this appendix, we record for reference a useful technical lemma. The proof of this lemma is based on classical methods from paradifferential calculus and paraproduct decompositions.

Lemma A.1. *Let $\chi_1 \in \mathcal{S}(\mathbb{R}^{n-1})$ and $\chi_2 \in \mathcal{S}(\mathbb{R})$. For each $i \in \mathbb{N}$ and $L > 0$, we define*

$$h_i^L(x) = \chi_1\left(\frac{x'}{L}\right)\chi_2\left(\frac{2^i x_n}{L}\right),$$

where $x' = (x_1,\dots,x_{n-1}) \in \mathbb{R}^{n-1}$. Further consider fixed parameters $s > 0$ and $0 < \lambda < 1$.

Then, for all $g \in L^\infty(\mathbb{R}^n)$, it holds that

$$\|(1-\Delta)^{-\frac{s}{2}}((\Delta_{2^k} g)h_i^L)\|_{L^2(\mathbb{R}^n)} \le \frac{CL^{\frac{n}{2}}}{2^{\frac{i}{2}}2^{ks}}\|g\|_{L^\infty(\mathbb{R}^n)}$$

for every $i,k \in \mathbb{N}$ such that $2^i \le L2^{\lambda k}$, where $C > 0$ is independent of i, k and L.

Proof. We first write a standard paraproduct decomposition (see (5-3) for the definition of dyadic blocks and the Littlewood–Paley decomposition):

$$(\Delta_{2^k} g)h_i^L = \Delta_{2^k} g\left(\Delta_0 h_i^L + \sum_{j=0}^{k-3}\Delta_{2^j} h_i^L\right) + \Delta_{2^k} g\sum_{j=k-2}^{k+2}\Delta_{2^j} h_i^L + \sum_{j=k+3}^{\infty}\Delta_{2^k} g\Delta_{2^j} h_i^L.$$

It is then easy to see that in the right-hand side above

(1) the first term has frequencies localized inside an annulus of inner radius 2^{k-2} and outer radius $9\cdot 2^{k-2}$,

(2) the second term has frequencies localized inside a ball of radius $5\cdot 2^{k+1}$,

(3) each summand in the third term has frequencies localized inside an annulus of inner radius 2^{j-2} and outer radius $9\cdot 2^{j-2}$.

Accordingly, we estimate that

$$
\begin{aligned}
&\|(1-\Delta)^{-\frac{s}{2}}((\Delta_{2^k} g)h_i^L)\|_{L^2}\\
&\qquad\le \frac{C}{2^{ks}}\|g\|_{L^\infty}\|h_i^L\|_{L^2} + C\|g\|_{L^\infty}\sum_{j=k-2}^{k+2}\|\Delta_{2^j} h_i^L\|_{L^2} + C\Big(\sum_{j=k+3}^{\infty}\frac{1}{2^{js}}\Big)\|g\|_{L^\infty}\|h_i^L\|_{L^2}\\
&\qquad\le \frac{CL^{\frac{n}{2}}}{2^{\frac{i}{2}}2^{ks}}\|g\|_{L^\infty} + C\|g\|_{L^\infty}\sum_{j=k-2}^{k+2}\|\Delta_{2^j} h_i^L\|_{L^2}.
\end{aligned}
\tag{A-1}
$$

There only remains to control the terms $\|\Delta_{2^j} h_i^L\|_{L^2}$ above. To this end, noticing that

$$
\begin{aligned}
\psi(\eta) &= \psi(\eta)(\mathbb{1}_{\{|\eta'|\ge\frac14\}} + \mathbb{1}_{\{|\eta'|<\frac14\}})\\
&\le \mathbb{1}_{\{\frac14\le|\eta'|\le 2\}}\mathbb{1}_{\{|\eta_n|\le 2\}} + \mathbb{1}_{\{|\eta'|<\frac14\}}\mathbb{1}_{\{\frac14\le|\eta_n|\le 2\}} \le \mathbb{1}_{\{\frac14\le|\eta'|\le 2\}} + \mathbb{1}_{\{\frac14\le|\eta_n|\le 2\}},
\end{aligned}
$$

where $\eta' = (\eta_1,\dots,\eta_{n-1})\in\mathbb{R}^{n-1}$, and using Plancherel's theorem, we obtain

$$
\begin{aligned}
\|\Delta_{2^j} h_i^L\|_{L^2} &= \frac{L^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}2^i}\left\|\psi\Big(\frac{\eta}{L2^j}\Big)\hat\chi_1(\eta')\hat\chi_2\Big(\frac{\eta_n}{2^i}\Big)\right\|_{L^2_\eta}\\
&\le \frac{L^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}2^i}\|\mathbb{1}_{\{L2^{j-2}\le|\eta'|\le L2^{j+1}\}}\hat\chi_1(\eta')\|_{L^2_{\eta'}}\left\|\hat\chi_2\Big(\frac{\eta_n}{2^i}\Big)\right\|_{L^2_{\eta_n}}\\
&\qquad\qquad + \frac{L^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}2^i}\|\hat\chi_1(\eta')\|_{L^2_{\eta'}}\left\|\mathbb{1}_{\{L2^{j-2}\le|\eta_n|\le L2^{j+1}\}}\hat\chi_2\Big(\frac{\eta_n}{2^i}\Big)\right\|_{L^2_{\eta_n}}\\
&\le C\frac{L^{\frac{n}{2}}(L2^j)^{\frac{n-1}{2}}}{2^{\frac{i}{2}}}\|\mathbb{1}_{\{\frac14\le|\eta'|\le 2\}}\hat\chi_1(L2^j\eta')\|_{L^2_{\eta'}}\\
&\qquad\qquad + C\frac{L^{\frac{n}{2}}(L2^j)^{\frac12}}{2^i}\left\|\mathbb{1}_{\{\frac14\le|\eta_n|\le 2\}}\hat\chi_2\Big(\frac{L2^j}{2^i}\eta_n\Big)\right\|_{L^2_{\eta_n}}.
\end{aligned}
$$

Hence, recalling that both $\hat\chi_1$ and $\hat\chi_2$ decay faster than any inverse power at infinity, we find, for any given large $N_1, N_2 > 0$,

$$
\begin{aligned}
\|\Delta_{2^j} h_i^L\|_{L^2} &\le C\frac{L^{\frac{n}{2}}}{2^{\frac{i}{2}}(L2^j)^{N_1-\frac{n-1}{2}}} + C\frac{L^{\frac{n}{2}}2^{(N_2-\frac12)i}}{2^{\frac{i}{2}}(L2^j)^{N_2-\frac12}}\\
&\le C\frac{L^{\frac{n}{2}}}{2^{\frac{i}{2}}}\bigg(\Big(\frac{2^i}{L2^j}\Big)^{N_1-\frac{n-1}{2}} + \Big(\frac{2^i}{L2^j}\Big)^{N_2-\frac12}\bigg)\\
&\le C\frac{L^{\frac{n}{2}}}{2^{\frac{i}{2}}}\bigg(\frac{1}{2^{(N_1-\frac{n-1}{2})(1-\lambda)k}} + \frac{1}{2^{(N_2-\frac12)(1-\lambda)k}}\bigg),
\end{aligned}
$$

so that, choosing N_1 and N_2 such that

$$N_1 - \frac{n-1}{2} = N_2 - \frac{1}{2} \geq \frac{s}{1-\lambda},$$

we get

$$\|\Delta_{2^j} h_i^L\|_{L^2} \leq C \frac{L^{\frac{n}{2}}}{2^{\frac{i}{2}} 2^{ks}}, \tag{A-2}$$

where $C > 0$ is independent of i, k and L.

On the whole, incorporating (A-2) into (A-1) yields

$$\|(1-\Delta)^{-\frac{s}{2}}((\Delta_{2^k} g) h_i^L)\|_{L^2} \leq \frac{C L^{\frac{n}{2}}}{2^{\frac{i}{2}} 2^{ks}} \|g\|_{L^\infty}. \qquad \square$$

Appendix B. Boundedness of Fourier transforms in L^p, with $1 \leq p < 2$

For reference, we show here a few handy criteria for establishing the boundedness in Lebesgue spaces L^p, with $1 \leq p < 2$, of Fourier transforms of given functions.

Lemma B.1. *Let $f(x) \in C^\alpha(\mathbb{R}^n)$, for some given $\alpha \in \mathbb{N}$, be such that*

$$\sup_{\substack{\gamma \in \mathbb{N}^n \\ |\gamma| \leq \alpha}} |x|^{|\gamma|} |\partial_x^\gamma f(x)| \leq \frac{C}{(1+|x|)^\lambda} \quad \textit{for all } x \in \mathbb{R}^n, \tag{B-1}$$

for some $\lambda > 0$.

Then, the Fourier transform $\hat{f}$ belongs to $L^p(\mathbb{R}^n)$ for any $1 \leq p < 2$ satisfying

$$\alpha > n\Big(\frac{1}{p} - \frac{1}{2}\Big) \quad \textit{and} \quad \lambda > n\Big(1 - \frac{1}{p}\Big). \tag{B-2}$$

In particular, for any given $1 \leq p < 2$ and any $\beta > \frac{1}{p'}$, the Fourier transform of $(1+|x|^2)^{-\frac{n\beta}{2}}$ belongs to $L^p(\mathbb{R}^n)$.

Proof. Let $\psi_0(x), \psi(x) \in C_c^\infty(\mathbb{R}^n)$ be compactly supported smooth cutoff functions, whose supports satisfy

$$\operatorname{supp} \psi_0 \subset \{|x| \leq 1\} \quad \text{and} \quad \operatorname{supp} \psi \subset \big\{\tfrac{1}{2} \leq |x| \leq 2\big\},$$

and such that

$$\psi_0(x) + \sum_{j=0}^{\infty} \psi\left(\frac{x}{2^j}\right) = 1 \quad \text{for all } x \in \mathbb{R}^n.$$

We define $g(\eta), h_j(\eta) \in \mathcal{S}(\mathbb{R}^n)$, for each $j \in \mathbb{N}$, by the inverse Fourier transforms

$$\tilde{g}(x) = \psi_0(x) f(x) \quad \text{and} \quad \tilde{h}_j(x) = \psi\left(\frac{x}{2^j}\right) f(x),$$

so that

$$\hat{f}(\eta) = g(\eta) + \sum_{j=0}^{\infty} h_j(\eta) \quad \text{in } \mathcal{S}'(\mathbb{R}^n). \tag{B-3}$$

Then, for any $1 \le p < 2$ satisfying (B-2), so that $(2p\alpha)/(2-p) > n$, and by Hölder's inequality, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} |g(\eta)|^p \, d\eta &= \int_{\mathbb{R}^n} (1+|\eta|)^{p\alpha} |g(\eta)|^p \frac{1}{(1+|\eta|)^{p\alpha}} \, d\eta \\
&\le \left(\int_{\mathbb{R}^n} (1+|\eta|)^{2\alpha} |g(\eta)|^2 \, d\eta \right)^{\frac{p}{2}} \left(\int_{\mathbb{R}^n} \frac{1}{(1+|\eta|)^{\frac{2p\alpha}{2-p}}} \, d\eta \right)^{1-\frac{p}{2}} \\
&\le C \left(\int_{\mathbb{R}^n} (1+|\eta|)^{2\alpha} |g(\eta)|^2 \, d\eta \right)^{\frac{p}{2}},
\end{aligned}$$

and, similarly,

$$\begin{aligned}
\int_{\mathbb{R}^n} |h_j(\eta)|^p \, d\eta &= \int_{\mathbb{R}^n} (1+2^j|\eta|)^{p\alpha} |h_j(\eta)|^p \frac{1}{(1+2^j|\eta|)^{p\alpha}} \, d\eta \\
&\le \left(\int_{\mathbb{R}^n} (1+2^j|\eta|)^{2\alpha} |h_j(\eta)|^2 \, d\eta \right)^{\frac{p}{2}} \left(\int_{\mathbb{R}^n} \frac{1}{(1+2^j|\eta|)^{\frac{2p\alpha}{2-p}}} \, d\eta \right)^{1-\frac{p}{2}} \\
&\le C 2^{-jn\left(1-\frac{p}{2}\right)} \left(\int_{\mathbb{R}^n} (1+2^j|\eta|)^{2\alpha} |h_j(\eta)|^2 \, d\eta \right)^{\frac{p}{2}},
\end{aligned}$$

where $C > 0$ only depends on p, α and the dimension.

Next, since

$$(1+2^j|\eta|)^\alpha \le C\left(1 + 2^{j\alpha} \sum_{i=1}^n |\eta_i|^\alpha\right),$$

it follows from Plancherel's theorem that

$$\|g(\eta)\|_{L^p} \le C\|(1+|\eta|)^\alpha g(\eta)\|_{L^2} \le C \sum_{\substack{\gamma\in\mathbb{N}^n \\ |\gamma|\le\alpha}} \|\partial_x^\gamma(\psi_0(x)f(x))\|_{L^2} \le C \sum_{\substack{\gamma\in\mathbb{N}^n \\ |\gamma|\le\alpha}} \|\mathbb{1}_{\{|x|\le 1\}}\partial_x^\gamma f(x)\|_{L^2} < \infty.$$

and, further using (B-1),

$$\begin{aligned}
2^{jn\left(\frac{1}{p}-\frac{1}{2}\right)} \|h_j(\eta)\|_{L^p} &\le C\|(1+2^j|\eta|)^\alpha h_j(\eta)\|_{L^2} \\
&\le C \sum_{\substack{\gamma\in\mathbb{N}^n \\ |\gamma|\le\alpha}} 2^{j|\gamma|} \left\| \partial_x^\gamma \left(\psi\left(\frac{x}{2^j}\right) f(x) \right) \right\|_{L^2} \\
&\le C \sum_{\substack{\gamma\in\mathbb{N}^n \\ |\gamma|\le\alpha}} 2^{j|\gamma|} \left\| \mathbb{1}_{\{2^{j-1}\le|x|\le 2^{j+1}\}} \partial_x^\gamma f(x) \right\|_{L^2} \\
&\le C \sum_{\substack{\gamma\in\mathbb{N}^n \\ |\gamma|\le\alpha}} \left\| \mathbb{1}_{\{2^{j-1}\le|x|\le 2^{j+1}\}} |x|^{|\gamma|} \partial_x^\gamma f(x) \right\|_{L^2} \le C \frac{2^{j\frac{n}{2}}}{1+2^{j\lambda}}.
\end{aligned}$$

Hence, for any large $N \in \mathbb{N}$, since $\lambda - n\left(1-\frac{1}{p}\right) > 0$,

$$\sup_{N\in\mathbb{N}} \left\| g(\eta) + \sum_{j=0}^N h_j(\eta) \right\|_{L^p} \le C \sum_{j=0}^\infty \frac{2^{jn\left(1-\frac{1}{p}\right)}}{1+2^{j\lambda}} < \infty.$$

Therefore, according to (B-3), we deduce that the tempered distribution $\hat{f}$ coincides with the weak limit of functions uniformly bounded in L^p, which implies that $\hat{f} \in L^p$ for any $1 \le p < 2$ satisfying (B-2). $\square$

Lemma B.2. *Let $f(x) \in C^\alpha(\mathbb{R}^n \setminus \{0\})$, for some given $\alpha \in \mathbb{N}$, be such that*

$$\sup_{\substack{\gamma\in\mathbb{N}^n\\|\gamma|\le\alpha}} |x|^{|\gamma|}|\partial_x^\gamma f(x)| \le \frac{C|x|^\sigma}{(1+|x|)^{\lambda+\sigma}} \quad \textit{for all } x \in \mathbb{R}^n \setminus \{0\}, \tag{B-4}$$

for some $\lambda > 0$ and $\sigma > -\lambda$.

Then, the Fourier transform $\hat{f}$ belongs to $L^p(\mathbb{R}^n)$ for any $1 \le p < 2$ satisfying

$$\alpha > n\Big(\frac{1}{p} - \frac{1}{2}\Big) \quad \textit{and} \quad \lambda > n\Big(1 - \frac{1}{p}\Big) > -\sigma. \tag{B-5}$$

In particular, for any given $1 \le p < 2$ and any $\beta > \frac{1}{p'} > -\delta$, the Fourier transform of $|x|^{n\delta}(1+|x|^2)^{-\frac{n(\beta+\delta)}{2}}$ belongs to $L^p(\mathbb{R}^n)$.

Proof. Let $\psi(x) \in C_c^\infty(\mathbb{R}^n)$ be a compactly supported smooth cutoff function whose support satisfies

$$\operatorname{supp}\psi \subset \big\{\tfrac{1}{2} \le |x| \le 2\big\},$$

and such that

$$\sum_{j\in\mathbb{Z}} \psi\Big(\frac{x}{2^j}\Big) = 1 \quad \text{for all } x \in \mathbb{R}^n \setminus \{0\}.$$

We define $h_j(\eta) \in \mathcal{S}(\mathbb{R}^n)$, for each $j \in \mathbb{Z}$, by the inverse Fourier transforms

$$\tilde{h}_j(x) = \psi\Big(\frac{x}{2^j}\Big) f(x),$$

so that

$$\hat{f}(\eta) = \sum_{j\in\mathbb{Z}} h_j(\eta) \quad \text{in } \mathcal{S}'(\mathbb{R}^n). \tag{B-6}$$

Then, for any $1 \le p < 2$ satisfying (B-5), so that $(2p\alpha)/(2-p) > n$, and by Hölder's inequality, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} |h_j(\eta)|^p \, d\eta &= \int_{\mathbb{R}^n} (1+2^j|\eta|)^{p\alpha}|h_j(\eta)|^p \frac{1}{(1+2^j|\eta|)^{p\alpha}} \, d\eta \\
&\le \Big(\int_{\mathbb{R}^n} (1+2^j|\eta|)^{2\alpha}|h_j(\eta)|^2 \, d\eta\Big)^{\frac{p}{2}} \Big(\int_{\mathbb{R}^n} \frac{1}{(1+2^j|\eta|)^{\frac{2p\alpha}{2-p}}} \, d\eta\Big)^{1-\frac{p}{2}} \\
&\le C 2^{-jn(1-\frac{p}{2})} \Big(\int_{\mathbb{R}^n} (1+2^j|\eta|)^{2\alpha}|h_j(\eta)|^2 \, d\eta\Big)^{\frac{p}{2}},
\end{aligned}$$

where $C > 0$ only depends on p, α and the dimension.

Next, since

$$(1+2^j|\eta|)^\alpha \le C\Big(1 + 2^{j\alpha}\sum_{i=1}^n |\eta_i|^\alpha\Big),$$

it follows from Plancherel's theorem and (B-4) that

$$\begin{aligned}
2^{jn(\frac{1}{p}-\frac{1}{2})}\|h_j(\eta)\|_{L^p} &\le C\|(1+2^j|\eta|)^\alpha h_j(\eta)\|_{L^2}\\
&\le C\sum_{\substack{\gamma\in\mathbb{N}^n\\|\gamma|\le\alpha}} 2^{j|\gamma|}\left\|\partial_x^\gamma\left(\psi\left(\frac{x}{2^j}\right)f(x)\right)\right\|_{L^2}\\
&\le C\sum_{\substack{\gamma\in\mathbb{N}^n\\|\gamma|\le\alpha}} 2^{j|\gamma|}\left\|\mathbb{1}_{\{2^{j-1}\le|x|\le 2^{j+1}\}}\partial_x^\gamma f(x)\right\|_{L^2}\\
&\le C\sum_{\substack{\gamma\in\mathbb{N}^n\\|\gamma|\le\alpha}} \left\|\mathbb{1}_{\{2^{j-1}\le|x|\le 2^{j+1}\}}|x|^{|\gamma|}\partial_x^\gamma f(x)\right\|_{L^2} \le C\frac{2^{j(\frac{n}{2}+\sigma)}}{1+2^{j(\lambda+\sigma)}}.
\end{aligned}$$

Hence, for any large $N\in\mathbb{N}$, since $\lambda-n\big(1-\frac{1}{p}\big)>0$ and $\sigma+n\big(1-\frac{1}{p}\big)>0$,

$$\sup_{N\in\mathbb{N}}\left\|\sum_{j=-N}^{N}h_j(\eta)\right\|_{L^p}\le C\sum_{j\in\mathbb{Z}}\frac{2^{j(n(1-\frac{1}{p})+\sigma)}}{1+2^{j(\lambda+\sigma)}}<\infty.$$

Therefore, according to (B-6), we deduce that the tempered distribution $\hat f$ coincides with the weak limit of functions uniformly bounded in L^p, which implies that $\hat f\in L^p$ for any $1\le p<2$ satisfying (B-5). □

Lemma B.3. *Let $f(x)\in C^1(\mathbb{R}\setminus\{0\})$ be such that*

$$|f(x)|,\ |xf'(x)|\le\frac{C}{|x|^\lambda}\quad \textit{for all } x\in\mathbb{R}^n\setminus\{0\},$$

for some $0<\lambda<1$.

Then, the Fourier transform $\hat f$ belongs to L^1+L^∞ and satisfies

$$|\hat f(\eta)|\le\frac{C}{|\eta|^{1-\lambda}}\quad \textit{for almost every } \eta\in\mathbb{R}^n,$$

for some independent constant $C>0$.

Proof. Consider a cutoff $\chi\in C_c^\infty(\mathbb{R})$ such that $\mathbb{1}_{\{|x|\le1\}}\le\chi(x)\le\mathbb{1}_{\{|x|\le2\}}$. Then, on the one hand, the function $\chi(x)f(x)$ clearly is integrable so that its Fourier transform is bounded pointwise almost everywhere. On the other hand, the function $(1-\chi)(x)f(x)$ clearly verifies the hypotheses of Lemma B.1 so that its Fourier transform always coincides with an integrable function. This establishes that $\hat f\in L^1+L^\infty$.

Next, for any $t>0$, we have the estimate

$$\begin{aligned}
|\hat f(\eta)| &= \left|\int_{\mathbb{R}}e^{-i\eta x}f(x)\,dx\right|\\
&\le\left|\int_{\mathbb{R}}e^{-i\eta x}\chi\left(\frac{x}{t}\right)f(x)\,dx\right|+\left|\int_{\mathbb{R}}e^{-i\eta x}(1-\chi)\left(\frac{x}{t}\right)f(x)\,dx\right|
\end{aligned}$$

$$
\begin{aligned}
&\leq \int_{\{|x|\leq 2t\}} |f(x)|\,dx + \frac{1}{|\eta|}\left|\int_{\mathbb{R}} e^{-i\eta x}\left((1-\chi)\left(\frac{x}{t}\right)f(x)\right)' dx\right| \\
&\leq C\int_{\{|x|\leq 2t\}} \frac{1}{|x|^{\lambda}}\,dx + \frac{C}{t|\eta|}\int_{\{t\leq |x|\leq 2t\}} \frac{1}{|x|^{\lambda}}\,dx + \frac{C}{|\eta|}\int_{\{|x|\geq t\}} \frac{1}{|x|^{\lambda+1}}\,dx \leq C\left(t^{1-\lambda} + \frac{1}{|\eta| t^{\lambda}}\right).
\end{aligned}
$$

Therefore, optimizing the preceding estimate in $t > 0$, which amounts to setting $t = \frac{1}{|\eta|}$ above, yields

$$
|\hat{f}(\eta)| \leq \frac{C}{|\eta|^{1-\lambda}}. \qquad \square
$$

Lemma B.4. *Let $f \in L^{\infty}(\mathbb{R})$ be such that*

$$
|f(x)| \leq \frac{C}{1+|x|^{\alpha}} \quad \textit{for all } x \in \mathbb{R},
$$

for some $0 \leq \alpha < 1$, and consider the convolution

$$
f_R(x) = \int_{\mathbb{R}} R\chi(R(x-y))\, f(y)\, dy
$$

for any $R > 0$, where $\chi \in \mathcal{S}(\mathbb{R})$.

Then, the convolution f_R also satisfies

$$
|f_R(x)| \leq \frac{C}{1+|x|^{\alpha}} \quad \textit{for all } x \in \mathbb{R},
$$

for some constant $C > 0$ independent of R.

Proof. Note first that

$$
\|f_R\|_{L^\infty} \leq \|\chi\|_{L^1}\,\|f\|_{L^\infty}.
$$

Therefore, we only have to consider values $|x| \geq 1$, say. Furthermore, by possibly replacing χ and f by $|\chi|$ and $|f|$, respectively, we may assume that χ and f are both nonnegative.

Then, for any $N > 1$, we estimate that

$$
\begin{aligned}
f_R(x) &= \int_{\{|y|\geq \frac{|x|}{2}\}} R\chi(R(x-y))\, f(y)\, dy + \int_{\{|y|< \frac{|x|}{2}\}} R\chi(R(x-y))\, f(y)\, dy \\
&\leq \int_{\{|y|\geq \frac{|x|}{2}\}} R\chi(R(x-y)) \frac{C}{1+|y|^{\alpha}}\, dy + \int_{\{|y|< \frac{|x|}{2}\}} \frac{CR}{1+R^N|x-y|^N} f(y)\, dy \\
&\leq \frac{C}{1+|x|^{\alpha}} \int_{\{|y|\geq \frac{|x|}{2}\}} R\chi(R(x-y))\, dy + \frac{CR}{1+R^N|x|^N} \int_{\{|y|< \frac{|x|}{2}\}} \frac{1}{1+|y|^{\alpha}}\, dy \\
&\leq \frac{C}{1+|x|^{\alpha}} + \frac{CR}{1+R^N|x|^N} |x|^{1-\alpha}.
\end{aligned}
$$

Further noticing that

$$
\sup_{R>0} \frac{R}{1+R^N|x|^N} \leq \max\left\{ \sup_{0<R\leq \frac{1}{|x|}} R,\ \sup_{R>\frac{1}{|x|}} \frac{1}{R^{N-1}|x|^N} \right\} \leq \frac{1}{|x|},
$$

we deduce

$$
f_R(x) \leq \frac{C}{|x|^{\alpha}}. \qquad \square
$$

Appendix C. Velocities restricted to a manifold of codimension 1

In this final independent appendix section, we explore the connection between averaging lemmas for the stationary transport equation (1-2) for velocities in the Euclidean space $v \in \mathbb{R}^n$ and averaging lemmas for the same equation for velocities lying in an appropriate manifold of codimension 1 in $\mathbb{R}^n$. Here, for simplicity, we only consider the case $v \in \mathbb{S}^{n-1}$. However, the elementary methods developed here can be used to establish similar connections with settings in other manifolds of codimension 1. In particular, this approach includes the time dependent case (1-1) where $(t,x) \in \mathbb{R}^{n+1}$ and $v \in \mathbb{R}^n$ and, thus, allows us to translate the main results contained in this work to several other interesting and relevant situations.

Proposition C.1. *Let $n \geq 2$, $s > 0$ and $1 \leq p, q, r \leq \infty$ be such that*

$$p \leq r, \quad s + n\Big(\frac{1}{p} - \frac{1}{r}\Big) \leq 1 \quad \textit{and} \quad \frac{1}{q} + s + \Big(\frac{1}{p} - \frac{1}{r}\Big) \leq 1, \tag{C-1}$$

and suppose that, for any $\varphi \in L^\infty_c(\mathbb{R}^n)$, there exists $C > 0$ such that one has the estimate

$$\Big\| \int_{\mathbb{R}^n} f\varphi \, dv \Big\|_{W^{s,r}_x} \leq C(\|f\|_{L^p_x L^q_v} + \|g\|_{L^p_x L^q_v}) \tag{C-2}$$

for any $f, g \in L^p_x(\mathbb{R}^n; L^q_v(\mathbb{R}^n))$ such that (1-2) holds true.

Then, for some other constant $C > 0$, one has the estimate

$$\Big\| \int_{\mathbb{S}^{n-1}} f \, dv \Big\|_{W^{s,r}_x} \leq C(\|f\|_{L^p_x L^q_v} + \|g\|_{L^p_x L^q_v}) \tag{C-3}$$

for any $f, g \in L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))$ such that (1-2) holds true.

It is possible to show, though a dimensional analysis, that the restrictions (C-1) are in fact necessary in order that (C-2) may hold; see [Arsénio 2015, Section 4] for details.

Proof. We employ a strategy similar to the one used in [Arsénio and Saint-Raymond 2011, Appendix C] to go from the stationary case to a time-dependent setting. To this end, for any $f, g \in L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))$ such that (1-2) holds true, we introduce an artificial radial dimension by defining, for all $(x,v) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\tilde{f}(x,v) = f\Big(x, \frac{v}{|v|}\Big)\chi(|v|) \quad \text{and} \quad \tilde{g}(x,v) = |v| g\Big(x, \frac{v}{|v|}\Big)\chi(|v|)$$

for some given nonnegative cutoff function $\chi \in L^\infty_c(\mathbb{R})$.

Assuming that $\varphi(v) \equiv 1$ on the support of $\chi(|v|)$, it is then readily seen that

$$\int_{\mathbb{R}^n} \tilde{f}\varphi \, dv = \int_0^\infty \chi(r) r^{n-1} \, dr \int_{\mathbb{S}^{n-1}} f \, dv,$$

and

$$\begin{aligned} \|\tilde{f}\|_{L^q_v(\mathbb{R}^n)} &= \|r^{\frac{n-1}{q}} \chi(r)\|_{L^q_r([0,\infty))} \|f\|_{L^q_v(\mathbb{S}^{n-1})}, \\ \|\tilde{g}\|_{L^q_v(\mathbb{R}^n)} &= \|r^{\frac{n-1}{q}+1} \chi(r)\|_{L^q_r([0,\infty))} \|g\|_{L^q_v(\mathbb{S}^{n-1})}. \end{aligned}$$

Further observe that (1-2) also holds with $\tilde{f}$ and $\tilde{g}$ in place of f and g, respectively. Therefore, by plugging $\tilde{f}$ and $\tilde{g}$ into (C-2), we deduce the validity of estimate (C-3). □

A converse to the preceding proposition is also available.

Proposition C.2. *Let $n \geq 2$, $s > 0$ and $1 \leq p, q, r \leq \infty$ be such that*

$$p \leq r, \quad s + n\Big(\frac{1}{p} - \frac{1}{r}\Big) \leq 1 \quad \text{and} \quad \frac{1}{q} + s + \Big(\frac{1}{p} - \frac{1}{r}\Big) \leq 1, \tag{C-4}$$

and suppose that there exists $C > 0$ such that one has the estimate

$$\Big\| \int_{\mathbb{S}^{n-1}} f\, dv \Big\|_{W^{s,r}_x} \leq C(\|f\|_{L^p_x L^q_v} + \|g\|_{L^p_x L^q_v}) \tag{C-5}$$

for any $f, g \in L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))$ such that (1-2) *holds true.*

Then, for any $\varphi \in L^\infty_c(\mathbb{R}^n)$, there exists $C > 0$ such that one has the estimate

$$\begin{aligned} \Big\| \int_{\mathbb{R}^n} f\varphi\, dv \Big\|_{W^{s,r}_x} &\leq C(\|f\|_{L^p_x L^q_v} + \|g\|_{L^p_x L^q_v}) \quad \text{if } p \leq q, \\ \Big\| \int_{\mathbb{R}^n} f\varphi\, dv \Big\|_{W^{s,r}_x} &\leq C(\|f\|_{L^q_v L^p_x} + \|g\|_{L^q_v L^p_x}) \quad \text{if } p \geq q \end{aligned} \tag{C-6}$$

for any $f, g \in L^p_x(\mathbb{R}^n; L^q_v(\mathbb{R}^n))$ if $p \leq q$, or $f, g \in L^q_v(\mathbb{R}^n; L^p_x(\mathbb{R}^n))$ if $p \leq q$, such that (1-2) *holds true.*

Proof. We first assume that $p \leq q$. For any $f, g \in L^p_x(\mathbb{R}^n; L^q_v(\mathbb{R}^n))$ such that (1-2) holds true, we define, for all $(x, v) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$, $\lambda > 0$ and $\varphi \in L^\infty_c(\mathbb{R}^n)$,

$$\tilde{f}_\lambda(x, v) = f(\lambda x, \lambda v)\, \varphi(\lambda v) \quad \text{and} \quad \tilde{g}_\lambda(x, v) = g(\lambda x, \lambda v)\, \varphi(\lambda v).$$

It is then readily seen that

$$\begin{aligned} \Big\| \int_{\mathbb{S}^{n-1}} \tilde{f}_\lambda\, dv \Big\|_{L^r_x} &= \lambda^{-\frac{n}{r}} \Big\| \int_{\mathbb{S}^{n-1}} f(x, \lambda v)\, \varphi(\lambda v)\, dv \Big\|_{L^r_x}, \\ \Big\| \int_{\mathbb{S}^{n-1}} \tilde{f}_\lambda\, dv \Big\|_{\dot{W}^{s,r}_x} &= \lambda^{s-\frac{n}{r}} \Big\| \int_{\mathbb{S}^{n-1}} f(x, \lambda v)\, \varphi(\lambda v)\, dv \Big\|_{\dot{W}^{s,r}_x} \end{aligned}$$

and

$$\begin{aligned} \|\tilde{f}_\lambda\|_{L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))} &= \lambda^{-\frac{n}{p}} \|f(x, \lambda v)\, \varphi(\lambda v)\|_{L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))}, \\ \|\tilde{g}_\lambda\|_{L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))} &= \lambda^{-\frac{n}{p}} \|g(x, \lambda v)\, \varphi(\lambda v)\|_{L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))}. \end{aligned}$$

Further observe that (1-2) also holds with $\tilde{f}_\lambda$ and $\tilde{g}_\lambda$ in place of f and g, respectively. Therefore, by plugging $\tilde{f}$ and $\tilde{g}$ into (C-5), we deduce that

$$\begin{aligned} &\lambda^{n(\frac{1}{p}-\frac{1}{r})} \Big\| \int_{\mathbb{S}^{n-1}} f(x, \lambda v)\, \varphi(\lambda v)\, dv \Big\|_{L^r_x} + \lambda^{s+n(\frac{1}{p}-\frac{1}{r})} \Big\| \int_{\mathbb{S}^{n-1}} f(x, \lambda v)\, \varphi(\lambda v)\, dv \Big\|_{\dot{W}^{s,r}_x} \\ &\qquad \leq C\big(\|f(x, \lambda v)\, \varphi(\lambda v)\|_{L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))} + \|g(x, \lambda v)\, \varphi(\lambda v)\|_{L^p_x(\mathbb{R}^n; L^q_v(\mathbb{S}^{n-1}))}\big). \end{aligned}$$

Next, recalling that φ is compactly supported within some large ball $B(0,R)$, say, noticing that $\lambda^{\frac{n-1}{q'}-s-n\left(\frac{1}{p}-\frac{1}{r}\right)} \in L^{q'}_\lambda([0,R])$, by (C-4), and then integrating the preceding estimate in λ over $[0,R]$, we find

$$\begin{aligned}
\left\| \int_{\mathbb{R}^n} f(x,v)\,\varphi(v)\,dv \right\|_{W^{s,r}_x} &\le \int_0^R \lambda^{n-1} \left\| \int_{\mathbb{S}^{n-1}} f(x,\lambda v)\,\varphi(\lambda v)\,dv \right\|_{W^{s,r}_x} d\lambda \\
&\le C \int_0^R \lambda^{n-1-s-n\left(\frac{1}{p}-\frac{1}{r}\right)} \| f(x,\lambda v)\,\varphi(\lambda v)\|_{L^p_x(\mathbb{R}^n;L^q_v(\mathbb{S}^{n-1}))}\,d\lambda \\
&\qquad + C \int_0^R \lambda^{n-1-s-n\left(\frac{1}{p}-\frac{1}{r}\right)} \| g(x,\lambda v)\,\varphi(\lambda v)\|_{L^p_x(\mathbb{R}^n;L^q_v(\mathbb{S}^{n-1}))}\,d\lambda \\
&\le C \left(\int_0^R \| f(x,\lambda v)\,\varphi(\lambda v)\|^q_{L^p_x(\mathbb{R}^n;L^q_v(\mathbb{S}^{n-1}))} \lambda^{n-1}\,d\lambda \right)^{\frac{1}{q}} \\
&\qquad + C \left(\int_0^R \| g(x,\lambda v)\,\varphi(\lambda v)\|^q_{L^p_x(\mathbb{R}^n;L^q_v(\mathbb{S}^{n-1}))} \lambda^{n-1}\,d\lambda \right)^{\frac{1}{q}} \\
&\le C \big(\|f\|_{L^p_x(\mathbb{R}^n;L^q_v(\mathbb{R}^n))} + \|g\|_{L^p_x(\mathbb{R}^n;L^q_v(\mathbb{R}^n))} \big),
\end{aligned}$$

which concludes the proof of (C-6), when $p \le q$.

The case $p \ge q$ is obtained similarly. □

References

[Agoshkov 1984] V. I. Agoshkov, "Spaces of functions with differential-difference characteristics and the smoothness of solutions of the transport equation", *Dokl. Akad. Nauk SSSR* **276**:6 (1984), 1289–1293. In Russian; translated in *Sov. Math. Dokl.* **29** (1984), 662–666. MR Zbl

[Arsénio 2015] D. Arsénio, "Recent progress in velocity averaging", exposé no. 1, 17 pp. in *Journées équations aux dérivées partielles* (Roscoff, France, 2015), Cedram, Grenoble, 2015.

[Arsénio and Masmoudi 2014] D. Arsénio and N. Masmoudi, "A new approach to velocity averaging lemmas in Besov spaces", *J. Math. Pures Appl.* (9) **101**:4 (2014), 495–551. MR Zbl

[Arsénio and Saint-Raymond 2011] D. Arsénio and L. Saint-Raymond, "Compactness in kinetic transport equations and hypoellipticity", *J. Funct. Anal.* **261**:10 (2011), 3044–3098. MR Zbl

[Bergh and Löfström 1976] J. Bergh and J. Löfström, *Interpolation spaces: an introduction*, Grundlehren der Mathematischen Wissenschaften **223**, Springer, 1976. MR Zbl

[Bézard 1994] M. Bézard, "Régularité L^p précisée des moyennes dans les équations de transport", *Bull. Soc. Math. France* **122**:1 (1994), 29–76. MR Zbl

[Bouchut 2002] F. Bouchut, "Hypoelliptic regularity in kinetic equations", *J. Math. Pures Appl.* (9) **81**:11 (2002), 1135–1159. MR Zbl

[Bournaveas and Perthame 2001] N. Bournaveas and B. Perthame, "Averages over spheres for kinetic transport equations; hyperbolic Sobolev spaces and Strichartz inequalities", *J. Math. Pures Appl.* (9) **80**:5 (2001), 517–534. MR Zbl

[Castella and Perthame 1996] F. Castella and B. Perthame, "Estimations de Strichartz pour les équations de transport cinétique", *C. R. Acad. Sci. Paris Sér. I Math.* **322**:6 (1996), 535–540. MR Zbl

[DeVore and Petrova 2001] R. DeVore and G. Petrova, "The averaging lemma", *J. Amer. Math. Soc.* **14**:2 (2001), 279–296. MR Zbl

[DiPerna and Lions 1989] R. J. DiPerna and P.-L. Lions, "On the Cauchy problem for Boltzmann equations: global existence and weak stability", *Ann. of Math.* (2) **130**:2 (1989), 321–366. MR Zbl

[DiPerna et al. 1991] R. J. DiPerna, P.-L. Lions, and Y. Meyer, "L^p regularity of velocity averages", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **8**:3-4 (1991), 271–287. MR Zbl

[Golse and Perthame 2013] F. Golse and B. Perthame, "Optimal regularizing effect for scalar conservation laws", *Rev. Mat. Iberoam.* **29**:4 (2013), 1477–1504. MR Zbl

[Golse and Saint-Raymond 2004] F. Golse and L. Saint-Raymond, "The Navier–Stokes limit of the Boltzmann equation for bounded collision kernels", *Invent. Math.* **155**:1 (2004), 81–161. MR Zbl

[Golse et al. 1985] F. Golse, B. Perthame, and R. Sentis, "Un résultat de compacité pour les équations de transport et application au calcul de la limite de la valeur propre principale d'un opérateur de transport", *C. R. Acad. Sci. Paris Sér. I Math.* **301**:7 (1985), 341–344. MR Zbl

[Golse et al. 1988] F. Golse, P.-L. Lions, B. Perthame, and R. Sentis, "Regularity of the moments of the solution of a transport equation", *J. Funct. Anal.* **76**:1 (1988), 110–125. MR Zbl

[Grafakos 2008] L. Grafakos, *Classical Fourier analysis*, 2nd ed., Graduate Texts in Mathematics **249**, Springer, 2008. MR Zbl

[Grafakos 2009] L. Grafakos, *Modern Fourier analysis*, 2nd ed., Graduate Texts in Mathematics **250**, Springer, 2009. MR Zbl

[Jabin and Vega 2003] P.-E. Jabin and L. Vega, "Averaging lemmas and the X-ray transform", *C. R. Math. Acad. Sci. Paris* **337**:8 (2003), 505–510. MR Zbl

[Jabin and Vega 2004] P.-E. Jabin and L. Vega, "A real space method for averaging lemmas", *J. Math. Pures Appl.* (9) **83**:11 (2004), 1309–1351. MR Zbl

[Lions 1995] P.-L. Lions, "Régularité optimale des moyennes en vitesses", *C. R. Acad. Sci. Paris Sér. I Math.* **320**:8 (1995), 911–915. MR Zbl

[Lions et al. 1994a] P.-L. Lions, B. Perthame, and E. Tadmor, "A kinetic formulation of multidimensional scalar conservation laws and related equations", *J. Amer. Math. Soc.* **7**:1 (1994), 169–191. MR Zbl

[Lions et al. 1994b] P.-L. Lions, B. Perthame, and E. Tadmor, "Kinetic formulation of the isentropic gas dynamics and p-systems", *Comm. Math. Phys.* **163**:2 (1994), 415–431. MR Zbl

[Runst and Sickel 1996] T. Runst and W. Sickel, *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, De Gruyter Series in Nonlinear Analysis and Applications **3**, de Gruyter, Berlin, 1996. MR Zbl

[Westdickenberg 2002] M. Westdickenberg, "Some new velocity averaging results", *SIAM J. Math. Anal.* **33**:5 (2002), 1007–1032. MR Zbl

Received 16 Jan 2017. Revised 4 Dec 2017. Accepted 14 Mar 2018.

DIOGO ARSÉNIO: diogo.arsenio@imj-prg.fr
Institut de Mathématiques de Jussieu–Paris Rive Gauche, Université Paris Diderot, Paris, France
Current address: *New York University Abu Dhabi, Abu Dhabi, United Arab Emirates*

NADER MASMOUDI: masmoudi@cims.nyu.edu
Courant Institute of Mathematical Sciences, New York University, New York, NY, United States
Current address: *New York University Abu Dhabi, Abu Dhabi, United Arab Emirates*

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.389

ON THE EXISTENCE AND STABILITY OF BLOWUP FOR WAVE MAPS INTO A NEGATIVELY CURVED TARGET

ROLAND DONNINGER AND IRFAN GLOGIĆ

We consider wave maps on $(1+d)$-dimensional Minkowski space. For each dimension $d \geq 8$ we construct a negatively curved, d-dimensional target manifold that allows for the existence of a self-similar wave map which provides a stable blowup mechanism for the corresponding Cauchy problem.

1. Introduction

We consider the Cauchy problem for a wave map from the Minkowski spacetime $(\mathbb{R}^{1,d}, \eta)$ into a warped product manifold $N^d = \mathbb{R}^+ \times_g \mathbb{S}^{d-1}$ with metric h; see, e.g., [O'Neill 1983; Tachikawa 1985] for a definition. The metric h has the form

$$h = du^2 + g(u)^2\, d\theta^2, \tag{1-1}$$

where $(u, \theta) \in \mathbb{R}^+ \times \mathbb{S}^{d-1}$ is the natural polar coordinate system on N^d, $d\theta^2$ is the standard metric on $\mathbb{S}^{d-1}$ and

$$g \in C^\infty(\mathbb{R}), \quad g \text{ is odd}, \quad g'(0) = 1, \quad g > 0 \text{ on } (0, \infty). \tag{1-2}$$

Furthermore, we endow the Minkowski space with standard spherical coordinates $(t, r, \omega) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{S}^{d-1}$. The metric η thereby becomes

$$\eta = -dt^2 + dr^2 + r^2\, d\omega^2. \tag{1-3}$$

In this setting, a map $U : (\mathbb{R}^{1,d}, \eta) \to (N^d, h)$ can be written as

$$U(t, r, \omega) = (u(t, r, \omega), \theta(t, r, \omega)).$$

We restrict our attention to the special subclass of so-called 1-equivariant or corotational maps where

$$u(t, r, \omega) = u(t, r) \quad \text{and} \quad \theta(t, r, \omega) = \omega.$$

Under this ansatz the wave maps equation for U reduces to the single semilinear radial wave equation

$$\left(\partial_t^2 - \partial_r^2 - \frac{d-1}{r}\,\partial_r\right) u(t, r) + \frac{d-1}{r^2} g(u(t, r)) g'(u(t, r)) = 0; \tag{1-4}$$

see, e.g., [Shatah and Tahvildar-Zadeh 1994].

It is not hard to see that the Cauchy problem for (1-4) is locally well-posed for sufficiently smooth data and even the low-regularity theory is well understood [Shatah and Tahvildar-Zadeh 1994]. Consequently,

MSC2010: primary 35L71, 58E20, 58J45; secondary 35B35, 35B44.
Keywords: corotational wave maps, self-similar solutions, similarity coordinates, stable blowup.

the interesting questions concern the global Cauchy problem and in particular, the formation of singularities in finite time. There is by now a sizable literature on blowup for wave maps which we cannot review here in its entirety. Let it suffice to say that the energy-critical case $d = 2$ attracted particular attention; see, e.g., [Bizoń et al. 2001; Struwe 2003; Krieger et al. 2008; Rodnianski and Sterbenz 2010; Raphaël and Rodnianski 2012; Sterbenz and Tataru 2010a; 2010b; Krieger and Schlag 2012; Côte et al. 2015a; 2015b; Côte 2015; Gao and Krieger 2015; Lawrie and Oh 2016] for recent contributions. In supercritical dimensions $d \geq 3$ the existence of self-similar solutions is typical [Shatah 1988; Turok and Spergel 1990; Cazenave et al. 1998; Bizoń 2000; Bizoń and Biernat 2015] and stability results for blowup were obtained in [Bizoń et al. 2000; Donninger 2011; Donninger et al. 2012; Bizoń and Biernat 2015; Biernat et al. 2017; Chatzikaleas et al. 2017]. For nonexistence of type II blowup see [Dodson and Lawrie 2015]. Note, however, that there exists nonself-similar blowup in sufficiently high dimensions [Ghoul et al. 2018].

According to a heuristic principle, one typically has finite-time blowup if the curvature of the target is positive. For negatively curved targets, on the other hand, one expects global well-posedness. A notable exception to that rule is provided by the construction of a self-similar solution for a negatively curved target for $d = 7$ in [Cazenave et al. 1998], which indicates that the situation is more subtle. Here we show that the example from that paper is not a peculiarity. We construct suitable target manifolds for any dimension $d \geq 8$ that allow for the existence of an *explicit* self-similar solution. Moreover, we claim that the corresponding self-similar blowup is nonlinearly asymptotically stable under small perturbations of the initial data. In the case $d = 9$ we prove this claim rigorously. This provides the first example of *stable* blowup for wave maps into a negatively curved target.

1A. ***Self-similar solutions.*** In order to look for self-similar solutions, we first observe that (1-4) has the natural scaling symmetry

$$u(t,r) \mapsto u_\lambda(t,r) := u\left(\frac{t}{\lambda}, \frac{r}{\lambda}\right), \quad \lambda > 0, \tag{1-5}$$

in the sense that if u solves (1-4) then u_λ solves it, too. Consequently, it is natural to look for solutions of the form $u(t,r) = \phi(r/t)$. Taking into account the time-translation and reflection symmetries of (1-4), we arrive at the slightly more general ansatz

$$u(t,r) = \phi(\rho), \quad \rho = \frac{r}{T-t}, \tag{1-6}$$

where the free parameter $T > 0$ is the blowup time. By plugging the ansatz (1-6) into (1-4) we obtain the ordinary differential equation

$$(1-\rho^2)\phi''(\rho) + \left(\frac{d-1}{\rho} - 2\rho\right)\phi'(\rho) - \frac{(d-1)g(\phi(\rho))g'(\phi(\rho))}{\rho^2} = 0. \tag{1-7}$$

By recasting (1-7) into an integral equation and then using a fixed-point argument, one can show that any solution to (1-7) that vanishes together with its first derivative at $\rho = 0$ is identically zero near $\rho = 0$. Therefore, any nontrivial smooth solution ϕ to (1-7) for which $\phi(0) = 0$ must have $\phi'(0) \neq 0$, and since

$$\frac{\partial}{\partial r}\phi\left(\frac{r}{T-t}\right)\bigg|_{r=0} = \frac{\phi'(0)}{T-t}, \tag{1-8}$$

such a ϕ gives rise to a smooth solution of (1-4) which suffers a gradient blowup at the origin in finite time. Furthermore, due to finite speed of propagation, this type of singularity arises from smooth, compactly supported initial data. In the following, we restrict ourselves to the study of the solution in the backward lightcone of the singularity,

$$\mathcal{C}_T := \{(t,r) : t \in [0,T),\ r \in [0, T-t]\}. \tag{1-9}$$

Note that in terms of the coordinate ρ, $\mathcal{C}_T$ corresponds to the interval [0, 1]. Consequently, we look for solutions of (1-7) that belong to $C^\infty[0,1]$.

2. Existence of blowup for a negatively curved target manifold

We construct for every $d \geq 8$ a negatively curved d-dimensional Riemannian manifold (N^d, h) which allows for a wave map $U : (\mathbb{R}^{1,d}, \eta) \to (N^d, h)$ that starts off smooth and blows up in finite time. We do this by a suitable choice of the function g that defines the metric on N^d by means of (1-1). To begin with, we restrict ourselves to small u and set

$$g(u) := u\sqrt{1 + 7u^2 - (23d - 170)u^4}. \tag{2-1}$$

Clearly, g is odd and smooth locally around the origin. Furthermore, $g(u) > 0$ for small $u > 0$ and $g'(0) = 1$; see (1-2). In addition, for $d \geq 8$, the metric (1-1) makes the manifold N^d negatively curved locally around $u = 0$; see Proposition 2.1. Next, (1-4) takes the form

$$\left(\partial_t^2 - \partial_r^2 - \frac{d-1}{r}\partial_r\right)u(t,r) + \frac{(d-1)[u(t,r) + 14u(t,r)^3 - 3(23d-170)u(t,r)^5]}{r^2} = 0, \tag{2-2}$$

and the corresponding ordinary differential equation (1-7) becomes

$$(1-\rho^2)\phi''(\rho) + \left(\frac{d-1}{\rho} - 2\rho\right)\phi'(\rho) - \frac{(d-1)[\phi(\rho) + 14\phi(\rho)^3 - 3(23d-170)\phi(\rho)^5]}{\rho^2} = 0. \tag{2-3}$$

As already discussed, any nonzero function $\phi \in C^\infty[0,1]$ that solves (2-3) and vanishes at $\rho = 0$ yields a classical solution to (2-2) that blows up in finite time. In fact, (2-3) has an *explicit* formal solution

$$\phi_0(\rho) = \frac{a\rho}{\sqrt{b-\rho^2}}, \tag{2-4}$$

where

$$a = \sqrt{\frac{d}{E(d)}}, \quad b = 1 + \frac{d}{2} - \frac{7d(d-1)}{E(d)}$$

and

$$E(d) = \sqrt{(46d^2 - 291d - 49)(d-1)} + 7(d-1).$$

Furthermore, if $d \geq 8$ then $E(d)$ is positive and $b > 1$, which makes ϕ_0 a smooth and increasing function on $[0,1]$. Now we have the following result.

Proposition 2.1. *For each $d \geq 8$ there exists an $\varepsilon > 0$ and a function $g : \mathbb{R} \to \mathbb{R}$ satisfying* (1-2) *such that $g(u) = u\sqrt{1+7u^2-(23d-170)u^4}$ for $|u| < \phi_0(1)+\varepsilon$ and the manifold N^d with metric given by* (1-1) *has all sectional curvatures negative.*

The proof is somewhat lengthy but elementary and therefore postponed to Appendix A.

We define

$$u^T(t,r) := \phi_0\left(\frac{r}{T-t}\right), \quad (t,r) \in \mathcal{C}_T. \tag{2-5}$$

Note that $|\phi_0(\rho)| \leq \phi_0(1)$ for all $\rho \in [0,1]$ and thus,

$$U^T(t,r,\omega) := (u^T(t,r),\omega)$$

is a wave map from $\mathcal{C}_T \subset \mathbb{R}^{1,d}$ to (N^d,h). By finite speed of propagation we obtain the following result.

Theorem 2.2. *For every $d \geq 8$ there exists a d-dimensional, negatively curved Riemannian manifold N^d such that the Cauchy problem for wave maps from Minkowski space $\mathbb{R}^{1,d}$ into N^d admits a solution which develops from smooth Cauchy data of compact support and forms a singularity in finite time.*

Remark 2.3. Our focus in this work was on functions g which lead to polynomial-type nonlinearities gg' in (1-4). Since

$$\frac{d}{du}(g(u)^2) = 2g(u)g'(u),$$

this is equivalent to g^2 being an even polynomial. The lowest-degree even polynomial g^2 which, through the metric (1-1), yields negative curvature (locally around the pole $u=0$) on the target manifold N^d is of the form

$$g(u)^2 = u^2 + c_1u^4 + c_2u^6 \tag{2-6}$$

for an appropriate choice of $c_1, c_2 \in \mathbb{R}$. Furthermore, the function g given by (2-6) gives rise to a (formal) solution to (1-7) of the form (2-4). This solution in turn yields a bona fide self-similar blowup in the corresponding wave maps equation only if it is smooth on $[0,1]$ and the corresponding function u^T from (2-5) stays inside the negatively curved neighborhood of the pole $u=0$ whose metric is given by (2-6). The construction of such solutions, by a proper choice of coefficients c_1 and c_2 in (2-6), is in fact possible only for $d \geq 8$. There is, of course, some freedom in the choice of c_1 and c_2, and the one we made in (2-1) was led by the objective of "minimizing" their dependence on d by allowing them to depend at most linearly on it.

In order to determine the role of the solution u^T for generic evolutions, it is necessary to investigate its stability under perturbations. In fact, we claim that for any $d \geq 8$, the self-similar solution (2-5) exhibits *stable* blowup; i.e., there is an open set of radial initial data that give rise to solutions which approach u^T in $\mathcal{C}_T$ as $t \to T^-$. The rest of the paper is devoted to the proof of this stability property. We emphasize that the fact that the solutions we constructed are explicit is crucial for our approach to their stability analysis (see the proof of Proposition 3.7). Due to certain technical difficulties (see Remark 3.17) we restrict ourselves to the lowest odd dimension $d=9$.

3. Stability of blowup

From now on we fix $d = 9$. In view of (1-4) and (2-1), we consider the Cauchy problem

$$\begin{cases} \left(\partial_t^2 - \partial_r^2 - \frac{8}{r}\,\partial_r\right) u(t,r) + \dfrac{8[u(t,r) + 14u(t,r)^3 - 111u(t,r)^5]}{r^2} = 0, & (t,r) \in \mathcal{C}_T, \\ u(0,r) = u_0(r), \quad \partial_0 u(0,r) = u_1(r), & r \in [0,T]. \end{cases} \tag{3-1}$$

The restriction to the backward lightcone $\mathcal{C}_T$ is possible and natural by finite speed of propagation. Furthermore, to ensure regularity of the solution at the origin $r = 0$, we impose the boundary condition

$$u(t,0) = 0 \quad \text{for } t \in [0,T). \tag{3-2}$$

The blowup solution (2-4) now becomes

$$u^T(t,r) = \phi_0(\rho) = \frac{3\rho}{\sqrt{2(155 - 74\rho^2)}}, \quad \text{where } \rho = \frac{r}{T-t}. \tag{3-3}$$

Note that by construction, the wave map evolution for the target manifold N^9 is given by (3-1), provided that $|u(t,r)| \le \phi_0(1) + \varepsilon_1$ for some small $\varepsilon_1 > 0$. We are only interested in the evolution in the backward lightcone of the point of blowup and therefore study (3-1) with no a priori restriction on the size of u. A posteriori we show that the solutions we construct stay below $\phi_0(1) + \varepsilon_1$.

Note further that (3-1) can be viewed as a nonlinear wave equation with polynomial nonlinearity. Indeed, the boundary condition (3-2) allows for a change of variable $u(t,r) = rv(t,r)$ which leads to an eleven-dimensional radial wave equation in v,

$$\left(\partial_t^2 - \partial_r^2 - \frac{10}{r}\,\partial_r\right) v(t,r) = -8[14v(t,r)^3 - 111r^2 v(t,r)^5]. \tag{3-4}$$

In fact, this is the point of view we adopt here. In particular, the nonlinear term in (3-1) becomes smooth and therefore admits a uniform Lipschitz estimate needed for a contraction mapping argument; see Lemma 3.12. We also remark that (3-4), in spite of its defocusing character (at least for small values of v), admits an explicit self-similar blowup solution. This is in stark contrast to the cubic defocusing wave equation

$$\left(\partial_t^2 - \partial_r^2 - \frac{10}{r}\partial_r\right) v(t,r) = -v(t,r)^3$$

for which no self-similar solutions exist. The self-similar blowup in (3-4) can therefore be understood as a consequence of the presence of the focusing quintic term which dominates the dynamics for large initial data.

3A. ***Main result.*** We start by intuitively describing the main result. We fix $T_0 > 0$ and prescribe initial data $u[0]$ that are close to $u^{T_0}[0]$ on a ball of radius slightly larger than T_0. Here and throughout the paper we use the abbreviation $u[t] := (u(t,\cdot), \partial_t u(t,\cdot))$. Then we prove the existence of a particular T near T_0 for which the solution u converges to u^T inside the backward lightcone $\mathcal{C}_T$ in a norm adapted to the blowup behavior of u^T. For the precise statement of the main result we use Definitions 3.4 and 3.5.

Theorem 3.1. *Fix $T_0 > 0$. There exist constants $M, \delta, \varepsilon > 0$ such that for any radial initial data $u[0]$ satisfying*

$$\big\||\cdot|^{-1}\big(u[0](|\cdot|) - u^{T_0}[0](|\cdot|)\big)\big\|_{H^6(\mathbb{B}^{11}_{T_0+\delta})\times H^5(\mathbb{B}^{11}_{T_0+\delta})} \le \frac{\delta}{M} \tag{3-5}$$

the following statements hold:

(i) *The blowup time at the origin $T := T_{u[0]}$ belongs to the interval $[T_0-\delta, T_0+\delta]$.*

(ii) *The solution $u : \mathcal{C}_T \to \mathbb{R}$ to* (3-1) *satisfies*

$$(T-t)^{-\frac{9}{2}+k}\big\||\cdot|^{-1}(u(t,|\cdot|) - u^T(t,|\cdot|))\big\|_{\dot H^k(\mathbb{B}^{11}_{T-t})} \le \delta(T-t)^{\varepsilon}, \tag{3-6}$$

$$(T-t)^{-\frac{7}{2}+l}\big\||\cdot|^{-1}(\partial_t u(t,|\cdot|) - \partial_t u^T(t,|\cdot|))\big\|_{\dot H^l(\mathbb{B}^{11}_{T-t})} \le \delta(T-t)^{\varepsilon} \tag{3-7}$$

for integers $0 \le k \le 6$ and $0 \le l \le 5$. Furthermore,

$$\|u(t,\cdot) - u^T(t,\cdot)\|_{L^\infty(0,T-t)} \le \delta(T-t)^{\varepsilon}. \tag{3-8}$$

Remark 3.2. The normalizing factor on the left-hand side of (3-6) and (3-7) appears naturally as it reflects the behavior of the self-similar solution u^T in the respective Sobolev norm; i.e.,

$$\big\||\cdot|^{-1}u^T(t,|\cdot|)\big\|_{\dot H^k(\mathbb{B}^{11}_{T-t})} = \left\||\cdot|^{-1}\phi_0\left(\frac{|\cdot|}{T-t}\right)\right\|_{\dot H^k(\mathbb{B}^{11}_{T-t})} = (T-t)^{\frac{9}{2}-k}\big\||\cdot|^{-1}\phi_0(|\cdot|)\big\|_{\dot H^k(\mathbb{B}^{11}_1)}$$

and

$$\big\||\cdot|^{-1}\,\partial_t u_T(t,|\cdot|)\big\|_{\dot H^l(\mathbb{B}^{11}_{T-t})} = (T-t)^{-2}\left\|\phi_0'\left(\frac{|\cdot|}{T-t}\right)\right\|_{\dot H^l(\mathbb{B}^{11}_{T-t})} = (T-t)^{\frac{7}{2}-l}\|\phi_0'(|\cdot|)\|_{\dot H^l(\mathbb{B}^{11}_1)}.$$

Remark 3.3. Since ϕ_0 is monotonically increasing on [0, 1], we have

$$\|u^T(t,\cdot)\|_{L^\infty(0,T-t)} = \max_{\rho\in[0,1]}\|\phi_0(\rho)\| = \phi_0(1). \tag{3-9}$$

Therefore, given $\varepsilon_1 > 0$, it follows from (3-8) and (3-9) that δ can be chosen small enough so that

$$\|u(t,\cdot)\|_{L^\infty(0,T-t)} \le \|u(t,\cdot) - u^T(t,\cdot)\|_{L^\infty(0,T-t)} + \|u^T(t,\cdot)\|_{L^\infty(0,T-t)} \le \varepsilon_1 + \phi_0(1).$$

Hence, for $t < T$ the solution $u(t,r)$ stays inside a neighborhood of $u = 0$ where the metric is given by (2-1); i.e., the portion of the target manifold that participates in the dynamics of the blowup solution is described by the metric (2-1).

3B. ***Outline of the proof.*** We use the method developed in the series of papers [Donninger 2011; 2014; 2017; Donninger and Schörkhuber 2012; 2014; 2016; 2017; Costin et al. 2017]. First, we introduce the rescaled variables

$$v_1(t,r) := \frac{T-t}{r}\,u(t,r), \quad v_2(t,r) := \frac{(T-t)^2}{r}\,\partial_t u(t,r). \tag{3-10}$$

Division by r is justified by the boundary condition (3-2) and the presence of the prefactors involving $T-t$ has to do with the change of variables we subsequently introduce. That is, we introduce similarity coordinates (τ, ρ) defined by

$$\tau := -\log(T-t) + \log T, \quad \rho := \frac{r}{T-t}, \tag{3-11}$$

and set

$$\psi_j(\tau, \rho) := v_j(T(1-e^{-\tau}), Te^{-\tau}\rho) \tag{3-12}$$

for $j = 1, 2$. As a consequence, (3-4) can be written as an abstract evolution equation,

$$\partial_\tau \Psi(\tau) = \boldsymbol{L}_0 \Psi(\tau) + \boldsymbol{M}(\Psi(\tau)), \tag{3-13}$$

where $\Psi(\tau) = (\psi_1(\tau, \cdot), \psi_2(\tau, \cdot))$, $\boldsymbol{L}_0$ is the spatial part of the radial wave operator in the new coordinates, and $\boldsymbol{M}(\Psi(\tau))$ consists of the remaining nonlinear terms. The benefit of passing to the new variables (3-11) and (3-12) is that the backward lightcone $\mathcal{C}_T$ is transformed into a cylinder

$$\mathcal{C} := \{(\tau, \rho) : \tau \in [0, \infty),\ \rho \in [0, 1]\},$$

the rescaled self-similar blowup solution u^T becomes a τ-independent function Ψ_{res} (this justifies the presence of t-dependent prefactors in (3-10)), and the problem of stability of blowup transforms into the problem of asymptotic stability of a static solution. We subsequently follow the standard approach for studying the stability of steady-state solutions and plug the ansatz $\Psi(\tau) = \Psi_{\text{res}} + \Phi(\tau)$ into (3-13). This leads to an evolution equation in Φ,

$$\partial_\tau \Phi(\tau) = \boldsymbol{L}_0 \Phi(\tau) + \boldsymbol{L}' \Phi(\tau) + \boldsymbol{N}(\Phi(\tau)), \tag{3-14}$$

where $\boldsymbol{L}'$ is the Fréchet derivative of $\boldsymbol{M}$ at Ψ_{res} and $\boldsymbol{N}(\Phi(\tau))$ is the nonlinear remainder. We then proceed by studying (3-14) as an ordinary differential equation in a Hilbert space with the norm

$$\|\boldsymbol{u}\|^2 = \|(u_1, u_2)\|^2 := \|u_1(|\cdot|)\|^2_{H^6(\mathbb{B}^{11})} + \|u_2(|\cdot|)\|^2_{H^5(\mathbb{B}^{11})}. \tag{3-15}$$

However, passing to new variables also comes with a price. Namely, the radial wave operator $\boldsymbol{L}_0$ is not self-adjoint. Nonetheless, we establish well-posedness of the linearized problem (that is, (3-14) with $\boldsymbol{N}$ removed) by using methods from semigroup theory. In particular, we use a norm equivalent to (3-15) and the Lumer–Phillips theorem to show that $\boldsymbol{L}_0$ generates a semigroup $(\boldsymbol{S}_0(\tau))_{\tau \geq 0}$ with a negative growth bound. This in particular allows for locating the spectrum of $\boldsymbol{L}_0$. Furthermore, $\boldsymbol{L}'$ is compact so $\boldsymbol{L} := \boldsymbol{L}_0 + \boldsymbol{L}'$ generates a strongly continuous semigroup $(\boldsymbol{S}(\tau))_{\tau \geq 0}$ and well-posedness of the linearized problem follows.

The stability of the solution u^T follows from a decay estimate on the semigroup $\boldsymbol{S}(\tau)$. To obtain such an estimate we exploit the relation between the growth bound of a semigroup and the location of the spectrum of its generator. We therefore study $\sigma(\boldsymbol{L})$ which, thanks to the compactness of $\boldsymbol{L}'$, amounts to studying the eigenvalue problem $(\lambda - \boldsymbol{L})\boldsymbol{u} = 0$. We subsequently show that $\sigma(\boldsymbol{L})$ is contained in the left half-plane except for the point $\lambda = 1$. However, this unstable eigenvalue corresponds to an apparent instability and we later use it to fix the blowup time. We therefore proceed by defining a spectral projection $\boldsymbol{P}$ onto the unstable space and study the semigroup $\boldsymbol{S}(\tau)$ restricted to $\operatorname{rg}(1 - \boldsymbol{P})$. Furthermore, we establish a uniform bound on the resolvent $\boldsymbol{R_L}(\lambda)$ and invoke the Gearhart–Prüss theorem to obtain a negative growth bound on $(1-\boldsymbol{P})\boldsymbol{S}(\tau)$.

Appealing to Duhamel's principle, we rewrite (3-14) in the integral form

$$\Phi(\tau)=\boldsymbol{S}(\tau)\boldsymbol{U}(\boldsymbol{v},T)+\int_0^\tau \boldsymbol{S}(\tau-s)\boldsymbol{N}(\Phi(s))\,ds, \tag{3-16}$$

where $\boldsymbol{U}(\boldsymbol{v},T)$ represents the rescaled initial data. We remark that the parameter T does not appear in the equation itself but in the initial data only. To obtain a decaying solution to (3-16) we suppress the unstable part of $\boldsymbol{S}(\tau)$ by introducing a correction term

$$\boldsymbol{C}(\Phi,\boldsymbol{U}(\boldsymbol{v},T)):=\boldsymbol{P}\left(\boldsymbol{U}(\boldsymbol{v},T)+\int_0^\infty e^{-s}\boldsymbol{N}(\Phi(s))\,ds\right)$$

into (3-16). That is, we consider the modified equation

$$\Phi(\tau)=\boldsymbol{S}(\tau)\big(\boldsymbol{U}(\boldsymbol{v},T)-\boldsymbol{C}(\Phi,\boldsymbol{U}(\boldsymbol{v},T))\big)+\int_0^\tau \boldsymbol{S}(\tau-s)\boldsymbol{N}(\Phi(s))\,ds. \tag{3-17}$$

We subsequently prove that for a fixed T_0 and small enough initial data $\boldsymbol{v}$, every T close to T_0 yields a unique solution to (3-17) that decays to zero at the linear decay rate. In other words, we prove the existence of a solution curve to (3-17) parametrized by T inside a small neighborhood of T_0, provided $\boldsymbol{v}$ is small enough.

Finally, we use the very presence of the unstable eigenvalue $\lambda=1$ to prove the existence of a particular T near T_0 for which $\boldsymbol{C}(\Phi,\boldsymbol{U}(\boldsymbol{v},T))=0$ and hence obtain a decaying solution to (3-16) which, when translated back to the original coordinates, implies the main result.

3C. ***Notation.*** We denote by $\mathbb{B}^d_R$ the d-dimensional open ball of radius R centered at the origin. For brevity we let $\mathbb{B}^d:=\mathbb{B}^d_1$. We write 2-component vector quantities in boldface, e.g., $\boldsymbol{u}=(u_1,u_2)$. By $\mathcal{B}(\mathcal{H})$ we denote the space of bounded operators on the Hilbert space $\mathcal{H}$. We denote by $\sigma(\boldsymbol{L})$ and $\sigma_p(\boldsymbol{L})$ the spectrum and the point spectrum, respectively, of a linear operator $\boldsymbol{L}$. Also, we denote by $\rho(\boldsymbol{L})$ the resolvent set $\mathbb{C}\setminus\sigma(\boldsymbol{L})$ and use the convention $\boldsymbol{R_L}(\lambda):=(\lambda-\boldsymbol{L})^{-1}$, $\lambda\in\rho(\boldsymbol{L})$, for the resolvent operator. We use the symbol $\lesssim$ with the standard meaning: $a\lesssim b$ if there exists a positive constant c, independent of a,b, such that $a\le cb$. Also, $a\simeq b$ means that both $a\lesssim b$ and $b\lesssim a$ hold.

3D. ***Similarity coordinates and cylinder formulation.*** After introducing the similarity coordinates

$$\tau:=-\log(T-t)+\log T,\quad \rho:=\frac{r}{T-t},$$

and the rescaled variables

$$v_1(t,r):=\frac{T-t}{r}\,u(t,r),\quad v_2(t,r):=\frac{(T-t)^2}{r}\,\partial_t u(t,r),$$
$$\psi_j(\tau,\rho)=v_j(T(1-e^{-\tau}),T\rho e^{-\tau}),\quad j=1,2,$$

we obtain from (3-1) the first-order system

$$\begin{bmatrix}\partial_\tau\psi_1\\ \partial_\tau\psi_2\end{bmatrix}=\begin{bmatrix}-\rho\,\partial_\rho\psi_1-\psi_1+\psi_2\\ \partial_\rho^2\psi_1+(10/\rho)\,\partial_\rho\psi_1-\rho\partial_\rho\psi_2-2\psi_2\end{bmatrix}-\begin{bmatrix}0\\ 8(14\psi_1^3-111\rho^2\psi_1^5)\end{bmatrix} \tag{3-18}$$

for $(\tau,\rho)\in\mathcal{C}$. Furthermore, the initial data become

$$\begin{bmatrix}\psi_1(0,\rho)\\ \psi_2(0,\rho)\end{bmatrix}=\frac{1}{\rho}\begin{bmatrix}u_0(T\rho)\\ Tu_1(T\rho)\end{bmatrix}=\frac{1}{\rho}\begin{bmatrix}u^{T_0}(0,T\rho)\\ T\partial_0 u^{T_0}(0,T\rho)\end{bmatrix}+\frac{1}{\rho}\begin{bmatrix}F(T\rho)\\ TG(T\rho)\end{bmatrix}, \tag{3-19}$$

where T_0 is a fixed parameter and

$$F:=u_0-u^{T_0}(0,\cdot),\quad G:=u_1-\partial_0 u^{T_0}(0,\cdot).$$

In addition, we have the regularity conditions

$$\partial_\rho\psi_1(\tau,\rho)|_{\rho=0}=\partial_\rho\psi_2(\tau,\rho)|_{\rho=0}=0$$

for $\tau\geq 0$. Note further that we are studying the dynamics around u^{T_0} for a fixed T_0 and thus, it is natural to split the initial data as in (3-19). The parameter T is assumed to be close to T_0 and will be fixed later. As a consequence, the proximity of the initial data to $u^{T_0}[0]$ is measured by $\boldsymbol{v}:=(F,G)$.

3E. ***Perturbations of the blowup solution.*** For convenience, we set

$$\Psi(\tau)(\rho):=\begin{bmatrix}\psi_1(\tau,\rho)\\ \psi_2(\tau,\rho)\end{bmatrix}.$$

In the rescaled variables the blowup solution u^T becomes τ-independent, i.e.,

$$\begin{bmatrix}((T-t)/r)u^T(t,r)\\ ((T-t)^2/r)\partial_t u^T(t,r)\end{bmatrix}=\begin{bmatrix}(1/\rho)\phi_0(\rho)\\ \phi_0'(\rho)\end{bmatrix}=:\Psi_{\text{res}}(\tau)(\rho).$$

We proceed by studying the dynamics of (3-18) around Ψ_{res}. Our aim is to prove the asymptotic stability of Ψ_{res}, which in turn translates into the appropriate notion of stability of u^T. We therefore follow the standard method and plug the ansatz $\Psi=\Psi_{\text{res}}+\Phi$ into (3-18), where $\Phi(\tau)(\rho):=(\varphi_1(\tau,\rho),\varphi_2(\tau,\rho))$. This leads to an evolution equation for the perturbation Φ,

$$\begin{cases}\partial_\tau\Phi(\tau)=\widetilde{\boldsymbol{L}}\Phi(\tau)+\boldsymbol{N}(\Phi(\tau)),\\ \quad\Phi(0)=\boldsymbol{U}(\boldsymbol{v},T),\end{cases} \tag{3-20}$$

where $\widetilde{\boldsymbol{L}}$ and $\boldsymbol{N}$ are spatial operators and $\boldsymbol{U}(\boldsymbol{v},T)$ are the initial data. More precisely, $\widetilde{\boldsymbol{L}}:=\widetilde{\boldsymbol{L}}_0+\boldsymbol{L}'$, where

$$\widetilde{\boldsymbol{L}}_0\boldsymbol{u}(\rho):=\begin{bmatrix}-\rho u_1'(\rho)-u_1(\rho)+u_2(\rho)\\ u_1''(\rho)+(10/\rho)u_1'(\rho)-\rho u_2'(\rho)-2u_2(\rho)\end{bmatrix}, \tag{3-21}$$

$$\boldsymbol{L}'\boldsymbol{u}(\rho):=\begin{bmatrix}0\\ W(\rho,\phi_0(\rho))u_1(\rho)\end{bmatrix}, \tag{3-22}$$

$$\boldsymbol{N}(\boldsymbol{u})(\rho):=\begin{bmatrix}0\\ N(\rho,u_1(\rho))\end{bmatrix} \tag{3-23}$$

for a 2-component function $\boldsymbol{u}(\rho) = (u_1(\rho), u_2(\rho))$, where

$$\begin{aligned} N(\rho, u_1(\rho)) &= -\frac{8}{\rho^3}[n(\phi_0(\rho)+\rho u_1(\rho)) - n(\phi_0(\rho)) - n'(\phi_0(\rho))\rho u_1(\rho)], \\ W(\rho, \phi_0(\rho)) &= -\frac{8}{\rho^2} n'(\phi_0(\rho)) \quad \text{for } n(x) = 14x^3 - 111x^5. \end{aligned} \tag{3-24}$$

Also, we write the initial data as

$$\Phi(0)(\rho) = \boldsymbol{U}(\boldsymbol{v}, T)(\rho) = \begin{bmatrix} (1/\rho)\phi_0((T/T_0)\rho) \\ (T^2/T_0^2)\phi_0'((T/T_0)\rho) \end{bmatrix} - \begin{bmatrix} (1/\rho)\phi_0(\rho) \\ \phi_0'(\rho) \end{bmatrix} + \boldsymbol{V}(\boldsymbol{v}, T)(\rho), \tag{3-25}$$

where

$$\boldsymbol{V}(\boldsymbol{v}, T)(\rho) := \begin{bmatrix} (1/\rho)F(T\rho) \\ (T/\rho)G(T\rho) \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} F \\ G \end{bmatrix}.$$

3F. *Strong lightcone solutions and blowup time at the origin.* To proceed, we need the notion of a solution to the problem (3-20). In Section 3G we introduce the space

$$\mathcal{H} := H^6_{\text{rad}}(\mathbb{B}^{11}) \times H^5_{\text{rad}}(\mathbb{B}^{11})$$

and prove that the closure of the operator $\widetilde{\boldsymbol{L}}$, defined on a suitable domain, generates a strongly continuous semigroup $\boldsymbol{S}(\tau)$ on $\mathcal{H}$. Consequently, we formulate the problem (3-20) as an abstract integral equation via Duhamel's formula,

$$\Phi(\tau) = \boldsymbol{S}(\tau)\boldsymbol{U}(\boldsymbol{v}, T) + \int_0^\tau \boldsymbol{S}(\tau - s)\boldsymbol{N}(\Phi(s))\, ds. \tag{3-26}$$

This in particular establishes the well-posedness of the problem (3-20) in $\mathcal{H}$. We are now in the position to introduce the following definitions.

Definition 3.4. We say that $u : \mathcal{C}_T \to \mathbb{R}$ is a *solution* to (3-1) if the corresponding $\Phi : [0, \infty) \to \mathcal{H}$ belongs to $C([0, \infty); \mathcal{H})$ and satisfies (3-26) for all $\tau \geq 0$.

Definition 3.5. For the radial initial data (u_0, u_1) we define $\mathcal{T}(u_0, u_1)$ as the set of all $T > 0$ such that there exists a solution $u : \mathcal{C}_T \to \mathbb{R}$ to (3-1). We call

$$T_{(u_0,u_1)} := \sup(\mathcal{T}(u_0, u_1) \cup \{0\}) \tag{3-27}$$

the *blowup time at the origin*.

3G. *Functional setting.* We consider radial Sobolev functions $\hat{u} : \mathbb{B}^{11}_R \to \mathbb{C}$, i.e., $\hat{u}(\xi) = u(|\xi|)$ for $\xi \in \mathbb{B}^{11}_R$ and some $u : [0, R) \to \mathbb{C}$. We furthermore define

$$u \in H^m_{\text{rad}}(\mathbb{B}^{11}_R) \quad \text{if and only if} \quad \hat{u} \in H^m(\mathbb{B}^{11}_R) := W^{m,2}(\mathbb{B}^{11}_R).$$

With the norm

$$\|u\|_{H^m_{\text{rad}}(\mathbb{B}^{11}_R)} := \|\hat{u}\|_{H^m(\mathbb{B}^{11}_R)},$$

$H^m_{\text{rad}}(\mathbb{B}^{11}_R)$ becomes a Banach space. In the rest of this paper we do not distinguish between u and $\hat{u}$. Now we define the Hilbert space

$$\mathcal{H} := H^6_{\text{rad}}(\mathbb{B}^{11}) \times H^5_{\text{rad}}(\mathbb{B}^{11}),$$

with the induced norm

$$\|\boldsymbol{u}\|^2 = \|(u_1, u_2)\|^2 := \|u_1\|^2_{H^6_{\text{rad}}(\mathbb{B}^{11})} + \|u_2\|^2_{H^5_{\text{rad}}(\mathbb{B}^{11})}.$$

3H. ***Well-posedness of the linearized equation.*** To establish well-posedness of the problem (3-20) we start by defining the domain of the free operator $\widetilde{\boldsymbol{L}}_0$; see (3-21). We follow [Donninger and Schörkhuber 2017] and let

$$\mathcal{D}(\widetilde{\boldsymbol{L}}_0) := \{\boldsymbol{u} \in C^\infty(0,1)^2 \cap \mathcal{H} : w_1 \in C^3[0,1],\ w_1''(0) = 0,\ w_2 \in C^2[0,1]\},$$

where

$$w_j(\rho) := D_{11} u_j(\rho) := \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^4 (\rho^9 u_j(\rho)) = \sum_{n=0}^{4} c_n \rho^{n+1} u_j^{(n)}(\rho)$$

for certain positive constants c_n, $\rho \in [0,1]$, and $j = 1, 2$. Since $C^\infty(\overline{\mathbb{B}^{11}})$ is dense in $H^m(\mathbb{B}^{11})$,

$$C^\infty_{\text{even}}[0,1]^2 := \{\boldsymbol{u} \in C^\infty[0,1]^2 : \boldsymbol{u}^{(2k+1)}(0) = 0,\ k = 0, 1, 2, \dots\} \subset \mathcal{D}(\widetilde{\boldsymbol{L}}_0)$$

is dense in $\mathcal{H}$, which in turn implies that $\widetilde{\boldsymbol{L}}_0$ is densely defined on $\mathcal{H}$. Furthermore, we have the following result.

Proposition 3.6. *The operator $\widetilde{\boldsymbol{L}}_0 : \mathcal{D}(\widetilde{\boldsymbol{L}}_0) \subset \mathcal{H} \to \mathcal{H}$ is closable and its closure $\boldsymbol{L}_0 : \mathcal{D}(\boldsymbol{L}_0) \subset \mathcal{H} \to \mathcal{H}$ generates a strongly continuous one-parameter semigroup $(\boldsymbol{S}_0(\tau))_{\tau \geq 0}$ of bounded operators on $\mathcal{H}$ satisfying the growth estimate*

$$\|\boldsymbol{S}_0(\tau)\| \leq M e^{-\tau} \tag{3-28}$$

for all $\tau \geq 0$ and some $M > 0$. Furthermore, the operator $\boldsymbol{L} := \boldsymbol{L}_0 + \boldsymbol{L}' : \mathcal{D}(\boldsymbol{L}) \subset \mathcal{H} \to \mathcal{H}$, $\mathcal{D}(\boldsymbol{L}) = \mathcal{D}(\boldsymbol{L}_0)$, is the generator of a strongly continuous semigroup $(\boldsymbol{S}(\tau))_{\tau \geq 0}$ on $\mathcal{H}$ and $\boldsymbol{L}' : \mathcal{H} \to \mathcal{H}$ is compact.

Proof. The proof essentially follows the one of Proposition 3.1 in [Chatzikaleas et al. 2017] for $d = 9$. □

3I. ***The spectrum of the free operator.*** By exploiting the relation between the growth bound of a semigroup and the spectral bound of its generator, we can locate the spectrum of the operator $\boldsymbol{L}_0$. Namely, according to [Engel and Nagel 2000, p. 55, Theorem 1.10] the estimate (3-28) implies

$$\sigma(\boldsymbol{L}_0) \subseteq \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda \leq -1\}. \tag{3-29}$$

3J. ***The spectrum of the full linear operator.*** To understand the properties of the semigroup $\boldsymbol{S}(\tau)$ we investigate the spectrum of the full linear operator $\boldsymbol{L}$. First of all, we remark that $\lambda = 1$ is an eigenvalue of $\boldsymbol{L}$ (see Section 3K), which is an artifact of the freedom of choice of the parameter T; see, e.g., [Costin et al. 2017] for a discussion on this. What is more, $\lambda = 1$ is the only spectral point of $\boldsymbol{L}$ with a nonnegative real part. To prove this we first focus on the point spectrum.

Proposition 3.7. *We have*

$$\sigma_p(\boldsymbol{L}) \subseteq \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < 0\} \cup \{1\}. \tag{3-30}$$

Proof. We argue by contradiction and assume there exists a $\lambda \in \sigma_p(\boldsymbol{L}) \setminus \{1\}$ with $\operatorname{Re}\lambda \geq 0$. This means that there exists a $\boldsymbol{u} = (u_1, u_2) \in \mathcal{D}(\boldsymbol{L}) \setminus \{0\}$ such that $\boldsymbol{u} \in \ker(\lambda - \boldsymbol{L})$. The spectral equation $(\lambda - \boldsymbol{L})\boldsymbol{u} = 0$ implies that the first component u_1 satisfies the equation

$$(1-\rho^2)u_1''(\rho) + \left(\frac{10}{\rho} - 2(\lambda+2)\rho\right)u_1'(\rho) - (\lambda+1)(\lambda+2)u_1(\rho) - V(\rho)u_1(\rho) = 0 \tag{3-31}$$

for $\rho \in (0,1)$, where

$$V(\rho) := -W(\rho, \phi_0(\rho)) = \frac{8n'(\phi_0(\rho))}{\rho^2} = -\frac{54(3737\rho^2 - 4340)}{(155 - 74\rho^2)^2}.$$

Since $\boldsymbol{u} \in \mathcal{H}$, we know u_1 must be an element of $H^6_{\text{rad}}(\mathbb{B}^{11})$. From the smoothness of the coefficients in (3-31) we have an a priori regularity $u_1 \in C^\infty(0,1)$. In fact, we claim that $u_1 \in C^\infty[0,1]$. To show this, we use the Frobenius method. Namely, both $\rho = 0$ and $\rho = 1$ are regular singularities of (3-31) and Frobenius' theory gives a series form of solutions locally around singular points.

The Frobenius indices at $\rho = 0$ are $s_1 = 0$ and $s_2 = -9$. Therefore, two independent solutions of (3-31) have the form

$$u_1^1(\rho) = \sum_{i=0}^{\infty} a_i \rho^i \quad \text{and} \quad u_1^2(\rho) = C\log(\rho)u_1^1(\rho) + \rho^{-9}\sum_{i=0}^{\infty} b_i \rho^i$$

for some constant $C \in \mathbb{C}$ and $a_0 = b_0 = 1$. Since $u_1^1(\rho)$ is analytic at $\rho = 0$ and $u_1^2(\rho)$ does not belong to $H^6_{\text{rad}}(\mathbb{B}^{11})$, we conclude that u_1 is a multiple of u_1^1 and therefore, $u_1 \in C^\infty[0,1)$.

The Frobenius indices at $\rho = 1$ are $s_1 = 0$ and $s_2 = 4 - \lambda$, and we distinguish different cases. If $4 - \lambda \notin \mathbb{Z}$ then the two linearly independent solutions are

$$u_1^1(\rho) = \sum_{i=0}^{\infty} a_i (1-\rho)^i \quad \text{and} \quad u_1^2(\rho) = (1-\rho)^{4-\lambda}\sum_{i=0}^{\infty} b_i (1-\rho)^i$$

with $a_0 = b_0 = 1$. Since $u_1^1(\rho)$ is analytic at $\rho = 1$ and u_1^2 does not belong to $H^6_{\text{rad}}(\mathbb{B}^{11})$, we conclude that $u_1 \in C^\infty[0,1]$. If $4 - \lambda \in \mathbb{N}_0$, then the fundamental solutions around $\rho = 1$ are of the form

$$u_1^1(\rho) = (1-\rho)^{4-\lambda}\sum_{i=0}^{\infty} a_i (1-\rho)^i \quad \text{and} \quad u_1^2(\rho) = \sum_{i=0}^{\infty} b_i (1-\rho)^i + C\log(1-\rho)u_1^1(\rho),$$

with $a_0 = b_0 = 1$. Since $u_1^1(\rho)$ is analytic at $\rho = 1$ and u_1^2 does not belong to $H^6_{\text{rad}}(\mathbb{B}^{11})$ unless $C = 0$, we again conclude that $u_1 \in C^\infty[0,1]$. Finally, if $4 - \lambda$ is a negative integer, the linearly independent solutions around $\rho = 1$ are

$$u_1^1(\rho) = \sum_{i=0}^{\infty} a_i (1-\rho)^i \quad \text{and} \quad u_1^2(\rho) = (1-\rho)^{4-\lambda}\sum_{i=0}^{\infty} b_i (1-\rho)^i + C\log(1-\rho)u_1^1(\rho),$$

with $a_0 = b_0 = 1$. Once again, since $u_1^1(\rho)$ is analytic at $\rho = 1$ and u_1^2 is not a member of $H^6_{\mathrm{rad}}(\mathbb{B}^{11})$, we infer that $u_1 \in C^\infty[0, 1]$.

To obtain the desired contradiction, it remains to prove that (3-31) does not have a solution in $C^\infty[0, 1]$ for $\operatorname{Re}\lambda \geq 0$ and $\lambda \neq 1$. This claim is called the *mode stability* of the solution u^T. A general approach to proving mode stability of explicit self-similar blowup solutions to nonlinear wave equations of the type (1-4) was developed in [Costin et al. 2016; 2017]. We argue here along the lines of [Costin et al. 2017]. Also, for the rest of the proof, we follow the terminology of that paper. Namely, we call $\lambda \in \mathbb{C}$ an *eigenvalue* if it yields a $C^\infty[0, 1]$ solution to the equation in question. Also, if an eigenvalue λ satisfies $\operatorname{Re}\lambda \geq 0$ we say it is *unstable*; otherwise we call it *stable*. Our aim is therefore to prove that, apart from $\lambda = 1$, there are no unstable eigenvalues of the problem (3-31).

First of all, we make the substitution $v(\rho) = \rho u_1(\rho)$. This leads to the equation

$$(1-\rho^2)v''(\rho) + \left(\frac{8}{\rho} - 2(\lambda+1)\rho\right)v'(\rho) - \lambda(\lambda+1)v(\rho) - \widehat{V}(\rho)v(\rho) = 0, \tag{3-32}$$

where

$$\widehat{V}(\rho) := -\frac{10(15799\rho^4 - 5084\rho^2 - 19220)}{\rho^2(155-74\rho^2)^2}.$$

Now we formulate the corresponding supersymmetric problem,

$$(1-\rho^2)\tilde{v}''(\rho) + \left(\frac{8}{\rho} - 2(\lambda+1)\rho\right)\tilde{v}'(\rho) - (\lambda+2)(\lambda-1)\tilde{v}(\rho) - \widetilde{V}(\rho)\tilde{v}(\rho) = 0, \tag{3-33}$$

where

$$\widetilde{V}(\rho) := -\frac{18(3737\rho^4 + 5735\rho^2 - 24025)}{\rho^2(155-74\rho^2)^2};$$

see [Costin et al. 2017, Section 3.2] for the derivation. We claim that, apart from $\lambda = 1$, (3-32) and (3-33) have the same set of unstable eigenvalues. This is proved by a straightforward adaptation of the proof of Proposition 3.1 in [Costin et al. 2017].

To establish the nonexistence of unstable eigenvalues of the supersymmetric problem (3-33) we follow the proof of Theorem 4.1 in [Costin et al. 2017]. We start by introducing the change of variables

$$x = \rho^2, \quad \tilde{v}(\rho) = \frac{x}{\sqrt{155-74x}}\,y(x). \tag{3-34}$$

Equation (3-33) transforms into Heun's equation in its canonical form,

$$y''(x) + \left(\frac{13}{2x} + \frac{\lambda-3}{x-1} - \frac{74}{74x-155}\right)y'(x) + \frac{74\lambda(\lambda+3)x - (155\lambda^2 + 775\lambda + 1656)}{4x(x-1)(74x-155)}y(x) = 0. \tag{3-35}$$

Note that (3-34) preserves the analyticity of solutions at 0 and 1, and consequently, (3-33) and (3-35) have the same set of eigenvalues. The Frobenius indices of (3-35) at $x = 0$ are $s_1 = 0$ and $s_2 = -\frac{11}{2}$, so its normalized analytic solution at $x = 0$ is given by the power series

$$\sum_{n=0}^{\infty} a_n(\lambda)x^n, \quad a_0(\lambda) = 1. \tag{3-36}$$

The strategy is to study the asymptotic behavior of the coefficients $a_n(\lambda)$ as $n \to \infty$. More precisely, we prove that if $\lambda \in \overline{\mathbb{H}}$[1] then $\lim_{n\to\infty} a_{n+1}(\lambda)/a_n(\lambda) = 1$. Since $x = 1$ is the only singular point of (3-35) on the unit circle, it follows that the solution given by the series (3-36) is not analytic at $x = 1$.

First, we obtain the recurrence relation for coefficients $\{a_n(\lambda)\}_{n\in\mathbb{N}_0}$. By inserting (3-36) into (3-35) we get

$$\begin{aligned}310(2n+15)(n+2)a_{n+2}(\lambda)\\ = [155\lambda(\lambda+4n+9)+2(458n^2+2357n+2727)]a_{n+1}(\lambda)-74(\lambda+2n+3)(\lambda+2n)a_n(\lambda),\end{aligned}$$

where $a_{-1}(\lambda) = 0$ and $a_0(\lambda) = 1$, or, written differently,

$$a_{n+2}(\lambda) = A_n(\lambda)\, a_{n+1}(\lambda) + B_n(\lambda)\, a_n(\lambda), \tag{3-37}$$

where

$$A_n(\lambda) = \frac{155\lambda(\lambda+4n+9)+2(458n^2+2357n+2727)}{310(2n+15)(n+2)}$$

and

$$B_n(\lambda) = \frac{-37(\lambda+2n+3)(\lambda+2n)}{155(2n+15)(n+2)}.$$

We now let

$$r_n(\lambda) = \frac{a_{n+1}(\lambda)}{a_n(\lambda)}, \tag{3-38}$$

and thereby transform (3-37) into

$$r_{n+1}(\lambda) = A_n(\lambda) + \frac{B_n(\lambda)}{r_n(\lambda)}, \tag{3-39}$$

with the initial condition

$$r_0(\lambda) = \frac{a_1(\lambda)}{a_0(\lambda)} = A_{-1}(\lambda) = \tfrac{1}{26}\lambda^2 + \tfrac{5}{26}\lambda + \tfrac{828}{2015}.$$

Analogous to Lemma 4.2 in [Costin et al. 2017] we have that, given $\lambda \in \overline{\mathbb{H}}$, either

$$\lim_{n\to\infty} r_n(\lambda) = 1 \tag{3-40}$$

or

$$\lim_{n\to\infty} r_n(\lambda) = \tfrac{74}{155}. \tag{3-41}$$

Our aim is to prove that (3-40) holds throughout $\overline{\mathbb{H}}$. We do that by approximately solving (3-39) for $\lambda \in \overline{\mathbb{H}}$. Namely, we define an approximate solution (also called a quasisolution)

$$\tilde{r}_n(\lambda) = \frac{\lambda^2}{4n^2+28n+27} + \frac{\lambda}{n+7} + \frac{2n+12}{2n+23}$$

to (3-39); see [Costin et al. 2016, Section 4.1] for a discussion on how to obtain such an expression. Subsequently, we let

$$\delta_n(\lambda) = \frac{r_n(\lambda)}{\tilde{r}_n(\lambda)} - 1 \tag{3-42}$$

[1] Here, as in [Costin et al. 2017], $\overline{\mathbb{H}}$ denotes the closed complex right half-plane.

and from (3-39) we get the recurrence relation

$$\delta_{n+1} = \varepsilon_n - C_n \frac{\delta_n}{1+\delta_n} \tag{3-43}$$

for δ_n, where

$$\varepsilon_n = \frac{A_n \tilde{r}_n + B_n}{\tilde{r}_n \tilde{r}_{n+1}} - 1 \quad \text{and} \quad C_n = \frac{B_n}{\tilde{r}_n \tilde{r}_{n+1}}. \tag{3-44}$$

Now, for all $\lambda \in \overline{\mathbb{H}}$ and $n \geq 7$ we have the bounds

$$|\delta_7(\lambda)| \leq \tfrac{1}{3}, \quad |\varepsilon_n(\lambda)| \leq \tfrac{1}{12}, \quad |C_n(\lambda)| \leq \tfrac{1}{2}. \tag{3-45}$$

The last two inequalities above are proved in the same way as the corresponding ones in Lemma 4.4 in [Costin et al. 2017]. However, the proof of the first one needs to be slightly adjusted and we provide it in the appendix; see Proposition B.1. Next, by a simple inductive argument we conclude from (3-43) and (3-45) that

$$|\delta_n(\lambda)| \leq \tfrac{1}{3} \quad \text{for all } n \geq 7 \text{ and } \lambda \in \overline{\mathbb{H}}. \tag{3-46}$$

Since for any fixed $\lambda \in \overline{\mathbb{H}}$ we have $\lim_{n\to\infty} \tilde{r}_n(\lambda) = 1$, (3-46) and (3-42) exclude the case (3-41). Hence, (3-40) holds throughout $\overline{\mathbb{H}}$ and we conclude that there are no unstable eigenvalues of the supersymmetric problem (3-33), thus arriving at a contradiction and thereby completing the proof of the proposition. ☐

Remark 3.8. Apart from $\lambda = 1$ the point spectrum of the operator $\boldsymbol{L}$ is completely contained in the open left half-plane. It is natural to try to locate the eigenvalues that are closest to the imaginary axis, as their location is typically related to the rate of convergence to the blowup solution u^T. Our numerical calculations indicate that $-0.98 \pm 3.76\,i$ is the approximate location of the pair of (complex conjugate) stable eigenvalues with the largest real parts. It is interesting to contrast this with the analogous spectral problems for equivariant wave maps into the sphere and Yang–Mills fields, where all eigenvalues appear to be real; see [Bizoń and Biernat 2015].

Corollary 3.9. *We have*

$$\sigma(\boldsymbol{L}) \subseteq \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < 0\} \cup \{1\}.$$

Proof. Assume there exists a $\lambda \in \sigma(\boldsymbol{L}) \setminus \{1\}$ with $\operatorname{Re}\lambda \geq 0$. From (3-29) we see that λ is contained in the resolvent set of $\boldsymbol{L}_0$. Therefore, we have the identity

$$\lambda - \boldsymbol{L} = [1 - \boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)](\lambda - \boldsymbol{L}_0). \tag{3-47}$$

This implies that $1 \in \sigma(\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda))$ and since $\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)$ is compact, it follows that $1 \in \sigma_p(\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda))$. Thus, there exists a nontrivial $\boldsymbol{f} \in \mathcal{H}$ such that $[1 - \boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)]\boldsymbol{f} = 0$. Consequently, $\boldsymbol{u} := \boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\boldsymbol{f} \neq 0$ satisfies $(\lambda - \boldsymbol{L})\boldsymbol{u} = 0$ and thus, $\lambda \in \sigma_p(\boldsymbol{L})$, but this is in conflict with Proposition 3.7. ☐

3K. ***The eigenspace of the isolated eigenvalue.*** In this section, we prove that the (geometric) eigenspace of the isolated eigenvalue $\lambda = 1$ for the full linear operator $\boldsymbol{L}$ is spanned by

$$\boldsymbol{g}(\rho) := \begin{bmatrix} g_1(\rho) \\ g_2(\rho) \end{bmatrix} = \begin{bmatrix} \phi_0'(\rho) \\ \rho\phi_0''(\rho) + 2\phi_0'(\rho) \end{bmatrix}. \tag{3-48}$$

Namely, we are looking for all $\boldsymbol{u} = (u_1, u_2) \in \mathcal{D}(\boldsymbol{L}) \setminus \{0\}$ which belong to $\ker(1 - \boldsymbol{L})$. A straightforward calculation shows that the spectral equation $(1-\boldsymbol{L})\boldsymbol{u} = 0$ is equivalent to the system of ordinary differential equations

$$\begin{cases} u_2(\rho) = \rho u_1'(\rho) + 2u_1(\rho), \\ (1-\rho^2)u_1''(\rho) + (10/\rho - 6\rho)u_1'(\rho) - \big(6 + (8/\rho^2)n'(\phi_0(\rho))\big)u_1(\rho) = 0 \end{cases} \tag{3-49}$$

for $\rho \in (0, 1)$. One can easily verify that a fundamental system of the second equation is given by the functions $\phi_0'(\rho)$ and $\rho^{-9}A(\rho)$, where $A(\rho)$ is analytic and nonvanishing at $\rho = 0$. We can therefore write the general solution to the second equation as

$$u_1(\rho) = C_1\phi_0'(\rho) + C_2\frac{A(\rho)}{\rho^9}.$$

The condition $\boldsymbol{u} \in \mathcal{D}(\boldsymbol{L})$ requires u_1 to lie in the Sobolev space $H^6_{\mathrm{rad}}(\mathbb{B}^{11})$. Since $\phi_0' \in C^\infty[0, 1]$, this requirement yields $C_2 = 0$ which, according to the first equation in (3-49), gives $\boldsymbol{u} = C_1\boldsymbol{g}$. In conclusion,

$$\ker(1 - \boldsymbol{L}) = \langle \boldsymbol{g} \rangle, \tag{3-50}$$

as initially claimed.

3L. ***Time evolution of the linearized problem.*** To get around the spurious instability on the linear level, we use the fact that $\lambda = 1$ is isolated to introduce a (nonorthogonal) spectral projection $\boldsymbol{P}$ and study the subspace semigroup $\boldsymbol{S}(\tau)(1 - \boldsymbol{P})$. From Corollary 3.9 we then infer that the spectrum of its generator is contained in the left half-plane. This does not necessarily imply the desired decay on $\boldsymbol{S}(\tau)(1 - \boldsymbol{P})$. We nonetheless establish such a decay by first proving uniform boundedness of the resolvent of $\boldsymbol{L}$ in a half-plane that strictly contains $\overline{\mathbb{H}}$ and then using the Gearhart–Prüss theorem. For this purpose, we define

$$\Omega_{\varepsilon,R} := \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda \geq -1 + \varepsilon, |\lambda| \geq R\}$$

for $\varepsilon, R > 0$.

Proposition 3.10. *Let $\varepsilon > 0$. Then there exists a constant $R_\varepsilon > 0$ such that the resolvent $\boldsymbol{R_L}$ exists on $\Omega_{\varepsilon,R_\varepsilon}$ and satisfies*

$$\|\boldsymbol{R_L}(\lambda)\| \leq \frac{2}{\varepsilon}$$

for all $\lambda \in \Omega_{\varepsilon,R_\varepsilon}$.

Proof. Fix $\varepsilon > 0$ and take $\lambda \in \Omega_{\varepsilon,R}$ for an arbitrary $R \geq 2$. Then $\lambda \in \rho(\boldsymbol{L}_0)$ and the identity (3-47) holds. The proof proceeds as follows. For large enough R, we show that the operator $1 - \boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)$ is invertible in $\Omega_{\varepsilon,R}$ and $\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)$ and $[1 - \boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)]^{-1}$ are uniformly norm bounded there. Via (3-47) this implies the desired bound on $\boldsymbol{R_L}(\lambda)$.

First of all, semigroup theory yields the estimate

$$\|\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\| \leq \frac{1}{\operatorname{Re}\lambda + 1}; \tag{3-51}$$

see [Engel and Nagel 2000, p. 55, Theorem 1.10]. Next, by a Neumann series argument, the operator $1-\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)$ is invertible if $\|\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\|<1$. To prove smallness of $\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)$, we recall the definition of $\boldsymbol{L}'$, (3-22),

$$\boldsymbol{L}'\boldsymbol{u}(\rho):=\begin{bmatrix}0\\ \widetilde{W}(\rho)u_1(\rho)\end{bmatrix},\qquad \widetilde{W}(\rho)=-\frac{8}{\rho^2}n'(\phi_0(\rho))\quad\text{for } n(x)=14x^3-111x^5.$$

Let $\boldsymbol{u}=\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\boldsymbol{f}$ or, equivalently, $(\lambda-\boldsymbol{L}_0)\boldsymbol{u}=\boldsymbol{f}$. The latter equation implies

$$(\lambda+1)u_1(\rho)=u_2(\rho)-\rho u_1'(\rho)+f_1(\rho).$$

Now we use Lemma 4.1 from [Donninger and Schörkhuber 2017] and $\|\widetilde{W}^{(k)}\|_{L^\infty(0,1)}\lesssim 1$ for all $k\in\{0,1,\dots,5\}$ to obtain

$$\begin{aligned}|\lambda+1|\|\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\boldsymbol{f}\|=|\lambda+1|\|\boldsymbol{L}'\boldsymbol{u}\|&\simeq\|\widetilde{W}(u_2-(\cdot)u_1'+f_1)\|_{H^5_{\mathrm{rad}}(\mathbb{B}^{11})}\\&\lesssim\|u_2\|_{H^5_{\mathrm{rad}}(\mathbb{B}^{11})}+\|(\cdot)u_1'\|_{H^5_{\mathrm{rad}}(\mathbb{B}^{11})}+\|f_1\|_{H^5_{\mathrm{rad}}(\mathbb{B}^{11})}\\&\lesssim\|u_2\|_{H^5_{\mathrm{rad}}(\mathbb{B}^{11})}+\|u_1\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})}+\|f_1\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})}\\&\lesssim\|\boldsymbol{u}\|+\|\boldsymbol{f}\|\lesssim\left(\frac{1}{\operatorname{Re}\lambda+1}+1\right)\|\boldsymbol{f}\|\lesssim\|\boldsymbol{f}\|,\end{aligned}$$

where we used (3-51). In other words,

$$\|\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\|\lesssim\frac{1}{|\lambda+1|}\le\frac{1}{|\lambda|-1}\le\frac{1}{R-1}$$

and by choosing R sufficiently large, we can achieve $\|\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\|\le\frac{1}{2}$. As a consequence, $[1-\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)]^{-1}$ exists for $\lambda\in\Omega_{\varepsilon,R_\varepsilon}$ and we obtain the bound

$$\begin{aligned}\|\boldsymbol{R}_{\boldsymbol{L}}(\lambda)\|&=\|\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)[1-\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)]^{-1}\|\\&\le\|\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\|\|[1-\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)]^{-1}\|\\&\le\|\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\|\sum_{n=0}^{\infty}\|\boldsymbol{L}'\boldsymbol{R}_{\boldsymbol{L}_0}(\lambda)\|^n\le\frac{2}{\varepsilon}.\end{aligned}$$ ◻

We now show the existence of a projection $\boldsymbol{P}$ which decomposes the Hilbert space $\mathcal{H}$ into a stable and an unstable subspace and furthermore prove that data from the stable subspace lead to solutions that decay exponentially in time. We also remark that it is crucial to ensure that rank $\boldsymbol{P}=1$, i.e., that $\boldsymbol{g}$ is the only unstable direction in $\mathcal{H}$.

Proposition 3.11. *There exists a projection operator*

$$\boldsymbol{P}\in\mathcal{B}(\mathcal{H}),\quad \boldsymbol{P}:\mathcal{H}\to\langle\boldsymbol{g}\rangle,$$

which commutes with the semigroup $(\boldsymbol{S}(\tau))_{\tau\ge0}$. In addition, we have

$$\boldsymbol{S}(\tau)\boldsymbol{P}\boldsymbol{f}=e^{\tau}\boldsymbol{P}\boldsymbol{f}\tag{3-52}$$

and there are constants $C, \varepsilon > 0$ such that

$$\|(1-\boldsymbol{P})\boldsymbol{S}(\tau)\boldsymbol{f}\| \leq Ce^{-\varepsilon\tau}\|(1-\boldsymbol{P})\boldsymbol{f}\| \tag{3-53}$$

for all $\boldsymbol{f} \in \mathcal{H}$ and $\tau \geq 0$.

Proof. By Proposition 3.7, the eigenvalue $\lambda = 1$ of the operator $\boldsymbol{L}$ is isolated. We therefore introduce the spectral projection

$$\boldsymbol{P} : \mathcal{H} \to \mathcal{H}, \quad \boldsymbol{P} := \frac{1}{2\pi i}\int_\gamma \boldsymbol{R}_{\boldsymbol{L}}(\mu)\, d\mu,$$

where γ is a positively oriented circle around $\lambda = 1$. The radius of the circle is chosen small enough so that γ is completely contained inside the resolvent set of $\boldsymbol{L}$ and such that the interior of γ contains no spectral points of $\boldsymbol{L}$ other than $\lambda = 1$. The projection $\boldsymbol{P}$ commutes with the operator $\boldsymbol{L}$ and therefore with the semigroup $\boldsymbol{S}(\tau)$. Moreover, the Hilbert space $\mathcal{H}$ is decomposed as $\mathcal{H} = \mathcal{M} \oplus \mathcal{N}$, where $\mathcal{M} := \operatorname{rg} \boldsymbol{P}$ and $\mathcal{N} := \operatorname{rg}(1-\boldsymbol{P}) = \ker \boldsymbol{P}$. Also, the spaces $\mathcal{M}$ and $\mathcal{N}$ reduce the operator $\boldsymbol{L}$, which is therefore decomposed into $\boldsymbol{L}_{\mathcal{M}}$ and $\boldsymbol{L}_{\mathcal{N}}$. The spectra of these operators are given by

$$\sigma(\boldsymbol{L}_{\mathcal{N}}) = \sigma(\boldsymbol{L}) \setminus \{1\}, \quad \sigma(\boldsymbol{L}_{\mathcal{M}}) = \{1\}. \tag{3-54}$$

We refer the reader to [Kato 1980, Chapter III, Section 6.4] for these standard results.

To proceed with the proof we show that $\operatorname{rank} \boldsymbol{P} := \dim \operatorname{rg} \boldsymbol{P} < +\infty$. We argue by contradiction and assume that $\operatorname{rank} \boldsymbol{P} = +\infty$. This means that $\lambda = 1$ belongs to the essential spectrum of $\boldsymbol{L}$; see [Kato 1980, p. 239, Theorem 5.28]. But according to Proposition 3.6 the operator $\boldsymbol{L}_0 = \boldsymbol{L} - \boldsymbol{L}'$ is a compact perturbation of $\boldsymbol{L}$, and due to the stability of the essential spectrum under compact perturbations we conclude that $\lambda = 1$ is a spectral point of $\boldsymbol{L}_0$. However, this is in conflict with (3-29), and therefore $\operatorname{rank} \boldsymbol{P} < +\infty$.

Now we prove that $\langle \boldsymbol{g} \rangle = \operatorname{rg} \boldsymbol{P}$. From the definition of the projection $\boldsymbol{P}$ we have $\boldsymbol{P}\boldsymbol{g} = \boldsymbol{g}$. Therefore $\langle \boldsymbol{g} \rangle \subseteq \operatorname{rg} \boldsymbol{P}$ and it remains to prove the reverse inclusion. From the fact that the operator $1 - \boldsymbol{L}_{\mathcal{M}}$ acts on the finite-dimensional Hilbert space $\mathcal{M} = \operatorname{rg} \boldsymbol{P}$ and (3-54) we infer that $\lambda = 0$ is the only spectral point of $1 - \boldsymbol{L}_{\mathcal{M}}$. Hence, $1 - \boldsymbol{L}_{\mathcal{M}}$ is nilpotent; i.e., there exists a $k \in \mathbb{N}$ such that

$$(1 - \boldsymbol{L}_{\mathcal{M}})^k \boldsymbol{u} = 0$$

for all $\boldsymbol{u} \in \operatorname{rg} \boldsymbol{P}$ and we assume k to be minimal. Due to (3-50) the claim follows immediately for $k = 1$. We therefore assume that $k \geq 2$. This implies the existence of a nontrivial function $\boldsymbol{u} \in \operatorname{rg} \boldsymbol{P} \subseteq \mathcal{D}(\boldsymbol{L})$ such that $(1 - \boldsymbol{L}_{\mathcal{M}})\boldsymbol{u}$ is nonzero and belongs to $\ker(1 - \boldsymbol{L}_{\mathcal{M}}) \subseteq \ker(1 - \boldsymbol{L}) = \langle \boldsymbol{g} \rangle$. Therefore $(1 - \boldsymbol{L})\boldsymbol{u} = \alpha \boldsymbol{g}$ for some $\alpha \in \mathbb{C} \setminus \{0\}$. For convenience and without loss of generality, we set $\alpha = -1$. By a straightforward computation we see that the first component of $\boldsymbol{u}$ satisfies the differential equation

$$(1-\rho^2)u_1''(\rho) + \left(\frac{10}{\rho} - 6\rho\right)u_1'(\rho) - \left(6 + \frac{8}{\rho^2}n'(\phi_0(\rho))\right)u_1(\rho) = G(\rho) \tag{3-55}$$

for $\rho \in (0, 1)$, where

$$G(\rho) := 2\rho\phi_0''(\rho) + 5\phi_0'(\rho), \quad \rho \in [0, 1].$$

To find a general solution to (3-55) we first observe that

$$\hat{u}_1(\rho) := g_1(\rho) = \phi_0'(\rho), \quad \rho \in (0,1),$$

is a particular solution to the homogeneous equation

$$(1-\rho^2)u_1''(\rho) + \left(\frac{10}{\rho} - 6\rho\right)u_1'(\rho) - \left(6 + \frac{8}{\rho^2}n'(\phi_0(\rho))\right)u_1(\rho) = 0;$$

see (3-48) and (3-49). Note that the Wronskian for the equation above is

$$\mathcal{W}(\rho) := \frac{(1-\rho^2)^2}{\rho^{10}}.$$

Therefore, another linearly independent solution is

$$\hat{u}_2(\rho) := \hat{u}_1(\rho) \int_\rho^1 \frac{(1-x^2)^2}{x^{10}} \frac{1}{\phi_0'(x)^2}\, dx$$

for all $\rho \in (0,1)$. Note that near $\rho = 0$ we have the expansion

$$\hat{u}_2(\rho) = \frac{1}{\rho^9} \sum_{j=0}^\infty a_j \rho^j, \quad a_0 \neq 0,$$

as already indicated in Section 3K. Furthermore, we have

$$\hat{u}_2(\rho) = (1-\rho)^3 \sum_{j=0}^\infty b_j (1-\rho)^j, \quad b_0 \neq 0,$$

near $\rho = 1$. Now, by the variation-of-constants formula we see that the general solution to (3-55) can be written as

$$u_1(\rho) = c_1\hat{u}_1(\rho) + c_2\hat{u}_2(\rho) + \hat{u}_2(\rho)\int_0^\rho \frac{\hat{u}_1(y)G(y)y^{10}}{(1-y^2)^3}\, dy - \hat{u}_1(\rho)\int_0^\rho \frac{\hat{u}_2(y)G(y)y^{10}}{(1-y^2)^3}\, dy$$

for some constants $c_1, c_2 \in \mathbb{C}$ and for all $\rho \in (0,1)$. The fact that $u_1 \in H^6_{\text{rad}}(\mathbb{B}^{11})$ implies $c_2 = 0$ as $\hat{u}_2$ has a ninth-order pole at $\rho = 0$. Therefore

$$u_1(\rho) = c_1\hat{u}_1(\rho) + \hat{u}_2(\rho)\int_0^\rho \frac{\hat{u}_1(y)G(y)y^{10}}{(1-y^2)^3}\, dy - \hat{u}_1(\rho)\int_0^\rho \frac{\hat{u}_2(y)G(y)y^{10}}{(1-y^2)^3}\, dy. \tag{3-56}$$

The last term in (3-56) is smooth on $[0,1]$. To analyze the second term, we set

$$\mathcal{I}(\rho) := \hat{u}_2(\rho)\int_0^\rho \frac{F(y)}{(1-y)^3}\, dy,$$

where

$$F(y) := \frac{\hat{u}_1(y)G(y)y^{10}}{(1+y)^3} = \frac{y^{10}(2y\phi_0'(y)\phi_0''(y) + 5\phi_0'(y)^2)}{(1+y)^3}.$$

By a direct calculation we get $F''(1) \neq 0$ and thus, the expansion of $\mathcal{I}(\rho)$ near $\rho = 1$ contains a term of the form $(1-\rho)^3\log(1-\rho)$. Consequently, $\mathcal{I}^{(4)} \notin L^2\left(\frac{1}{2}, 1\right)$, which is a contradiction to $u_1 \in H^6_{\text{rad}}(\mathbb{B}^{11})$.

Finally we prove (3-52) and (3-53). Note that (3-52) follows from the fact that $\lambda = 1$ is an eigenvalue of the operator $\boldsymbol{L}$ with eigenfunction $\boldsymbol{g}$ and $\operatorname{rg} \boldsymbol{P} = \langle \boldsymbol{g} \rangle$. Next, from Corollary 3.9 and Proposition 3.10 we deduce the existence of constants $D, \varepsilon > 0$ such that

$$\|\boldsymbol{R_L}(\lambda)(1 - \boldsymbol{P})\| \leq D$$

for all complex λ with $\operatorname{Re} \lambda > -\varepsilon$. Thus, (3-53) follows from the Gearhart–Prüss theorem; see [Engel and Nagel 2000, p. 302, Theorem 1.11]. □

3M. ***Estimates for the nonlinearity.*** In the next section we employ a fixed-point argument to prove the existence of decaying solutions to (3-26) for small initial data. To accomplish that, we need a Lipschitz-type estimate for the nonlinear operator $\boldsymbol{N}$; see (3-23). We first define

$$\mathcal{B}_\delta := \{\boldsymbol{u} \in \mathcal{H} : \|\boldsymbol{u}\| = \|(u_1, u_2)\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11}) \times H^5_{\mathrm{rad}}(\mathbb{B}^{11})} \leq \delta\}.$$

Lemma 3.12. *Let $\delta > 0$. For $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{B}_\delta$, we have*

$$\|\boldsymbol{N}(\boldsymbol{u}) - \boldsymbol{N}(\boldsymbol{v})\| \lesssim (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)\|\boldsymbol{u} - \boldsymbol{v}\|. \tag{3-57}$$

Remark 3.13. From this lemma we infer the estimate

$$\|\boldsymbol{N}(\boldsymbol{u}) - \boldsymbol{N}(\boldsymbol{v})\| \lesssim \delta \|\boldsymbol{u} - \boldsymbol{v}\|. \tag{3-58}$$

Therefore, the implied Lipschitz constant in (3-58) can be made as small as needed by adjusting the size of δ.

Proof. Based on (3-23) and (3-3), the difference $N(\rho, u) - N(\rho, v)$ can be written as

$$N(\rho, u) - N(\rho, v) = \sum_{j=1}^{4} n_j(\rho^2)(u^{j+1} - v^{j+1}), \tag{3-59}$$

where $n_j \in C^\infty[0, 1]$. For $\delta > 0$, we have $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{B}_\delta$, and due to the bilinear estimate

$$\|f_1 f_2\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})} \lesssim \|f_1\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})} \|f_2\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})},$$

we have

$$\begin{aligned}
\|\boldsymbol{N}(\boldsymbol{u}) - \boldsymbol{N}(\boldsymbol{v})\| &= \|N(\cdot, u_1) - N(\cdot, v_1)\|_{H^5_{\mathrm{rad}}(\mathbb{B}^{11})} \\
&\leq \|N(\cdot, u_1) - N(\cdot, v_1)\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})} \\
&\lesssim \sum_{j=1}^{4} \|n_j((\cdot)^2)\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})} \|u_1^{j+1} - v_1^{j+1}\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})} \\
&\lesssim (\|u_1\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})} + \|v_1\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})}) \|u_1 - v_1\|_{H^6_{\mathrm{rad}}(\mathbb{B}^{11})} \\
&\leq (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)\|\boldsymbol{u} - \boldsymbol{v}\|.
\end{aligned}$$

□

3N. ***The abstract nonlinear Cauchy problem.*** In this section we treat the existence and uniqueness of solutions to (3-20) for small initial data. According to Definition 3.4 we study the integral equation

$$\boldsymbol{\Phi}(\tau) = \boldsymbol{S}(\tau)\boldsymbol{U}(\boldsymbol{v}, T) + \int_0^\tau \boldsymbol{S}(\tau - s)\boldsymbol{N}(\Phi(s))\, ds \tag{3-60}$$

for $\tau \geq 0$ and $\boldsymbol{v} \in \mathcal{H}$ small. In order to employ a fixed-point argument, we introduce the necessary definitions. First, we define a Banach space

$$\mathcal{X} := \{\Phi \in C([0, \infty), \mathcal{H}) : \|\Phi\|_{\mathcal{X}} := \sup_{\tau>0} e^{\varepsilon\tau}\|\Phi(\tau)\| < \infty\}, \tag{3-61}$$

where ε is sufficiently small and fixed. We denote by $\mathcal{X}_\delta$ the closed ball in $\mathcal{X}$ with radius δ; that is,

$$\mathcal{X}_\delta := \{\Phi \in \mathcal{X} : \|\Phi\|_{\mathcal{X}} \leq \delta\}. \tag{3-62}$$

Finally, we define the correction term

$$\boldsymbol{C}(\Phi, \boldsymbol{u}) := \boldsymbol{P}\left(\boldsymbol{u} + \int_0^\infty e^{-s}\boldsymbol{N}(\Phi(s))\, ds\right),$$

and set

$$\boldsymbol{K}(\Phi, \boldsymbol{u})(\tau) := \boldsymbol{S}(\tau)(\boldsymbol{u} - \boldsymbol{C}(\Phi, \boldsymbol{u})) + \int_0^\tau \boldsymbol{S}(\tau - s)\boldsymbol{N}(\Phi(s))\, ds.$$

The correction term serves the purpose of suppressing the exponential growth of the semigroup $\boldsymbol{S}(\tau)$ on the unstable space. We have the following result.

Theorem 3.14. *There exist constants $\delta, C > 0$ such that for every $\boldsymbol{u} \in \mathcal{H}$ which satisfies $\|\boldsymbol{u}\| \leq \delta/C$, there exists a unique $\Phi_{\boldsymbol{u}} \in \mathcal{X}_\delta$ such that*

$$\Phi_{\boldsymbol{u}} = \boldsymbol{K}(\Phi_{\boldsymbol{u}}, \boldsymbol{u}). \tag{3-63}$$

In addition, the solution $\Phi_{\boldsymbol{u}}$ is unique in the whole space $\mathcal{X}$ and the solution map $\boldsymbol{u} \mapsto \Phi_{\boldsymbol{u}}$ is Lipschitz continuous.

The proof coincides with the one of Theorem 3.7 in [Chatzikaleas et al. 2017].

We now study the initial data $\boldsymbol{U}(\boldsymbol{v}, T)$, see (3-25), and prove its continuity in T near T_0. For that reason we define

$$\mathcal{H}^R := H^6_{\text{rad}} \times H^5_{\text{rad}}(\mathbb{B}^{11}_R),$$

with the induced norm

$$\|\boldsymbol{w}\|^2_{\mathcal{H}^R} = \|w_1(|\cdot|)\|^2_{H^6(\mathbb{B}^{11}_R)} + \|w_2(|\cdot|)\|^2_{H^5(\mathbb{B}^{11}_R)}.$$

Lemma 3.15. *Fix $T_0 > 0$. Let $|\cdot|^{-1}\boldsymbol{v} \in \mathcal{H}^{T_0+\delta}$ for δ positive and sufficiently small. Then the map*

$$T \mapsto \boldsymbol{U}(\boldsymbol{v}, T) : [T_0 - \delta, T_0 + \delta] \to \mathcal{H}$$

is continuous. Furthermore, for all $T \in [T_0 - \delta, T_0 + \delta]$,

$$\||\cdot|^{-1}\boldsymbol{v}\|_{\mathcal{H}^{T_0+\delta}} \leq \delta \quad \Longrightarrow \quad \|\boldsymbol{U}(\boldsymbol{v}, T)\| \lesssim \delta.$$

Proof. We prove the result for $T_0 = 1$ only, as the general case is treated similarly. Assume $|\cdot|^{-1}\boldsymbol{v} \in \mathcal{H}^{1+\delta}$ for δ positive but less than $\frac{1}{2}T_0 = \frac{1}{2}$. We first introduce some auxiliary facts. Namely, by scaling we see that for $f \in H^6_{\text{rad}}(\mathbb{B}^{11}_{1+\delta})$ and $T \in [1-\delta, 1+\delta]$,

$$\|f(|T\,\cdot|)\|_{H^6(\mathbb{B}^{11}_1)} \lesssim \|f(|\cdot|)\|_{H^6(\mathbb{B}^{11}_{1+\delta})}.$$

Furthermore, from the density of $C^\infty_{\text{even}}[0, 1+\delta]$ in $H^6_{\text{rad}}(\mathbb{B}^{11}_{1+\delta})$ we conclude that given $\varepsilon > 0$, there exists a $\tilde{v}_1 \in C^\infty_{\text{even}}[0, 1+\delta]$ such that $\||\cdot|^{-1}v_1(|\cdot|) - \tilde{v}_1(|\cdot|)\|_{H^6_{\text{rad}}(\mathbb{B}^{11}_{1+\delta})} < \varepsilon$. Also, the functions $(1/T\rho)\phi_0(T\rho)$ and $\tilde{v}(T\rho)$ are smooth on $[0, 1]$ for $T \in [1-\delta, 1+\delta]$. Therefore,

$$\lim_{T\to\widetilde{T}} \big\||T\,\cdot|^{-1}\big(\phi_0(|T\,\cdot|) - \phi_0(|\widetilde{T}\,\cdot|)\big)\big\|_{H^6(\mathbb{B}^{11})} + \|\tilde{v}_1(|T\,\cdot|) - \tilde{v}_1(|\widetilde{T}\,\cdot|)\|_{H^6(\mathbb{B}^{11})} = 0. \tag{3-64}$$

Using these facts, we prove the continuity of the first component of the map $T \to \boldsymbol{U}(\boldsymbol{v}, T)$. Namely, given $\varepsilon > 0$, there exists a $\tilde{v}_1 \in C^\infty_{\text{even}}[0, 1+\delta]$ such that for $T, \widetilde{T} \in [1-\delta, 1+\delta]$ we have

$$\begin{aligned}
&\big\|[\boldsymbol{U}(\boldsymbol{v}, T)]_1 - [\boldsymbol{U}(\boldsymbol{v}, \widetilde{T})]_1\big\|_{H^6(\mathbb{B}^{11})} \\
&\qquad = \big\||\cdot|^{-1}\phi_0(|T\,\cdot|) + |\cdot|^{-1}v_1(|T\,\cdot|) - |\cdot|^{-1}\phi_0(|\widetilde{T}\,\cdot|) - |\cdot|^{-1}v_1(|\widetilde{T}\,\cdot|)\big\|_{H^6(\mathbb{B}^{11})} \\
&\qquad \lesssim \big\||\cdot|^{-1}(\phi_0(|T\,\cdot|) - \phi_0(|\widetilde{T}\,\cdot|))\big\|_{H^6(\mathbb{B}^{11})} + \big\||T\,\cdot|^{-1}v_1(|T\,\cdot|) - \tilde{v}_1(|T\,\cdot|)\big\|_{H^6(\mathbb{B}^{11})} \\
&\qquad\qquad + \big\|\tilde{v}_1(|T\,\cdot|) - \tilde{v}_1(|\widetilde{T}\,\cdot|)\big\|_{H^6(\mathbb{B}^{11})} + \big\||T\,\cdot|^{-1}\tilde{v}_1(|\widetilde{T}\,\cdot|) - v_1(|\widetilde{T}\,\cdot|)\big\|_{H^6(\mathbb{B}^{11})} \\
&\qquad \lesssim \big\||T\,\cdot|^{-1}(\phi_0(|T\,\cdot|) - \phi_0(|\widetilde{T}\,\cdot|))\big\|_{H^6(\mathbb{B}^{11})} + \big\||\cdot|^{-1}v_1(|\cdot|) - \tilde{v}_1(|\cdot|)\big\|_{H^6(\mathbb{B}^{11}_{1+\delta})} \\
&\qquad\qquad + \big\|\tilde{v}_1(|T\,\cdot|) - \tilde{v}_1(|\widetilde{T}\,\cdot|)\big\|_{H^6(\mathbb{B}^{11})} \\
&\qquad \le \big\||T\,\cdot|^{-1}\big(\phi_0(|T\,\cdot|) - \phi_0(|\widetilde{T}\,\cdot|)\big)\big\|_{H^6(\mathbb{B}^{11})} + \big\|\tilde{v}_1(|T\,\cdot|) - \tilde{v}_1(|\widetilde{T}\,\cdot|)\big\|_{H^6(\mathbb{B}^{11})} + \varepsilon,
\end{aligned}$$

This together with (3-64) implies that $[\boldsymbol{U}(\boldsymbol{v}, T)]_1$ is continuous. The second component is treated analogously. Now, given $\||\cdot|^{-1}\boldsymbol{v}\|_{\mathcal{H}^{1+\delta}} \le \delta$ and $T \in [1-\delta, 1+\delta]$, we have

$$\begin{aligned}
\|[\boldsymbol{U}(\boldsymbol{v}, T)]_1\|_{H^6(\mathbb{B}^{11})} &= \big\||\cdot|^{-1}\phi_0(|T\,\cdot|) - |\cdot|^{-1}\phi_0(|\cdot|) + |\cdot|^{-1}v_1(|T\,\cdot|)\big\|_{H^6(\mathbb{B}^{11})} \\
&\lesssim |T-1| + \||\cdot|^{-1}v_1\|_{H^6(\mathbb{B}^{11}_{1+\delta})} \lesssim \delta.
\end{aligned}$$

We obtain a similar estimate for the second component and finally deduce that

$$\|\boldsymbol{U}(\boldsymbol{v}, T)\| \lesssim \delta. \qquad \square$$

As already mentioned, the unstable eigenvalue $\lambda = 1$ is present due to the freedom of choice of the parameter T, and is therefore not considered a "real" instability of the linear problem. The following theorem is the precise version of this statement. Namely, for a given T_0 and small enough initial data $\boldsymbol{v}$, there exists a $T_{\boldsymbol{v}}$ close to T_0 that makes the correction term $\boldsymbol{C}(\Phi_{\boldsymbol{U}(\boldsymbol{v},T_{\boldsymbol{v}})}, \boldsymbol{U}(\boldsymbol{v}, T_{\boldsymbol{v}}))$ vanish. This in turn allows for proving the existence and uniqueness of an exponentially decaying solution to (3-60).

Theorem 3.16. *Fix $T_0 > 0$. Then there exist $\delta, M > 0$ such that for any $\boldsymbol{v}$ that satisfies*

$$\||\cdot|^{-1}\boldsymbol{v}\|_{\mathcal{H}^{T_0+\delta}} \le \frac{\delta}{M}$$

there exists a $T \in [T_0 - \delta, T_0 + \delta]$ *and a function* $\Phi \in \mathcal{X}_\delta$ *which satisfies*

$$\Phi(\tau) = \boldsymbol{S}(\tau)\boldsymbol{U}(\boldsymbol{v}, T) + \int_0^\tau \boldsymbol{S}(\tau - s)\boldsymbol{N}(\Phi(s))\, ds \tag{3-65}$$

for all $\tau > 0$. *Moreover*, Φ *is the unique solution of this equation in* $C([0, \infty), \mathcal{H})$.

Proof. Let $T_0 > 0$ be fixed. We first prove that for any T in a small neighborhood of T_0 and small enough initial data $\boldsymbol{v}$ there exists a unique solution to (3-63) for $\boldsymbol{u} = \boldsymbol{U}(\boldsymbol{v}, T)$. From Lemma 3.15 we deduce the existence of sufficiently small δ and sufficiently large $M > 0$ so that for every $T \in [T_0 - \delta, T_0 + \delta]$, $\| |\cdot|^{-1}\boldsymbol{v}\|_{\mathcal{H}^{T_0+\delta}} \le \delta/M$ implies $\|\boldsymbol{U}(\boldsymbol{v}, T)\|_{\mathcal{H}} \le \delta/C$ for a large enough $C > 0$. Via Theorem 3.14 this yields the unique solution to (3-63) for every T in the designated range. It remains to show that for small enough $\boldsymbol{v}$, there exists a particular $T_{\boldsymbol{v}} \in [T_0 - \delta, T_0 + \delta]$ that makes the correction term vanish, i.e., $\boldsymbol{C}(\Phi_{\boldsymbol{U}(\boldsymbol{v}, T_{\boldsymbol{v}})}, \boldsymbol{U}(\boldsymbol{v}, T_{\boldsymbol{v}})) = 0$. Since $\boldsymbol{C}$ has values in rg $\boldsymbol{P} = \langle \boldsymbol{g} \rangle$, the latter is equivalent to the existence of a $T_{\boldsymbol{v}} \in [T_0 - \delta, T_0 + \delta]$ such that

$$\big\langle \boldsymbol{C}(\Phi_{\boldsymbol{U}(\boldsymbol{v}, T_{\boldsymbol{v}})}, \boldsymbol{U}(\boldsymbol{v}, T_{\boldsymbol{v}})), \boldsymbol{g} \big\rangle_{\mathcal{H}} = 0. \tag{3-66}$$

By definition, we have

$$\partial_T \begin{bmatrix} (1/\rho)\phi_0((T/T_0)\rho) \\ (T^2/T_0^2)\phi_0'((T/T_0)\rho) \end{bmatrix}\Bigg|_{T=T_0} = \frac{\boldsymbol{g}(\rho)}{T_0}$$

and this yields the expansion

$$\big\langle \boldsymbol{C}(\Phi_{\boldsymbol{U}(\boldsymbol{v}, T_{\boldsymbol{v}})}, \boldsymbol{U}(\boldsymbol{v}, T)), \boldsymbol{g} \big\rangle_{\mathcal{H}} = \frac{\|\boldsymbol{g}\|^2}{T_0}(T - T_0) + O((T - T_0)^2) + O\bigg(\frac{\delta}{M} T^0\bigg) + O(\delta^2 T^0).$$

A simple fixed-point argument now proves (3-66); see [Donninger and Schörkhuber 2017, Theorem 4.15] for full details. ☐

Proof of Theorem 3.1. Fix $T_0 > 0$ and assume the radial initial data $u[0]$ satisfy

$$\big\| |\cdot|^{-1}(u[0] - u^{T_0}[0]) \big\|_{H^6(\mathbb{B}^{11}_{T_0+\delta}) \times H^5(\mathbb{B}^{11}_{T_0+\delta})} \le \frac{\delta}{M_0^2}$$

with $\delta, M_0 > 0$ to be chosen later. We set $\boldsymbol{v} := u[0] - u^{T_0}[0]$; see Section 3E. Then we have

$$\| |\cdot|^{-1}\boldsymbol{v}\|_{\mathcal{H}^{T_0+\delta}} = \big\| |\cdot|^{-1}(u[0] - u^{T_0}[0]) \big\|_{\mathcal{H}^{T_0+\delta}} \le \frac{\delta}{M_0^2}.$$

Now, upon choosing $\delta > 0$ sufficiently small and $M_0 > 0$ sufficiently large, Theorem 3.16 yields a $T \in [T_0 - \delta/M_0, T_0 + \delta/M_0] \subset [1 - \delta, 1 + \delta]$ such that there exists a unique solution $\Phi = (\varphi_1, \varphi_2) \in \mathcal{X}$ to (3-65) with $\|\Phi(\tau)\| \le (\delta/M_0)e^{-2\varepsilon\tau}$ for all $\tau \ge 0$ and some $\varepsilon > 0$. Therefore, by construction,

$$u(t, r) = u^T(t, r) + \frac{r}{T - t}\varphi_1\bigg(\log\frac{T}{T - t}, \frac{r}{T - t}\bigg)$$

solves the original wave maps (3-1). Moreover,

$$\partial_t u(t, r) = \partial_t u^T(t, r) + \frac{r}{(T - t)^2}\varphi_2\bigg(\log\frac{T}{T - t}, \frac{r}{T - t}\bigg).$$

Therefore,

$$\begin{aligned}(T-t)^{k-\frac{9}{2}}\big\||\cdot|^{-1}(u(t,|\cdot|)-u^T(t,|\cdot|))\big\|_{\dot H^k(\mathbb{B}^{11}_{T-t})} &= (T-t)^{k-\frac{11}{2}}\left\|\varphi_1\left(\log\frac{T}{T-t},\frac{|\cdot|}{T-t}\right)\right\|_{\dot H^k(\mathbb{B}^{11}_{T-t})}\\ &= \left\|\varphi_1\left(\log\frac{T}{T-t},|\cdot|\right)\right\|_{\dot H^k(\mathbb{B}^{11})} \le \left\|\Phi\left(\log\frac{T}{T-t}\right)\right\|_{\mathcal{H}}\\ &\le \frac{\delta}{M_0}(T-t)^{2\varepsilon}\end{aligned}$$

for all $t\in[0,T)$ and any integer $0\le k\le 6$. Furthermore,

$$\begin{aligned}&(T-t)^{l-\frac{7}{2}}\big\||\cdot|^{-1}(\partial_t u(t,|\cdot|)-\partial_t u^T(t,|\cdot|))\big\|_{\dot H^l(\mathbb{B}^{11}_{T-t})}\\ &\qquad= (T-t)^{l-\frac{11}{2}}\left\|\varphi_2\left(\log\frac{T}{T-t},\frac{|\cdot|}{T-t}\right)\right\|_{\dot H^l(\mathbb{B}^{11}_{T-t})}\\ &\qquad= \left\|\varphi_2\left(\log\frac{T}{T-t},|\cdot|\right)\right\|_{\dot H^l(\mathbb{B}^{11})} \le \left\|\Phi\left(\log\frac{T}{T-t}\right)\right\|_{\mathcal{H}} \le \frac{\delta}{M_0}(T-t)^{2\varepsilon}\end{aligned}$$

for all $l=0,1,\dots,5$. Finally, by Sobolev embedding we infer

$$\begin{aligned}\|u(t,\cdot)-u^T(t,\cdot)\|_{L^\infty(0,T-t)} &\le (T-t)\big\||\cdot|^{-1}(u(t,|\cdot|)-u^T(t,|\cdot|))\big\|_{L^\infty(0,T-t)}\\ &\lesssim (T-t)\big\||\cdot|^{-1}(u(t,|\cdot|)-u^T(t,|\cdot|))\big\|_{H^{11/2+\varepsilon}(\mathbb{B}^{11}_{T-t})}\\ &\lesssim \frac{\delta}{M_0}(T-t)^{\varepsilon}\end{aligned}$$

and this finishes the proof by setting $M:=M_0^2$. ☐

Remark 3.17. Based on [Donninger and Schörkhuber 2017; Chatzikaleas et al. 2017], the analogue of Theorem 3.1 in any odd dimension $d\ge 11$ follows from the mode stability of the solution u^T. However, a nontrivial adjustment of the method of the proof of Proposition 3.7 is required in order to establish the analogous result for all higher odd d simultaneously. This will be addressed in a forthcoming publication.

Appendix A: Proof of Proposition 2.1

A straightforward computation shows that all sectional curvatures of the manifold N^d are given by either

$$\text{(i)}\ \ \frac{-g''(u)}{g(u)}\quad\text{or}\quad\text{(ii)}\ \ \frac{1-g'(u)^2}{g(u)^2}. \tag{A-1}$$

We first show that the two expressions above are negative provided $d\ge 8$ and $u\in I:=[0,\phi_0(1)]$. For convenience we let $d=e+8$. We now have

$$\frac{g''(u)}{g(u)}=\frac{6(23e+14)^2u^6-63(23e+14)u^4-2(115e+21)u^2+21}{[(23e+14)u^4-7u^2-1]^2}. \tag{A-2}$$

Denote the numerator in the above expression by $N(e,u)$. To show that the first quantity in (A-1) is negative it suffices to prove that $N(e,u)>0$ for $(e,u)\in[0,\infty)\times I$. To that end, it is enough to show that for any fixed $e \geq 0$ the following inequalities hold:

$$\text{(i)}\ N(e,0) > 0, \quad \text{(ii)}\ N(e,\phi_0(1)) > 0 \quad \text{and} \quad \text{(iii)}\ \partial_u^2 N(e,u) < 0 \ \text{ for } u \in I. \tag{A-3}$$

We start by proving the third claim above. Note that it is enough to show that

$$\text{(i)}\ \partial_u^2 N(e,0) < 0, \quad \text{and} \quad \text{(ii)}\ \partial_u^3 N(e,u) \leq 0 \ \text{ for } u \in I. \tag{A-4}$$

To establish (A-4) we need the following:

$$\partial_u^2 N(e,u) = 4[45(23e+14)^2 u^4 - 189(23e+14)u^2 - 115e - 21], \tag{A-5}$$

$$\partial_u^3 N(e,u) = 72(23e+14)u[10(23e+14)u^2 - 21], \tag{A-6}$$

$$\partial_u^5 N(e,u) = 4320u(23e+14)^2. \tag{A-7}$$

Equation (A-5) gives $\partial_u^2 N(e,0) = -4(115e+21)$ and the first claim in (A-4) follows. From (A-7) we see that $\partial_u^3 N(e,u)$ is convex for $u \in I$. Therefore, since $\partial_u^3 N(e,0) = 0$ it is enough to show that

$$\partial_u^3 N(e,\phi_0(1)) \leq 0 \tag{A-8}$$

for the second claim in (A-4) to hold. To establish this inequality, we first use definition (2-4) to compute

$$\phi_0(1) = \left(\frac{2}{\sqrt{(e+7)(46e^2+445e+567)} - 7(e+7)}\right)^{\frac{1}{2}}.$$

Now, according to (A-6), it is enough to prove that $10(23e+14)\phi_0(1)^2 - 21 < 0$ for (A-8) to hold. This inequality is equivalent to $441e^2 - 925e + 1316 > 0$, which clearly holds for all $e \geq 0$. This concludes the proof of the third claim in (A-3). Since the first claim in (A-3) is obviously true, it is left to prove that $N(e,\phi_0(1)) > 0$. To that end we first compute

$$N(e,\phi_0(1)) = \frac{2(P(e)\sqrt{Q(e)} - R(e))}{[\sqrt{Q(e)} - 7(e+7)]^3}, \tag{A-9}$$

where

$$P(e) = 7(69e^3 + 1831e^2 + 11500e + 17094),$$
$$Q(e) = (e+7)(46e^2 + 445e + 567),$$
$$R(e) = 20723e^4 + 433338e^3 + 3077307e^2 + 8566502e + 7537866.$$

The denominator in (A-9) is positive if and only of $Q(e)^2 - 49(e+7)^2 > 0$. This is equivalent to $2(e+8)(e+7)(23e+14) > 0$, which is manifestly true for $e \geq 0$. The numerator in (A-9) is positive if and only if $P(e)^2 Q(e) - R(e)^2 > 0$, which is equivalent to $2(23e+14)^2 S(e) > 0$, where

$$S(e) = 10143e^7 + 289189e^6 + 2979735e^5 + 12402439e^4 + 11046366e^3 - 30567884e^2 + 15651132e + 22614480.$$

The positivity of $S(e)$ is easily shown; for example we have

$$12402439e^4 + 22614480 > 30567884e^2.$$

The positivity of $N(e,\phi_0(1))$ follows.

Now we turn to proving that the second expression in (A-1) is negative for $d \geq 8$ and $u \in I$. Since $g''(u)/g(u)$ is positive for $u \in I$ and $g(u) > 0$ for small positive values of u, we conclude that both g'' and g are positive on $(0, \phi_0(1)]$. Consequently

$$g'(u) - 1 = g'(u) - g'(0) = \int_0^u g''(t)\,dt > 0 \quad \text{for } u \in (0, \phi_0(1)].$$

Hence $g'(u)^2 - 1 > 0$ and therefore

$$\frac{1 - g'(u)^2}{g(u)^2} < 0$$

for $u \in (0, \phi_0(1)]$. Additionally, by direct computation we see that

$$\frac{1 - g'(0)^2}{g(0)^2} = -21 < 0.$$

Finally, for each $d \geq 8$ we infer the existence of $\varepsilon > 0$ for which both expressions in (A-1) are negative provided $|u| < \phi_0(1) + \varepsilon$. For $|u| \geq \phi_0(1) + \varepsilon$, the function $g(u)$ can be easily modified so that it satisfies (1-2) and both expressions in (A-1) remain negative.

Appendix B: Estimate for δ_7

Proposition B.1. *For δ_7 defined in* (3-42) *and $\lambda \in \overline{\mathbb{H}}$ we have*

$$|\delta_7(\lambda)| \leq \tfrac{1}{3}. \tag{B-1}$$

Proof. Following the proof of Lemma 4.3 in [Costin et al. 2017] we show that r_7 and $(\tilde{r}_7)^{-1}$ are analytic in $\overline{\mathbb{H}}$. This implies that δ_7 is also analytic there. Furthermore, being a rational function, δ_7 is evidently polynomially bounded in $\overline{\mathbb{H}}$. Therefore, according to the Phragmén–Lindelöf principle,[2] it suffices to prove that (B-1) holds on the imaginary line, i.e.,

$$|\delta_7(is)|^2 \leq \tfrac{1}{9} \quad \text{for } s \in \mathbb{R}. \tag{B-2}$$

Note that the function $s \mapsto |\delta_7(is)|^2$ is even. It is therefore enough to prove (B-2) for nonnegative s only. We show that for $t \geq 0$,

$$\left|\delta_7\left(\frac{4t}{t+1}i\right)\right|^2 \leq \tfrac{1}{9} \quad \text{and} \quad |\delta_7((t+4)i)|^2 \leq \tfrac{1}{9}. \tag{B-3}$$

The first estimate above proves (B-2) for $s \in [0, 4)$, while the second one covers the complementary interval $[4, \infty)$. We prove both estimates in (B-3) in the same way and therefore illustrate the proof of the second one only. Note that

$$|\delta_7((t+4)i)|^2 = \frac{Q_1(t)}{Q_2(t)},$$

[2] We use the sectorial formulation of this principle; see, for example, [Titchmarsh 1939, p. 177].

where $Q_j(t) \in \mathbb{Z}[t]$, $\deg Q_j = 32$ and Q_2 has all positive coefficients. Therefore, $|\delta_7((t+4)i)|^2 \leq \frac{1}{9}$ is equivalent to $Q_2 - 9Q_1 \geq 0$ and a direct calculation shows that the polynomial $Q_2 - 9Q_1$ has manifestly positive coefficients. □

References

[Biernat et al. 2017] P. Biernat, P. Bizoń, and M. Maliborski, "Threshold for blowup for equivariant wave maps in higher dimensions", *Nonlinearity* **30**:4 (2017), 1513–1522. MR Zbl

[Bizoń 2000] P. Bizoń, "Equivariant self-similar wave maps from Minkowski spacetime into 3-sphere", *Comm. Math. Phys.* **215**:1 (2000), 45–56. MR Zbl

[Bizoń and Biernat 2015] P. Bizoń and P. Biernat, "Generic self-similar blowup for equivariant wave maps and Yang–Mills fields in higher dimensions", *Comm. Math. Phys.* **338**:3 (2015), 1443–1450. MR Zbl

[Bizoń et al. 2000] P. Bizoń, T. Chmaj, and Z. Tabor, "Dispersion and collapse of wave maps", *Nonlinearity* **13**:4 (2000), 1411–1423. MR Zbl

[Bizoń et al. 2001] P. Bizoń, T. Chmaj, and Z. Tabor, "Formation of singularities for equivariant (2+1)-dimensional wave maps into the 2-sphere", *Nonlinearity* **14**:5 (2001), 1041–1053. MR Zbl

[Cazenave et al. 1998] T. Cazenave, J. Shatah, and A. S. Tahvildar-Zadeh, "Harmonic maps of the hyperbolic space and development of singularities in wave maps and Yang–Mills fields", *Ann. Inst. H. Poincaré Phys. Théor.* **68**:3 (1998), 315–349. MR Zbl

[Chatzikaleas et al. 2017] A. Chatzikaleas, R. Donninger, and I. Glogić, "On blowup of co-rotational wave maps in odd space dimensions", *J. Differential Equations* **263**:8 (2017), 5090–5119. MR Zbl

[Costin et al. 2016] O. Costin, R. Donninger, I. Glogić, and M. Huang, "On the stability of self-similar solutions to nonlinear wave equations", *Comm. Math. Phys.* **343**:1 (2016), 299–310. MR Zbl

[Costin et al. 2017] O. Costin, R. Donninger, and I. Glogić, "Mode stability of self-similar wave maps in higher dimensions", *Comm. Math. Phys.* **351**:3 (2017), 959–972. MR Zbl

[Côte 2015] R. Côte, "On the soliton resolution for equivariant wave maps to the sphere", *Comm. Pure Appl. Math.* **68**:11 (2015), 1946–2004. MR Zbl

[Côte et al. 2015a] R. Côte, C. E. Kenig, A. Lawrie, and W. Schlag, "Characterization of large energy solutions of the equivariant wave map problem, I", *Amer. J. Math.* **137**:1 (2015), 139–207. MR Zbl

[Côte et al. 2015b] R. Côte, C. E. Kenig, A. Lawrie, and W. Schlag, "Characterization of large energy solutions of the equivariant wave map problem, II", *Amer. J. Math.* **137**:1 (2015), 209–250. MR Zbl

[Dodson and Lawrie 2015] B. Dodson and A. Lawrie, "Scattering for radial, semi-linear, super-critical wave equations with bounded critical norm", *Arch. Ration. Mech. Anal.* **218**:3 (2015), 1459–1529. MR Zbl

[Donninger 2011] R. Donninger, "On stable self-similar blowup for equivariant wave maps", *Comm. Pure Appl. Math.* **64**:8 (2011), 1095–1147. MR Zbl

[Donninger 2014] R. Donninger, "Stable self-similar blowup in energy supercritical Yang–Mills theory", *Math. Z.* **278**:3-4 (2014), 1005–1032. MR Zbl

[Donninger 2017] R. Donninger, "Strichartz estimates in similarity coordinates and stable blowup for the critical wave equation", *Duke Math. J.* **166**:9 (2017), 1627–1683. MR Zbl

[Donninger and Schörkhuber 2012] R. Donninger and B. Schörkhuber, "Stable self-similar blow up for energy subcritical wave equations", *Dyn. Partial Differ. Equ.* **9**:1 (2012), 63–87. MR Zbl

[Donninger and Schörkhuber 2014] R. Donninger and B. Schörkhuber, "Stable blow up dynamics for energy supercritical wave equations", *Trans. Amer. Math. Soc.* **366**:4 (2014), 2167–2189. MR Zbl

[Donninger and Schörkhuber 2016] R. Donninger and B. Schörkhuber, "On blowup in supercritical wave equations", *Comm. Math. Phys.* **346**:3 (2016), 907–943. MR Zbl

[Donninger and Schörkhuber 2017] R. Donninger and B. Schörkhuber, "Stable blowup for wave equations in odd space dimensions", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **34**:5 (2017), 1181–1213. MR Zbl

[Donninger et al. 2012] R. Donninger, B. Schörkhuber, and P. C. Aichelburg, "On stable self-similar blow up for equivariant wave maps: the linearized problem", *Ann. Henri Poincaré* **13**:1 (2012), 103–144. MR Zbl

[Engel and Nagel 2000] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics **194**, Springer, 2000. MR Zbl

[Gao and Krieger 2015] C. Gao and J. Krieger, "Optimal polynomial blow up range for critical wave maps", *Commun. Pure Appl. Anal.* **14**:5 (2015), 1705–1741. MR Zbl

[Ghoul et al. 2018] T. Ghoul, S. Ibrahim, and V. T. Nguyen, "Construction of type II blowup solutions for the 1-corotational energy supercritical wave maps", *J. Differential Equations* **265**:7 (2018), 2968–3047. MR Zbl

[Kato 1980] T. Kato, *Perturbation theory for linear operators*, Grundlehren der Mathematischen Wissenschaften **132**, Springer, 1980. Zbl

[Krieger and Schlag 2012] J. Krieger and W. Schlag, *Concentration compactness for critical wave maps*, European Mathematical Society, Zürich, 2012. MR Zbl

[Krieger et al. 2008] J. Krieger, W. Schlag, and D. Tataru, "Renormalization and blow up for charge one equivariant critical wave maps", *Invent. Math.* **171**:3 (2008), 543–615. MR Zbl

[Lawrie and Oh 2016] A. Lawrie and S.-J. Oh, "A refined threshold theorem for (1+2)-dimensional wave maps into surfaces", *Comm. Math. Phys.* **342**:3 (2016), 989–999. MR Zbl

[O'Neill 1983] B. O'Neill, *Semi-Riemannian geometry: with applications to relativity*, Pure and Applied Mathematics **103**, Academic, New York, 1983. MR Zbl

[Raphaël and Rodnianski 2012] P. Raphaël and I. Rodnianski, "Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang–Mills problems", *Publ. Math. Inst. Hautes Études Sci.* **115** (2012), 1–122. MR Zbl

[Rodnianski and Sterbenz 2010] I. Rodnianski and J. Sterbenz, "On the formation of singularities in the critical O(3) σ-model", *Ann. of Math.* (2) **172**:1 (2010), 187–242. MR Zbl

[Shatah 1988] J. Shatah, "Weak solutions and development of singularities of the SU(2) σ-model", *Comm. Pure Appl. Math.* **41**:4 (1988), 459–469. MR Zbl

[Shatah and Tahvildar-Zadeh 1994] J. Shatah and A. S. Tahvildar-Zadeh, "On the Cauchy problem for equivariant wave maps", *Comm. Pure Appl. Math.* **47**:5 (1994), 719–754. MR Zbl

[Sterbenz and Tataru 2010a] J. Sterbenz and D. Tataru, "Energy dispersed large data wave maps in $2+1$ dimensions", *Comm. Math. Phys.* **298**:1 (2010), 139–230. MR Zbl

[Sterbenz and Tataru 2010b] J. Sterbenz and D. Tataru, "Regularity of wave-maps in dimension $2+1$", *Comm. Math. Phys.* **298**:1 (2010), 231–264. MR Zbl

[Struwe 2003] M. Struwe, "Equivariant wave maps in two space dimensions", *Comm. Pure Appl. Math.* **56**:7 (2003), 815–823. Dedicated to the memory of Jürgen K. Moser. MR Zbl

[Tachikawa 1985] A. Tachikawa, "Rotationally symmetric harmonic maps from a ball into a warped product manifold", *Manuscripta Math.* **53**:3 (1985), 235–254. MR Zbl

[Titchmarsh 1939] E. C. Titchmarsh, *The theory of functions*, 2nd ed., Oxford University Press, 1939. MR Zbl

[Turok and Spergel 1990] N. Turok and D. Spergel, "Global texture and the microwave background", *Phys. Rev. Lett.* **64**:23 (1990), 2736–2739.

Received 22 May 2017. Revised 23 Nov 2017. Accepted 14 May 2018.

ROLAND DONNINGER: donninge@math.uni-bonn.de
roland.donninger@univie.ac.at
Rheinische Friedrich-Wilhelms-Universität Bonn, Mathematisches Institut, Bonn, Germany
and
Universität Wien, Fakultät für Mathematik, Vienna, Austria

IRFAN GLOGIĆ: glogic.1@osu.edu
irfan.glogic@univie.ac.at
Department of Mathematics, The Ohio State University, Columbus, OH, United States
and
Universität Wien, Fakultät für Mathematik, Vienna, Austria

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.417

FRACTURE WITH HEALING: A FIRST STEP TOWARDS A NEW VIEW OF CAVITATION

GILLES FRANCFORT, ALESSANDRO GIACOMINI AND OSCAR LOPEZ-PAMIES

Recent experimental evidence on rubber has revealed that the internal cracks that arise out of the process, often referred to as cavitation, can actually heal.

We demonstrate that crack healing can be incorporated into the variational framework for quasistatic brittle fracture evolution that has been developed in the last twenty years. This will be achieved for two-dimensional linearized elasticity in a topological setting, that is, when the putative cracks are closed sets with a preset maximum number of connected components.

Other important features of cavitation in rubber, such as near incompressibility and the evolution of the fracture toughness as a function of the cumulative history of fracture and healing, have yet to be addressed even in the proposed topological setting.

1. Introduction

A simplistic model for cavitation. Ever since the 1930s, ample experimental evidence points to the specificity of the initiation and propagation of fracture in rubber, or more generally in soft organic solids; see, e.g., [Busse 1938; Gent and Lindley 1959; Gent and Park 1984]. While metals, ceramics, and, more generally, crystalline and glassy solids show well-defined crack patterns when subject to extreme loading processes, fracture in rubber tends to initiate through the growth of microscopic defects arising in regions under sufficiently high hydrostatic stress. Because of its fluidic elder counterpart, the phenomenon has become known as cavitation.

It was initially thought that cavitation could be explained on pure elastic ground. In the mechanical universe, the most notorious proponents of elastic cavitation were undoubtedly A. N. Gent and P. B. Lindley [1959]. In their footsteps, J. M. Ball [1982] pioneered the first mathematical translation of that idea. There he posited that hyperelasticity can, in and of itself, create cavities through solutions of the type $x/|x|$ that are good Sobolev functions, provided that the growth at infinity of the elastic energy be subcritical, that is, less than the spatial dimension. In a more classical framework an equivalent viewpoint posits incipient point defects that balloon up to cavities. This insight generated a slew of mathematical studies that did show promise.

However, the spectacle of cavitation as a purely elastic phenomenon is in our opinion unrealistic. On pure theoretical grounds, it strikes us as somewhat peculiar that an innate sense of self would raise material awareness of its energetic elastic health under very large stretches, a prerequisite for any cogent statement

MSC2010: 74R10, 35Q74, 49J45, 28A75, 47J35.
Keywords: minimizing evolutions, fracture, free discontinuity problems.

of its growth. On more practical grounds, it was recently shown in [Lefèvre et al. 2015] that, in the classical poker-chip experiments of Gent and Lindley, as well as for a different experiment that uses a rubber reinforced by filler particles [Poulain et al. 2017],[1] a mere accounting of the elastic properties of the solids, while leading to a superficially adequate qualitative agreement with a number of experimental observations, fails to provide a complete qualitative and, most importantly quantitative, rendering of the evolution.

Our guiding principle is therefore that elasticity alone cannot account for the full complexity of the phenomenon of cavitation in rubber. From a macroscopic point of view, one should at the least introduce new internal surfaces within the solid to adequately describe the actual microscopic mechanisms behind fracture, be it the spatial rearrangement of the underlying macromolecules, or the breakage of chemical bonds. Such a viewpoint would seem to promote a fracture-type model in the vein of those adopted for brittle solids, albeit in the context of finite elasticity, see, e.g., [Dal Maso et al. 2005], and with the additional accounting of near or full incompressibility.[2]

Incompressibility notwithstanding, a refined fracture model was recently advocated in various mathematical works of D. Henao and C. Mora-Corral [Henao and Mora-Corral 2010; Mora-Corral 2014]. There, a surface energy proportional to the perimeter of the cavities in the deformed configuration is considered, in the spirit of surface tension. It is then added to the elastic energy and subsequently viewed, at least in [Mora-Corral 2014], as a conservative contribution. Adopting for a moment a common terminology in the mechanics community,[3] the only source of dissipation is born out of the irreversible creation of a countable number of point discontinuities that will grow into cavities.

The idea of endowing created surfaces with an energy is original and potentially fruitful. This refined viewpoint — or even a classical fracture viewpoint for that matter — may provide a good fit for some of the poker-chip experiments. But both will most likely become exercises in alt-reality when it comes to the *filler particle experiments*. Recent such experiments, carried out at high spatiotemporal resolution in [Poulain et al. 2017], showed that some of the created cavities actually vanish during the loading process, while others migrate away from the particles. Traditional or revamped theories of fracture do not sustain disappearance or migration and, while arguably predicting the final location of the cavities, completely fail in their depiction of the path that would lead to the final migrated state.

The full picture of the *filler particle experiment* is actually more intricate. The experiments in [Poulain et al. 2017] have also shown that the regions of the rubber that experience healing appear to acquire different fracture properties from those of the original rubber, thereby hinting at an evolution of the underlying molecular rearrangement and/or chemical bonding due to the healing process.

A full account of such observations is not our purpose at this point. It would certainly involve a healing process, together with a hardening or softening process in the fracture toughness, if such a notion makes

[1] We refer to that experiment as the *filler-particle experiment.*

[2] The addition of an incompressibility constraint is a huge mathematical hurdle from the standpoint of the variational theory of (brittle) fracture and the reader should be alerted to the absence of any mathematically significant result that encompasses both incompressibility and fracture.

[3] While a prevailing one, the postulate that fracture, or cavitation, should be described in terms of entropy production due to some kind of dissipation is just that, see, a contrario, [Il'iushin 1961], and our casting of the cavitation model in those terms is mere abidance by the majority view.

sense. Further, near or full incompressibility would certainly be a major partner, although its role has yet to be scripted.

Rather we propose in this contribution to focus solely on healing. The above quoted experiment notwithstanding, there is ample independent evidence that healing does take place in soft organic solids; see, e.g., [Madsen et al. 2016; Blaiszik et al. 2010; Cordier et al. 2008]. Now of course, as far as rubber is concerned, healing and near incompressibility should not be viewed as independent agents. We will woefully ignore their relationship in the following study. Mathematical impotence, rather than spite, motivates our choice.

So, as an admittedly childish first step, we propose to incorporate healing in A. A. Griffith's theory of fracture [1921], suitably re-engineered through a variational lens [Francfort and Marigo 1998; Bourdin et al. 2008], for two-dimensional linear elasticity. At first glance such a task would seem simple enough, at least from a modeling standpoint and provided that one is willing to view the healing process as rate-independent, which is most likely not so.[4] The naive recipe would be to dissipate some amount of surface energy for crack repair. In other words one would pay, say $c_1 \times$ length of $\Gamma \setminus K,\ c_1 > 0$, for changing the crack K to a different crack Γ and would also pay $c_2 \times$ length of $K \setminus \Gamma,\ c_2 > 0$, for repairing some of K with Γ.

Such petulance must be tempered with the recognition that doing so would result in a model for which healing would never take place because a healed part of the crack would increase the elastic energy while dissipating some surface energy through healing. Thus the healing process, if rate-independent and proportional to the length of the healed part must actually decrease the dissipated energy. A formal account will be given at the onset of Section 2.

For now, just think of a preset connected crack path Γ in a domain Ω and of a connected crack $\Gamma(\ell)$ of length ℓ *starting from a set point* — say the origin — along Γ (which should also contain the origin). Denote by $\mathcal{W}(\ell)$ the potential energy associated to the elastic equilibrium of $\Omega \setminus \Gamma(\ell)$ — the uncracked part of the domain — under the current loading at time t. Then we impose fealty of the dual fracture/healing process to that of Griffith's fracture [1921].

It is thus assumed that the energy dissipated through any putative advancement of the crack is proportional to the add-crack length with c_1 as fracture toughness; similarly that gained through healing is proportional to the subtract-crack length with c_2 as healing toughness. Of course $c_1 > c_2$ so that there indeed be a net dissipation.

To determine $\ell(t)$, a two-pronged formulation is espoused:

- First, a stability criterion à la Griffith is imposed: the energy release rate must satisfy
$$c_2 \le -\frac{\partial \mathcal{W}}{\partial \ell}(\ell(t)) \le c_1.$$
- Then the crack cannot extend unless the second inequality is an equality, while it cannot shrink unless the first one is an equality.

[4] Rearranging the molecular structure of the rubber and/or forming new chemical bonds are in all likelihood viscosity-driven processes that will shatter rate independence while potentially still variationally tractable; see the recent approach of viscoplasticity using energy-dissipation-balance solutions [Mielke et al. 2018]. As for the problem at hand, the precise nature of viscosity is very unclear as of yet.

Further, because irreversibility is de facto abandoned, there is no impediment to surface energy contributing to internal energy as well. In the above cartoon picture of the evolution, this amounts to adding a term like $c\ell,\ c \geq 0$, to the elastic energy $\mathcal{W}(\ell)$.

Sections 2–4 investigate the setting of antiplane shear linear elasticity, which is undoubtedly the simplest available framework for fracture evolution. The resulting model is presented in Section 2 in its variational reformulation. Section 3 is devoted to the proof of a stability result which is essential in the success of the limit process when passing from a time-incremental to a time-continuous formulation. Section 4 establishes the existence result for an evolution where both cracking and healing are allowed. In Section 5 we generalize the results of Section 4 to the setting of planar elasticity (plane strain or plane stress) in the footsteps of similar work on the fracture only case [Chambolle 2003].

From a mathematical standpoint, the first existence results for the variational theory of fracture were obtained in [Dal Maso and Toader 2002] in the antiplane shear case under the topological restriction that the cracks should have no more than m connected components, m being a preset connectivity threshold. This restriction was subsequently alleviated in [Francfort and Larsen 2003]. The present study unfortunately forces us to return to the topological setting of [Dal Maso and Toader 2002], mainly because we do not know how to prove energy conservation in the fully "variational" framework, that is, with no restriction on the topology of the cracks (see in particular Remark 1.5 below).

There is by now a vast literature on various aspects of the variational theory of fracture. We trust that the potential readership for this work is well versed in the main tenet of that theory and consequently refrain from any detailed explanation of the expounded formulation. We refer newcomers to [Bourdin et al. 2008] for an exposition of that theory and in particular to Section 2 of that work, where the link between the variational theory and the above two-pronged formulation is unraveled.

At the close of this introduction, we see it fit to put forth the following disclaimer: the model that is advocated below is not meant to be viewed as the final adjudication of cavitation. In view of recent experimental evidence, we merely assert that fracture and healing are essential partners in the cavitation process. We then proceed to incorporate healing into the variational theory of fracture in the mathematically simplest possible manner. Doing so at this time does not preclude subsequent refinements or modifications of the model. The paper [Kumar et al. 2018] presents a much more intricate phase field model that strives to account for both incompressibility and hardening on top of healing.

But it would be presumptuous on our part to pretend that we know how to address the mathematical hurdles that would accompany a rigorous analysis of more complex cavitation models such as that offered in [Kumar et al. 2018]. So, from a mathematical standpoint, the analysis below is the sum total of what lies within our reach for now.

Notation. Given $x \in \mathbb{R}^2,\ r > 0$ and $\nu \in \mathbb{R}^2$, we denote by $Q_\nu(x,r) \subset \mathbb{R}^2$ the square of center x with one side orthogonal to ν and length r. When ν is vertical, we will write simply $Q(x,r)$. $B(x,r)$ will denote the disk of center x and radius r.

Given two sets $A, B \subseteq \mathbb{R}^2$, we denote their symmetric difference by $A \Delta B$, while $A \Subset B$ will mean $\overline{A} \subseteq B$.

In all that follows $\mathrm{M}^2_{\mathrm{sym}}$ and $\mathrm{M}^2_{\mathrm{skew}}$ denote the families of symmetric and antisymmetric 2×2-matrices, respectively, while $\mathcal{L}_s(\mathrm{M}^2_{\mathrm{sym}})$ stands for the space of symmetric endomorphisms of $\mathrm{M}^2_{\mathrm{sym}}$.

For any mapping $u:\mathbb{R}^2\mapsto\mathbb{R}^2$, $e(u)$ denotes the symmetrized gradient of u, that is, $e(u):=\frac{1}{2}(\nabla u+\nabla u^T)$.

Also, for any open set A, we define $\mathscr{L}\mathscr{D}(A):=\{u\in L^2_{\mathrm{loc}}(A;\mathbb{R}^2):e(u)\in L^2(A;\mathrm{M}^2_{\mathrm{sym}})\}$.

Finally, we use standard notation for Sobolev spaces and for Hausdorff measures, specifically denoting by $\|\ \|$ the L^2-norm and by $\|\ \|_\infty$ the L^∞-norm. Also, for a Banach space X, we denote by $\mathrm{AC}([0,T];X)$ the space of X-valued absolutely continuous functions.

Mathematical preliminaries*: *Hausdorff convergence of compact sets. In the sequel, Hausdorff convergence will play an essential role. For the reader's convenience, we recall a few properties that will be used throughout.

The family $\mathcal{K}(\mathbb{R}^N)$ of closed sets in $\mathbb{R}^N$ can be endowed with the Hausdorff metric d_H defined by

$$d_H(K_1,K_2):=\max\{\sup\nolimits_{x\in K_1}\mathrm{dist}(x,K_2),\ \sup\nolimits_{y\in K_2}\mathrm{dist}(y,K_1)\},$$

with the conventions $\mathrm{dist}(x,\varnothing)=+\infty$ and $\sup_\varnothing=0$, so that $d_H(\varnothing,K)=0$ if $K=\varnothing$ and $d_H(\varnothing,K)=+\infty$ if $K\neq\varnothing$.

The Hausdorff metric has good compactness properties; see [Ambrosio and Tilli 2004, Theorem 4.4.15].

Proposition 1.1 (compactness). *Let $(K_n)_{n\in\mathbb{N}}$ be a sequence of compact sets contained in a fixed compact set of $\mathbb{R}^N$. Then there exists a compact set $K\subseteq\mathbb{R}^N$ such that up to a subsequence*

$$K_n\to K\quad\textit{in the Hausdorff metric}.$$

We will repeatedly make use of the following property due to Gołąb; for the proof we refer the reader to [Falconer 1986, Theorem 3.18; Ambrosio and Tilli 2004, Theorem 4.4.17].

Theorem 1.2 (Gołąb). *Let $(K_n)_{n\in\mathbb{N}}$ be a sequence of compact connected sets in $\mathbb{R}^N$ such that*

$$K_n\to K\quad\textit{in the Hausdorff metric.}$$

Then K is connected and for every open set $A\subseteq\mathbb{R}^N$

$$\mathcal{H}^1(K\cap A)\le\liminf_{n\to\infty}\mathcal{H}^1(K_n\cap A).$$

Remark 1.3. The lower semicontinuity of Gołąb's theorem still holds when K_n has a uniformly bounded number of connected components.

Lemma 1.4. *Let $(K_n)_{n\in\mathbb{N}}$ and $(H_n)_{n\in\mathbb{N}}$ be two sequences of compact sets in $\mathbb{R}^N$, each with a uniformly bounded number of connected components. Assume that*

$$K_n\to K\quad\textit{and}\quad H_n\to H\quad\textit{in the Hausdorff metric}.$$

Then, for any open set $A\subseteq\mathbb{R}^N$,

$$\mathcal{H}^1((K\setminus H)\cap A)\le\liminf_n\mathcal{H}^1((K_n\setminus H_n)\cap A).\tag{1-1}$$

Proof. Let $V \subseteq \mathbb{R}^N$ be an open neighborhood of H. For n large enough we have $H_n \subseteq V$, so that by Gołąb's theorem

$$\begin{aligned}\mathcal{H}^1((K \setminus \bar{V}) \cap A) &\leq \liminf_n \mathcal{H}^1((K_n \setminus \bar{V}) \cap A) \\ &\leq \liminf_n \mathcal{H}^1((K_n \setminus H_n) \cap A).\end{aligned}$$

Since V is arbitrary, the conclusion follows. □

Remark 1.5. The topological setting for the cracks adopted in the paper, i.e., cracks which are closed and with a preset number of connected components, is motivated precisely by Lemma 1.4. A larger class of admissible cracks, as that adopted in [Dal Maso et al. 2005] where cracks are just rectifiable, requires suitable convergences of variational type, under which inequality (1-1) is known to fail. But that inequality is in particular an essential ingredient in the proof of the energy inequality (4-19) below to the extent that it establishes that (4-16) holds true.

A simple example for which inequality (1-1) is violated under the variational convergences of [Dal Maso et al. 2005] is the following: Let K be a segment of unit length, and let H_n be the dotted segment of length $\frac{1}{2}$ obtained from K by dividing it into 2^n equal parts and retaining only every other subsegment. It is easily proved that $H_n \to \varnothing$ in the variational sense, see [Dal Maso et al. 2005, Section 4.1], so that choosing $K_n = K$,

$$\mathcal{H}^1(K \setminus H) = \mathcal{H}^1(K) = 1, \quad \text{while} \quad \mathcal{H}^1(K_n \setminus H_n) = \tfrac{1}{2}.$$

2. Setting of the problem

The reference configuration is an open bounded set $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary.

Admissible cracks. Let $m \in \mathbb{N}$ with $m \geq 1$ be given. The class of admissible cracks is given by

$$\mathcal{K}^f_m(\bar{\Omega}) := \{K \subset \bar{\Omega} : K \text{ is compact, with at most } m \text{ connected components and } \mathcal{H}^1(K) < +\infty\}. \tag{2-1}$$

Admissible configurations. Let $\partial_D \Omega \subseteq \partial \Omega$ be open in the relative topology. The class of admissible boundary displacements g is given by the space $H^1(\Omega) \cap L^\infty(\Omega)$. We say that the pair (u, K) is an admissible configuration of our system for g if

$$K \in \mathcal{K}^f_m(\bar{\Omega})$$

and

$$u \in H^1(\Omega \setminus K) \quad \text{with } u = g \text{ on } \partial_D \Omega \setminus K.$$

We will write $(u, K) \in \mathcal{A}(g)$. Note that the pair $(\nabla u, u)$ can be thought of as an element of $L^2(\Omega; \mathbb{R}^3)$ since K has null Lebesgue measure.

The following compactness result will be used several times.

Lemma 2.1. *Let $g_n, g \in H^1(\Omega)$ be such that*

$$g_n \to g \quad \textit{strongly in } H^1(\Omega).$$

Assume that $(u_n, K_n) \in \mathcal{A}(g_n)$ *with*

$$\begin{aligned} (\nabla u_n, u_n) \rightharpoonup (\Phi, u) &\quad \textit{weakly in } L^2(\Omega; \mathbb{R}^3), \\ K_n \to K &\quad \textit{in the Hausdorff metric}. \end{aligned}$$

Then $(u, K) \in \mathcal{A}(g)$, *and* $\Phi = \nabla u$ *on* $\Omega \setminus K$.

Proof. Let $\varphi \in C_c^\infty(\Omega \setminus K)$. Then, for n large,

$$\varphi \in C_c^\infty(\Omega \setminus K_n).$$

We can thus write, for $i = 1, 2$,

$$\int_{\Omega \setminus K} \Phi_i \varphi \, dx = \lim_n \int_{\Omega \setminus K_n} \partial_i u_n \varphi \, dx = -\lim_n \int_{\Omega \setminus K_n} u_n \, \partial_i \varphi \, dx = -\int_{\Omega \setminus K} u \, \partial_i \varphi \, dx.$$

We deduce that $u \in H^1(\Omega \setminus K)$ with $\nabla u = \Phi$. Let us check that $(u, K) \in \mathcal{A}(g)$. Lest the result be trivial, it is not restrictive to assume that

$$\partial_D \Omega \setminus K \neq \varnothing.$$

Since $\partial_D \Omega$ is open in the relative topology, for every $x_0 \in \partial_D \Omega \setminus K$ we can find an open neighborhood $U \subset \mathbb{R}^2$ of x_0 such that $\operatorname{dist}(U, K) > 0$ and $U \cap \Omega$ has a Lipschitz boundary in U given by $\partial_D \Omega \cap U$. Since $K_n \cap U = \varnothing$ for n large, we infer that $u_n \in H^1(\Omega \cap U)$ with

$$u_n \rightharpoonup u \quad \text{weakly in } H^1(\Omega \cap U),$$

so $u = g$ on $\partial_D \Omega \cap U$. □

Remark 2.2. The choice of $H^1(\Omega) \cap L^\infty(\Omega)$ as the class of admissible displacements allows one to work in $H^1(\Omega \setminus K)$ when dealing with the variational constructions of Section 4. Without an L^∞-bound, the arguments can be adapted provided that we choose the displacements in $L^{1,2}(\Omega \setminus K)$, a Deny–Lions-type space [1954]. Such will not be the case in Section 5 below (see Remark 5.5).

Energies. We associate to an admissible configuration (u, K) the *elastic energy*

$$\|\nabla u\|^2 = \int_\Omega |\nabla u|^2 \, dx.$$

Here ∇u is viewed as an element of $L^2(\Omega; \mathbb{R}^2)$.

Assume that the system goes from the configuration (u, K) to the configuration (v, Γ). Then

$$\begin{cases} \Gamma \setminus K \text{ is the add-crack,} \\ K \setminus \Gamma \text{ is the healed zone.} \end{cases}$$

We assume the energy dissipated through such a process is

$$c_1 \mathcal{H}^1_D(\Gamma \setminus K) - c_2 \mathcal{H}^1_D(K \setminus \Gamma),$$

with $c_1, c_2 > 0$.

In the expression above and throughout the rest of the paper $\mathcal{H}^1_D$ stands for $\mathcal{H}^1 \lfloor_{\Omega \cup \partial_D \Omega}$. This is so because no energy should be dissipated for the part of the crack that lies on the free boundary $\partial\Omega \setminus \partial_D\Omega$.

Summing up, the passage from (u, K) to (v, Γ) involves a change in energy of the form

$$\{\|\nabla v\|^2 - \|\nabla u\|^2\} + c_1 \mathcal{H}^1_D(\Gamma \setminus K) - c_2 \mathcal{H}^1_D(K \setminus \Gamma)\}.$$

Notice that the expression can be rewritten in the form

$$\mathcal{E}(v, \Gamma) - \mathcal{E}(u, K) + (c_1 - c_2)\, \mathcal{H}^1_D(\Gamma \setminus K),$$

where

$$\mathcal{E}(v, \Gamma) := \|\nabla v\|^2 + c_2 \mathcal{H}^1_D(\Gamma). \tag{2-2}$$

Indeed,

$$\mathcal{H}^1_D(K \setminus \Gamma) = \mathcal{H}^1_D(K) - \mathcal{H}^1_D(K \cap \Gamma) = \mathcal{H}^1_D(K) - (\mathcal{H}^1_D(\Gamma) - \mathcal{H}^1_D(\Gamma \setminus K))$$

so that

$$c_1 \mathcal{H}^1_D(\Gamma \setminus K) - c_2 \mathcal{H}^1_D(K \setminus \Gamma) = c_2(\mathcal{H}^1_D(\Gamma) - \mathcal{H}^1_D(K)) + (c_1 - c_2)\, \mathcal{H}^1_D(\Gamma \setminus K).$$

In view of this new expression, we will assume that

$$c_1 > c_2 > 0. \tag{2-3}$$

See Remark 2.4 below for the case $c_1 = c_2$.

Quasistatic evolutions. Let $T > 0$ and

$$g \in \mathrm{AC}([0, T]; H^1(\Omega)), \quad \|g(t)\|_\infty \le C,\ t \in [0, T],$$

be a given time-dependent boundary displacement.

Given $t \mapsto K(t) \in \mathcal{K}^f_m(\overline{\Omega})$ we set, for $t \le T$,

$$\mathrm{Diss}(t) := (c_1 - c_2) \sup\bigl\{\textstyle\sum_{i=0}^n \mathcal{H}^1_D(K(s_{i+1}) \setminus K(s_i)) : 0 = s_0 < s_1 < \cdots < s_{n+1} = t\bigr\}.$$

Definition 2.3 (quasistatic evolution). We say that $\{t \mapsto (u(t), K(t)) \in \mathcal{A}(g(t)) : t \in [0, T]\}$ is a quasistatic evolution provided that for every $t \in [0, T]$ the following items hold true:

(a) (global stability) For every $(v, \Gamma) \in \mathcal{A}(g(t))$

$$\mathcal{E}(u(t), K(t)) \le \mathcal{E}(v, \Gamma) + (c_1 - c_2)\, \mathcal{H}^1_D(\Gamma \setminus K(t)), \tag{2-4}$$

where $\mathcal{E}$ is defined in (2-2).

(b) (energy balance) We have

$$\mathcal{E}(u(t), K(t)) + \mathrm{Diss}(t) = \mathcal{E}(u(0), K(0)) + 2\int_0^t \int_\Omega \nabla u(\tau) \cdot \nabla \dot{g}(\tau)\, dx\, d\tau.$$

Remark 2.4. In the spirit of our introductory remarks, we could modify the definition of $\mathcal{E}$ in (2-2) through addition of a term of the form $c\mathcal{H}^1_D(\Gamma)$ with $c \geq 0$, that is, a stored surface energy term. The analysis performed in the rest of the paper and Theorems 4.1, 5.4 would remain unchanged in this enlarged setting.

If, in lieu of (2-3), $c_1 = c_2$, the quasistatic evolution is conservative and consists in a time-parametrized set of independent minimization problems: the term $(c_1 - c_2)\,\mathcal{H}^1_D(\Gamma \setminus K(t))$ disappears in the global stability statement, while $\mathrm{Diss}(t)$ disappears in the energy balance statement of Definition 2.3. The existence proofs leading to Theorems 4.1, 5.4 become straightforward.

3. Stability of the global minimality property

A crucial step in the proof of the existence of a quasistatic evolution concerns the stability of the global minimality property (2-4) under Hausdorff convergence for the cracks. The proof is based on a topological version of the jump transfer construction in [Francfort and Larsen 2003]. Similar ideas have been put forth in [Acanfora and Ponsiglione 2006] in the case of the fracture problem for a flexural linear plate.

Theorem 3.1 (stability of the global minimality property). *Let c, c' be fixed positive constants. Let $g_n, g \in H^1(\Omega)$ be such that*

$$g_n \to g \quad \text{strongly in } H^1(\Omega).$$

Assume that $(u_n, K_n) \in \mathcal{A}(g_n)$ satisfy the following global stability condition: for every $(v, \Gamma) \in \mathcal{A}(g_n)$,

$$\|\nabla u_n\|^2 + c\mathcal{H}^1_D(K_n) \leq \|\nabla v\|^2 + c\mathcal{H}^1_D(\Gamma) + c'\mathcal{H}^1_D(\Gamma \setminus K_n)$$

and assume further that

$$K_n \to K \qquad \text{in the Hausdorff metric},$$
$$\nabla u_n \rightharpoonup \nabla u \quad \text{weakly in } L^2(\Omega; \mathbb{R}^2)$$

for some $(u, K) \in \mathcal{A}(g)$. Then (u, K) is a globally stable configuration, that is, for every $(v, \Gamma) \in \mathcal{A}(g)$,

$$\|\nabla u\|^2 + c\mathcal{H}^1_D(K) \leq \|\nabla v\|^2 + c\mathcal{H}^1_D(\Gamma) + c'\mathcal{H}^1_D(\Gamma \setminus K).$$

In order to prove Theorem 3.1, we need two geometric results concerning the blow-up behavior of sets in the family $\mathcal{K}^f_1(\mathbb{R}^2)$ of compact connected sets in $\mathbb{R}^2$ with finite length.

Theorem 3.2. *Let $K \in \mathcal{K}^f_1(\mathbb{R}^2)$. The following items hold true:*

(a) *K is countably $\mathcal{H}^1$-rectifiable with*

$$K = K_0 \cup \bigcup_{n=1}^{\infty} \gamma_n(I_n),$$

where $I_n \subseteq \mathbb{R}$ is an open interval, $\gamma_n : I_n \to \mathbb{R}^2$ are Lipschitz curves and $\mathcal{H}^1(K_0) = 0$. Further, there exists $N \subseteq K$ with $\mathcal{H}^1(N) = 0$ such that, for every $x \notin N$, K admits an approximate tangent line l_x at x with normal ν_x.

(b) *Take $x \in K \setminus N$. Then for $r \to 0^+$*

$$K_{x,r} := \frac{K-x}{r} \to l_x \quad \textit{locally in the Hausdorff metric.} \tag{3-1}$$

(c) *There exists $N_1 \subseteq K$ with $N \subseteq N_1$ and $\mathcal{H}^1(N_1) = 0$ such that the following property holds. Take $x \in K \setminus N_1$. Then, for every $\varepsilon > 0$, there exists $r_0 > 0$ such that for every $r < r_0$ the rectangles*

$$R^+_{\varepsilon,r} := Q_{\nu_x}(x,r) \cap \{y \in \mathbb{R}^2 : (y-x)\cdot \nu_x > \varepsilon r\},$$
$$R^-_{\varepsilon,r} := Q_{\nu_x}(x,r) \cap \{y \in \mathbb{R}^2 : (y-x)\cdot \nu_x < -\varepsilon r\}$$

belong to different connected components of $Q_{\nu_x}(x,r) \setminus K$.

Proof. The rectifiability property of point (a) is proved in [Falconer 1986, Lemma 3.13]. From the general theory of rectifiable sets, we know that K admits an approximate tangent line l_x at $\mathcal{H}^1$-a.e. $x \in K$; see [Ambrosio et al. 2000, Theorem 2.83].

Now for point (b). Up to an isometry, we may assume $x = 0$ and that the approximate tangent line l is horizontal. Then, by the very definition of an approximate tangent line,

$$\mathcal{H}^1 \lfloor K_r \overset{*}{\rightharpoonup} \mathcal{H}^1 \lfloor l \quad \text{locally weakly}^* \text{ in } \mathcal{M}_b(\mathbb{R}^2) \tag{3-2}$$

as $r \to 0^+$, where $K_r := \frac{1}{r} K$.

We claim that, for every $R > 0$,

$$K_r \cap \overline{Q}(0,R) \to l \cap \overline{Q}(0,R) \quad \text{in the Hausdorff metric.} \tag{3-3}$$

Indeed, given any sequence $r_n \to 0$, the compactness of Hausdorff convergence and a diagonal argument imply the existence of a subsequence $(r_{n_h})_{h\in\mathbb{N}}$ such that for every $m \in \mathbb{N}$, $m \geq 1$,

$$K_{r_{n_h}} \cap \overline{Q}(0,m) \to K_0^m \quad \text{in the Hausdorff metric.}$$

It is readily checked that, for every $m \geq 1$,

$$K_0^m \subseteq K_0^{m+1} \quad \text{and} \quad K_0^m \cap \overline{Q}(0,m) = K_0^{m+1} \cap \overline{Q}(0,m). \tag{3-4}$$

Set $K_0 := \bigcup_{m=1}^{\infty} K_0^m$. We claim that

$$K_0 = l. \tag{3-5}$$

First, $K_0 \subseteq l$. Indeed, assume by contradiction that $\xi \in K_0 \setminus l$ with $\overline{B}_\eta(\xi) \cap l = \varnothing$ for some $\eta > 0$. Using the measure convergence (3-2), we obtain that

$$\mathcal{H}^1(K_{r_{n_h}} \cap \overline{B}_\eta(\xi)) \to 0. \tag{3-6}$$

But $K_{r_{n_h}}$ is connected by arcs, see [Falconer 1986, Lemma 3.12], so that, taking $\xi_{n_h} \in K_{r_{n_h}}$ such that $\xi_{n_h} \to \xi$, we have ξ_{n_h} is connected to 0 through an arc contained in $K_{r_{n_h}}$ so, for h large enough,

$$\mathcal{H}^1(K_{r_{n_h}} \cap \overline{B}_{\eta/2}(\xi_{n_h})) \geq \tfrac{1}{4}\eta.$$

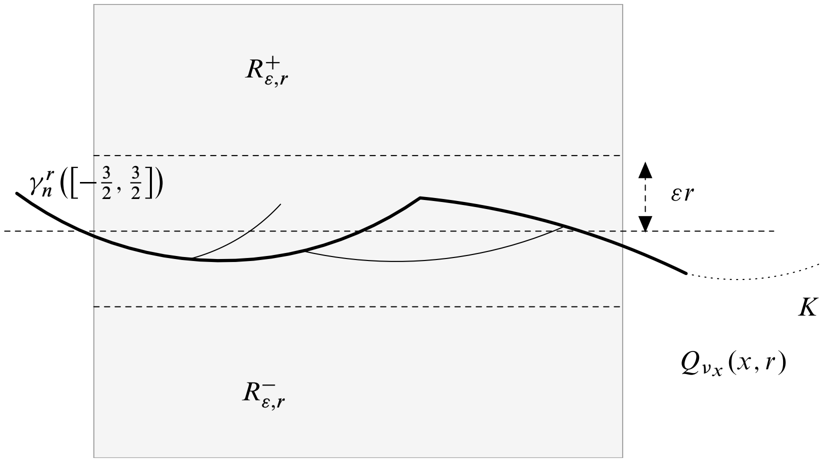

Figure 1. Illustration of item (c) in Theorem 3.2; the thick curve is $\gamma_n^r\big(\big[-\frac{3}{2},\frac{3}{2}\big]\big)$.

Thus

$$\liminf_{h\to\infty}\mathcal{H}^1(K_{r_{n_h}}\cap\overline{B}_\eta(\xi))\geq\liminf_{h\to\infty}\mathcal{H}^1(K_{r_{n_h}}\cap\overline{B}_{\eta/2}(\xi_{n_h}))\geq\tfrac{1}{4}\eta,$$

in contradiction with (3-6).

Conversely, $l\subseteq K_0$. Indeed, assume by contradiction that $\xi\in l\setminus K_0$. Then there exists $\eta>0$ such that $K_{r_{n_h}}\cap B_\eta(\xi)=\varnothing$ for h large, against (3-2).

In view of (3-4) and (3-5) we deduce that for $\varepsilon\to 0$ and for every $R>0$

$$K_r\cap\overline{Q}(0,R)\to l\cap\overline{Q}(0,R)\quad\text{in the Hausdorff metric,}$$

that is, (3-3). This means that the local convergence of (3-1) holds true, and point (b) is proved.

Let us come to point (c). See Figure 1.

Notice that we can reparametrize each Lipschitz curve γ_n by arc length. As a consequence, we may assume that for a.e. $t\in I_n$

$$\gamma_n\text{ is differentiable at }t\text{ with }|\gamma_n'(t)|=1. \tag{3-7}$$

From point (a), we deduce that there exists $N_1\subseteq K$ with $\mathcal{H}^1(N_1)=0$, $N\subseteq N_1$, and such that if $x\in K\setminus N_1$, then $x=\gamma_n(t_0)$ for some n, with t_0 satisfying (3-7). It is not restrictive to assume that $x=0$ with a horizontal tangent line l, and that $t_0=0$. By differentiability, for $r\to 0^+$,

$$\gamma_n^r(s):=\frac{1}{r}\gamma_n(rs)\to\gamma_n'(0)s\quad\text{locally uniformly in }s\in\mathbb{R}. \tag{3-8}$$

In view of (3-1), $\gamma_n'(0)$ is horizontal, and we can assume that $\gamma_n'(0)=(1,0)$.

Let $\varepsilon>0$. Because of (3-8) and since

$$\gamma_n^r\big(-\tfrac{3}{2}\big)\to\big(-\tfrac{3}{2},0\big)\quad\text{and}\quad\gamma_n^r\big(\tfrac{3}{2}\big)\to\big(\tfrac{3}{2},0\big),$$

we infer that, for r small enough, the (connected) arc $\gamma_n^r\big(\big[-\frac{3}{2},\frac{3}{2}\big]\big)$ satisfies

$$\gamma_n^r\big(\big[-\tfrac{3}{2},\tfrac{3}{2}\big]\big) \subseteq \{(x_1,x_2)\in\mathbb{R}^2 : |x_2|<\varepsilon\}$$

and that $Q(0,1)\setminus\gamma_n^r\big(\big[-\frac{3}{2},\frac{3}{2}\big]\big)$ is disconnected. We deduce that the open rectangles

$$R_\varepsilon^+ := Q(0,1)\cap\{(x_1,x_2)\in\mathbb{R}^2 : x_2>\varepsilon\},$$
$$R_\varepsilon^- := Q(0,1)\cap\{(x_1,x_2)\in\mathbb{R}^2 : x_2<-\varepsilon\}$$

belong to different connected components of $Q(0,1)\setminus\big(\frac{1}{r}K\big)$. The conclusion of point (c) now follows by rescaling. □

The following result shows that the topological property of point (c) of Theorem 3.2 is essentially stable under Hausdorff convergence. We will need this property for our topological version of the jump transfer.

Proposition 3.3. *Let $(K_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{K}_1^f(\mathbb{R}^2)$ and $K\in\mathcal{K}_1^f(\mathbb{R}^2)$ be such that*

$$K_n \to K \quad \text{in the Hausdorff metric.}$$

Let $N_1\subseteq K$ with $\mathcal{H}^1(N_1)=0$ be as in Theorem 3.2. For every $x\notin N_1$ and $\varepsilon>0$ we can find $r_0>0$ and $\nu_x\in\mathbb{R}^2$ with $|\nu_x|=1$ such that for every $r<r_0$ there exists $n_0\in\mathbb{N}$ and $(\widehat{K}_n)_{n\in\mathbb{N}}$ a sequence in $\mathcal{K}_1^f(\mathbb{R}^2)$ with

$$K_n\subseteq\widehat{K}_n,\quad \widehat{K}_n\setminus K_n\subseteq Q_{\nu_x}(x,r),\quad \mathcal{H}^1(\widehat{K}_n\setminus K_n)\le 3\varepsilon r$$

such that for $n\ge n_0$ the rectangles

$$R_{\varepsilon,r}^+ := Q_{\nu_x}(x,r)\cap\{y\in\mathbb{R}^2 : (y-x)\cdot\nu_x>\varepsilon r\}, \tag{3-9}$$
$$R_{\varepsilon,r}^- := Q_{\nu_x}(x,r)\cap\{y\in\mathbb{R}^2 : (y-x)\cdot\nu_x<-\varepsilon r\} \tag{3-10}$$

belong to different connected components of $Q_{\nu_x}(x,r)\setminus\widehat{K}_n$.

Proof. In view of Theorem 3.2, for every $x\notin N_1$ points (b) and (c) hold true.

Let us fix $x\notin N_1$ and $\varepsilon>0$, and let $r_0>0$ and $\nu_x\in\mathbb{R}^2$ be associated to x according to point (c) of Theorem 3.2. Up to a rototranslation, we may assume

$$x=0,\quad \nu_x=(0,1),\quad l_x=\{x=(x_1,x_2):x_2=0\}.$$

Notice that, in view of item (b) in Theorem 3.2, we may also assume that

$$K\setminus\overline{Q}(0,r_0)\neq\varnothing.$$

Since $K_n\to K$ in the Hausdorff metric, from the corresponding property of K we deduce that there exists $n_0>0$ such that for every $n\ge n_0$

$$K_n\cap\overline{Q}(0,r)\subset\{(x_1,x_2)\in\mathbb{R}^2 : |x_2|<\varepsilon r\} \tag{3-11}$$

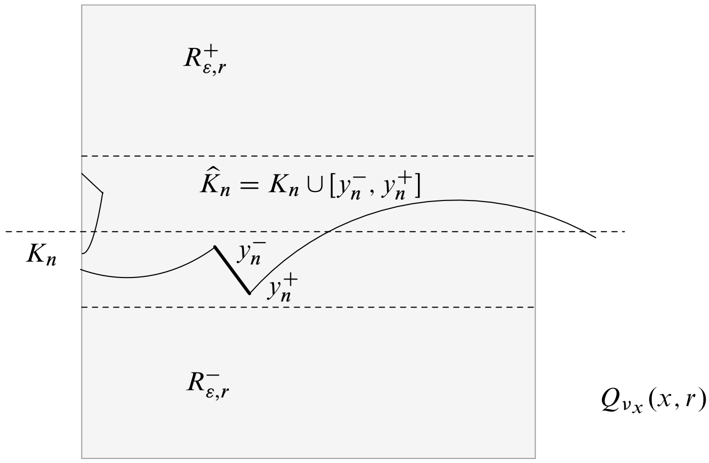

Figure 2. Construction of $\widehat{K}_n$ in Proposition 3.3.

and

$$K_n \setminus \overline{Q}(0, r_0) \neq \varnothing. \tag{3-12}$$

Let $z_n \in K_n \setminus \overline{Q}(0, r_0)$.

Since K_n is connected by arcs, given $x \in K_n \cap Q(0, r)$, we can find an arc contained in K_n with extremes x and z_n. In view of (3-11), (3-12), this arc has to intersect either S_r^- or S_r^+, where $S_r^\pm$ are the vertical segments

$$S_r^\pm := \{\pm r\} \times [-\varepsilon r, \varepsilon r].$$

Modulo reparametrization, we thus infer that there exists (at least) one arc $\gamma^+_{x,r} : [0, 1] \to \mathbb{R}^2$ or one $\gamma^-_{x,r} : [0, 1] \to \mathbb{R}^2$ with image contained in $K_n \cap \overline{Q}(0, r)$ such that

$$\gamma^\pm_{x,r}(0) = x \quad \text{and} \quad \gamma^\pm_{x,r}(1) \in S_r^\pm.$$

Let us consider the intervals contained in $[-r, r]$ given by

$$J^-_{n,r} := \bigcup_{x \in K_n \cap \overline{Q}(0,r)} \pi_1(\gamma^-_{x,r}([0, 1])) \quad \text{and} \quad J^+_{n,r} := \bigcup_{x \in K_n \cap \overline{Q}(0,r)} \pi_1(\gamma^+_{x,r}([0, 1])),$$

obtained by projecting the curves constructed above onto the horizontal axis.

We claim that we can find $\alpha_n^\pm \in J^\pm_{n,r}$ such that

$$|\alpha_n^+ - \alpha_n^-| \to 0. \tag{3-13}$$

If this is the case, since by definition there exists

$$y_n^\pm = (\alpha_n^\pm, \beta_n^\pm) \in K_n \cap \overline{Q(0, r)},$$

we then define $\widehat{K}_n$ to be (see Figure 2)

$$\widehat{K}_n = K_n \cup [y_n^-, y_n^+],$$

where $[y_n^-, y_n^+]$ is the segment joining y_n^- and y_n^+. In view of (3-11), we have

$$\limsup_{n\to\infty} \mathcal{H}^1([y_n^-, y_n^+]) \le 2r\varepsilon.$$

Finally, since

$$\gamma^-_{y_n^-,r}([0,1]) \cup [y_n^-, y_n^+] \cup \gamma^+_{y_n^+,r}([0,1]) \subset \widehat{K}_n,$$

we deduce that $\widehat{K}_n \in \mathcal{K}_1^f(\mathbb{R}^2)$ satisfy the conclusion of the theorem.

Let us prove claim (3-13). If the relation is not satisfied, we get for n large

$$\inf J^+_{n,r} - \sup J^-_{n,r} \ge \eta > 0.$$

Since $K_n \to K$ in the Hausdorff metric, we would infer that the projection of $K \cap \overline{Q}(0,r)$ onto the horizontal axis is composed of two distinct intervals contained in $[-r, r]$, against the fact that K disconnects $\overline{Q}(0,r)$. $\square$

Remark 3.4. Let $\Omega \subseteq \mathbb{R}^2$ be open, bounded and with a Lipschitz boundary. Assume that the sets K_n, K of Proposition 3.3 are such that $K_n, K \subseteq \overline{\Omega}$. Notice that, for $\mathcal{H}^1$-a.e. $x \in K \cap \partial\Omega$, the tangent lines to K and $\partial\Omega$ at the point x coincide, so that the topological blow-up properties of Theorem 3.2 at the point x hold simultaneously for K and $\partial\Omega$. Consequently, the proof of Proposition 3.3 shows that $\widehat{K}_n$ can be chosen such that in addition $\widehat{K}_n \subseteq \overline{\Omega}$.

We can now prove Theorem 3.1.

Proof of Theorem 3.1. The global stability we need to prove can be rewritten in the form

$$\|\nabla u\|^2 + c\,\mathcal{H}^1_D(K \setminus \Gamma) \le \|\nabla v\|^2 + (c + c')\,\mathcal{H}^1_D(\Gamma \setminus K)$$

for every $(v, \Gamma) \in \mathcal{A}(g)$ (see the computations in Section 2).

We divide the proof into two steps.

Step 1: Let us assume that $K \in \mathcal{K}_1^f(\overline{\Omega})$. Thanks to [Dal Maso and Toader 2002, Lemma 3.6], there exists $H_n \in \mathcal{K}_1^f(\overline{\Omega})$ with $K_n \subseteq H_n$,

$$\mathcal{H}^1(H_n \setminus K_n) \to 0 \quad \text{and} \quad H_n \to K \quad \text{in the Hausdorff metric.} \tag{3-14}$$

We need to introduce the connected sets H_n because it might be the case that, although K is connected, the K_n are not since they are only restricted to be elements of $\mathcal{K}_m^f(\overline{\Omega})$.

Let $V \subseteq \mathbb{R}^2$ be open with $\Gamma \subseteq V$. Let then $U \subset V$ be open with $\overline{U} \subset V$ and $\Gamma \cap K \subset U$. Let also $\varepsilon > 0$ be fixed. See Figure 3.

Note that, for $\mathcal{H}^1$-a.e. $x \in \Gamma \cap K$, the tangent lines to Γ and K at the point x coincide. We can thus find $N \subseteq \Gamma \cap K$ with $\mathcal{H}^1(N) = 0$ and such that for $x \in (\Gamma \cap K) \setminus N$ the conclusions of point (c) in Theorem 3.2 hold true with respect to both K and Γ simultaneously.

For $x \in (\Gamma \cap K) \setminus N$, let $r_0(x) > 0$ and $\nu_x \in \mathbb{R}^2$ be given by Proposition 3.3. We may assume in addition that

$$Q_{\nu_x}(x, r_0(x)) \subset U$$

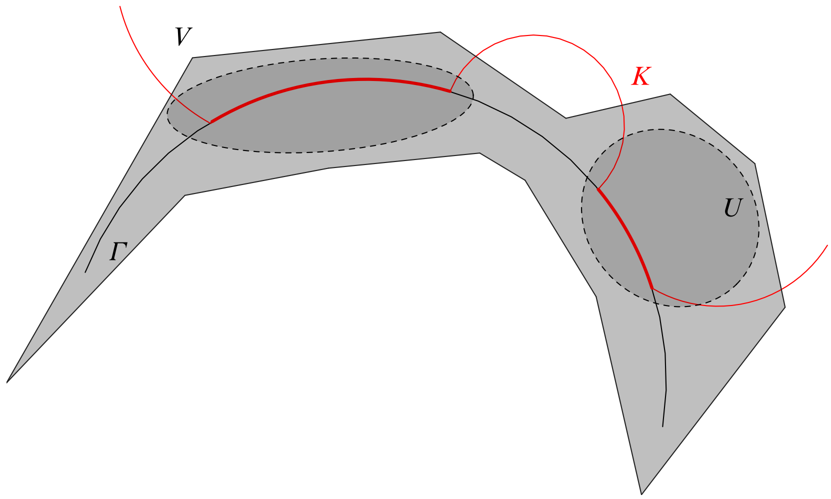

Figure 3. Setting the geometry for the proof of Theorem 3.1.

and also, thanks to, e.g., [Ambrosio et al. 2000, Theorem 2.83(i)], that, for every $r < r_0(x)$,

$$(1-\varepsilon)r \le \mathcal{H}^1(Q_{\nu_x}(x,r) \cap (K \cap \Gamma)) \le (1+\varepsilon)r. \tag{3-15}$$

By the Vitali–Besicovitch lemma, see, e.g., Theorem 2.19 of the above-quoted work, we can find a finite number of disjoint such squares $\{Q_{\nu_j(x_j,r_j)}\}_{j=1,\dots,m}$ with $x_j \in K \cap \Gamma$, $\nu_j := \nu_{x_j}$, $r_j < r_0(x_j)$, such that

$$\mathcal{H}^1\Bigl((K \cap \Gamma) \setminus \bigcup_{j=1}^{m} Q_{\nu_j}(x_j,r_j)\Bigr) < \varepsilon. \tag{3-16}$$

It is no restriction to assume that either $Q_{\nu_j}(x_j,r_j) \Subset \Omega$ or $x_j \in \partial\Omega$, with $\partial\Omega \cap Q_{\nu_j}(x_j,r_j)$ given by the graph of a Lipschitz function with respect to a reference frame with ν_j as vertical direction.

We modify H_n in each square according to Proposition 3.3 and Remark 3.4 and find $\widehat{H}_n \in \mathcal{K}_1^f(\overline{\Omega})$ with $H_n \subseteq \widehat{H}_n$, such that for n large

$$\widehat{H}_n = H_n \quad \text{outside} \bigcup_{j=1}^{m} Q_{\nu_j}(x_j,r_j),$$

and

$$\mathcal{H}^1(\widehat{H}_n \setminus H_n) \le 3\varepsilon \sum_{i=1}^{m} r_i. \tag{3-17}$$

Moreover, we can assume that the rectangles $R_j^{\pm}$ associated to $Q_{\nu_j}(x_j,r_j)$ according to (3-9) and (3-10) belong to different connected components $A_{j,n}^{\pm}$ of $Q_{\nu_j}(x_j,r_j) \setminus \widehat{H}_n$. Let us denote by

$$v_j^{\pm} \in H^1(Q_{\nu_j}(x_j,r_j)) \tag{3-18}$$

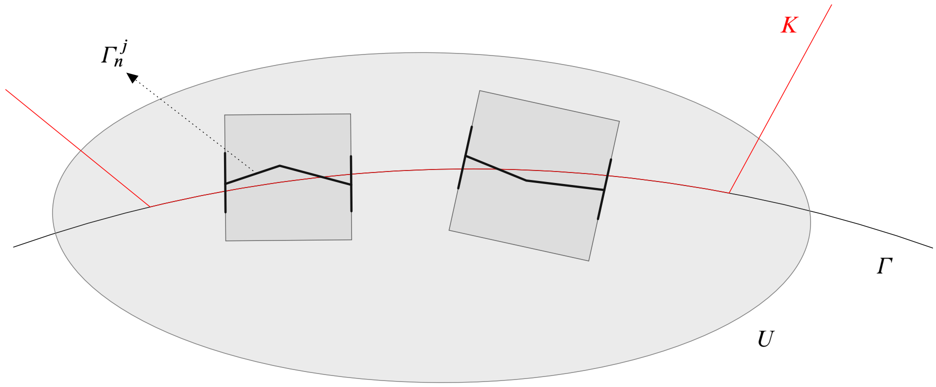

Figure 4. The sets Γ_n defined in (3-19).

the extension of $v\lfloor R_j^\pm$ obtained through a reflection across the line $l_{x_j} \pm \varepsilon r \nu_j$: notice that the Sobolev regularity of $v_j^\pm$ is ensured because, by construction,

$$\Gamma \cap Q_{\nu_j}(x_j, r_j) \subseteq \{x \in \mathbb{R}^2 : |(x - x_j)\cdot \nu_j| < \varepsilon r\}.$$

Let us set

$$\Gamma_n := \left(\Gamma \setminus \bigcup_{j=1}^m Q_{\nu_j}(x_j, r_j)\right) \cup \bigcup_{j=1}^m \Gamma_n^j, \tag{3-19a}$$

with

$$\Gamma_n^j := \big(\widehat{H}_n \cap \overline{Q}_{\nu_j}(x_j, r_j)\big) \cup \big(\partial Q_{\nu_j}(x_j, r_j) \ \cap \{|(y - x_j)\cdot \nu_j| \le \varepsilon r_j\} \cap \overline{\Omega}\big). \tag{3-19b}$$

See Figure 4.

Notice that $\Gamma_n \in \mathcal{K}_1^f(\overline{\Omega})$. Moreover thanks to (3-15)–(3-17),

$$\begin{aligned}
\mathcal{H}^1_D(\Gamma_n \setminus K_n) &\le \mathcal{H}^1_D(\Gamma_n \setminus \widehat{H}_n) + \mathcal{H}^1(\widehat{H}_n \setminus H_n) + \mathcal{H}^1(H_n \setminus K_n)\\
&\le \mathcal{H}^1_D\left(\Gamma \setminus \bigcup_{j=1}^m Q_{\nu_j}(x_j, r_j)\right) + 7\varepsilon \sum_{j=1}^m r_j + \mathcal{H}^1(H_n \setminus K_n)\\
&\le \mathcal{H}^1_D(\Gamma \setminus K) + \varepsilon + 7\varepsilon \frac{1}{1-\varepsilon}\mathcal{H}^1(\Gamma) + \mathcal{H}^1(H_n \setminus K_n),
\end{aligned} \tag{3-20}$$

and, since $\Gamma_n \subseteq V$,

$$\mathcal{H}^1_D(K_n \setminus \Gamma_n) \ge \mathcal{H}^1_D(K_n \setminus \overline{V}).$$

Let us define v_n as follows:

(a) $v_n = v$ outside $\bigcup_{j=1}^m Q_{\nu_j}(x_j, r_j)$;

(b) $v_n := \begin{cases} v_j^+ & \text{in } A_{j,n}^+, \\ v_j^- & \text{else} \end{cases}$ in each cube $Q_{\nu_j}(x_j, r_j) \Subset \Omega$, where the functions $v_j^\pm$ were defined in (3-18);

(c) $v_n := \begin{cases} v_j^+ & \text{in } A_{j,n}^+, \\ g & \text{otherwise} \end{cases}$ in each boundary cube $Q_{\nu_j}(x_j, r_j)$ (that is, those with $x_j \in \partial\Omega$).

Remark that, by construction, $(v_n, \Gamma_n) \in \mathcal{A}(g)$. Moreover,

$$\|\nabla v_n\|^2 \le \|\nabla v\|^2 + 2\sum_{j=1}^{m} \int_{Q_{\nu_j}(x_j,r_j)\cap\bar{\Omega}} |\nabla v|^2\,dx + \sum_{j=1}^{m} \int_{Q_{\nu_j}(x_j,r_j)\cap\bar{\Omega}} |\nabla g|^2\,dx$$
$$\le \|\nabla v\|^2 + 2\int_{U\cap\bar{\Omega}} |\nabla v|^2\,dx + \int_{U\cap\bar{\Omega}} |\nabla g|^2\,dx. \tag{3-21}$$

Let us compare (u_n, K_n) to $(v_n - g + g_n, \Gamma_n) \in \mathcal{A}(g_n)$. Since

$$\|\nabla u_n\|^2 + c\mathcal{H}^1_D(K_n \setminus \Gamma_n) \le \|\nabla v_n - \nabla g + \nabla g_n\|^2 + (c+c')\mathcal{H}^1_D(\Gamma_n \setminus K_n)$$
$$= \|\nabla v_n\|^2 + (c+c')\,\mathcal{H}^1_D(\Gamma_n \setminus K_n) + e_n,$$

where

$$|e_n| \le \|\nabla v_n\|\|\nabla g_n - \nabla g\| + \|\nabla g_n - \nabla g\|^2 \to 0,$$

we infer in view of (3-20)–(3-21) that

$$\|\nabla u_n\|^2 + c\mathcal{H}^1_D(K_n \setminus \bar{V}) \le \|\nabla v\|^2 + 2\int_{U\cap\bar{\Omega}} |\nabla v|^2\,dx + \int_{U\cap\bar{\Omega}} |\nabla g|^2\,dx + e_n$$
$$+ (c+c')\Big[\mathcal{H}^1_D(\Gamma \setminus K) + \varepsilon + 7\varepsilon\frac{1}{1-\varepsilon}\mathcal{H}^1(\Gamma) + \mathcal{H}^1(H_n \setminus K_n)\Big].$$

Passing to the limit, we obtain, thanks to Gołąb's theorem and to (3-14),

$$\|\nabla u\|^2 + c\mathcal{H}^1_D(K \setminus \bar{V}) \le \|\nabla v\|^2 + (c+c')\mathcal{H}^1_D(\Gamma \setminus K)$$
$$+ (c+c')\Big[\varepsilon + \frac{7\varepsilon}{1-\varepsilon}\mathcal{H}^1(\Gamma)\Big] + 2\int_{U\cap\bar{\Omega}} |\nabla v|^2\,dx + \int_{U\cap\bar{\Omega}} |\nabla g|^2\,dx.$$

Since V, U and ε are arbitrary, we conclude that

$$\|\nabla u\|^2 + c\mathcal{H}^1_D(K \setminus \Gamma) \le \|\nabla v\|^2 + (c+c')\mathcal{H}^1_D(\Gamma \setminus K),$$

so that the minimality condition follows.

Step 2: Let us consider the general case $K \in \mathcal{K}^f_m(\bar{\Omega})$. If $K^1, \dots K^p$ with $p \le m$ are the connected components of K, thanks to [Dal Maso and Toader 2002, Lemma 3.6] we can find $H_n \in \mathcal{K}^f_m(\bar{\Omega})$ with exactly p connected components $H^1_n, \dots, H^p_n$ such that $K_n \subseteq H_n$,

$$H^j_n \to K^j \quad \text{in the Hausdorff metric} \tag{3-22}$$

and

$$\mathcal{H}^1(H_n \setminus K_n) \to 0.$$

Since the K^j are compact and disjoint, and

$$\Gamma \cap K = \bigcup_{j=1}^{p} (\Gamma \cap K^j),$$

we can operate on each $\Gamma \cap K^j$ as in Step 1 using the approximation (3-22) and localizing on disjoint neighborhoods U_j of $\Gamma \cap K^j$. The modifications of Γ and v which take place on the family of squares contained in U_j are independent from those taking place in the squares contained in U_i with $i \neq j$, so we can glue them together to get an approximating configuration $(v_n - g + g_n, \Gamma_n) \in \mathcal{A}(g_n)$ and deduce as in Step 1 the global minimality of (u, K). ☐

4. Existence of a quasistatic evolution

In this section we derive the main result of the paper.

Theorem 4.1 (existence of a quasistatic evolution). *Let $\Omega \subseteq \mathbb{R}^2$ be open, bounded, with Lipschitz boundary, and let $\partial_D \Omega \subseteq \partial\Omega$ be open in the relative topology. Assume* (2-3) *and let $g \in \mathrm{AC}([0, T]; H^1(\Omega))$ be such that*

$$\sup_{t\in[0,T]} \|g(t)\|_\infty < +\infty. \tag{4-1}$$

Let finally $(u_0, K_0) \in \mathcal{A}(g(0))$ be a globally stable configuration, i.e., satisfying property (2-4).

Then there exists a quasistatic evolution $\{t \mapsto (u(t), K(t)) : t \in [0, T]\}$ in the sense of Definition 2.3 such that $(u(0), K(0)) = (u_0, K_0)$.

Remark 4.2. The existence of at least one globally stable initial configuration $(u_0, K_0) \in \mathcal{A}(g(0))$ is straightforward. It is enough to minimize $\mathcal{E}(v, \Gamma)$ over $\mathcal{A}(g(0))$ following, e.g., an argument identical to that expounded in the proof of Lemma 4.3 below.

As usual, the existence of a quasistatic evolution is obtained by time discretization, establishing the existence of a discrete-in-time evolution through the direct method of the calculus of variations, and then studying its limit as the time-step discretization parameter vanishes.

Let $\delta > 0$ be given, and let

$$0 = t_0^\delta < t_1^\delta < \cdots < t_{N_\delta}^\delta = T$$

be a subdivision of the time interval $[0, T]$ with

$$\max_{i=0,\dots,N_\delta-1} (t_{i+1}^\delta - t_i^\delta) < \delta.$$

We set

$$g_i^\delta := g(t_i^\delta) \quad \text{and} \quad (u_0^\delta, K_0^\delta) := (u_0, K_0).$$

The following lemma deals with the existence of incremental configurations.

Lemma 4.3 (incremental configurations). *Assume* (2-3) *and* (4-1). *Then for $i = 1, \dots, N_\delta$ there exists $(u_i^\delta, K_i^\delta) \in \mathcal{A}(g_i^\delta)$ with $\|u_i^\delta\|_\infty \leq \|g_i^\delta\|_\infty$, $(u_0^\delta, K_0^\delta) = (u_0, K_0)$, such that*

$$(u_i^\delta, K_i^\delta) \in \mathrm{Argmin}\{\mathcal{E}(v, \Gamma) + (c_1 - c_2)\,\mathcal{H}_D^1(\Gamma \setminus K_{i-1}^\delta) : (v, \Gamma) \in \mathcal{A}(g_i^\delta)\}.$$

Proof. We proceed by induction, assuming that $(u_{i-1}^\delta, K_{i-1}^\delta)$ has been constructed, and showing the existence of (u_i^δ, K_i^δ).

Set

$$\mathcal{F}_i^\delta(u,\Gamma) := \mathcal{E}(v,\Gamma) + (c_1-c_2)\,\mathcal{H}^1_D(\Gamma\setminus K^\delta_{i-1}).$$

and let $\{(v_n,\Gamma_n)\}_{n\in\mathbb{N}}$ be a minimizing sequence for $\mathcal{F}_i^\delta$ on $\mathcal{A}(g_i^\delta)$, that is,

$$I_i^\delta := \inf_{\mathcal{A}(g_i^\delta)} \mathcal{F}_i^\delta \le \mathcal{F}_i^\delta(v_n,\Gamma_n) \le I_i^\delta + \frac{1}{n}.$$

By truncation, it is not restrictive to assume

$$\|v_n\|_\infty \le \|g_i^\delta\|_\infty.$$

Comparing with the admissible configuration $(g_i^\delta,\varnothing)$ we get

$$\mathcal{E}(v_n,\Gamma_n) + (c_1-c_2)\,\mathcal{H}^1_D(\Gamma_n\setminus K^\delta_{i-1}) \le \|\nabla g_i^\delta\|^2.$$

As a consequence, up to a subsequence we may assume

$$\begin{aligned} (\nabla v_n, v_n) &\rightharpoonup (\Phi, v) &&\text{weakly in } L^2(\Omega;\mathbb{R}^3),\\ \Gamma_n &\to \Gamma &&\text{in the Hausdorff metric.} \end{aligned}$$

Thanks to Gołąb's theorem, we infer $\Gamma\in\mathcal{K}_f^m(\overline{\Omega})$, and, by Lemma 2.1, we deduce that $(v,\Gamma)\in\mathcal{A}(g_i^\delta)$, with $\Phi=\nabla v$ on $\Omega\setminus\Gamma$. In particular

$$\nabla v_n \rightharpoonup \nabla v \quad \text{weakly in } L^2(\Omega;\mathbb{R}^2).$$

Moreover, in view of Lemma 1.4

$$\mathcal{H}^1_D(\Gamma) \le \liminf_n \mathcal{H}^1_D(\Gamma_n) \quad\text{and}\quad \mathcal{H}^1_D(\Gamma\setminus K^\delta_{i-1}) \le \liminf_n \mathcal{H}^1_D(\Gamma_n\setminus K^\delta_{i-1}),$$

so that

$$\mathcal{F}_i^\delta(v,\Gamma) = I_i^\delta.$$

The thesis follows by setting $(u_i^\delta, K_i^\delta) := (v,\Gamma)$. □

For $t_i^\delta \le t < t_{i+1}^\delta$, $i=0,\dots,N_\delta$, we set

$$u^\delta(t) := u_i^\delta, \quad g^\delta(t) := g_i^\delta \quad\text{and}\quad K^\delta(t) := K_i^\delta. \tag{4-2}$$

We denote by $i^\delta(t)$ the index such that $t^\delta_{i^\delta(t)} \le t < t^\delta_{i^\delta(t)+1}$.

The properties below follow directly from the construction of the incremental configurations.

Lemma 4.4. *For every $t\in[0,T]$ the following items hold true*:

(a) $(u^\delta(0), K^\delta(0)) = (u_0, K_0)$.

(b) *The pair $(u^\delta(t), K^\delta(t))\in\mathcal{A}(g^\delta(t))$ satisfies the global stability condition* (2-4).

(c) *Setting*

$$\mathrm{Diss}^\delta(t) := (c_1-c_2)\sum_{j=1}^{i^\delta(t)} \mathcal{H}^1_D(K_j^\delta\setminus K^\delta_{j-1}), \tag{4-3}$$

we have the energy inequality

$$\mathcal{E}(u^\delta(t), K^\delta(t)) + \mathrm{Diss}^\delta(t) \leq \mathcal{E}(u_0, K_0) + 2\int_0^{t_i^\delta}\int_\Omega \nabla u^\delta(\tau)\cdot\nabla\dot{g}(\tau)\,dx\,d\tau + e(\delta), \tag{4-4}$$

where $e(\delta) \to 0$ *as* $\delta \to 0$.

Proof. Point (a) follows since $(u^\delta(0), K^\delta(0)) = (u_0^\delta, K_0^\delta) = (u_0, K_0)$.

On to point (b). By construction, for every $i = 1, \dots, N_\delta$ and $(v, \Gamma) \in \mathcal{A}(g_i^\delta)$,

$$\mathcal{E}(u_i^\delta, K_i^\delta) + (c_1 - c_2)\,\mathcal{H}_D^1(K_i^\delta \setminus K_{i-1}^\delta) \leq \mathcal{E}(v, \Gamma) + (c_1 - c_2)\,\mathcal{H}_D^1(\Gamma \setminus K_{i-1}^\delta).$$

Since

$$\mathcal{H}_D^1(\Gamma \setminus K_{i-1}^\delta) \leq \mathcal{H}_D^1(\Gamma \setminus K_i^\delta) + \mathcal{H}_D^1(K_i^\delta \setminus K_{i-1}^\delta),$$

we deduce

$$\mathcal{E}(u_i^\delta, K_i^\delta) \leq \mathcal{E}(v, \Gamma) + (c_1 - c_2)\,\mathcal{H}_D^1(\Gamma \setminus K_i^\delta),$$

from which the global stability condition (2-4) follows.

Let us come to point (c). In view of Lemma 4.3 we may write, for every $i = 1, \dots, N_\delta$,

$$\begin{aligned}\mathcal{E}(u_i^\delta, K_i^\delta) &+ (c_1 - c_2)\,\mathcal{H}_D^1(K_i^\delta \setminus K_{i-1}^\delta)\\ &\leq \mathcal{E}(u_{i-1}^\delta + g_i^\delta - g_{i-1}^\delta, K_{i-1}^\delta)\\ &\leq \mathcal{E}(u_{i-1}^\delta, K_{i-1}^\delta) + 2\int_{t_{i-1}^\delta}^{t_i^\delta}\int_\Omega \nabla u^\delta(\tau)\cdot\nabla\dot{g}(\tau)\,dx\,d\tau + (t_i^\delta - t_{i-1}^\delta)\int_{t_{i-1}^\delta}^{t_i^\delta} \|\nabla\dot{g}(\tau)\|^2\,d\tau.\end{aligned}$$

Iterating this estimate we obtain for every $t \in [0, T]$

$$\mathcal{E}(u^\delta(t), K^\delta(t)) + (c_1 - c_2)\sum_{j=1}^{i^\delta(t)} \mathcal{H}_D^1(K_j^\delta \setminus K_{j-1}^\delta) \leq \mathcal{E}(u_0, K_0) + 2\int_0^{t_i^\delta}\int_\Omega \nabla u^\delta(\tau)\cdot\nabla\dot{g}(\tau)\,dx\,d\tau + e(\delta),$$

with $e(\delta) \to 0$ as $\delta \to 0$, which is precisely (4-4). □

In order to pass to the continuous-in-time evolution, we need the following bounds.

Lemma 4.5 (a priori bounds). *Let* $\{t \mapsto (u^\delta(t), K^\delta(t)) : t \in [0, T]\}$ *be the discrete-in-time evolution given by* (4-2). *There exists* $C > 0$ *independent of* δ *such that, for every* $t \in [0, T]$,

$$\|\nabla u^\delta(t)\| + \|u^\delta(t)\|_\infty + \mathcal{H}_D^1(K^\delta(t)) + z^\delta(t) \leq C, \tag{4-5}$$

where

$$z^\delta(t) := \sum_{j=1}^{i^\delta(t)} \mathcal{H}_D^1(K_j^\delta \Delta K_{j-1}^\delta). \tag{4-6}$$

Proof. Since by construction and global minimality

$$\|\nabla u^\delta(t)\| \leq \|\nabla g^\delta(t)\| \quad \text{and} \quad \|u^\delta(t)\|_\infty \leq \|g^\delta(t)\|_\infty, \tag{4-7}$$

we deduce from (4-4) that

$$\mathcal{H}^1_D(K^\delta(t)) + \sum_{j=1}^{i^\delta(t)} \mathcal{H}^1_D(K^\delta_j \setminus K^\delta_{j-1}) \le C_1 \tag{4-8}$$

for some $C_1 > 0$. Since $\mathcal{H}^1_D(B \setminus A) - \mathcal{H}^1_D(A \setminus B) = \mathcal{H}^1_D(B) - \mathcal{H}^1_D(A)$,

$$\sum_{j=1}^{i^\delta(t)} \mathcal{H}^1_D(K^\delta_j \setminus K^\delta_{j-1}) - \sum_{j=1}^{i^\delta(t)} \mathcal{H}^1_D(K^\delta_{j-1} \setminus K^\delta_j) = \mathcal{H}^1_D(K^\delta(t)) - \mathcal{H}^1_D(K_0),$$

we also obtain from (4-8) that

$$\sum_{j=1}^{i^\delta(t)} \mathcal{H}^1_D(K^\delta_{j-1} \setminus K^\delta_j) \le C_2 \tag{4-9}$$

for some $C_2 > 0$. The conclusion follows gathering (4-7)–(4-9). □

A crucial step in the $\delta \searrow 0$-analysis is the following

Proposition 4.6 (compactness of the cracks). *There exist a sequence $\delta_n \to 0$ and a map $\{t \mapsto K(t) \in \mathcal{K}^f_m(\overline{\Omega}) : t \in [0, T]\}$ such that, if*

$$K_n(t) := K^{\delta_n}(t), \quad t \in [0, T],$$

then, for every $t \in [0, T]$, any limit point H of $(K_n(t))_{n\in\mathbb{N}}$ in the Hausdorff metric is such that

$$\mathcal{H}^1_D(H \Delta K(t)) = 0.$$

Proof. Let $\delta_n \to 0$ be such that

$$z_n := z^{\delta_n} \to z \quad \text{pointwise on } [0, T],$$

where z^δ is given in (4-6) and $z : [0, T] \to \mathbb{R}$ is a suitable increasing function. The existence of $(\delta_n)_{n\in\mathbb{N}}$ is a consequence of the bound (4-5) and of Helly's theorem.

Let $D \subseteq [0, T]$ be a countable and dense set containing 0 and the discontinuity points of the function z. Up to a further subsequence (that we will not relabel), we may assume, in view of the compactness of the Hausdorff metric and of the bound (4-5), that for every $t \in D$ there exists $K(t) \in \mathcal{K}^f_m(\overline{\Omega})$ such that

$$K_n(t) \to K(t) \quad \text{in the Hausdorff metric.}$$

Let now $s \notin D$, and let H be a limit point of the sequence $(K_n(s))_{n\in\mathbb{N}}$, that is,

$$K_{n_k}(s) \to H \quad \text{in the Hausdorff metric}$$

for a suitable subsequence $(n_k)_{k\in\mathbb{N}}$. By the definition of z_n, for every $t < s$ and $t \in D$,

$$\mathcal{H}^1_D(K_{n_k}(s) \Delta K_{n_k}(t)) \le z_{n_k}(s) - z_{n_k}(t).$$

Sending $k \to +\infty$ and using Lemma 1.4 we obtain

$$\mathcal{H}^1_D(H \Delta K(t)) \le z(s) - z(t).$$

Let now $t_k \nearrow s$ with $t_k \in D$ and such that

$$K(t_k) \to \widetilde{K}(s) \quad \text{in the Hausdorff metric.}$$

Recalling that s is a continuity point for z, we infer (using again Lemma 1.4) that

$$\mathcal{H}^1_D(H \Delta \widetilde{K}(s)) = 0. \tag{4-10}$$

Since $(t_k)_{k\in\mathbb{N}}$ is arbitrary, we deduce that any limit point $\widetilde{K}(s)$ of the family $\{K(t) : t \in D\}$ for $t \to s^-$ satisfies (4-10). The proof now follows by choosing $K(s)$ as one of these limit points. □

Remark 4.7. Let $H, K \in \mathcal{K}^f_m(\overline{\Omega})$ be such that

$$\mathcal{H}^1_D(K \Delta H) = 0. \tag{4-11}$$

Then:

(i) K and H differ by at most m points on $\Omega \cup \partial_D \Omega$.

(ii) If $(v, H) \in \mathcal{A}(g)$, then also $(v, K) \in \mathcal{A}(g)$.

Indeed let H_j be a connected component of H which contains a point x not in K. Since H_j is connected by arcs, (4-11) implies that H_j reduces to the point x, which proves point (i).

As far as point (ii) is concerned, we know that $(\nabla v, v)$ can be interpreted as an element of $L^2(\Omega; \mathbb{R}^3)$. Let us first check that $v \in W^{1,2}(\Omega \setminus K)$ with gradient on $\Omega \setminus K$ given by ∇v. We can proceed locally near every point $x \in \Omega \setminus K$:

(a) If $x \notin H$, since $u \in W^{1,2}(\Omega \setminus H)$ we deduce $u \in W^{1,2}(B(x, r))$ for some $r > 0$ small enough, with gradient given by ∇v.

(b) If $x \in H$, then, according to point (i), the connected component H_j of H that contains x reduces to the point x. From $u \in W^{1,2}(\Omega \setminus H)$ we then deduce that for some $r > 0$ small enough

$$u \in W^{1,2}(B(x, r) \setminus H_j) = W^{1,2}(B(x, r) \setminus \{x\}) = W^{1,2}(B(x, r)),$$

with gradient given by ∇v.

Concerning the boundary condition, since $u = g$ on $\partial_D \Omega \setminus H$ in the sense of traces, (4-11) then entails that the equality also holds true on $\partial_D \Omega \setminus K$. We thus conclude that $(u, K) \in \mathcal{A}(g)$.

We are now in a position to prove Theorem 4.1.

Proof of Theorem 4.1. Let $\delta_n \to 0$ and $\{t \mapsto K(t) : t \in [0, T]\}$ be given by Proposition 4.6. Set

$$(u_n(t), K_n(t)) := (u^{\delta_n}(t), K^{\delta_n}(t)) \quad \text{and} \quad \mathrm{Diss}_n(t) := \mathrm{Diss}^{\delta_n}(t).$$

Up to a further subsequence, the a priori bounds of Lemma 4.5, imply that

$$\mathrm{Diss}_n \to D \quad \text{pointwise on } [0, T] \tag{4-12}$$

for some increasing function $D : [0, T] \to [0, +\infty)$.

For every $t \in [0,T]$ take $u(t) \in H^1(\Omega \setminus K(t))$ to be a minimizer of

$$\min_{(v,K(t))\in\mathcal{A}(g(t))} \|\nabla v\|^2.$$

By strict convexity, $\nabla u(t)$ is uniquely determined by $K(t)$ and $g(t)$, while $u(t)$ is well-defined up to a constant on the connected components of $\Omega \setminus K(t)$ which do not touch $\partial_D\Omega$.

We now prove that

$$\{t \mapsto (u(t), K(t)) : t \in [0,T]\}$$

is a quasistatic evolution for the boundary displacement g such that $(u(0), K(0)) = (u_0, K_0)$ according to Definition 2.3.

<u>Step 1</u>: global stability. Let us check that, for every $t \in [0,T]$, the pair $(u(t), K(t))$ satisfies the global stability condition (2-4), which reads

$$\|\nabla u(t)\|^2 + c_2\mathcal{H}^1_D(K(t)) \le \|\nabla v\|^2 + c_2\mathcal{H}^1_D(\Gamma) + (c_1-c_2)\,\mathcal{H}^1_D(\Gamma \setminus K(t)). \tag{4-13}$$

In view of the bound (4-5), by Lemma 2.1 and by the compactness of the Hausdorff convergence, we may assume that, up to a subsequence,

$$\begin{aligned} K_n(t) &\to H \in \mathcal{K}^f_m(\overline{\Omega}) && \text{in the Hausdorff metric,}\\ (\nabla u_n(t), u_n(t)) &\rightharpoonup (\nabla u, u) && \text{weakly in } L^2(\Omega;\mathbb{R}^3) \end{aligned}$$

for some $(u,H) \in \mathcal{A}(g(t))$.

From item (b) in Lemma 4.4 and Theorem 3.1 we infer that (u,H) satisfies the global stability condition

$$\|\nabla u\|^2 + c_2\mathcal{H}^1_D(H) \le \|\nabla v\|^2 + c_2\mathcal{H}^1_D(\Gamma) + (c_1-c_2)\,\mathcal{H}^1_D(\Gamma \setminus H) \tag{4-14}$$

for every $(v,\Gamma) \in \mathcal{A}(g(t))$. Note that, by Proposition 4.6,

$$\mathcal{H}^1_D(H \Delta K(t)) = 0.$$

Then Remark 4.7 implies that $(u, K(t)) \in \mathcal{A}(g(t))$, so that the minimality property (4-14) becomes

$$\|\nabla u\|^2 + c_2\mathcal{H}^1_D(K(t)) \le \|\nabla v\|^2 + c_2\mathcal{H}^1_D(\Gamma) + (c_1-c_2)\,\mathcal{H}^1_D(\Gamma \setminus K(t))$$

for every $(v,\Gamma) \in \mathcal{A}(g(t))$. Comparing with the admissible configuration $(u(t), K(t))$ yields

$$\|\nabla u\|^2 \le \|\nabla u(t)\|^2,$$

so that, by the very definition of $u(t)$, we get $\nabla u(t) = \nabla u$ and conclude that (4-13) is satisfied.

From the arguments above, passing to subsequences is not necessary and we infer that

$$\nabla u_n(t) \rightharpoonup \nabla u(t) \quad \text{weakly in } L^2(\Omega;\mathbb{R}^2) \tag{4-15}$$

for every $t \in [0,T]$.

Step 2: energy balance. Let us first prove that, for every $t \in [0, T]$,

$$\mathrm{Diss}(t) \leq D(t). \tag{4-16}$$

Indeed, for every $0 = s_0 < s_1 < \cdots < s_{k+1} = t$,

$$(c_1 - c_2) \sum_{h=0}^{k} \mathcal{H}^1_D(K_n(s_{h+1}) \setminus K_n(s_h)) \leq \mathrm{Diss}_n(t). \tag{4-17}$$

According to Proposition 4.6, up to a further subsequence, we have

$$K_n(s_j) \to H(s_j) \quad \text{in the Hausdorff metric,}$$

with

$$\mathcal{H}^1_D(H(s_j) \Delta K(s_j)) = 0. \tag{4-18}$$

Then, with the help of Lemma 1.4 and of (4-18) we pass to the limit in (4-17) and obtain, in view of (4-12),

$$(c_1 - c_2) \sum_{h=0}^{k} \mathcal{H}^1_D(K(s_{h+1}) \setminus K(s_h)) \leq D(t),$$

from which (4-16) easily follows.

Thanks to (4-15),(4-16) and to Gołąb's theorem, we can pass to the limit in the discrete energy inequality (4-4) and obtain

$$\mathcal{E}(u(t), K(t)) + \mathrm{Diss}(t) \leq \mathcal{E}(u_0, K_0) + 2\int_0^t \int_\Omega \nabla u(\tau) \cdot \nabla \dot{g}(\tau)\, dx\, d\tau. \tag{4-19}$$

The opposite inequality in (4-19) holds true, thanks to a by now standard *Riemann sum argument*; see [Dal Maso et al. 2005, Section 4.4]. In a nutshell, the argument consists in choosing a specific sequence of partitions $\{s_i^n\}_{i=0,\dots,k(n)}$ with $k(n) \overset{n}{\nearrow} \infty$ of the interval $[0, t]$ such that the Riemann sums

$$\sum_{i=0}^{k(n)-1} \int_{s_i^n}^{s_{i+1}^n} \int_\Omega \nabla u(s_{i+1}^n) \cdot \dot{g}(s)\, dx\, ds$$

converge to

$$\int_0^t \int_\Omega \nabla u(s) \cdot \dot{g}(s)\, dx\, ds.$$

Then one writes the minimality condition for $(u(s_i^n), K(s_i^n)) \in \mathcal{A}(g(s_i^n))$ established in Step 1, tested against $(u(s_{i+1}^n) - g(s_{i+1}^n) + g(s_i^n), K(s_{i+1}^n)) \in \mathcal{A}(g(s_i^n))$ and adds all resulting contributions for $i = 0, \dots, k(n) - 1$; see [Dal Maso et al. 2005, Section 4.4] for the details.

The energy balance

$$\mathcal{E}(u(t), K(t)) + \mathrm{Diss}(t) = \mathcal{E}(u_0, K_0) + 2\int_0^t \int_\Omega \nabla u(\tau) \cdot \nabla \dot{g}(\tau)\, dx\, d\tau$$

follows. We conclude that $t \mapsto (u(t), K(t))$ is a quasistatic evolution. □

Remark 4.8 (improved convergences). The proof of Theorem 4.1 shows that, for every $t \in [0, T]$,

$$\nabla u_n(t) \to \nabla u(t) \quad \text{strongly in } L^2(\Omega; \mathbb{R}^2), \tag{4-20}$$

$$\mathcal{H}^1_D(K_n(t)) \to \mathcal{H}^1_D(K(t)), \tag{4-21}$$

$$\mathrm{Diss}_n(t) \to \mathrm{Diss}(t).$$

Indeed from the arguments of Step 2 and (4-4) we have

$$\begin{aligned}
\mathcal{E}(u_0, K_0) + 2\int_0^t \int_\Omega \nabla u(\tau) \cdot \nabla \dot{g}(\tau)\, dx\, d\tau& \\
&= \mathcal{E}(u(t), K(t)) + \mathrm{Diss}(t) \\
&\le \liminf_n [\mathcal{E}(u_n(t), K_n(t)) + \mathrm{Diss}_n(t)] \le \limsup_n [\mathcal{E}(u_n(t), K_n(t)) + \mathrm{Diss}_n(t)] \\
&\le \limsup_n \Big[\mathcal{E}(u_0, K_0) + \int_0^t \int_\Omega \nabla u_n(\tau) \cdot \nabla \dot{g}(\tau)\, dx\, d\tau + e(\delta_n) \Big] \\
&= \mathcal{E}(u_0, K_0) + 2\int_0^t \int_\Omega \nabla u(\tau) \cdot \nabla \dot{g}(\tau)\, dx\, d\tau,
\end{aligned}$$

from which

$$\lim_n [\mathcal{E}(u_n(t), K_n(t)) + \mathrm{Diss}_n(t)] = \mathcal{E}(u(t), K(t)) + \mathrm{Diss}(t).$$

We thus deduce that

$$\lim_n \mathcal{E}(u_n(t), K_n(t)) = \mathcal{E}(u(t), K(t)) \quad \text{and} \quad \lim_n \mathrm{Diss}_n(t) = \mathrm{Diss}(t),$$

and the first convergence gives immediately (4-20) and (4-21).

Remark 4.9 (the connected case). In the connected case, loss of Hausdorff convergence only takes place at *healing times*, i.e., when $K(t)$ reduces to a point (or is the empty set) on $\Omega \cup \partial_D \Omega$. Indeed, assume the existence of two different subsequences $K_{n_k}(t)$, $K_{\tilde{n}_k}(t)$, with

$$\begin{aligned}
K_{n_k}(t) &\to H_1 \quad \text{in the Hausdorff metric,} \\
K_{\tilde{n}_k}(t) &\to H_2 \quad \text{in the Hausdorff metric,}
\end{aligned}$$

with $H_1 \neq H_2$. Since, in view of Proposition 4.6,

$$\mathcal{H}^1_D(H_1 \Delta K(t)) = \mathcal{H}^1_D(H_1 \Delta K(t)) = 0,$$

it must be that $\mathcal{H}^1_D(H_1 \Delta H_2) = 0$.

According to point (i) in Remark 4.7, those two sets, which are connected, must then reduce to at most a single point on $\Omega \cup \partial_D \Omega$. Since $K(t)$ is also connected, it in turn reduces to at most a single point on $\Omega \cup \partial_D \Omega$.

Finally, taking into account Remark 4.8, at such a time,

$$\mathcal{H}^1_D(K_n(t)) \to 0$$

because $\mathcal{H}^1_D(K(t)) = 0$. So, if Hausdorff convergence does not take place at time t, the approximating cracks are actually vanishing in length.

The argument fails if $m > 1$. In that case, using similar arguments, we can merely assert the existence of a subsequence of $K_n(t)$ such that one of its connected component heals in the limit, which is not much. . . .

5. The case of two-dimensional elasticity

In this section, we show how to modify the previous arguments in the case of linearized two-dimensional elasticity.

Admissible configurations. Let the reference configuration be an open bounded set $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary, while we consider $H^1(\Omega;\mathbb{R}^2)$ as the class of admissible boundary displacements.

Given $\partial_D\Omega \subseteq \partial\Omega$ open in the relative topology, we say that the pair (u, K) is an admissible configuration for the boundary displacement $g \in H^1(\Omega;\mathbb{R}^2)$ if

$$K \in \mathcal{K}_m^f(\overline{\Omega}) \quad \text{and} \quad u \in \mathscr{LD}(\Omega \setminus K), \quad \text{with } u = g \text{ on } \partial_D\Omega \setminus K,$$

where $m \geq 1$ is a fixed number, and $\mathcal{K}_m^f(\overline{\Omega})$ is given in (2-1). We will write $(u, K) \in \mathcal{A}(g)$. The pair $(u, e(u))$ can be thought of as an element of $L^2_{\mathrm{loc}}(\Omega;\mathbb{R}^2) \times L^2(\Omega;\mathrm{M}^2_{\mathrm{sym}})$ since K has null Lebesgue measure.

Remark 5.1. Let $(u, K) \in \mathcal{A}(g)$, and let $H \in \mathcal{K}_m^f(\overline{\Omega})$ be such that $\mathcal{H}^1_D(K\Delta H) = 0$. Then $(u, H) \in \mathcal{A}(g)$. The proof follows precisely that in Remark 4.7: indeed the local arguments can be reproduced because, in view of Korn's inequality, elements of $\mathscr{LD}(\Omega \setminus K)$ are locally in H^1.

The following compactness result plays the role of Lemma 2.1 in our context.

Lemma 5.2. *Let $g_n, g \in H^1(\Omega;\mathbb{R}^2)$ be such that*

$$g_n \to g \quad \textit{strongly in } H^1(\Omega;\mathbb{R}^2).$$

Assume that $(u_n, K_n) \in \mathcal{A}(g_n)$ with

$$\begin{aligned} e(u_n) \rightharpoonup \Phi &\quad \textit{weakly in } L^2(\Omega;\mathrm{M}^2_{\mathrm{sym}}),\\ K_n \to K &\quad \textit{in the Hausdorff metric}. \end{aligned}$$

Then there exists $(u, K) \in \mathcal{A}(g)$ such that $\Phi = e(u)$ on $\Omega \setminus K$.

Proof. Let A be a connected component of $\Omega \setminus K$, and let $B \Subset A$ be a disk. Consider

$$\mathscr{R} := \left\{v \in \mathscr{LD}(\Omega \setminus K) : \textstyle\int_B v \cdot r \, dx = 0 \text{ for all } r \in \mathcal{R}\right\},$$

where $\mathcal{R}$ is the set of infinitesimally rigid motions, i.e.,

$$\mathcal{R} := \{Mx + b : M \in \mathrm{M}^2_{\mathrm{skew}}, b \in \mathbb{R}^2\}.$$

Define $\hat{u}_n$ to be the $L^2(B)$-orthogonal projection of u_n onto $\mathscr{R}$; clearly $e(\hat{u}_n) = e(u_n)$.

Since K_n Hausdorff-converges to K, any open Lipschitz connected subdomain G compactly embedded in A and containing B is also included, for n large enough, in $\Omega \setminus K_n$. Thus, according to Korn's inequality, $\hat{u}_n \in H^1(G;\mathbb{R}^2)$ and there exists $C_{G,B} > 0$ such that

$$\|\hat{u}_n\|_{L^2(G;\mathbb{R}^2)} \le C_{G,B}\|e(u_n)\|_{L^2(G;\mathrm{M}^2_{\mathrm{sym}})} \le C$$

for some C depending on G, B; hence, up to a subsequence,

$$\hat{u}_n \rightharpoonup u_G \quad \text{weakly in } H^1(G;\mathbb{R}^2),$$

with

$$e(u_G) = \Phi. \tag{5-1}$$

But u_G also belongs to $\mathscr{R}$. In view of (5-1), it is thus uniquely defined so that the whole sequence $\hat{u}_n$ converges to u_G weakly in $H^1(G;\mathbb{R}^2)$, and hence strongly in $L^2(G;\mathbb{R}^2)$. Then taking G to be an increasing sequence of Lipschitz connected open sets with union A, we immediately conclude that $u_G \equiv u$ independent of G with $u \in L^2_{\mathrm{loc}}(A;\mathbb{R}^2)$ and $e(u) = \Phi$. Since A is an arbitrary connected component of $\Omega \setminus K$, we infer that $u \in \mathscr{LD}(\Omega \setminus K)$.

The proof that $u = g$ on $\partial_D\Omega \setminus K$ is identical to that in Lemma 2.1 upon renewed use of Korn's inequality. □

Quasistatic evolutions. Let the Hooke's law be given by an element $\mathbb{C} \in L^\infty(\Omega;\mathcal{L}_s(\mathrm{M}^2_{\mathrm{sym}}))$ such that

$$a_1|M|^2 \le \mathbb{C}(x)M\cdot M \le a_2|M|^2 \quad \text{for every } M \in \mathrm{M}^2_{\mathrm{sym}}, \tag{5-2}$$

with $a_1, a_2 > 0$. Here $\cdot$ denotes the standard Frobenius matrix inner product.

We associate to an admissible configuration (u, K) the *elastic energy*

$$\mathcal{Q}(e(u)) := \frac{1}{2}\int_\Omega \mathbb{C}(x)e(u)(x)\cdot e(u)(x)\,dx.$$

As in Section 2, let $T > 0$ and $g \in \mathrm{AC}([0,T]; H^1(\Omega;\mathbb{R}^2))$ be a given time-dependent boundary displacement, and let

$$c_1 > c_2 > 0 \tag{5-3}$$

be two given constants. In analogy with the scalar case (see Definition 2.3), we define a quasistatic evolution in the case of linearized elasticity as follows.

Definition 5.3 (quasistatic evolution). We say that $\{t \mapsto (u(t), K(t)) \in \mathcal{A}(g(t)) : t \in [0,T]\}$ is a quasistatic evolution provided that for every $t \in [0,T]$ the following items hold true:

(a) (global stability) For every $(v, \Gamma) \in \mathcal{A}(g(t))$

$$\mathcal{E}(u(t), K(t)) \le \mathcal{E}(v, \Gamma) + (c_1 - c_2)\,\mathcal{H}^1_D(\Gamma \setminus K(t)),$$

where, for $(u, K) \in \mathcal{A}(g)$,

$$\mathcal{E}(u, K) := \mathcal{Q}(e(u)) + c_2\mathcal{H}^1_D(K).$$

(b) (energy balance) We have

$$\mathcal{E}(u(t), K(t)) + \mathrm{Diss}(t) = \mathcal{E}(u(0), K(0)) + \int_0^t \int_\Omega \mathbb{C}e(u(\tau)) \cdot e(\dot{g}(\tau))\, dx\, d\tau,$$

where

$$\mathrm{Diss}(t) := (c_1 - c_2) \sup\Big\{\textstyle\sum_{i=0}^n \mathcal{H}^1_D(K(s_{i+1}) \setminus K(s_i)) : 0 = s_0 < s_1 < \cdots < s_{n+1} = t\Big\}.$$

Existence of quasistatic evolutions. The main result of the section is the following.

Theorem 5.4 (existence of a quasistatic evolution for two-dimensional elasticity). *Let $\Omega \subseteq \mathbb{R}^2$ be an open, bounded Lipschitz domain and $\partial_D\Omega \subseteq \partial\Omega$ be open in the relative topology. Let $g \in \mathrm{AC}([0, T]; H^1(\Omega; \mathbb{R}^2))$ and assume* (5-2) *and* (5-3) *hold true. Let finally $(u_0, K_0) \in \mathcal{A}(g(0))$ be a globally stable configuration*; *i.e., satisfying property* (2-4).

Then, there exists a quasistatic evolution $\{t \mapsto (u(t), K(t)) : t \in [0, T]\}$ in the sense of Definition 5.3 such that $(u(0), K(0)) = (u_0, K_0)$.

Proof. We proceed as in Section 4 by constructing incremental configurations $(u_i^\delta, K_i^\delta) \in \mathcal{A}(g_i^\delta)$. We consider

$$(u_i^\delta, K_i^\delta) \in \mathrm{Argmin}\{\mathcal{E}(v, \Gamma) + (c_1 - c_2)\,\mathcal{H}^1_D(\Gamma \setminus K_{i-1}^\delta) : (v, \Gamma) \in \mathcal{A}(g_i^\delta)\}. \tag{5-4}$$

The variational problems are well-posed thanks to Lemma 5.2 and to Gołąb's theorem.

Interpolating in time, we obtain the discrete-in-time evolution

$$\{t \mapsto (u^\delta(t), K^\delta(t)) : t \in [0, T]\}$$

such that, defining Diss^δ as in (4-3),

$$\mathcal{E}(u^\delta(t), K^\delta(t)) + \mathrm{Diss}^\delta(t) \le \mathcal{E}(u_0, K_0) + \int_0^{t_i^\delta} \int_\Omega \mathbb{C}e(u^\delta(\tau)) \cdot e(\dot{g}(\tau))\, dx\, d\tau + e(\delta),$$

with $e(\delta) \to 0$ as $\delta \to 0$. In view of (5-2), this inequality yields the uniform bound

$$\|e(u^\delta(t))\| + \mathcal{H}^1_D(K^\delta(t)) + z^\delta(t) \le C,$$

where z^δ is defined as in (4-6).

Thanks to Lemma 5.2, the proof is now completely analogous to that of Theorem 4.1, provided that we adapt Theorem 3.1 to our context.

Specifically, it suffices to prove the following. Let $c, c' \ge 0$, and let $g_n, g \in H^1(\Omega; \mathbb{R}^2)$ be such that

$$g_n \to g \quad \text{strongly in } H^1(\Omega; \mathbb{R}^2).$$

Assume that $(u_n, K_n) \in \mathcal{A}(g_n)$ satisfy the following global stability condition: for every $(v, \Gamma) \in \mathcal{A}(g_n)$,

$$\mathcal{Q}(e(u_n)) + c\,\mathcal{H}^1_D(K_n) \le \mathcal{Q}(e(v)) + c\,\mathcal{H}^1_D(\Gamma) + c'\mathcal{H}^1_D(\Gamma \setminus K_n)$$

and assume further that

$$\begin{aligned} K_n &\to K && \text{in the Hausdorff metric,} \\ e(u_n) &\rightharpoonup e(u) && \text{weakly in } L^2(\Omega; \mathrm{M}^2_{\mathrm{sym}}) \end{aligned}$$

for some $(u, K) \in \mathcal{A}(g)$. Then (u, K) is a globally stable configuration, that is that, for every $(v, \Gamma) \in \mathcal{A}(g)$,

$$\mathcal{Q}(e(u)) + c\,\mathcal{H}^1_D(K) \le \mathcal{Q}(e(v)) + c\,\mathcal{H}^1_D(\Gamma) + c'\mathcal{H}^1_D(\Gamma \setminus K). \tag{5-5}$$

Notice that, in view of [Chambolle 2003, Theorem 1], there exists $v_m \in H^1(\Omega \setminus \Gamma; \mathbb{R}^2)$ with $v_m = g$ on $\partial_D \Omega$ and such that

$$e(v_m) \to e(v) \quad \text{strongly in } L^2(\Omega; \mathrm{M}^2_{\mathrm{sym}}).$$

As a consequence, it is sufficient to establish (5-5) in the case $(v, \Gamma) \in \mathcal{A}(g)$ with

$$v \in H^1(\Omega \setminus \Gamma; \mathbb{R}^2). \tag{5-6}$$

This is a great simplification, since we can employ the same construction as that in the proof of Theorem 3.1 working on each component.

Specifically, if $v := (v^1, v^2)$, we fix U, V, ε as in Step 1 of the proof of Theorem 3.1, and construct the associated Γ_n, v^1_n, v^2_n (approximations of the scalar functions v^1, v^2). The crucial estimate (3-21) now reads as follows (we can estimate in the squares the symmetrized gradient by the full gradient thanks to (5-6)):

$$\begin{aligned}\mathcal{Q}(e(v_n)) &\le \mathcal{Q}(e(v)) + 2a_2 \sum_{j=1}^m \int_{Q_{\nu_j}(x_j, r_j) \cap \bar{\Omega}} |\nabla v|^2 \, dx + a_2 \sum_{j=1}^m \int_{Q_{\nu_j}(x_j, r_j) \cap \bar{\Omega}} |\nabla g|^2 \, dx \\ &\le \mathcal{Q}(e(v)) + 2a_2 \int_{U \cap \bar{\Omega}} |\nabla v|^2 \, dx + \int_{U \cap \bar{\Omega}} |\nabla g|^2 \, dx,\end{aligned}$$

where a_2 is the coercivity constant in (5-2). Comparing (u_n, K_n) with $(v_n - g + g_n, \Gamma_n) \in \mathcal{A}(g_n)$ and using the previous inequality we deduce that

$$\begin{aligned}\mathcal{Q}(e(u)) + c\mathcal{H}^1_D(K \setminus \bar{V}) &\le \mathcal{Q}(e(v)) + (c + c')\mathcal{H}^1_D(\Gamma \setminus K) \\ &\quad + (c + c')\left[\varepsilon + \frac{7\varepsilon}{1-\varepsilon}\mathcal{H}^1(\Gamma)\right] + 2a_2 \int_{U \cap \bar{\Omega}} |\nabla v|^2 \, dx + a_2 \int_{U \cap \bar{\Omega}} |\nabla g|^2 \, dx,\end{aligned}$$

so that the global stability follows since V, U and ε are arbitrary. □

Remark 5.5. Notice that even if an L^∞-bound for the boundary displacement g is assumed, the functional framework for the displacement u^δ_i in the incremental problems (5-4) cannot reduce to $H^1(\Omega \setminus K^\delta_i)$ since truncation fails in the case of energies that depend on the symmetrized gradient.

Acknowledgements

This research has been supported by the National Science Foundation Grants DMS-1615661 and DMS-1615839 and was completed in part while A. Giacomini was visiting the Courant Institute in March 2017 with the support of the grant DMS-1615839. A. Giacomini is also member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

References

[Acanfora and Ponsiglione 2006] F. Acanfora and M. Ponsiglione, "Quasi static growth of brittle cracks in a linearly elastic flexural plate", *Ann. Mat. Pura Appl.* (4) **185**:2 (2006), 293–317. MR Zbl

[Ambrosio and Tilli 2004] L. Ambrosio and P. Tilli, *Topics on analysis in metric spaces*, Oxford Lecture Series in Math. and Its Appl. **25**, Oxford Univ. Press, 2004. MR Zbl

[Ambrosio et al. 2000] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Univ. Press, 2000. MR Zbl

[Ball 1982] J. M. Ball, "Discontinuous equilibrium solutions and cavitation in nonlinear elasticity", *Philos. Trans. Roy. Soc. London Ser. A* **306**:1496 (1982), 557–611. MR Zbl

[Blaiszik et al. 2010] B. J. Blaiszik, S. L. B. Kramer, S. C. Olugebefola, J. S. Moore, N. R. Sottos, and S. R. White, "Self-healing polymers and composites", *Ann. Rev. Mater. Res.* **40** (2010), 179–211.

[Bourdin et al. 2008] B. Bourdin, G. A. Francfort, and J.-J. Marigo, "The variational approach to fracture", *J. Elasticity* **91**:1-3 (2008), 5–148. MR Zbl

[Busse 1938] W. F. Busse, "Physics of rubber as related to the automobile", *J. Appl. Phys.* **9**:7 (1938), 438–451.

[Chambolle 2003] A. Chambolle, "A density result in two-dimensional linearized elasticity, and applications", *Arch. Ration. Mech. Anal.* **167**:3 (2003), 211–233. MR Zbl

[Cordier et al. 2008] P. Cordier, F. Tournilhac, C. Soulié-Ziakovic, and L. Leibler, "Self-healing and thermoreversible rubber from supramolecular assembly", *Nature* **451** (2008), 977–980.

[Dal Maso and Toader 2002] G. Dal Maso and R. Toader, "A model for the quasi-static growth of brittle fractures: existence and approximation results", *Arch. Ration. Mech. Anal.* **162**:2 (2002), 101–135. MR Zbl

[Dal Maso et al. 2005] G. Dal Maso, G. A. Francfort, and R. Toader, "Quasistatic crack growth in nonlinear elasticity", *Arch. Ration. Mech. Anal.* **176**:2 (2005), 165–225. MR Zbl

[Deny and Lions 1954] J. Deny and J. L. Lions, "Les espaces du type de Beppo Levi", *Ann. Inst. Fourier, Grenoble* **5** (1954), 305–370. MR Zbl

[Falconer 1986] K. J. Falconer, *The geometry of fractal sets*, Cambridge Tracts in Math. **85**, Cambridge Univ. Press, 1986. MR Zbl

[Francfort and Larsen 2003] G. A. Francfort and C. J. Larsen, "Existence and convergence for quasi-static evolution in brittle fracture", *Comm. Pure Appl. Math.* **56**:10 (2003), 1465–1500. MR Zbl

[Francfort and Marigo 1998] G. A. Francfort and J.-J. Marigo, "Revisiting brittle fracture as an energy minimization problem", *J. Mech. Phys. Solids* **46**:8 (1998), 1319–1342. MR Zbl

[Gent and Lindley 1959] A. N. Gent and P. B. Lindley, "Internal rupture of bonded rubber cylinders in tension", *Proc. Roy. Soc. Lond. A* **249**:1257 (1959), 195–205.

[Gent and Park 1984] A. N. Gent and B. Park, "Failure processes in elastomers at or near a rigid spherical inclusion", *J. Mater. Sci.* **19**:6 (1984), 1947–1956.

[Griffith 1921] A. A. Griffith, "The phenomena of rupture and flow in solids", *Philos. Trans. Roy. Soc. London Ser. A* **221** (1921), 163–198.

[Henao and Mora-Corral 2010] D. Henao and C. Mora-Corral, "Invertibility and weak continuity of the determinant for the modelling of cavitation and fracture in nonlinear elasticity", *Arch. Ration. Mech. Anal.* **197**:2 (2010), 619–655. MR Zbl

[Il'iushin 1961] A. A. Il'iushin, "On the postulate of plasticity", *Prikl. Mat. Meh.* **25**:3 (1961), 503–507. In Russian; translated in *J. Appl. Math. Mech.* **25**:3 (1961), 746–752. MR Zbl

[Kumar et al. 2018] A. Kumar, G. A. Francfort, and O. Lopez-Pamies, "Fracture and healing of elastomers: a phase-transition theory and numerical implementation", *J. Mech. Phys. Solids* **112** (2018), 523–551. MR

[Lefèvre et al. 2015] V. Lefèvre, K. Ravi-Chandar, and O. Lopez-Pamies, "Cavitation in rubber: an elastic instability or a fracture phenomenon?", *Int. J. Fracture* **192**:1 (2015), 1–23.

[Madsen et al. 2016] F. B. Madsen, L. Yu, and A. L. Skov, "Self-healing, high-permittivity silicone dielectric elastomer", *ACS Macro Lett.* **5**:11 (2016), 1196–1200.

[Mielke et al. 2018] A. Mielke, R. Rossi, and G. Savaré, "Global existence results for viscoplasticity at finite strain", *Arch. Ration. Mech. Anal.* **227**:1 (2018), 423–475. MR Zbl

[Mora-Corral 2014] C. Mora-Corral, "Quasistatic evolution of cavities in nonlinear elasticity", *SIAM J. Math. Anal.* **46**:1 (2014), 532–571. MR Zbl

[Poulain et al. 2017] X. Poulain, V. Lefèvre, O. Lopez-Pamies, and K. Ravi-Chandar, "Damage in elastomers: nucleation and growth of cavities, micro-cracks, and macro-cracks", *Int. J. Fracture* **205**:1 (2017), 1–21.

Received 2 Oct 2017. Revised 20 Apr 2018. Accepted 29 Jun 2018.

GILLES FRANCFORT: gilles.francfort@univ-paris13.fr, gilles.francfort@cims.nyu.edu
Laboratoire Analyse, Géométrie et Applications, Université Paris-Nord, Villetaneuse, France

and

Courant Institute of Mathematical Sciences, New York, NY, USA

ALESSANDRO GIACOMINI: alessandro.giacomini@unibs.it
DICATAM, Sezione di Matematica, Università degli Studi di Brescia, Brescia, Italy

OSCAR LOPEZ-PAMIES: pamies@illinois.edu
Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Urbana, IL, United States

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.449

GENERAL CLARK MODEL FOR FINITE-RANK PERTURBATIONS

CONSTANZE LIAW AND SERGEI TREIL

All unitary (contractive) perturbations of a given unitary operator U by finite-rank-d operators with fixed range can be parametrized by $d \times d$ unitary (contractive) matrices Γ; this generalizes unitary rank-one ($d = 1$) perturbations, where the Aleksandrov–Clark family of unitary perturbations is parametrized by the scalars on the unit circle $\mathbb{T} \subset \mathbb{C}$.

For a strict contraction Γ the resulting perturbed operator T_Γ is (under the natural assumption about star cyclicity of the range) a completely nonunitary contraction, so it admits the functional model.

We investigate the Clark operator, i.e., a unitary operator that intertwines T_Γ (written in the spectral representation of the nonperturbed operator U) and its model. We make no assumptions on the spectral type of the unitary operator U; an absolutely continuous spectrum may be present.

We first find a universal representation of the adjoint Clark operator in the coordinate-free Nikolski–Vasyunin functional model; the word "universal" means that it is valid in any transcription of the model. This representation can be considered to be a special version of the vector-valued Cauchy integral operator.

Combining the theory of singular integral operators with the theory of functional models, we derive from this abstract representation a concrete formula for the adjoint of the Clark operator in the Sz.-Nagy–Foiaş transcription. As in the scalar case, the adjoint Clark operator is given by a sum of two terms: one is given by the boundary values of the vector-valued Cauchy transform (postmultiplied by a matrix-valued function) and the second one is just the multiplication operator by a matrix-valued function.

Finally, we present formulas for the direct Clark operator in the Sz.-Nagy–Foiaş transcription.

The work of C. Liaw is supported by the National Science Foundation DMS-1802682. Work of S. Treil is supported by the National Science Foundation under the grants DMS-1301579, DMS-1600139. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

MSC2010: 44A15, 47A20, 47A55.

Keywords: finite-rank perturbations, Clark theory, dilation theory, functional model, normalized Cauchy transform.

0. Introduction

The contractive (or unitary) perturbations $U + K$ of a unitary operator U on a Hilbert space H by operators K of finite rank d with fixed range are parametrized by the $d \times d$ contractive (resp. unitary) matrices Γ. Namely, if $\operatorname{Ran} K \subset \mathfrak{R}$, where $\mathfrak{R} \subset H$, $\dim \mathfrak{R} = d$, is fixed, and $\boldsymbol{B} : \mathbb{C}^d \to \mathfrak{R}$ is a fixed unitary operator (which we call the coordinate operator), then K is represented as $K = \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^* U$, where Γ is a contraction (resp. a unitary operator) on $\mathbb{C}^d$. Therefore, all such perturbations with $\operatorname{Ran} K \subset \mathfrak{R}$ are represented as $T_\Gamma = U + \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^* U$, where Γ runs over all $d \times d$ contractive (resp. unitary) matrices.

Recall that T being a *contraction* (contractive) means that $\|T\| \le 1$.

Focusing on the nontrivial part of the perturbation, we can assume that $\operatorname{Ran} \boldsymbol{B} = \mathfrak{R}$ is a star-cyclic subspace for U, i.e., $H = \overline{\operatorname{span}}\{U^k \mathfrak{R}, (U^*)^k \mathfrak{R} : k \in \mathbb{Z}_+\}$. Below we will show that star-cyclicity together with the assumption that Γ is a pure contraction ensures that the operator T_Γ is what is called a completely nonunitary contraction, meaning that T_Γ does not have a nontrivial unitary part. Model theory informs us that such T_Γ is unitarily equivalent to its functional model $\mathcal{M}_\theta$, $\theta = \theta_\Gamma$, that is, the compression of the shift operator on the model space $\mathcal{K}_\theta$ with the characteristic function $\theta = \theta_\Gamma$ of T_Γ.

In this paper we investigate the so-called Clark operator, i.e., a unitary operator Φ that intertwines the contraction T_Γ (in the spectral representation of the unperturbed operator U) with its model: $\mathcal{M}_\theta \Phi = \Phi T_\Gamma$, $\theta = \theta_\Gamma$. The case of rank-one perturbations ($d = 1$) was treated by D. Clark [1972] when θ is inner, and later by D. Sarason [1994] under the assumption that θ is an extreme point of the unit ball of H^∞. For finite-rank perturbations with inner characteristic matrix-valued functions θ, V. Kapustin and A. Poltoratski [2006] studied boundary convergence of functions in the model space $\mathcal{K}_\theta$. The setting of inner characteristic functions corresponds to the operators U that have purely singular spectrum (no a.c. component); see, e.g., [Douglas and Liaw 2013].

In [Liaw and Treil 2016] we completely described the general case of rank-one perturbations (when the measure can have absolutely continuous part, or equivalently, the characteristic function is not necessarily inner).

In the present paper we extend the results from [Liaw and Treil 2016] to finite-rank perturbations with general matrix-valued characteristic functions. We first find a universal representation of the adjoint Clark operator, which features a special case of a matrix-valued Cauchy integral operator. By "universal" we mean that our formula is valid in any transcription of the functional model. This representation is a pretty straightforward, albeit more algebraically involved, generalization of the corresponding result from [Liaw and Treil 2016]; it might look like "abstract nonsense", since it is proved under the assumption that we pick a model operator that "agrees" with the Clark model (more precisely that the corresponding coordinate/parametrizing operators agree).

However, by careful investigation of the construction of the functional model, using the coordinate-free Nikolski–Vasyunin model we were able to present a formula giving the parametrizing operators for the model that agree with given coordinate operators for a general contraction T; see Lemma 3.2. Moreover, for the Sz.-Nagy–Foiaș transcription of the model we get explicit formulas for the parametrizing operators in terms of the characteristic function; see Lemma 3.3. Similar formulas can be obtained for other transcriptions of the model.

We also compute the characteristic function of the perturbed operator T_Γ; the formula involves the Cauchy integral of the matrix-valued measure.

For the Sz.-Nagy–Foiaş transcription of the model we give a more concrete representation of the adjoint Clark operator in terms of the vector-valued Cauchy transform; see Theorem 8.1. This representation looks more natural when one considers spectral representations of the nonperturbed operator U defined with the help of matrix-valued measures; see Theorem 8.7.

0A. ***Plan of the paper.*** In Section 1 we set the stage by introducing finite-rank perturbations and studying some of their basic properties. In particular, we discuss the concept of a star-cyclic subspace and find a measure-theoretic characterization for it.

The main result of Section 2 is the universal representation formula for the adjoint Clark operator; see Theorem 2.4. In this section we also introduce the notion of agreement of the coordinate/parametrizing operators and make some preliminary observations about such an agreement.

Section 3 is devoted to the detailed investigation of the agreement of the coordinate/parametrizing operators. Careful analysis of the construction of the model from the coordinate-free point of view of Nikolski–Vasyunin allows us to get for a general contraction T-formulas for the parametrizing operators for the model that agree with the coordinate operators; see Lemma 3.2. Explicit formulas (in terms of the characteristic function) are presented for the case of Sz.-Nagy–Foiaş transcription; see Lemma 3.3.

The characteristic function θ_Γ of the perturbed operator T_Γ is the topic of Sections 4 and 5. Theorem 4.2 gives a formula for θ_Γ in terms of a Cauchy integral of a matrix-valued measure. In Section 5 we show that, similarly to the rank-one case, the characteristic functions θ_Γ and $\theta_{\mathbf{0}}$ are related via a special linear fractional transformation. Relations between defect functions $\Delta_{\mathbf{0}}$ and Δ_Γ are also described.

Section 6 contains a brief heuristic overview of what subtle techniques are to come in Sections 7 and 8.

In Section 7 we present results about regularizations of the Cauchy transform, and about uniform boundedness of such generalizations, which we need to get the representation formulas in Section 8.

In Section 8 we give a formula for the adjoint Clark operator in the Sz.-Nagy–Foiaş transcription of the model. As in the scalar case, the adjoint Clark operator is given by the sum of two terms: one is in essence a vector-valued Cauchy transform (postmultiplied by a matrix-valued function), and the second one is just a multiplication operator by a matrix-valued function; see Theorem 8.1. In the case of inner characteristic functions (purely singular spectral measure of U) the second term disappears, and the adjoint Clark operator is given by what can be considered a matrix-valued analogue of the scalar *normalized Cauchy transform*; see Section 8E.

Section 9 is devoted to a description of the Clark operator Φ; see Theorem 9.2.

1. Preliminaries

Consider the family of rank-d perturbations $U + K$ of a unitary operator U on a separable Hilbert space H. If we fix a subspace $\mathfrak{R} \subset H$, $\dim \mathfrak{R} = d$, such that $\operatorname{Ran} K \subset \mathfrak{R}$, then all unitary perturbations of $U + K$

of U can be parametrized as

$$T = U + (X - \boldsymbol{I}_{\mathfrak{R}}) P_{\mathfrak{R}} U, \tag{1-1}$$

where X runs over all possible unitary operators in $\mathfrak{R}$.

It is more convenient to factorize the representation of X through the fixed space $\mathfrak{D} := \mathbb{C}^d$ by picking an isometric operator $\boldsymbol{B} : \mathfrak{D} \to H$, $\operatorname{Ran} \boldsymbol{B} = \mathfrak{R}$. Then any X in (1-1) can be represented as $X = \boldsymbol{B} \Gamma \boldsymbol{B}^*$, where $\Gamma : \mathfrak{D} \to \mathfrak{D}$ (i.e., Γ is a $d \times d$ matrix). The perturbed operator $T = T_\Gamma$ can be rewritten as

$$T = U + \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathfrak{D}}) \boldsymbol{B}^* U. \tag{1-2}$$

If we decompose the space H treated as the domain as $H = U^* \mathfrak{R} \oplus (U^* \mathfrak{R})^\perp$, and the same space treated as the target space as $H = \mathfrak{R} \oplus \mathfrak{R}^\perp$, then the operator T can be represented with respect to this decomposition as

$$T = \begin{pmatrix} \boldsymbol{B} \Gamma \boldsymbol{B}^* U & 0 \\ 0 & T_1 \end{pmatrix}, \tag{1-3}$$

where block T_1 is unitary.

From the above decomposition we can immediately see that if Γ is a contraction then T is a contraction (and if Γ is unitary then T is unitary).

In this formula we slightly abuse notation, since formally the operator $\boldsymbol{B} \Gamma \boldsymbol{B}^* U$ is defined on the whole space H. However, this operator clearly annihilates $(U^* \mathfrak{R})^\perp$, and its range belongs to $\mathfrak{R}$, so we can restrict its domain and target space to $U^* \mathfrak{R}$ and $\mathfrak{R}$ respectively. So when such operators appear in the block decomposition we will assume that its domain and target space are restricted.

In this paper we assume that the isometry $\boldsymbol{B}$ is fixed and that all the perturbations are parametrized by the $d \times d$ matrix Γ.

1A. ***Spectral representation of U.*** By the spectral theorem the operator U is unitarily equivalent to the multiplication M_ξ by the independent variable ξ in the von Neumann direct integral

$$\mathcal{H} = \int_{\mathbb{T}}^{\oplus} E(\xi) \, \mathrm{d}\mu(\xi), \tag{1-4}$$

where μ is a finite Borel measure on $\mathbb{T}$ (without loss of generality we can assume that μ is a probability measure, $\mu(\mathbb{T}) = 1$).

Let us recall the construction of the direct integral; we present not the most general one, but one that is sufficient for our purposes. Let E be a separable Hilbert space with an orthonormal basis $\{e_n\}_{n=1}^\infty$, and let $N : \mathbb{T} \to \mathbb{N} \cup \{\infty\}$ be a measurable function (the so-called *dimension function*). Define

$$E(\xi) = \overline{\operatorname{span}}\{e_n \in E : 1 \le n \le N(\xi)\}.$$

Then the direct integral $\mathcal{H}$ is the subspace of the E-valued space $L^2(\mu; E) = L^2(\mathbb{T}, \mu; E)$ consisting of the functions f such that $f(\xi) \in E(\xi)$ for μ-a.e. ξ.

Note, that the dimension function N and the spectral type $[\mu]$ of μ (i.e., the collection of all measures that are mutually absolutely continuous with μ) are spectral invariants of U, meaning that they define the operator U up to unitary equivalence.

So, without loss of generality, we assume that U is the multiplication M_ξ by the independent variable ξ in the direct integral (1-4).

An important particular case is the case when U is star-cyclic, meaning that there exists a vector $h \in H$ such that $\overline{\operatorname{span}}\{U^n h : n \in \mathbb{Z}\} = H$. In this case $N(\xi) \equiv 1$, and the operator U is unitarily equivalent to the multiplication operator M_ξ in the scalar space $L^2(\mu) = L^2(\mathbb{T}, \mu)$.

In the representation of U in the direct integral it is convenient to give a "matrix" representation of the isometry $\boldsymbol{B}$. Namely, for $k = 1, 2, \dots, d$ define functions $b_k \in \mathcal{H} \subset L^2(\mu; E)$ by $b_k := \boldsymbol{B} e_k$; here $\{e_k\}_{k=1}^d$ is the standard orthonormal basis in $\mathbb{C}^d$.

In this notation the operator $\boldsymbol{B}$, if we follow the standard rules of the linear algebra, is the multiplication by a row B of vector-valued functions,

$$B(\xi) = (b_1(\xi), b_2(\xi), \dots, b_d(\xi)).$$

If we represent $b_k(\xi)$ in the standard basis in E that we used to construct the direct integral (1-4), then $\boldsymbol{B}$ is just the multiplication by the matrix-valued function of size $(\dim E) \times d$.

1B. *Star-cyclic subspaces and completely nonunitary contractions.*

Definition 1.1. As was previously mentioned, a subspace $\mathfrak{R}$ is said to be *star-cyclic* for an operator T on H if

$$H = \overline{\operatorname{span}}\{T^k \mathfrak{R}, (T^*)^k \mathfrak{R} : k \in \mathbb{Z}_+\}.$$

For a perturbation (not necessarily unitary) $T = T_\Gamma$ of the unitary operator U given by (1-2) the subspace

$$\mathcal{E} = \overline{\operatorname{span}}\{U^k \mathfrak{R}, (U^*)^k \mathfrak{R} : k \in \mathbb{Z}_+\} = \overline{\operatorname{span}}\{U^k \mathfrak{R} : k \in \mathbb{Z}\} \tag{1-5}$$

is a reducing subspace for both U and T_Γ (i.e., $\mathcal{E}$ and $\mathcal{E}^\perp$ are invariant for both U and T_Γ).

Since $T_\Gamma|_{\mathcal{E}^\perp} = U|_{\mathcal{E}^\perp}$, the perturbation does not influence the action of T_Γ on $\mathcal{E}^\perp$, so nothing interesting for perturbation theory happens on $\mathcal{E}^\perp$; all action happens on $\mathcal{E}$. Therefore, we can restrict our attention to $T_\Gamma|_{\mathcal{E}}$, i.e., assume without loss of generality that $\mathfrak{R} = \operatorname{Ran} \boldsymbol{B}$ is a star-cyclic subspace for U.

We note the following result.

Lemma 1.2. *Let $\mathfrak{R} = \operatorname{Ran} \boldsymbol{B}$ be a star-cyclic subspace for U and let Γ be unitary. Then $\mathfrak{R}$ is also a star-cyclic subspace for all perturbed unitary operators $U_\Gamma = T_\Gamma$ given by* (1-2).

We postpone for a moment a proof of this well-known fact.

Definition 1.3. A contraction T in a Hilbert space H is called *completely nonunitary* (c.n.u. for short) if there is no nonzero reducing subspace on which T acts unitarily.

Recall that a contraction is called *strict* if $\|Tx\| < \|x\|$ for all $x \neq \mathbf{0}$.

Lemma 1.4. *If $\mathfrak{R} = \operatorname{Ran} \boldsymbol{B}$ is a star-cyclic subspace for U and Γ is a strict contraction, then T defined by* (1-2) *is a c.n.u. contraction.*

Proof. Since Γ is a strict contraction, we get that $\boldsymbol{B}\Gamma\boldsymbol{B}^*U|_{U^*\mathfrak{R}}$ is also a strict contraction. Therefore (1-3) implies

$$\|Tx\| = \|x\| \iff x \perp U^{-1}\mathfrak{R},$$
$$\|T^*x\| = \|x\| \iff x \perp \mathfrak{R}.$$

Moreover, we can see from (1-3) that if $x \perp U^{-1}\mathfrak{R}$ then $Tx = Uf$ and if $x \perp \mathfrak{R}$ then $T^*x = U^{-1}x$.

Consider a reducing subspace G for T such that $T|_G$ is unitary. Then the above observations imply $G \perp \mathfrak{R}$ and $G \perp U^{-1}\mathfrak{R}$, and that for any $x \in G$

$$T^n x = U^n x \quad \text{as well as} \quad (T^*)^n x = U^{-n}x.$$

Since G is a reducing subspace for T it follows that $U^k x \in G$ for all integers k. But this implies that $U^n x \perp \mathfrak{R}$, or equivalently $x \perp U^n\mathfrak{R}$ for all $n \in \mathbb{Z}$. But $\mathfrak{R}$ is a star-cyclic subspace for U, so we get a contradiction. □

Proof of Lemma 1.2. Assume now that for unitary Γ, the subspace $\operatorname{Ran}\boldsymbol{B}$ is not a star-cyclic subspace for $U_\Gamma = T_\Gamma$ (but is a star-cyclic subspace for U). Consider the perturbation $T_{\mathbf{0}}$,

$$T_{\mathbf{0}} = U + \boldsymbol{B}(\mathbf{0} - \boldsymbol{I}_{\mathfrak{D}})\boldsymbol{B}^*U.$$

We will show that

$$T_{\mathbf{0}} = U_\Gamma + \boldsymbol{B}(\mathbf{0} - \boldsymbol{I}_{\mathfrak{D}})\boldsymbol{B}^*U_\Gamma. \tag{1-6}$$

By Lemma 1.4 the operator $T_{\mathbf{0}}$ is a c.n.u. contraction.

But, as we discussed in the beginning of this subsection, if $\operatorname{Ran}\boldsymbol{B}$ is not star-cyclic for U, then for $\mathcal{E}$ defined by (1-5) the subspace $\mathcal{E}^\perp$ is a reducing subspace for T_Γ (with any Γ) on which T_Γ acts unitarily.

Since by (1-6) the operator $T_{\mathbf{0}}$ is a perturbation of form (1-2) of the unitary operator T_Γ, we conclude that the operator $T_{\mathbf{0}}$ has a nontrivial unitary part, and arrive at a contradiction.

To prove (1-6) we notice that

$$T_{\mathbf{0}} = U - \boldsymbol{B}\boldsymbol{B}^*U = U_\Gamma - \boldsymbol{B}\Gamma\boldsymbol{B}^*U. \tag{1-7}$$

Direct computations show that

$$U_\Gamma U^*\boldsymbol{B} = UU^*\boldsymbol{B} + \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathfrak{D}})\boldsymbol{B}^*UU^*\boldsymbol{B} = \boldsymbol{B} + \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathfrak{D}}) = \boldsymbol{B}\Gamma.$$

Taking the adjoint of this identity we get $\boldsymbol{B}^*UU_\Gamma^* = \Gamma^*\boldsymbol{B}^*$, and so $\Gamma\boldsymbol{B}^*U = \boldsymbol{B}^*U_\Gamma$. Substituting $\boldsymbol{B}^*U_\Gamma$ instead of $\Gamma\boldsymbol{B}^*U$ in (1-7) we get (1-6). □

1C. ***Characterization of star-cyclic subspaces.*** Recall that for an isometry $\boldsymbol{B} : \mathcal{D} \to \mathcal{H}$, where $\mathcal{H}$ is the direct integral (1-4), we denoted by $b_k \in \mathcal{H}$ the "columns" of $\boldsymbol{B}$,

$$b_k = \boldsymbol{B}e_k,$$

where $e_1, e_2, \dots, e_d$ is the standard basis in $\mathbb{C}^d$.

Lemma 1.5. *Let U be the multiplication M_ξ by the independent variable ξ in the direct integral $\mathcal{H}$ given by* (1-4), *and let $\boldsymbol{B} : \mathbb{C}^d \to \mathcal{H}$ be as above. The space* $\operatorname{Ran} \boldsymbol{B} = \operatorname{span}\{b_k : 1 \le k \le d\}$ *is star-cyclic for U if and only if* $\overline{\operatorname{span}}\{b_k(\xi) : 1 \le k \le d\} = E(\xi)$ *for μ-a.e. ξ.*

Proof. First assume that $\operatorname{Ran} \boldsymbol{B}$ is not a star-cyclic subspace for U. Then there exists $f \in \mathcal{H} \subset L^2(\mu; E)$, $f \ne 0$ μ-a.e., such that

$$U^l f \perp b_k \quad \text{for all } l \in \mathbb{Z} \text{ and } k = 1, \dots, d,$$

or, equivalently

$$\int_{\mathbb{T}} (f(\xi), b_k(\xi))_E \, \xi^l \, \mathrm{d}\mu(\xi) = 0 \quad \text{for all } l \in \mathbb{Z} \text{ and } k = 1, \dots, d.$$

But that means for all $k = 1, 2, \dots, d$ we have

$$(f(\xi), b_k(\xi))_E = 0 \quad \mu\text{-a.e.},$$

so on some set of positive μ-measure (where $f(\xi) \ne \mathbf{0}$) we have

$$\overline{\operatorname{span}}\{b_k(\xi) : 1 \le k \le d\} \subsetneq E(\xi). \tag{1-8}$$

Vice versa, assume that (1-8) holds on some Borel subset $A \subset \mathbb{T}$ with $\mu(A) > 0$. For $n = 1, 2, \dots, \infty$ define sets $A_n := \{\xi \in A : \dim E(\xi) = n\}$. Then $\mu(A_n) > 0$ for some n. Fix this n and denote the corresponding space $E(\xi)$, $\xi \in A_n$, by E_n.

We know that $\overline{\operatorname{span}}\{b_k(\xi) : 1 \le k \le d\} \subsetneq E_n$ on A_n, so there exists $e \in E_n$ such that

$$e \notin \overline{\operatorname{span}}\{b_k(\xi) : 1 \le k \le d\}$$

on a set of positive measure in A_n.

Trivially, if $f \in \overline{\operatorname{span}}\{U^k \operatorname{Ran} \boldsymbol{B} : k \in \mathbb{Z}\}$ then

$$f(\xi) \in \overline{\operatorname{span}}\{b_k(\xi) : 1 \le k \le d\} \quad \mu\text{-a.e.},$$

and therefore $f = \mathbf{1}_{A_n} e$ is not in $\overline{\operatorname{span}}\{U^k \operatorname{Ran} \boldsymbol{B} : k \in \mathbb{Z}\}$. □

1D. *The case of star-cyclic U.* If U is star-cyclic (i.e., it has a one-dimensional star-cyclic subspace/vector), U is unitarily equivalent to the multiplication operator M_ξ in the scalar space $L^2(\mu)$; of course the scalar space $L^2(\mu)$ is a particular case of the direct integral, where all spaces $E(\xi)$ are one-dimensional.

Lemma 1.5 says that $\operatorname{Ran} \boldsymbol{B}$ is star-cyclic for U if and only if there is no measurable set A, $\mu(A) > 0$, on which all the functions b_k vanish. If we consider the case when U is star-cyclic, i.e., when it has a star-cyclic vector, we can ask the question:

Does a star-cyclic operator U have a star-cyclic vector that belongs to a prescribed (finite-dimensional) star-cyclic subspace?

The following lemma answers "yes" to that question. Moreover, it implies that if $\operatorname{Ran} \boldsymbol{B}$ is star-cyclic for $U = M_\xi$ on the scalar-valued space $L^2(\mu)$, then almost all vectors $b \in \operatorname{Ran} \boldsymbol{B}$ are star-cyclic for U. As the result is measure-theoretic in nature, we formulate it in a general context.

Lemma 1.6. *Consider a σ-finite scalar-valued measure τ on a measure space $\mathcal{X}$. Let $b_1, b_2, \dots, b_d \in L^2(\tau)$ be such that*

$$\sum_{k=1}^{d} |b_k| \neq 0 \quad \tau\text{-a.e.}$$

Then for almost all (with respect to the Lebesgue measure) $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{C}^d$ we have

$$\sum_{k=1}^{d} \alpha_k b_k \neq 0 \quad \tau\text{-a.e. on } \mathcal{X}.$$

Remark. The above lemma also holds for almost all $\alpha \in \mathbb{R}^d$.

Proof of Lemma 1.6. Consider first the case $\tau(\mathcal{X}) < \infty$.

We proceed by induction in d. Clearly, if $|b_1| \neq 0$ τ-a.e. on $\mathcal{X}$, then $\alpha b_1 \neq 0$ τ-a.e. on $\mathcal{X}$ for all $\alpha \in \mathbb{C} \setminus \{0\}$.

Now assume the statement of the lemma for $d = n$ for some $n \in \mathbb{N}$. Deleting a set of τ-measure 0, we can assume that $\sum_{k=1}^{n+1} |b_k| \neq 0$ on $\mathcal{X}$.

Let $\mathcal{Y} := \{x \in \mathcal{X} : \sum_{k=1}^{n} |b_k(x)| > 0\}$. By the induction assumption, for almost all $\alpha' = (\alpha_1, \alpha_2, \dots, \alpha_n)$

$$b(\alpha', x) := \sum_{k=1}^{n} \alpha_k b_k(x) \neq 0 \quad \text{on } \mathcal{Y}.$$

Fix $\alpha' = (\alpha_1, \alpha_2, \dots, \alpha_n)$ such that $b(\alpha', x) \neq 0$ on $\mathcal{Y}$. We will show that for any such fixed α'

$$\tau\left(\left\{x \in \mathcal{X} : \sum_{k=1}^{n+1} \alpha_k b_k(x) = 0\right\}\right) > 0 \tag{1-9}$$

only for countably many values of α_{n+1}.

To show this, define for $\beta = \alpha_{n+1} \in \mathbb{C}$ the set

$$\mathcal{X}_\beta := \{x \in \mathcal{X} : b(\alpha', x) + \beta b_{n+1}(x) = 0\}.$$

Let $\tilde{\beta} \in \mathbb{C} \setminus \{0\}$, $\tilde{\beta} \neq \beta$. We claim that the sets X_β and $X_{\tilde{\beta}}$ are disjoint.

Indeed, the assumption that $\sum_{k=1}^{n+1} |b_k| > 0$ implies $b_{n+1} \neq 0$ on $\mathcal{X} \setminus \mathcal{Y}$, so $\mathcal{X}_\beta, \mathcal{X}_{\tilde{\beta}} \in \mathcal{Y}$. Moreover, solving for b_{n+1} we get that if $\beta \neq 0$, then

$$\mathcal{X}_\beta = \{x \in \mathcal{Y} : b_{n+1}(x) = -b(\alpha', x)/\beta\},$$

and similarly for $\mathcal{X}_{\tilde{\beta}}$. Since $b(\alpha', x) \neq 0$ on $\mathcal{Y}$, we get

$$b(\alpha', x)/\beta \neq b(\alpha', x)/\tilde{\beta} \quad \text{for all } x \in \mathcal{Y},$$

so if $\beta \neq 0$, then $\mathcal{X}_\beta$ and $\mathcal{X}_{\tilde{\beta}}$ are disjoint as preimages of disjoint sets (points).

If $\beta = 0$, then $\mathcal{X}_0 = \mathcal{X} \setminus \mathcal{Y}$, so the sets $\mathcal{X}_{\tilde{\beta}}$ and $\mathcal{X}_0$ are disjoint.

The set $\mathcal{X}$ has finite measure, and $\mathcal{X}$ is the union of disjoint sets $\mathcal{X}_\beta$, $\beta \in \mathbb{C}$. So, only countably many sets $\mathcal{X}_\beta$ can satisfy $\tau(\mathcal{X}_\beta) > 0$. We have proved the lemma for $\tau(\mathcal{X}) < \infty$.

The rest can be obtained by Tonelli's theorem. Namely, define

$$\mathcal{A} := \left\{(x,\alpha) : x \in \mathcal{X},\ \alpha \in \mathbb{C}^{n+1},\ \sum_{k=1}^{n+1} \alpha_k b_k(x) = 0\right\}$$

and let $F = \mathbf{1}_{\mathcal{A}}$. From Tonelli's theorem we can see that

$$\int \mathbf{1}_{\mathcal{A}}(x,\alpha)\, \mathrm{d}m(\alpha)\, \mathrm{d}\tau(x) > 0 \tag{1-10}$$

if and only if for the set of $\alpha \in \mathbb{C}^{n+1}$ of positive Lebesgue measure

$$\tau\left(\left\{x \in \mathcal{X} : \sum_{k=1}^{n+1} \alpha_k b_k(x) = 0\right\}\right) > 0.$$

It follows from (1-9) that for almost all $\alpha' = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{C}^n$

$$\int \mathbf{1}_{\mathcal{A}}(x, \alpha', \alpha_{n+1})\, \mathrm{d}m(\alpha_{n+1})\, \mathrm{d}\tau(x) = 0,$$

so, by Tonelli, the integral in (1-10) equals 0. □

2. Abstract formula for the adjoint Clark operator

We now introduce necessary known facts about functional models and then give a general abstract formula for the adjoint Clark operator. To do this we need a new notion of coordinate/parametrizing operators for the model and their agreement: the abstract representation formula (Theorem 2.4) holds under the assumption that the coordinate operators $\boldsymbol{C}$ and $\boldsymbol{C}_*$ agree with the Clark model.

Later in Section 3 we construct the coordinate operators that agree with the Clark model, and in Section 4 we compute the characteristic function, so the abstract Theorem 2.4 will give us concrete, albeit complicated, formulas.

2A. *Functional models.*

Definition 2.1. Recall that for a contraction T its *defect operators* D_T and D_{T^*} are defined as

$$D_T := (\boldsymbol{I} - T^*T)^{1/2}, \quad D_{T^*} := (\boldsymbol{I} - TT^*)^{1/2}.$$

The *defect spaces* $\mathfrak{D}_T$ and $\mathfrak{D}_{T^*}$ are defined as

$$\mathfrak{D}_T := \operatorname{clos} \operatorname{Ran} D_T, \quad \mathfrak{D}_{T^*} := \operatorname{clos} \operatorname{Ran} D_{T^*}.$$

The characteristic function is an (explicitly computed from the contraction T) operator-valued function $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$, where $\mathfrak{D}$ and $\mathfrak{D}_*$ are Hilbert spaces of appropriate dimensions,

$$\dim \mathfrak{D} = \dim \mathfrak{D}_T, \quad \dim \mathfrak{D}_* = \dim \mathfrak{D}_{T^*}.$$

Using the characteristic function θ one can then construct the so-called *model space* $\mathcal{K}_\theta$, which is a subspace of a weighted L^2 space $L^2(\mathbb{T}, W; \mathfrak{D}_* \oplus \mathfrak{D}) = L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$ with an operator-valued weight W. The model operator $\mathcal{M}_\theta : \mathcal{K}_\theta \to \mathcal{K}_\theta$ is then defined as the *compression* of the multiplication M_z by the independent variable z,

$$\mathcal{M}_\theta f = P_{\mathcal{K}_\theta} M_z f, \quad f \in \mathcal{K}_\theta;$$

here $M_z f(z) = zf(z)$.

Let us remind the reader that the norm in the weighted space $L^2(\mathbb{T}, W; H)$ with an operator weight W is given by

$$\|f\|^2_{L^2(W;H)} = \int_{\mathbb{T}} (W(z)f(z), f(z))_H \, \mathrm{d}m(z);$$

in the case $\dim H = \infty$ there are some technical details, but in the finite-dimensional case considered in this paper everything is pretty straightforward.

The best-known example of a model is the Sz.-Nagy–Foiaş (transcription of a) model [Sz.-Nagy et al. 2010]. The Sz.-Nagy–Foiaş model space $\mathcal{K}_\theta$ is a subspace of a nonweighted space $L^2(\mathfrak{D}_* \oplus \mathfrak{D})$ (here $W \equiv \boldsymbol{I}$), given by

$$\mathcal{K}_\theta := \begin{pmatrix} H^2(\mathfrak{D}_*) \\ \operatorname{clos} \Delta L^2(\mathfrak{D}) \end{pmatrix} \ominus \begin{pmatrix} \theta \\ \Delta \end{pmatrix} H^2(\mathfrak{D}),$$

where

$$\Delta(z) := (\boldsymbol{I}_{\mathfrak{D}} - \theta^*(z)\theta(z))^{1/2} \quad \text{and} \quad \begin{pmatrix} \theta \\ \Delta \end{pmatrix} H^2(\mathfrak{D}) = \left\{ \begin{pmatrix} \theta f \\ \Delta f \end{pmatrix} : f \in H^2(\mathfrak{D}) \right\}.$$

In the literature, the case when the vector-valued characteristic function θ is *inner* (i.e., its boundary values are isometries for a.e. $z \in \mathbb{T}$) is often considered. Then $\Delta(z) = \mathbf{0}$ on $\mathbb{T}$, so in that case the second component of $\mathcal{K}_\theta$ collapses completely and the Sz.-Nagy–Foiaş model space reduces to the familiar space

$$\mathcal{K}_\theta = H^2(\mathfrak{D}_*) \ominus \theta H^2(\mathfrak{D}).$$

Also, in the literature, see [Sz.-Nagy et al. 2010], the characteristic function is defined up to multiplication by constant unitary factors from the right and from the left. Namely, two functions $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$ and $\tilde{\theta} \in H^\infty(\widetilde{\mathfrak{D}} \to \widetilde{\mathfrak{D}}_*)$ are equivalent if there exist unitary operators $U : \mathfrak{D} \to \widetilde{\mathfrak{D}}$ and $U_* : \mathfrak{D}_* \to \widetilde{\mathfrak{D}}_*$ such that $\tilde{\theta} = U_* \theta U^*$.

It is a well-known fact, see [Sz.-Nagy et al. 2010], that two c.n.u. contractions are unitarily equivalent if and only if their characteristic functions are equivalent as described above. So, usually in the literature the characteristic function is understood as the corresponding equivalence class, or an arbitrary representative in this class. However, in this paper, to get correct formulas it is essential to track which representative is chosen.

2B. ***Coordinate operators, parametrizing operators, and their agreement.*** Let $T : H \to H$ be a contraction, and let $\mathfrak{D}$, $\mathfrak{D}_*$ be Hilbert spaces, $\dim \mathfrak{D} = \dim \mathfrak{D}_T$, $\dim \mathfrak{D}_* = \dim \mathfrak{D}_{T^*}$. Unitary operators $V : \mathfrak{D}_T \to \mathfrak{D}$ and $V_* : \mathfrak{D}_{T^*} \to \mathfrak{D}_*$ will be called *coordinate operators* for the corresponding defect spaces; the reason for that name is that often spaces $\mathfrak{D}$ and $\mathfrak{D}_*$ are spaces with a fixed orthonormal basis (and one can introduce coordinates there), so the operators introduce coordinates on the defect spaces.

The inverse operators $V^* : \mathfrak{D} \to \mathfrak{D}_T$ and $V_*^* : \mathfrak{D}_* \to \mathfrak{D}_{T^*}$ will be called *parametrizing* operators. For a contraction T we will use the symbols V and V_* for the coordinate operators, but for its model $\mathcal{M}_\theta$ the parametrizing operators will be used, and we reserve letters $\boldsymbol{C}$ and $\boldsymbol{C}_*$ for these operators.

Let T be a c.n.u. contraction with characteristic function $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$, and let $\mathcal{M}_\theta : \mathcal{K}_\theta \to \mathcal{K}_\theta$ be its model. Let also $V : \mathfrak{D}_T \to \mathfrak{D}$ and $V_* : \mathfrak{D}_{T^*} \to \mathfrak{D}_*$ be coordinate operators for the defect spaces of T, and $\boldsymbol{C} : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta}$ and $\boldsymbol{C}_* : \mathfrak{D}_* \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ be the parametrizing operators for the defect spaces of $\mathcal{M}_\theta$ (this simply means that all four operators are unitary).

We say that the operators V, V_* agree with operators $\boldsymbol{C}$, $\boldsymbol{C}_*$ if there exists a unitary operator $\Phi : \mathcal{K}_\theta \to H$ intertwining T and $\mathcal{M}_\theta$,

$$T\Phi = \Phi\mathcal{M}_\theta,$$

and such that

$$\boldsymbol{C}^* = V\Phi|_{\mathfrak{D}_{\mathcal{M}_\theta}}, \quad \boldsymbol{C}_*^* = V_*\Phi|_{\mathfrak{D}_{\mathcal{M}_\theta^*}}. \tag{2-1}$$

The above identities simply mean that the diagrams below are commutative:

$$\begin{tikzcd}
D_T \arrow[rr, "V"] & & \mathfrak{D} \\
& \mathfrak{D}_{\mathcal{M}_\theta} \arrow[ul, "\Phi"] \arrow[ur, "\boldsymbol{C}^*"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathfrak{D}_* \arrow[rr, "V_*^*"] & & D_{T^*} \\
& \mathfrak{D}_{\mathcal{M}_\theta^*} \arrow[ul, "\boldsymbol{C}_*^*"] \arrow[ur, "\Phi"'] &
\end{tikzcd}$$

In this paper, when convenient, we always extend an operator between subspaces to the operator between the whole spaces, by extending it by 0 on the orthogonal complement of the domain; slightly abusing notation we will use the same symbol for both operators. Thus a unitary operator between subspaces E and F can be treated as a partial isometry with initial space E and final space F, and vice versa. With this convention (2-1) can be rewritten as

$$\boldsymbol{C}^* = V\Phi, \quad \boldsymbol{C}_*^* = V_*\Phi.$$

2C. *Clark operator.* Consider a contraction T given by (1-2) with Γ being a strict contraction. We also assume that Ran $\boldsymbol{B}$ is a star-cyclic subspace for U, so T is a c.n.u. contraction; see Lemma 1.4.

We assume that U is given in its spectral representation, so U is the multiplication operator M_ξ in the direct integral $\mathcal{H}$.

A Clark operator $\Phi : \mathcal{K}_\theta \to \mathcal{H}$ is a unitary operator, intertwining this special contraction T and its model $\mathcal{M}_\theta$, $\Phi\mathcal{M}_\theta = T\Phi$, or equivalently

$$\Phi^* T = \mathcal{M}_\theta \Phi^*. \tag{2-2}$$

We name it so after D. Clark, who in [Clark 1972] described it for rank-one perturbations of unitary operators with purely singular spectrum.

We want to describe the operator Φ (more precisely, its adjoint Φ^*) in our situation. In our case, $\dim \mathfrak{D}_T = \dim \mathfrak{D}_{T^*} = d$, and it will be convenient for us to consider models with $\mathfrak{D} = \mathfrak{D}_* = \mathbb{C}^d$.

As discussed above, it can be easily seen from the representation (1-3) that the operators $U^*\boldsymbol{B} : \mathfrak{D} = \mathbb{C}^d \to \mathfrak{D}_T$ and $\boldsymbol{B} : \mathfrak{D} = \mathbb{C}^d \to \mathfrak{D}_{T^*}$ are unitary operators canonically (for our setup) identifying $\mathfrak{D}$ with the corresponding defect spaces, i.e., the canonical parametrizing operators for these spaces. The corresponding coordinate operators are given by $V = \boldsymbol{B}^*U$, $V_* = \boldsymbol{B}^*$.

We say that parametrizing operators $\boldsymbol{C} : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta}$, $\boldsymbol{C}_* : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ *agree* with the Clark model if the above coordinate operators $V = \boldsymbol{B}^*U$, $V_* = \boldsymbol{B}^*$ agree with the parametrizing operators $\boldsymbol{C}$, $\boldsymbol{C}_*$ in the sense of Section 2B. In other words, they agree if there exists a Clark operator Φ such that the following diagram commutes:

$$\begin{tikzcd}
\mathfrak{D}_T \arrow[rr, "\boldsymbol{B}^*U"] \arrow[dr, "\Phi^*"'] & & \mathfrak{D} = \mathbb{C}^d \arrow[rr, "\boldsymbol{B}"] \arrow[dl, "\boldsymbol{C}"] \arrow[dr, "\boldsymbol{C}_*"'] & & \mathfrak{D}_{T^*} \arrow[dl, "\Phi^*"] \\
& \mathfrak{D}_{\mathcal{M}_\theta} & & \mathfrak{D}_{\mathcal{M}_\theta^*} &
\end{tikzcd} \tag{2-3}$$

Note, that in this diagram one can travel in both directions: to change the direction, one just needs to take the adjoint of the corresponding operator.

Slightly abusing notation, we use $\boldsymbol{C}$ to also denote the extension of $\boldsymbol{C}$ to the model space $\mathcal{K}_\theta$ by the zero operator, and similarly for $\boldsymbol{C}_*$.

Note that agreement of $\boldsymbol{C}$ and $\boldsymbol{C}_*$ with the Clark model can be rewritten as

$$\Phi^*(\boldsymbol{B}^*U)^* = \boldsymbol{C}, \quad \Phi^*\boldsymbol{B} = \boldsymbol{C}_*. \tag{2-4}$$

And by taking restrictions (where necessary) we find

$$\mathcal{M}_\theta \boldsymbol{C} = \boldsymbol{C}_*\Gamma \quad \text{and} \quad \mathcal{M}_\theta^* \boldsymbol{C}_* = \boldsymbol{C}\Gamma^*. \tag{2-5}$$

We express the action of the model operator and its adjoint in an auxiliary result. The result holds in any transcription of the model. We will need the following simple fact.

Lemma 2.2. *For a contraction T*

$$T\mathfrak{D}_T \subset \mathfrak{D}_{T^*}, \quad T^*\mathfrak{D}_{T^*} \subset \mathfrak{D}_T.$$

Proof. Since D_T is a strict contraction on $\mathfrak{D}_T$ we get

$$\|Tx\| = \|x\| \quad \Longleftrightarrow \quad x \perp \mathfrak{D}_T,$$

and similarly, since T^* is a strict contraction on $\mathfrak{D}_{T^*}$,

$$\|T^*x\| = \|x\| \quad \Longleftrightarrow \quad x \perp \mathfrak{D}_{T^*}. \tag{2-6}$$

Thus the operator T is an isometry on $\mathfrak{D}_T^\perp$, so the polarization identity implies $T^*Tx = x$ for all $x \in \mathfrak{D}_T^\perp$. Together with (2-6) this implies $T(\mathfrak{D}_T^\perp) \subset \mathfrak{D}_{T^*}^\perp$, which is equivalent to the inclusion $T^*\mathfrak{D}_{T^*} \subset \mathfrak{D}_T$.

Replacing T by T^* we get $T\mathfrak{D}_T \subset \mathfrak{D}_{T^*}$. □

Lemma 2.3. *Let T be as defined in* (1-2) *with Γ being a strict contraction. Assume also that* Ran $\boldsymbol{B}$ *is star-cyclic* (*so T is completely nonunitary*; *see Lemma 1.4*).

Let $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$, $\mathfrak{D} = \mathfrak{D}_* = \mathbb{C}^d$, be the characteristic function of T, and let $\mathcal{M}_\theta : \mathcal{K}_\theta \to \mathcal{K}_\theta$ be a model operator. Let $\boldsymbol{C} : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta}$ and $\boldsymbol{C}_* : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ be the parametrizing unitary operators that agree with the Clark model.

Then

$$\mathcal{M}_\theta = \boldsymbol{M}_z + (\boldsymbol{C}_*\Gamma - \boldsymbol{M}_z\boldsymbol{C})\boldsymbol{C}^* \quad \textit{and} \quad \mathcal{M}_\theta^* = \boldsymbol{M}_{\bar z} + (\boldsymbol{C}\Gamma^* - \boldsymbol{M}_{\bar z}\boldsymbol{C}_*)\boldsymbol{C}_*^*.$$

Proof. Since the operator $\mathcal{M}_\theta$ acts on $\mathcal{K}_\theta \ominus \mathfrak{D}_{\mathcal{M}_\theta}$ as the multiplication operator $\boldsymbol{M}_z$, we can trivially write

$$\mathcal{M}_\theta = \boldsymbol{M}_z(\boldsymbol{I} - P_{\mathfrak{D}_{\mathcal{M}_\theta}}) + \mathcal{M}_\theta P_{\mathfrak{D}_{\mathcal{M}_\theta}}.$$

Recalling that $\boldsymbol{C} : \mathfrak{D} \to \mathcal{K}_\theta$ is an isometry with range $\mathfrak{D}_{\mathcal{M}_\theta}$, we can see that $P_{\mathfrak{D}_{\mathcal{M}_\theta}} = \boldsymbol{C}\boldsymbol{C}^*$, so

$$\boldsymbol{M}_z(\boldsymbol{I} - P_{\mathfrak{D}_{\mathcal{M}_\theta}}) = \boldsymbol{M}_z(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{C}^*). \tag{2-7}$$

Using the identity $P_{\mathfrak{D}_{\mathcal{M}_\theta}} = \boldsymbol{C}\boldsymbol{C}^*$ and the first equation of (2-5) we get

$$\mathcal{M}_\theta P_{\mathfrak{D}_{\mathcal{M}_\theta}} = \mathcal{M}_\theta \boldsymbol{C}\boldsymbol{C}^* = \boldsymbol{C}_*\Gamma\boldsymbol{C}^*,$$

which together with (2-7) gives us the desired formula for $\mathcal{M}_\theta$.

To get the formula for $\mathcal{M}_\theta^*$ we represent it as

$$\mathcal{M}_\theta^* = \boldsymbol{M}_{\bar z}(\boldsymbol{I} - P_{\mathfrak{D}_{\mathcal{M}_\theta^*}}) + \mathcal{M}_\theta^* P_{\mathfrak{D}_{\mathcal{M}_\theta^*}}.$$

Using the identities

$$P_{\mathfrak{D}_{\mathcal{M}_\theta^*}} = \boldsymbol{C}_*\boldsymbol{C}_*^*, \quad \mathcal{M}_\theta^* P_{\mathfrak{D}_{\mathcal{M}_\theta^*}} = \boldsymbol{C}\Gamma^*\boldsymbol{C}_*^*$$

(the first holds because $\mathfrak{D}_{\mathcal{M}_\theta^*}$ is the range of the isometry $\boldsymbol{C}_*$, and the second one follows from the second equation in (2-5)), we get the desired formula. □

2D. ***Representation theorem.*** For a (general) model operator $\mathcal{M}_\theta$, $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$, the parametrizing operators $\boldsymbol{C} : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta}$, $\boldsymbol{C}_* : \mathfrak{D}_* \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ give rise to (uniquely defined) operator-valued functions C and C_* on $\mathbb{T}$, where $C(\xi) : \mathfrak{D} \to \mathfrak{D} \oplus \mathfrak{D}_*$, $C_*(\xi) : \mathfrak{D}_* \to \mathfrak{D} \oplus \mathfrak{D}_*$, $\xi \in \mathbb{T}$, such that for almost all $\xi \in \mathbb{T}$

$$(\boldsymbol{C}e)(\xi) = C(\xi)e \qquad \text{for all } e \in \mathfrak{D}, \tag{2-8}$$

$$(\boldsymbol{C}_*e_*)(\xi) = C_*(\xi)e_* \quad \text{for all } e_* \in \mathfrak{D}_*; \tag{2-9}$$

here $\boldsymbol{C}e, \boldsymbol{C}_*e_*$ are elements of $\mathcal{K}_\theta$, i.e., functions with values in $\mathfrak{D} \oplus \mathfrak{D}_*$, and $(\boldsymbol{C}e)(\xi)$, $(\boldsymbol{C}_*e_*)(\xi)$ are the values of these functions at $\xi \in \mathbb{T}$.

If we fix orthonormal bases in $\mathfrak{D}$ and $\mathfrak{D}_*$, then the k-th column of the matrix of $C(\xi)$ is defined as $(\boldsymbol{C}_*e_k)(\xi)$, where e_k is the k-th vector in the orthonormal basis in $\mathfrak{D}$, and similarly for C_*.

If $\mathcal{M}_\theta$ is a model for a contraction $T = T_\Gamma$ with Γ being a strict contraction on $\mathfrak{D} = \mathbb{C}^d$, we can see from (1-3) that $\dim \mathfrak{D}_T = \dim \mathfrak{D}_{T^*} = d$, so we can always pick a characteristic function $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$ (i.e., with $\mathfrak{D}_* = \mathfrak{D} = \mathbb{C}^d$).

The following formula for the adjoint Φ^* of the Clark operator Φ generalizes the "universal" representation theorem [Liaw and Treil 2016, Theorem 3.1] to higher-rank perturbations.

Theorem 2.4 (representation theorem). *Let T be as defined in* (1-2) *with Γ being a strict contraction and $U = M_\xi$ in $\mathcal{H} \subset L^2(\mu; E)$. Let $\theta = \theta_T$ be a characteristic function of T, and let $\mathcal{K}_\theta$ and $\mathcal{M}_\theta$ be the corresponding model space and model operator.*

Let $\boldsymbol{C} : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta}$ and $\boldsymbol{C}_* : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ be the parametrizing unitary operators[1] *that agree with Clark model*, *i.e.*, *such that* (2-4) *is satisfied for some Clark operator Φ. And let $C(z)$ and $C_*(z)$ be given by* (2-8) *and* (2-9), *respectively.*

Then the action of the adjoint Clark operator Φ^* is given by

$$(\Phi^* hb)(z) = h(z)C_*(z)\boldsymbol{B}^* b + (C_*(z) - zC(z)) \int_{\mathbb{T}} \frac{h(\xi) - h(z)}{1 - z\bar{\xi}} B^*(\xi) b(\xi)\, \mathrm{d}\mu(\xi) \tag{2-10}$$

for any $b \in \operatorname{Ran} \boldsymbol{B}$ and for all $h \in C^1(\mathbb{T})$; here

$$B^*(\xi) = \begin{pmatrix} b_1(\xi)^* \\ b_2(\xi)^* \\ \vdots \\ b_d(\xi)^* \end{pmatrix}$$

and $\boldsymbol{B}^* b = \int_{\mathbb{T}} B^*(\xi) b(\xi)\, \mathrm{d}\mu(\xi)$, as explained more thoroughly in the proof below.

Remark. The above theorem looks like abstract nonsense because right now it is not clear how to find the parametrizing operators $\boldsymbol{C}$ and $\boldsymbol{C}_*$ that agree with the Clark model. However, Theorem 4.2 below gives an explicit formula for the characteristic function θ (one of the representative in the equivalence class), and Lemma 3.3 gives explicit formulas for $\boldsymbol{C}$ and $\boldsymbol{C}_*$ in the Sz.-Nagy–Foiaş transcription that agree with the Clark model for our θ.

When $d = 1$ this formula agrees with the special case of the representation formula derived in [Liaw and Treil 2016]. While some of the ideas of the following proof were originally developed there, the current extension to rank-d perturbations requires several new ideas and a more abstract way of thinking.

Proof of Theorem 2.4. Recall that $U = M_\xi$, so $T = M_\xi + \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^* M_\xi$. The intertwining relation $\Phi^* T = \mathcal{M}_\theta \Phi^*$ then can be rewritten as

$$\Phi^* M_\xi + \Phi^* \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^* U = \Phi^* T = \mathcal{M}_\theta \Phi^* = [M_z + (\boldsymbol{C}_*\Gamma - M_z \boldsymbol{C})\boldsymbol{C}^*]\Phi^*; \tag{2-11}$$

here we used Lemma 2.3 to express the model operator in the right-hand side of (2-11).

By the commutation relations in (2-4), the term $\Phi^* \boldsymbol{B}\Gamma \boldsymbol{B}^* U$ on the left-hand side of (2-11) cancels with the term $\boldsymbol{C}_*\Gamma \boldsymbol{C}^*\Phi^*$ on the right-hand side of (2-11). Then (2-11) can be rewritten as

$$\begin{aligned} \Phi^* M_\xi &= M_z \Phi^* + \Phi^* \boldsymbol{B}\boldsymbol{I}_{\mathbb{C}^d} \boldsymbol{B}^* U - M_z \boldsymbol{C}\boldsymbol{C}^* \Phi^* \\ &= M_z \Phi^* + (\boldsymbol{C}_* - M_z \boldsymbol{C})\boldsymbol{B}^* M_\xi; \end{aligned} \tag{2-12}$$

the last identity holds because, by (2-4), we have $\Phi^* \boldsymbol{B} = \boldsymbol{C}_*$ and $\boldsymbol{C}^*\Phi^* = \boldsymbol{B}^* U = \boldsymbol{B}^* M_\xi$.

[1] Note that here we set $\mathfrak{D}_* = \mathfrak{D}$, which is possible because the dimensions of the defect spaces are equal.

Right-multiplying (2-12) by M_ξ and using (2-12) we get

$$\begin{aligned}\Phi^* M_\xi^2 &= M_z \Phi^* M_\xi + (\boldsymbol{C}_* - M_z \boldsymbol{C}) \boldsymbol{B}^* M_\xi^2 \\ &= M_z^2 \Phi^* + M_z (\boldsymbol{C}_* - M_z \boldsymbol{C}) \boldsymbol{B}^* M_\xi + (\boldsymbol{C}_* - M_z \boldsymbol{C}) \boldsymbol{B}^* M_\xi^2.\end{aligned}$$

Right-multiplying the above equation by M_ξ and using (2-12) again we get the identity

$$\Phi^* M_\xi^n = M_z^n \Phi^* + \sum_{k=1}^{n} M_z^{k-1} (\boldsymbol{C}_* - M_z \boldsymbol{C}) \boldsymbol{B}^* M_\xi^{n-k+1}, \tag{2-13}$$

with $n = 3$. Right-multiplying by M_ξ and applying (2-12) we get by induction that (2-13) holds for all $n \geq 0$. (The case $n = 0$ trivially reads as $\Phi^* = \Phi^*$, and (2-12) is precisely the case $n = 1$.)

We now apply (2-13) to some $b \in \operatorname{Ran} \boldsymbol{B}$. By commutative diagram (2-3) we get $\Phi^* b = \boldsymbol{C}_* \boldsymbol{B}^* b$, i.e., $(\Phi^* b)(z) = C_*(z) \boldsymbol{B}^* b$. Using this identity we get

$$\begin{aligned}(\Phi^* M_\xi^n b)(z) &= z^n (\Phi^* b)(z) + \sum_{k=1}^{n} z^{k-1} (C_*(z) - zC(z)) \boldsymbol{B}^* M_\xi^{n-k+1} b \\ &= z^n C_*(z) (\boldsymbol{B}^* b)(z) + (C_*(z) - zC(z)) \sum_{k=1}^{n} z^{k-1} \boldsymbol{B}^* M_\xi^{n-k+1} b.\end{aligned} \tag{2-14}$$

To continue, we recall that $\boldsymbol{B} : \mathbb{C}^d \to L^2(\mu; E)$ acts as multiplication by the matrix $B(\xi) = (b_1(\xi), b_2(\xi), \ldots, b_d(\xi))$, so its adjoint $\boldsymbol{B}^* : \mathcal{H} \subset L^2(\mu; E) \to \mathbb{C}^d$ is given by

$$\boldsymbol{B}^* f = \int_{\mathbb{T}} B^*(\xi) f(\xi) \, \mathrm{d}\mu(\xi) \quad \text{for } f \in \mathcal{H},$$

where the integral can be expanded as

$$\int_{\mathbb{T}} B^*(\xi) f(\xi) \, \mathrm{d}\mu(\xi) = \begin{pmatrix} \int_{\mathbb{T}} b_1(\xi)^* f(\xi) \, \mathrm{d}\mu(\xi) \\ \int_{\mathbb{T}} b_2(\xi)^* f(\xi) \, \mathrm{d}\mu(\xi) \\ \vdots \\ \int_{\mathbb{T}} b_d(\xi)^* f(\xi) \, \mathrm{d}\mu(\xi) \end{pmatrix}.$$

Using the sum of geometric progression formula we evaluate the sum in (2-14) to

$$\begin{aligned}\sum_{k=1}^{n} z^{k-1} \boldsymbol{B}^* M_\xi^{n-k+1} b &= \sum_{k=1}^{n} z^{k-1} \int_{\mathbb{T}} \xi^{n-k+1} B^*(\xi) b(\xi) \, \mathrm{d}\mu(\xi) \\ &= \int_{\mathbb{T}} \sum_{k=1}^{n} z^{k-1} \xi^{n-k+1} B^*(\xi) b(\xi) \, \mathrm{d}\mu(\xi) \\ &= \int_{\mathbb{T}} \frac{\xi^n - z^n}{1 - z\bar{\xi}} B^*(\xi) b(\xi) \, \mathrm{d}\mu(\xi).\end{aligned} \tag{2-15}$$

Thus, we have proved (2-10) for monomials $h(\xi) = \xi^n$, $n \geq 0$. And by the linearity of Φ^*, the representation (2-10) holds for (analytic) polynomials h in ξ.

The argument leading to the determination of the action of Φ^* on polynomials h in $\bar{\xi}$ is similar. But we found that the devil is in the details and therefore decided to include much of the argument.

First observe that the intertwining relation (2-2) is equivalent to $\mathcal{M}_\theta^*\Phi^* = \Phi^* T^*$. Recalling $T^* = U^* + U^*\boldsymbol{B}(\Gamma^* - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^*$ and the resolution of the adjoint model operator $\mathcal{M}_\theta^*$ (see second statement of Lemma 2.3), we obtain

$$M_{\bar{z}}\Phi^* + (\boldsymbol{C}\Gamma^* - \bar{z}\boldsymbol{C}_*)\boldsymbol{C}_*^*\Phi^* = \mathcal{M}_\theta^*\Phi^* = \Phi^* T^* = \Phi^* U^* - \Phi^* U^* \boldsymbol{B}(\Gamma^* - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^*.$$

The terms involving Γ^* on the left-hand side and the right-hand side cancel by the commutation relations in (2-4) (actually by their adjoints). Now, rearrangement and another application of the adjoints of the commutation relations in (2-4) yields

$$\begin{aligned}\Phi^* M_{\bar{\xi}} &= \Phi^* U^* = M_{\bar{z}}\Phi^* + \Phi^* U^* \boldsymbol{B}\boldsymbol{I}_{\mathbb{C}^d}\boldsymbol{B}^* - \bar{z}\boldsymbol{C}_*\boldsymbol{C}_*^*\Phi^* = M_{\bar{z}}\Phi^* + (\boldsymbol{C} - M_{\bar{z}}\boldsymbol{C}_*)\boldsymbol{B}^* \\ &= M_{\bar{z}}\Phi^* + M_{\bar{z}}(M_z\boldsymbol{C} - \boldsymbol{C}_*)\boldsymbol{B}^*. \end{aligned} \tag{2-16}$$

In analogy to the above, we right-multiply (2-16) by $M_{\bar{\xi}}$ and apply (2-16) twice to obtain

$$\Phi^* M_{\bar{\xi}}^2 = M_{\bar{z}}^2\Phi^* + \sum_{k=1}^{2} M_{\bar{z}}^k (M_z\boldsymbol{C} - \boldsymbol{C}_*)\boldsymbol{B}^* M_{\bar{\xi}}^{2-k}.$$

Inductively, we conclude

$$\Phi^* M_{\bar{\xi}}^n = M_{\bar{z}}^n\Phi^* - \sum_{k=1}^{n} M_{\bar{z}}^k (\boldsymbol{C}_* - M_z\boldsymbol{C})\boldsymbol{B}^* M_{\bar{\xi}}^{n-k},$$

which *differs in the exponents and in the sign from its counterpart expression* in (2-13).

Through an application of this identity to b and by the commutative diagram (2-3), we see

$$\begin{aligned}(\Phi^* M_{\bar{\xi}}^n b)(z) &= \bar{z}^n(\Phi^* b)(z) - \sum_{k=1}^{n} \bar{z}^k (C_*(z) - zC(z))\boldsymbol{B}^* M_{\bar{\xi}}^{n-k} b \\ &= \bar{z}^n C_*(z)(\boldsymbol{B}^* b)(z) - (C_*(z) - zC(z)) \sum_{k=1}^{n} \bar{z}^k \boldsymbol{B}^* M_{\bar{\xi}}^{n-k} b.\end{aligned}$$

As in (2-15), but here with the geometric progression

$$-\sum_{k=1}^{n} (\bar{z})^k (\bar{\xi})^{n-k} = \frac{(\bar{\xi})^n - (\bar{z})^n}{1 - \bar{\xi} z},$$

we can see (2-10) for monomials $\bar{\xi}^n$, $n \in \mathbb{N}$. And by the linearity of Φ^*, we obtain the *same formula* (2-10) for functions h that are polynomials in $\bar{\xi}$.

We have proved (2-10) for trigonometric polynomials f. The theorem now follows by a standard approximation argument, developed in [Liaw and Treil 2009]. The application of this argument to the current situation is a slight extension of the one used in [Liaw and Treil 2016]. Fix $f \in C^1(\mathbb{T})$ and let $\{p_k\}$ be a sequence of trigonometric polynomials with uniform-on-$\mathbb{T}$ approximations $p_k \rightrightarrows f$ and $p_k' \rightrightarrows f'$. In particular, we have $|p_k'|$ is bounded (with bound independent of k) and $p_k \to f$ as well as

$p_k b \to f b$ in $L^2(\mu; E)$. Since Φ^* is a unitary operator, it is bounded and therefore we have convergence on the left-hand side $\Phi^* p_k b \to \Phi^* f b$ in $\mathcal{K}_\theta$.

To investigate convergence on the right-hand side, first recall that the model space is a subspace of the weighted space $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$.

So convergence of the first term on the right-hand side happens, since $p_k \rightrightarrows f$ and the operator norm $\|\boldsymbol{C}_* \boldsymbol{B}^*\| = 1$ implies $p_k C_*(z) \boldsymbol{B}^* b = p_k \boldsymbol{C}_* \boldsymbol{B}^* b \to f \boldsymbol{C}_* \boldsymbol{B}^* b = f C_*(z) \boldsymbol{B}^* b$ in $\mathcal{K}_\theta$.

Lastly, to see convergence of the second term on the right-hand side, consider auxiliary functions $f_k := f - p_k$. We have $f_k \rightrightarrows 0$ and $f_k' \rightrightarrows 0$. Let $I_{\xi,z} \subset \mathbb{T}$ denote the shortest arc connecting ξ and z. Then by the intermediate value theorem

$$|f_k(\xi) - f_k(z)| \le \|f_k'\|_\infty |I_{\xi,z}| \quad \text{for all } \xi, z \in \mathbb{T}.$$

By virtue of the geometric estimate $|I_{\xi,z}| \le \frac{\pi}{2}|\xi - z|$, we obtain

$$\left| \frac{f_k(\xi) - f_k(z)}{1 - \bar{\xi} z} \right| \le \tfrac{\pi}{2} \|f_k'\|_\infty \to 0 \quad \text{as } k \to \infty.$$

And since $\boldsymbol{B}^*$ is bounded as a partial isometry, we conclude the componentwise uniform convergence

$$\int \frac{p_k(\xi) - p_k(z)}{1 - \bar{\xi} z} B^*(\xi) b(\xi) \, \mathrm{d}\mu(\xi) \ \rightrightarrows \ \int \frac{f(\xi) - f(z)}{1 - \bar{\xi} z} B^*(\xi) b(\xi) \, \mathrm{d}\mu(\xi), \quad z \in \mathbb{T}.$$

By Lemma 3.4 below, the functions $W^{1/2} C$ and $W^{1/2} C_*$ are bounded, and so is the function $W^{1/2} C_1$, $C_1(z) := C_*(z) - z C(z)$. That means the multiplication operator $f \mapsto C_1 f$ is a bounded operator from $L^2(\mathfrak{D})$ to $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$ (recall that in our case $\mathfrak{D} = \mathfrak{D}_*$ and we use $\mathfrak{D}_*$ here only for consistency with the general model notation).

The uniform convergence implies the convergence in $L^2(\mathfrak{D})$, so the boundedness of the multiplication by C_1 implies the convergence in norm in the second term in the right-hand side of (2-10) (in the norm of $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$). □

3. Model and agreement of operators

We want to explain how to get operators $\boldsymbol{C}$ and $\boldsymbol{C}_*$ that agree with each other.

To do that we need to understand in more detail how the model is constructed, and what operator gives the unitary equivalence of the function and its model.

Everything starts with the notion of unitary dilation. Recall that for a contraction T in a Hilbert space H its unitary dilation is a unitary operator $\mathcal{U}$ on a bigger space $\mathcal{H}$, $H \subset \mathcal{H}$, such that for all $n \ge 0$

$$T^n = P_H \mathcal{U}^n|_H. \tag{3-1}$$

Taking the adjoint of this identity we immediately get

$$(T^*)^n = P_H \mathcal{U}^{-n}|_H. \tag{3-2}$$

A dilation is called *minimal* if it is impossible to replace $\mathcal{U}$ by its restriction to a reducing subspace and still have the identities (3-1) and (3-2).

The structure of minimal unitary dilations is well known.

Theorem 3.1 [Nikolski and Vasyunin 1998, Theorem 1.4; Nikolski 2002b, Theorem 1.1.16]. *Let $\mathcal{U}: \mathcal{H} \to \mathcal{H}$ be a minimal unitary dilation of a contraction T. Then $\mathcal{H}$ can be decomposed as $\mathcal{H} = G_* \oplus H \oplus G$, and with respect to this decomposition $\mathcal{U}$ can be represented as*

$$\mathcal{U} = \begin{pmatrix} \mathcal{E}_*^* & 0 & 0 \\ D_{T^*} V_*^* & T & 0 \\ -VT^*V_*^* & VD_T & \mathcal{E} \end{pmatrix}, \tag{3-3}$$

where $\mathcal{E}: G \to G$ and $\mathcal{E}_*: G_* \to G_*$ are pure isometries, V is a partial isometry with initial space $\mathfrak{D}_T$ and final space $\ker \mathcal{E}^*$ *and V_* is a partial isometry with initial space $\mathfrak{D}_{T^*}$ and final space* $\ker \mathcal{E}_*^*$.

Moreover, any minimal unitary dilation of T can be obtained this way. Namely if we pick auxiliary Hilbert spaces G and G_* and isometries $\mathcal{E}$ and $\mathcal{E}_*$ there with $\dim\ker \mathcal{E}^* = \dim \mathfrak{D}_T$, $\dim\ker\mathcal{E}_*^* = \dim \mathfrak{D}_{T^*}$ *and then pick arbitrary partial isometries V and V_* with initial and final spaces as above, then* (3-3) *will give us a minimal unitary dilation of T.*

The construction of the model then goes as follows. We take auxiliary Hilbert spaces $\mathfrak{D}$ and $\mathfrak{D}_*$, $\dim \mathfrak{D} = \dim \mathfrak{D}_T$, $\dim \mathfrak{D}_* = \dim \mathfrak{D}_{T^*}$, and construct operators $\mathcal{E}$ and $\mathcal{E}_*$ such that $\ker \mathcal{E}^* = \mathfrak{D}$, $\ker \mathcal{E}_*^* = \mathfrak{D}_*$. We can do that by putting $G = \ell^2(\mathfrak{D}) = \ell^2(\mathbb{Z}_+; \mathfrak{D})$, and defining

$$\mathcal{E}(x_0, x_1, x_2, \dots) = (0, x_0, x_1, x_2, \dots), \quad x_k \in \mathfrak{D},$$

and similarly for $\mathcal{E}_*$.

Picking arbitrary partial isometries V and V_* with initial and final spaces as in (3-3), we get a minimal unitary dilation U of T given by (3-3).

Remark. Above, we were speaking a bit informally, identifying $x \in \mathfrak{D}$ with the sequence $(x, 0, 0, 0, \dots) \in \ell^2(\mathfrak{D})$, and $x_* \in \mathfrak{D}_*$ with $(x_*, 0, 0, 0, \dots) \in \ell^2(\mathfrak{D})$.

To be absolutely formal, we need to define canonical embeddings $\boldsymbol{e}: \mathfrak{D} \to G = \ell^2(\mathfrak{D})$, $\boldsymbol{e}_*: \mathfrak{D}_* \to G_* = \ell^2(\mathfrak{D}_*)$ with

$$\boldsymbol{e}(x) := (x, 0, 0, 0, \dots), \quad x \in \mathfrak{D}, \tag{3-4}$$

$$\boldsymbol{e}_*(x_*) := (x_*, 0, 0, 0, \dots), \quad x \in \mathfrak{D}_*. \tag{3-5}$$

Then, picking arbitrary unitary operators $V: \mathfrak{D}_T \to \mathfrak{D}$, $V_*: \mathfrak{D}_{T^*} \to \mathfrak{D}_*$, we rewrite (3-3) to define the corresponding unitary dilation as

$$\mathcal{U} = \begin{pmatrix} \mathcal{E}_*^* & 0 & 0 \\ D_{T^*} V_*^* \boldsymbol{e}_*^* & T & 0 \\ -\boldsymbol{e} VT^*V_*^* \boldsymbol{e}_*^* & \boldsymbol{e} VD_T & \mathcal{E} \end{pmatrix}. \tag{3-6}$$

The reason for being so formal is that if $\dim \mathfrak{D}_T = \dim \mathfrak{D}_{T^*}$ it is often convenient to put $\mathfrak{D} = \mathfrak{D}_*$, but we definitely want to be able to distinguish between the cases when $\mathfrak{D}$ is identified with $\ker \mathcal{E}$ and when it is identified with $\ker \mathcal{E}_*$.

We then define *functional embeddings* $\pi : L^2(\mathfrak{D}) \to \mathcal{H}$ and $\pi_* : L^2(\mathfrak{D}_*) \to \mathcal{H}$ by

$$\pi\Big(\sum_{k\in\mathbb{Z}} z^k e_k\Big) = \sum_{k\in\mathbb{Z}} \mathcal{U}^k \boldsymbol{e}(e_k), \qquad e_k \in \mathfrak{D},$$
$$\pi_*\Big(\sum_{k\in\mathbb{Z}} z^k e_k\Big) = \sum_{k\in\mathbb{Z}} \mathcal{U}^{k+1} \boldsymbol{e}_*(e_k), \quad e_k \in \mathfrak{D}_*.$$

We refer the reader to [Nikolski and Vasyunin 1998, Section 1.6] or to [Nikolski 2002b, Section 1.2] for the details. Note that there $\mathfrak{D}$ and $\mathfrak{D}_*$ were abstract spaces, $\dim \mathfrak{D} = \dim\ker \mathcal{E}^*$ and $\dim \mathfrak{D}_* = \dim\ker \mathcal{E}_*^*$, and the unitary operators $v : \mathfrak{D} \to \ker \mathcal{E}^*$, $v_* : \mathfrak{D}_* \to \ker \mathcal{E}_*^*$ used in the formulas there are just the canonical embeddings $\boldsymbol{e}$ and $\boldsymbol{e}_*$ in our case.

Note that π and π_* are isometries.

Note also that for $k \geq 0$

$$\mathcal{U}^k \boldsymbol{e}(e) = \mathcal{E}^k e, \quad e \in \mathfrak{D},$$
$$\mathcal{U}^{-k} \boldsymbol{e}_*(e_*) = \mathcal{E}_*^k e_*, \quad e_* \in \mathfrak{D}_*,$$

so

$$\pi(H^2(\mathfrak{D})) = G, \quad \pi_*(H^2_-(\mathfrak{D}_*)) = G_*.$$

The characteristic function is then defined as follows. We consider the operator $\boldsymbol{\theta} = \pi_*^*\pi : L^2(\mathfrak{D}) \to L^2(\mathfrak{D}_*)$. It is easy to check that $M_z\boldsymbol{\theta} = \boldsymbol{\theta} M_z$, so $\boldsymbol{\theta}$ is multiplication by a function $\theta \in L^\infty(\mathfrak{D} \to \mathfrak{D}_*)$. It is not hard to check that $\boldsymbol{\theta}$ is a contraction, so $\|\theta\|_\infty \leq 1$. Since

$$\pi(H^2(\mathfrak{D})) = G \perp G_* = \pi_*(H^2_-(\mathfrak{D}_*)),$$

we can conclude that $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$.

The characteristic function $\theta = \theta_T$ can be explicitly computed, see [Nikolski 2002b, Theorem 1.2.10],

$$\theta_T(z) = V_*(-T + zD_{T^*}(\boldsymbol{I}_{\mathcal{H}} - zT^*)^{-1}D_T)V^*|_{\mathfrak{D}}, \quad z \in \mathbb{D}. \tag{3-7}$$

Note that the particular representation of θ depends on the coordinate operators V and V_* identifying defect spaces $\mathfrak{D}_T$ and $\mathfrak{D}_{T^*}$ with the abstract spaces $\mathfrak{D}$ and $\mathfrak{D}_*$.

To construct a model (more precisely its particular transcription), we need to construct a unitary map Ψ between the space $\mathcal{H}$ of the minimal unitary dilation $\mathcal{U}$ and its spectral representation.

Namely, we represent $\mathcal{U}$ as a multiplication operator in some subspace $\widetilde{\mathcal{K}} = \widetilde{\mathcal{K}}_\theta$ of $L^2(\mathfrak{D}_* \oplus \mathfrak{D})$ or its weighted version.

We need to construct a unitary operator $\Psi : \mathcal{H} \to \widetilde{\mathcal{K}}$ intertwining $\mathcal{U}$ and M_z on $\widetilde{\mathcal{K}}$, i.e., such that

$$\Psi\mathcal{U} = M_z\Psi. \tag{3-8}$$

Note that if T is a completely nonunitary contraction, then $\pi(L^2(\mathfrak{D})) + \pi_*(L^2(\mathfrak{D}_*))$ is dense in $\mathcal{H}$.

So, for Ψ to be unitary it is necessary and sufficient that Ψ^* acts isometrically on $\pi(L^2(\mathfrak{D}))$ and on $\pi_*(L^2(\mathfrak{D}_*))$, and that for all $f \in L^2(\mathfrak{D})$, $g \in L^2(\mathfrak{D}_*)$

$$(\Psi^*\pi f, \Psi^*\pi g)_{\widetilde{\mathcal{K}}} = (\pi f, \pi_* g)_{\mathcal{H}} = (\boldsymbol{\theta} f, g)_{L^2(\mathfrak{D}_*)}; \tag{3-9}$$

the last equality here is just the definition of $\boldsymbol{\theta}$.

Of course, we need Ψ^* to be onto, but that can be easily accomplished by restricting the target space $\widetilde{\mathcal{K}}$ to $\operatorname{Ran}\Psi^*$.

Summing up, we have

$$\begin{array}{ccccccc} \mathcal{H} & = & G & \oplus & H & \oplus & G_* \\ \big\downarrow{\scriptstyle\Psi^*} & & \big\downarrow{\scriptstyle\Psi^*|_G} & & \big\downarrow{\scriptstyle\Psi^*|_H} & & \big\downarrow{\scriptstyle\Psi^*|_{G_*}} \\ \widetilde{\mathcal{K}} & = & \mathcal{G} & \oplus & \mathcal{K}_\theta & \oplus & \mathcal{G}_* \end{array}$$

3A. ***Pavlov transcription.*** Probably the easiest way to construct the model is to take $\widetilde{\mathcal{K}}$ to be the weighted space $L^2(\mathfrak{D}_*\oplus\mathfrak{D}, W)$, where the weight W is picked to make the simplest operator Ψ^* an isometry, and is given by

$$W(z) = \begin{pmatrix} \boldsymbol{I}_{\mathfrak{D}_*} & \theta(z) \\ \theta(z)^* & \boldsymbol{I}_{\mathfrak{D}} \end{pmatrix}. \tag{3-10}$$

Now the operator Ψ^* is defined on $\pi(L^2(\mathfrak{D}))$ and on $\pi_*(L^2(\mathfrak{D}_*))$ as

$$\begin{aligned} \Psi^*\Big(\sum_{k\in\mathbb{Z}} \mathcal{U}^k \boldsymbol{e}(e_k)\Big) &= \sum_{k\in\mathbb{Z}} z^k \begin{pmatrix} 0 \\ e_k \end{pmatrix}, \qquad e_k\in\mathfrak{D}, \\ \Psi^*\Big(\sum_{k\in\mathbb{Z}} \mathcal{U}^k \boldsymbol{e}_*(e_k)\Big) &= \sum_{k\in\mathbb{Z}} z^{k-1} \begin{pmatrix} e_k \\ 0 \end{pmatrix}, \quad e_k\in\mathfrak{D}_*, \end{aligned} \tag{3-11}$$

or equivalently

$$\begin{aligned} \Psi^*(\pi f) &= \begin{pmatrix} 0 \\ f \end{pmatrix}, \quad f\in L^2(\mathfrak{D}), \\ \Psi^*(\pi_* f) &= \begin{pmatrix} f \\ 0 \end{pmatrix}, \quad f\in L^2(\mathfrak{D}_*). \end{aligned}$$

The incoming and outgoing spaces $\mathcal{G}_* = \Psi^* G_*$, $\mathcal{G} = \Psi^* G$ are given by

$$\mathcal{G}_* := \operatorname{clos}_{\widetilde{\mathcal{K}}}\left\{\begin{pmatrix} f \\ 0 \end{pmatrix} : f\in H^2_-(\mathfrak{D}_*)\right\}, \quad \mathcal{G} := \operatorname{clos}_{\widetilde{\mathcal{K}}}\left\{\begin{pmatrix} 0 \\ f \end{pmatrix} : f\in H^2(\mathfrak{D})\right\},$$

and the model space $\mathcal{K} = \mathcal{K}_\theta$ is defined as

$$\mathcal{K}_\theta = \widetilde{\mathcal{K}} \ominus (\mathcal{G}_* \oplus \mathcal{G}).$$

3B. ***Sz.-Nagy–Foiaş transcription.*** This transcription appears when one tries to make the operator Ψ^* act into a nonweighted space $L^2(\mathfrak{D}_*\oplus\mathfrak{D})$. We make the action of the operator Ψ^* on $\pi_*(L^2(\mathfrak{D}_*))$ as simple as possible,

$$\Psi^*\Big(\sum_{k\in\mathbb{Z}} \mathcal{U}^k \boldsymbol{e}_*(e_k)\Big) = \sum_{k\in\mathbb{Z}} z^{k-1} \begin{pmatrix} e_k \\ 0 \end{pmatrix}, \quad e_k\in\mathfrak{D}_*; \tag{3-12}$$

this is exactly as in (3-11). The action of Ψ^* on $\pi(L^2(\mathfrak{D}))$ is defined as

$$\Psi^*\Big(\sum_{k\in\mathbb{Z}} \mathcal{U}^k \boldsymbol{e}(e_k)\Big) = \sum_{k\in\mathbb{Z}} z^{k} \begin{pmatrix} \theta e_k \\ \Delta e_k \end{pmatrix}, \quad e_k\in\mathfrak{D}, \tag{3-13}$$

where $\Delta(z) = (\boldsymbol{I} - \theta(z)^*\theta(z))^{1/2}$. Then (3-12) and (3-13) can clearly be rewritten as

$$\Psi^*(\pi f) = \begin{pmatrix} \theta f \\ \Delta f \end{pmatrix}, \quad f \in L^2(\mathfrak{D}), \tag{3-14}$$

$$\Psi^*(\pi_* f) = \begin{pmatrix} f \\ 0 \end{pmatrix}, \quad f \in L^2(\mathfrak{D}_*). \tag{3-15}$$

Note, that θ in the top entries in (3-13) and (3-14) is necessary to get (3-9); after (3-12), equivalently (3-15), is chosen, one does not have any choice here. The term Δ in the bottom entries of (3-13) and (3-14) is there to make Ψ^* act isometrically on $\pi(L^2(\mathfrak{D}))$. There is some freedom here; one can left-multiply Δ by any operator-valued function ϕ such that $\phi(z)$ acts isometrically on $\operatorname{Ran}\Delta(z)$. However, picking just Δ is the canonical choice for the Sz.-Nagy–Foiaş transcription, and we will follow it.

The incoming and outgoing spaces are given by

$$\mathcal{G}_* := \begin{pmatrix} H^2_-(\mathfrak{D}_*) \\ 0 \end{pmatrix}, \quad \mathcal{G} := \begin{pmatrix} \theta \\ \Delta \end{pmatrix} H^2(\mathfrak{D}).$$

The model space is given by

$$\mathcal{K}_\theta := \begin{pmatrix} L^2(\mathfrak{D}_*) \\ \operatorname{clos}\Delta L^2(\mathfrak{D}) \end{pmatrix} \ominus (\mathcal{G}_* \oplus \mathcal{G}) = \begin{pmatrix} H^2(\mathfrak{D}_*) \\ \operatorname{clos}\Delta L^2(\mathfrak{D}) \end{pmatrix} \ominus \begin{pmatrix} \theta \\ \Delta \end{pmatrix} H^2(\mathfrak{D}). \tag{3-16}$$

Remark. While the orthogonal projection from

$$\begin{pmatrix} L^2(\mathfrak{D}_*) \\ \operatorname{clos}\Delta L^2(\mathfrak{D}) \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} L^2(\mathfrak{D}_*) \\ \operatorname{clos}\Delta L^2(\mathfrak{D}) \end{pmatrix} \ominus \mathcal{G}_*$$

is rather simple, the one from

$$\begin{pmatrix} L^2(\mathfrak{D}_*) \\ \operatorname{clos}\Delta L^2(\mathfrak{D}) \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} L^2(\mathfrak{D}_*) \\ \operatorname{clos}\Delta L^2(\mathfrak{D}) \end{pmatrix} \ominus \mathcal{G}$$

involves the range of a Toeplitz operator.

3C. ***De Branges–Rovnyak transcription.*** This transcription looks very complicated, but its advantage is that both coordinates are analytic functions. To describe this transcription, we use the auxiliary weight $W = W(z)$ as in the Pavlov transcription; see (3-10). The model space is the subspace of $L^2(\mathfrak{D}_* \oplus \mathfrak{D}, W^{[-1]})$, where for a self-adjoint operator A the symbol $A^{[-1]}$ denotes its Moore–Penrose (pseudo)inverse, i.e., $A^{[-1]} = 0$ on $\operatorname{Ker} A$ and $A^{[-1]}$ is the left inverse of A on $(\operatorname{Ker} A)^\perp$.

The operator $\Psi^* : \mathcal{H} \to L^2(\mathfrak{D}_* \oplus \mathfrak{D}, W^{[-1]})$ is defined by

$$\Psi^*(\pi f) = W \begin{pmatrix} 0 \\ f \end{pmatrix} = \begin{pmatrix} \theta f \\ f \end{pmatrix}, \quad f \in L^2(\mathfrak{D}),$$

$$\Psi^*(\pi_* f) = W \begin{pmatrix} f \\ 0 \end{pmatrix} = \begin{pmatrix} f \\ \theta^* f \end{pmatrix}, \quad f \in L^2(\mathfrak{D}_*).$$

The incoming and outgoing spaces are

$$\mathcal{G}_* := \begin{pmatrix} \boldsymbol{I} \\ \theta^* \end{pmatrix} H^2(\mathfrak{D}_*), \quad \mathcal{G} := \begin{pmatrix} \theta \\ \boldsymbol{I} \end{pmatrix} H^2(\mathfrak{D}),$$

and the model space is defined as

$$\mathcal{K}_\theta := \left\{ \begin{pmatrix} f \\ g \end{pmatrix} : f \in H^2(\mathfrak{D}_*),\ g \in H^2_-(\mathfrak{D}),\ g - \theta^* f \in \Delta L^2(\mathfrak{D}) \right\};$$

see [Nikolski and Vasyunin 1998, Section 3.7] for the details (there is a typo in that paper; in the definition of $\mathcal{K}_\theta$ on p. 251 it should be $f \in H^2(E_*),\ g \in H^2(E)$).

3D. ***Parametrizing operators for the model, agreeing with coordinate operators.*** The parametrizing operators that agree with the coordinate operators V and V_* are described in the following lemma, which holds for any transcription of the model.

Let T be a c.n.u. contraction, and let $V : \mathfrak{D}_T \to \mathfrak{D}$ and $V_* : \mathfrak{D}_{T^*} \to \mathfrak{D}_*$ be coordinate operators for the defect spaces of T. Let $\theta = \theta_T = \theta_{T,V,V_*} \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$ be the characteristic function of T, defined by (3-7), and let $\mathcal{M}_\theta$ be the corresponding model operator (in any transcription).

Recall that Ψ is a unitary operator intertwining the minimal unitary dilation $\mathcal{U}$ of T and the multiplication operator M_z in the corresponding function space; see (3-8). The operator Ψ determines the transcription of the model, so for any particular transcription it is known.

Define

$$\tilde{\boldsymbol{e}} := \Psi^* \boldsymbol{e}, \quad \tilde{\boldsymbol{e}}_* := \Psi^* \boldsymbol{e}_*, \tag{3-17}$$

where the embeddings $\boldsymbol{e}$ and $\boldsymbol{e}_*$ are defined by (3-4), (3-5).

Lemma 3.2. *Under the above assumptions the parametrizing operators $\boldsymbol{C}_* : \mathfrak{D}_* \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ and $\boldsymbol{C} : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta}$ given by*

$$\boldsymbol{C}_* e_* = (D_{\mathcal{M}_\theta^*}|_{\mathfrak{D}_{\mathcal{M}_\theta^*}})^{-1} P_{\mathcal{K}_\theta} M_z \tilde{\boldsymbol{e}}_*(e_*), \quad e_* \in \mathfrak{D}_*, \tag{3-18}$$

$$\boldsymbol{C} e = (D_{\mathcal{M}_\theta}|_{\mathfrak{D}_{\mathcal{M}_\theta}})^{-1} P_{\mathcal{K}_\theta} M_{\bar{z}} \tilde{\boldsymbol{e}}(e), \qquad e \in \mathfrak{D}, \tag{3-19}$$

agree with the coordinate operators V and V_*.

Remark. It follows from (3-20) below that $P_{\mathcal{K}_\theta} M_z \tilde{\boldsymbol{e}}_*(e_*) \in \operatorname{Ran} D_{\mathcal{M}_\theta^*}$ as well as $P_{\mathcal{K}_\theta} M_{\bar{z}} \tilde{\boldsymbol{e}}(e) \in \operatorname{Ran} D_{\mathcal{M}_\theta}$, so everything in (3-18), (3-19) is well defined.

Proof of Lemma 3.2. Right- and left-multiplying (3-6) by Ψ and Ψ^* respectively, we get

$$\Psi^* \mathcal{U} \Psi = \begin{pmatrix} \widetilde{\mathcal{E}}_*^* & 0 & 0 \\ D_{\mathcal{M}_\theta^*} \boldsymbol{C}_* \tilde{\boldsymbol{e}}_*^* & \mathcal{M}_\theta & 0 \\ -\tilde{\boldsymbol{e}} \boldsymbol{C}^* \mathcal{M}_\theta^* \boldsymbol{C}_* \tilde{\boldsymbol{e}}_*^* & \tilde{\boldsymbol{e}} \boldsymbol{C}^* D_{\mathcal{M}_\theta} & \widetilde{\mathcal{E}} \end{pmatrix}, \tag{3-20}$$

where $\widetilde{\mathcal{E}} = \Psi^* \mathcal{E} \Psi,\ \widetilde{\mathcal{E}}_* = \Psi \mathcal{E}_* \Psi,\ \boldsymbol{C}^* = V \Psi,\ \boldsymbol{C}_*^* = V_* \Psi,\ \tilde{\boldsymbol{e}} = \Psi^* \boldsymbol{e},\ \tilde{\boldsymbol{e}}_* = \Psi^* \boldsymbol{e}_*$.

The operators $\tilde{\boldsymbol{e}}$ and $\tilde{\boldsymbol{e}}_*$ are the canonical embeddings of $\mathfrak{D}$ and $\mathfrak{D}_*$ into $\mathcal{G}$ and $\mathcal{G}_*$ that agree with the canonical embeddings $\boldsymbol{e}$ and $\boldsymbol{e}_*$. The operators $\boldsymbol{C}$ and $\boldsymbol{C}_*$ are the parametrizing operators for the defect spaces of the model operator $\mathcal{M}_\theta$ that agree with the coordinate operators V and V_* for the defect spaces of the operator T.

In any particular transcription of the model, the operator $\Psi^*\mathcal{U}\Psi$ is known (it is just the multiplication by z in an appropriate function space), so we get from the decomposition (3-20)

$$D_{\mathcal{M}_\theta^*}\boldsymbol{C}_*\tilde{\boldsymbol{e}}_*^* = P_{\mathcal{K}_\theta}M_z|_{\mathcal{G}_*}, \quad D_{\mathcal{M}_\theta}\boldsymbol{C}\tilde{\boldsymbol{e}}^* = P_{\mathcal{K}_\theta}M_{\bar{z}}|_{\mathcal{G}_*}.$$

Right- and left-multiplying the first identity by $\boldsymbol{e}_*$ and $(D_{\mathcal{M}_\theta^*}|_{\mathfrak{D}_{\mathcal{M}_\theta^*}})^{-1}$ respectively, we get (3-18). Similarly, to get (3-19) we just right- and left-multiply the second identity by $\boldsymbol{e}$ and $(D_{\mathcal{M}_\theta}|_{\mathfrak{D}_{\mathcal{M}_\theta}})^{-1}$. □

Applying the above Lemma 3.2 to a particular transcription of the model, we can get more concrete formulas for $\boldsymbol{C}$, $\boldsymbol{C}_*$ just in terms of the characteristic function θ. For example, the following lemma gives formulas for $\boldsymbol{C}$ and $\boldsymbol{C}_*$ in the Sz.-Nagy–Foiaş transcription.

Lemma 3.3. *Let T be a c.n.u. contraction, and let $\mathcal{M}_\theta$ be its model in the Sz.-Nagy–Foiaş transcription, with the characteristic function $\theta = \theta_{T,V,V_*}$, $\theta \in H^\infty(\mathfrak{D} \to \mathfrak{D}_*)$.*

Then the maps $\boldsymbol{C}_* : \mathfrak{D}_* \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ and $\boldsymbol{C} : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta}$ given by

$$\boldsymbol{C}_*e_* = \begin{pmatrix} \boldsymbol{I} - \theta(z)\theta^*(0) \\ -\Delta(z)\theta^*(0) \end{pmatrix}(\boldsymbol{I} - \theta(0)\theta^*(0))^{-1/2}e_*, \qquad e_* \in \mathfrak{D}_*, \tag{3-21}$$

$$\boldsymbol{C}e = \begin{pmatrix} z^{-1}(\theta(z) - \theta(0)) \\ z^{-1}\Delta(z) \end{pmatrix}(\boldsymbol{I} - \theta^*(0)\theta(0))^{-1/2}e, \quad e \in \mathfrak{D}, \tag{3-22}$$

agree with the coordinate operators V and V_*.

Proof. To prove (3-21) we will use (3-18). It follows from (3-12) that

$$\tilde{\boldsymbol{e}}_*(e_*) = z^{-1}\begin{pmatrix} e_* \\ 0 \end{pmatrix},$$

so by (3-18)

$$\boldsymbol{C}_*e_* = (\boldsymbol{I} - \mathcal{M}_\theta\mathcal{M}_\theta^*)|_{\mathfrak{D}_{\mathcal{M}_\theta^*}}^{-1/2} P_{\mathcal{K}_\theta}\begin{pmatrix} e_* \\ 0 \end{pmatrix}, \quad e_* \in \mathfrak{D}_*. \tag{3-23}$$

It is not hard to show that

$$P_{\mathcal{K}_\theta}\begin{pmatrix} e_* \\ 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} - \theta\theta(0)^* \\ -\Delta\theta(0)^* \end{pmatrix} e_*. \tag{3-24}$$

One also can compute

$$(\boldsymbol{I} - \mathcal{M}_\theta\mathcal{M}_\theta^*)\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} - \theta\theta(0)^* \\ -\Delta\theta(0)^* \end{pmatrix} f(0), \quad \begin{pmatrix} f \\ g \end{pmatrix} \in \mathcal{K}_\theta. \tag{3-25}$$

Combining the above identities we get

$$(\boldsymbol{I} - \mathcal{M}_\theta\mathcal{M}_\theta^*)P_{\mathcal{K}_\theta}\begin{pmatrix} e_* \\ 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} - \theta\theta(0)^* \\ -\Delta\theta(0)^* \end{pmatrix}(e_* - \theta(0)\theta^*(0)e_*). \tag{3-26}$$

As we discussed above just after (3-19), $P_{\mathcal{K}_\theta}\binom{e_*}{0} \in \operatorname{Ran} D_{\mathcal{M}_\theta^*}$, so in (3-26) we can replace $(\boldsymbol{I} - \mathcal{M}_\theta\mathcal{M}_\theta^*)$ by its restriction onto $\mathfrak{D}_{\mathcal{M}_\theta^*}$.

Applying $(\boldsymbol{I}-\mathcal{M}_\theta\mathcal{M}_\theta^*)|_{\mathfrak{D}_{\mathcal{M}_\theta^*}}$ to (3-26) (with $(\boldsymbol{I}-\mathcal{M}_\theta\mathcal{M}_\theta^*)$ replaced by its restriction onto $\mathfrak{D}_{\mathcal{M}_\theta}$) and using (3-25) we get

$$((\boldsymbol{I}-\mathcal{M}_\theta\mathcal{M}_\theta^*)|_{\mathfrak{D}_{\mathcal{M}_\theta^*}})^2 P_{\mathcal{K}_\theta}\begin{pmatrix} e_* \\ 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}-\theta\theta(0)^* \\ -\Delta\theta(0)^* \end{pmatrix}(\boldsymbol{I}_{\mathfrak{D}_*}-\theta(0)\theta^*(0))^2 e_*.$$

Applying $(\boldsymbol{I}-\mathcal{M}_\theta\mathcal{M}_\theta^*)|_{\mathfrak{D}_{\mathcal{M}_\theta^*}}$ to the above identity, and using again (3-25), we get by induction that

$$\varphi((\boldsymbol{I}-\mathcal{M}_\theta\mathcal{M}_\theta^*)|_{\mathfrak{D}_{\mathcal{M}_\theta^*}}) P_{\mathcal{K}_\theta}\begin{pmatrix} e_* \\ 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}-\theta\theta(0)^* \\ -\Delta\theta(0)^* \end{pmatrix}\varphi(\boldsymbol{I}_{\mathfrak{D}_*}-\theta(0)\theta^*(0)) e_* \tag{3-27}$$

for any monomial φ, $\varphi(x)=x^n$, $n\ge 0$ (the case $n=0$ is just the identity (3-24)).

Linearity implies that (3-27) holds for any polynomial φ. Using standard approximation reasoning we get that φ in (3-27) can be any measurable function. In particular, we can take $\varphi(x)=x^{-1/2}$, which together with (3-23) gives us (3-21).

To prove (3-22) we proceed similarly. Equation (3-13) implies

$$\tilde{\boldsymbol{e}}(e)=\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e,$$

so by (3-19)

$$\boldsymbol{C}e = ((\boldsymbol{I}-\mathcal{M}_\theta^*\mathcal{M}_\theta)|_{\mathfrak{D}_{\mathcal{M}_\theta}})^{-1/2} P_{\mathcal{K}_\theta} M_{\bar z}\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e, \quad e\in\mathfrak{D}. \tag{3-28}$$

One can see that

$$P_{\mathcal{K}_\theta} M_{\bar z}\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e = M_{\bar z}\begin{pmatrix} \theta-\theta(0) \\ \Delta \end{pmatrix} e,$$

so

$$\mathcal{M}_\theta P_{\mathcal{K}_\theta} M_{\bar z}\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e = P_{\mathcal{K}_\theta}\begin{pmatrix} \theta-\theta(0) \\ \Delta \end{pmatrix} e = -P_{\mathcal{K}_\theta}\begin{pmatrix} \theta(0) \\ 0 \end{pmatrix} e.$$

Combining this with (3-24), we get

$$\mathcal{M}_\theta P_{\mathcal{K}_\theta} M_{\bar z}\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e = \begin{pmatrix} \theta\theta(0)^*-\boldsymbol{I} \\ \Delta\theta(0)^* \end{pmatrix}\theta(0)e.$$

Using the fact that

$$\mathcal{M}_\theta^*\begin{pmatrix} f \\ g \end{pmatrix} = M_{\bar z}\begin{pmatrix} f-f(0) \\ g \end{pmatrix},$$

we arrive at

$$\mathcal{M}_\theta^*\mathcal{M}_\theta P_{\mathcal{K}_\theta} M_{\bar z}\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e = M_{\bar z}\begin{pmatrix} \theta-\theta(0) \\ \Delta \end{pmatrix}\theta(0)^*\theta(0)e,$$

so

$$(\boldsymbol{I}-\mathcal{M}_\theta^*\mathcal{M}_\theta) P_{\mathcal{K}_\theta} M_{\bar z}\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e = M_{\bar z}\begin{pmatrix} \theta-\theta(0) \\ \Delta \end{pmatrix}(\boldsymbol{I}-\theta(0)^*\theta(0))e.$$

Using the same reasoning as in the above proof of (3-21) we get

$$\varphi((\boldsymbol{I}-\mathcal{M}_\theta^*\mathcal{M}_\theta)|_{\mathfrak{D}_{\mathcal{M}_\theta}}) P_{\mathcal{K}_\theta} M_{\bar z}\begin{pmatrix} \theta \\ \Delta \end{pmatrix} e = M_{\bar z}\begin{pmatrix} \theta-\theta(0) \\ \Delta \end{pmatrix}\varphi(\boldsymbol{I}-\theta(0)^*\theta(0))e, \tag{3-29}$$

first with φ being a polynomial, and then any measurable function.

Using (3-29) with $\varphi(x)=x^{-1/2}$ and taking (3-28) into account, we get (3-22). $\square$

3E. ***An auxiliary lemma.*** We already used, and we will also need later, the following simple lemma.

Lemma 3.4. *Let $\mathcal{M} = \mathcal{M}_\theta$ be model operator on a model space $\mathcal{K}_\theta \subset L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$, and let $\boldsymbol{C} : \mathfrak{D}_* \to \mathfrak{D}_{\mathcal{M}_\theta}$, $\boldsymbol{C}_* : \mathfrak{D} \to \mathfrak{D}_{\mathcal{M}_\theta^*}$ be bounded operators.*

If C and C_* are the operator-valued functions defined by

$$C(z)e = \boldsymbol{C}e(z), \qquad z \in \mathbb{T},\ e \in \mathfrak{D},$$
$$C_*(z)e_* = \boldsymbol{C}_*e_*(z), \quad z \in \mathbb{T},\ e_* \in \mathfrak{D}_*,$$

then the functions $W^{1/2}C$ and $W^{1/2}C^*$ are bounded,

$$\|W^{1/2}C\|_{L^\infty} = \|\boldsymbol{C}\|, \quad \|W^{1/2}C_*\|_{L^\infty} = \|\boldsymbol{C}_*\|.$$

Proof. It is well known and is not hard to show that if T is a contraction and $\mathcal{U}$ is its unitary dilation, then the subspaces $\mathcal{U}^n\mathfrak{D}_T$, $n \in \mathbb{Z}$ (where recall $\mathfrak{D}_T$ is the defect space of T) are mutually orthogonal, and similarly for subspaces $\mathcal{U}^n\mathfrak{D}_{T^*}$, $n \in \mathbb{Z}$.

Therefore, the subspaces $z^n\mathfrak{D}_{\mathcal{M}}$, $n \in \mathbb{Z}$, are mutually orthogonal in $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$, and the same holds for the subspaces $z^n\mathfrak{D}_{\mathcal{M}^*}$, $n \in \mathbb{Z}$.

The subspaces $z^n\mathfrak{D} \subset L^2(\mathbb{T}; \mathfrak{D})$ are mutually orthogonal, and since

$$C(z)\sum_{n\in\mathbb{Z}} z^n \hat{f}(n) = \sum_{n\in\mathbb{Z}} z^n \boldsymbol{C} f_n, \quad \hat{f}(n) \in \mathfrak{D},$$

we conclude that the operator $f \mapsto Cf$ is a bounded operator acting from $L^2(\mathfrak{D})$ to $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$, and its norm is exactly $\|\boldsymbol{C}\|$.

But that means the multiplication operator $f \mapsto W^{1/2}f$ between the nonweighted spaces $L^2(\mathfrak{D})$ and $L^2(\mathfrak{D}_* \oplus \mathfrak{D})$ is bounded with the same norm, which immediately implies $\|W^{1/2}C\|_{L^\infty} = \|\boldsymbol{C}\|$.

The proof for C_* follows similarly. □

4. Characteristic function

We now derive formulas for the (matrix-valued) characteristic function θ_Γ; see Theorem 4.2 below.

4A. ***An inverse of a perturbation.*** We begin with an auxiliary result.

Lemma 4.1. *Let D be an operator in an auxiliary Hilbert space $\mathfrak{R}$ and let $B, C : \mathfrak{R} \to \mathcal{H}$. Then $\boldsymbol{I}_{\mathcal{H}} - CDB^*$ is invertible if and only if $\boldsymbol{I}_{\mathfrak{R}} - DB^*C$ is invertible, and if and only if $\boldsymbol{I}_{\mathfrak{R}} - B^*CD$ is invertible.*

Moreover, in this case

$$\begin{aligned}(\boldsymbol{I}_{\mathcal{H}} - CDB^*)^{-1} &= \boldsymbol{I}_{\mathcal{H}} + C(\boldsymbol{I}_{\mathfrak{R}} - DB^*C)^{-1}DB^* \\ &= \boldsymbol{I}_{\mathcal{H}} + CD(\boldsymbol{I}_{\mathfrak{R}} - B^*CD)^{-1}B^*. \end{aligned} \tag{4-1}$$

We will apply this lemma for $D : \mathbb{C}^d \to \mathbb{C}^d$, so in this case the inversion of $\boldsymbol{I}_{\mathcal{H}} - CDB$ is reduced to inverting a $d \times d$ matrix.

This lemma can be obtained from the Woodbury inversion formula [1950], although formally in [Woodbury 1950] only the matrix case was treated.

Proof of Lemma 4.1. First let us note that it is sufficient to prove the lemma with $D = \boldsymbol{I}_{\mathfrak{R}}$, because D can be incorporated either into C or into B^*.

One could guess the formula by writing the power series expansion of $\boldsymbol{I}_{\mathcal{H}} - CDB^*$, and we can get the result for the case when the series converges. This method can be made rigorous for finite-rank perturbations by considering the family $(\boldsymbol{I}_{\mathcal{H}} - \lambda CDB^*)^{-1}$, $\lambda \in \mathbb{C}$, and using analytic continuation.

However, the simplest way to prove the formula is just by performing multiplication,

$$\begin{aligned}(\boldsymbol{I}_{\mathcal{H}} - CB^*)(\boldsymbol{I}_{\mathcal{H}} + C(\boldsymbol{I}_{\mathfrak{R}} - B^*C)^{-1}B^*) &= \boldsymbol{I}_{\mathcal{H}} - CB^* + C(\boldsymbol{I}_{\mathfrak{R}} - B^*C)^{-1}B^* - CB^*C(\boldsymbol{I}_{\mathfrak{R}} - B^*C)^{-1}B^* \\ &= \boldsymbol{I}_{\mathcal{H}} + C(-\boldsymbol{I}_{\mathfrak{R}}(\boldsymbol{I}_{\mathfrak{R}} - B^*C) + \boldsymbol{I}_{\mathfrak{R}} - B^*C)(\boldsymbol{I}_{\mathfrak{R}} - B^*C)^{-1}B^* \\ &= \boldsymbol{I}_{\mathcal{H}}.\end{aligned}$$

Thus, when $\boldsymbol{I}_{\mathfrak{R}} - B^*C$ is invertible, the operator $\boldsymbol{I}_{\mathcal{H}} + C(\boldsymbol{I}_{\mathfrak{R}} - B^*C)^{-1}B^*$ is the right inverse of $\boldsymbol{I}_{\mathcal{H}} - CB^*$. To prove that it is also a left inverse we even do not need to perform the multiplication: we can just take the adjoint of the above identity and then interchange B and C.

So, the invertibility of $\boldsymbol{I}_{\mathfrak{R}} - B^*C$ implies the invertibility of $\boldsymbol{I}_{\mathcal{H}} - CB^*$ and the formula for the inverse. To prove the "if and only if" statement we just need to change the roles of $\mathcal{H}$ and $\mathfrak{R}$ and express, using the just proved formula, the inverse of $\boldsymbol{I}_{\mathfrak{R}} - B^*C$ in terms of $(\boldsymbol{I}_{\mathcal{H}} - CB^*)^{-1}$. □

4B. ***Computation of the characteristic function.*** We turn to computing the characteristic function of $T = U + \boldsymbol{B}(\Gamma - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^*U$, $\|\Gamma\| < 1$, where U is the multiplication operator M_ξ in $L^2(\mu; E)$.

We will use formula (3-7) with $V = \boldsymbol{B}^*U$, $V_* = \boldsymbol{B}^*$, $\mathfrak{D} = \mathfrak{D}_* = \mathbb{C}^d$.

Let us first calculate for $|z| < 1$,

$$\begin{aligned}(\boldsymbol{I}_{\mathcal{H}} - zT^*)^{-1} &= \big[(\boldsymbol{I}_{\mathcal{H}} - zU^*)\big(\boldsymbol{I}_{\mathcal{H}} - z(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1}U^*\boldsymbol{B}(\Gamma^* - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^*\big)\big]^{-1} \\ &= \big[\boldsymbol{I}_{\mathcal{H}} - z(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1}U^*\boldsymbol{B}(\Gamma^* - \boldsymbol{I}_{\mathbb{C}^d})\boldsymbol{B}^*\big]^{-1}(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1} \\ &=: X(z)(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1}.\end{aligned}$$

To compute the inverse $X(z)$ we use Lemma 4.1 with $z(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1}U^*\boldsymbol{B}$ instead of C, $\Gamma^* - \boldsymbol{I}_{\mathbb{C}^d}$ instead of D and $\boldsymbol{B}$ instead of B. Together with the first identity in (4-1) we get

$$X(z) = \boldsymbol{I}_{\mathcal{H}} + z(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1}U^*\boldsymbol{B}(\boldsymbol{I}_{\mathbb{C}^d} - zD\boldsymbol{B}^*(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1}U^*\boldsymbol{B})^{-1}D\boldsymbol{B}^*, \tag{4-2}$$

where $D = \Gamma^* - \boldsymbol{I}_{\mathbb{C}^d}$.

Now, let us express $z\boldsymbol{B}^*(\boldsymbol{I}_{\mathcal{H}} - zU^*)^{-1}U^*\boldsymbol{B}$ as a Cauchy integral of some matrix-valued measure. Recall that U is a multiplication by the independent variable ξ in $\mathcal{H} \subset L^2(\mu; E)$ and that $b_1, b_2, \dots, b_d \in \mathcal{H}$ denote the "columns" of $\boldsymbol{B}$ (i.e., $b_k = \boldsymbol{B}e_k$, where $e_1, e_2, \dots, e_d$ is the standard basis in $\mathbb{C}^d$), and $B(\xi) = (b_1(\xi), b_2(\xi), \dots, b_d(\xi))$ is the matrix with columns $b_k(\xi)$. Then

$$b_j^*(\boldsymbol{I}_{\mathbb{C}^d} - zU^*)^{-1}U^*b_k = \int_{\mathbb{T}} \frac{\bar{\xi}}{1 - z\bar{\xi}}\, b_j(\xi)^* b_k(\xi)\, \mathrm{d}\mu(\xi),$$

so

$$z\boldsymbol{B}^*(\boldsymbol{I}_{\mathcal{H}}-zU^*)^{-1}U^*\boldsymbol{B}=\int_{\mathbb{T}}\frac{z\bar{\xi}}{1-z\bar{\xi}}\,M(\xi)\,\mathrm{d}\mu(\xi)=:\mathcal{C}_1[M\mu](z)=:F_1(z), \tag{4-3}$$

where M is the matrix-valued function $M(\xi)=B(\xi)^*B(\xi)$, or equivalently $M_{j,k}(\xi)=b_j(\xi)^*b_k(\xi)$, $1\le j,k\le d$.

Using (4-3) and denoting $D:=\Gamma^*-\boldsymbol{I}_{\mathbb{C}^d}$ we get from the above calculations that

$$(\boldsymbol{I}_{\mathcal{H}}-zT^*)^{-1}=(\boldsymbol{I}_{\mathcal{H}}-zU^*)^{-1}+z(\boldsymbol{I}_{\mathcal{H}}-zU^*)^{-1}U^*\boldsymbol{B}(\boldsymbol{I}_{\mathbb{C}^d}-DF_1(z))^{-1}D\boldsymbol{B}^*(\boldsymbol{I}_{\mathcal{H}}-zU^*)^{-1}.$$

Applying formula (3-7), with $V=\boldsymbol{B}^*U,\ V_*=\boldsymbol{B}_*,\ \mathfrak{D}=\mathfrak{D}_*=\mathbb{C}^d$, we see that the characteristic function is an analytic function $\theta=\theta_T$ whose values are bounded linear operators acting on $\mathfrak{D}$, defined by the formula

$$\theta_T(z)=\boldsymbol{B}^*(-T+zD_{T^*}(\boldsymbol{I}_{\mathcal{H}}-zT^*)^{-1}D_T)U^*\boldsymbol{B}|_{\mathfrak{D}},\quad z\in\mathbb{D}. \tag{4-4}$$

We can see from (1-3) that the defect operators D_T and D_{T^*} are given by

$$D_T=U^*\boldsymbol{B}D_\Gamma\boldsymbol{B}^*U,\quad D_{T^*}=\boldsymbol{B}D_{\Gamma^*}\boldsymbol{B}^*.$$

We can also see from (1-3) that the term $-T$ in (4-4) contributes $-\Gamma$ to the matrix θ_T. The rest can be obtained from the above representation formula for $(\boldsymbol{I}_{\mathcal{H}}-zT^*)^{-1}$. Thus, recalling the definition (4-3) of $\mathcal{C}_1M\mu$ we get, defining $F_1(z):=(\mathcal{C}_1M\mu)(z)$, that

$$\begin{aligned}\theta_T(z)&=-\Gamma+D_{\Gamma^*}\big[F_1(z)+F_1(z)\big(\boldsymbol{I}_{\mathfrak{D}}-(\Gamma^*-\boldsymbol{I}_{\mathfrak{D}})F_1(z)\big)^{-1}(\Gamma^*-\boldsymbol{I}_{\mathfrak{D}})F_1(z)\big]D_\Gamma\\&=-\Gamma+D_{\Gamma^*}F_1(z)\big(\boldsymbol{I}_{\mathfrak{D}}-(\Gamma^*-\boldsymbol{I}_{\mathfrak{D}})F_1(z)\big)^{-1}D_\Gamma.\end{aligned}$$

In the above computation to compute $X(z)$ we can use the second formula in (4-1). We get instead of (4-2) an alternative representation

$$X(z)=\boldsymbol{I}_{\mathcal{H}}+z(\boldsymbol{I}_{\mathcal{H}}-zU^*)^{-1}U^*\boldsymbol{B}D\big(\boldsymbol{I}_{\mathfrak{D}}-z\boldsymbol{B}^*(\boldsymbol{I}_{\mathcal{H}}-zU^*)^{-1}U^*\boldsymbol{B}D\big)^{-1}\boldsymbol{B}^*.$$

Repeating the same computations as above we get another formula for θ_T,

$$\theta_T(z)=-\Gamma+D_{\Gamma^*}\big(\boldsymbol{I}_{\mathfrak{D}}-F_1(z)(\Gamma^*-\boldsymbol{I}_{\mathfrak{D}})\big)^{-1}F_1(z)D_\Gamma.$$

To summarize we have proved two representations of the characteristic operator-valued function.

Theorem 4.2. *Let $T=T_\Gamma$ be the operator given in* (1-3), *with Γ being a strict contraction. Then the characteristic function $\theta_T=\theta_{T_\Gamma}\in H^\infty(\mathfrak{D}\to\mathfrak{D}_*)$, with coordinate operators $V=\boldsymbol{B}^*U,\ V_*=\boldsymbol{B}^*$* (*and with $\mathfrak{D}=\mathfrak{D}_*=\mathbb{C}^d$*), *is given by*

$$\begin{aligned}\theta_{T_\Gamma}(z)&=-\Gamma+D_{\Gamma^*}F_1(z)\big(\boldsymbol{I}_{\mathfrak{D}}-(\Gamma^*-\boldsymbol{I}_{\mathfrak{D}})F_1(z)\big)^{-1}D_\Gamma\\&=-\Gamma+D_{\Gamma^*}\big(\boldsymbol{I}_{\mathfrak{D}}-F_1(z)(\Gamma^*-\boldsymbol{I}_{\mathfrak{D}})\big)^{-1}F_1(z)D_\Gamma,\end{aligned}$$

where $F_1(z)$ is the matrix-valued function given by (4-3).

In these formulas, the inverse is taken of a $d \times d$ matrix-valued function, which is much simpler than computing the inverse in (4-4).

4C. ***Characteristic function and the Cauchy integrals of matrix-valued measures.*** For a (possibly complex-valued) measure τ on $\mathbb{T}$ and $z \notin \mathbb{T}$, define the Cauchy-type transforms $\mathcal{C}$, $\mathcal{C}_1$ and $\mathcal{C}_2$,

$$\mathcal{C}\tau(z) := \int_{\mathbb{T}} \frac{\mathrm{d}\tau(\xi)}{1-\bar{\xi}z}, \quad \mathcal{C}_1\tau(z) := \int_{\mathbb{T}} \frac{\bar{\xi}z\,\mathrm{d}\tau(\xi)}{1-\bar{\xi}z}, \quad \mathcal{C}_2\tau(z) := \int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z}\,\mathrm{d}\tau(\xi).$$

Performing the Cauchy transforms componentwise we can define them for matrix-valued measures as well.

Thus F_1 from the above Theorem 4.2 is given by $F_1 = \mathcal{C}_1[M\mu]$, where $M(\xi) = B^*(\xi)B(\xi)$. We would like to give the representation of θ_{T_Γ} in terms of function $F_2 := \mathcal{C}_2[M\mu]$.

Slightly abusing notation we will write θ_Γ instead of θ_{T_Γ}.

Corollary 4.3. *For $\theta_{\mathbf{0}} := \theta_{T_{\mathbf{0}}}$ we have*

$$\theta_{\mathbf{0}}(z) = F_1(z)(\boldsymbol{I} + F_1(z))^{-1} = (\boldsymbol{I} + F_1(z))^{-1}F_1(z) \tag{4-5}$$

$$= (F_2(z)-\boldsymbol{I})(F_2(z)+\boldsymbol{I})^{-1} = (F_2(z)+\boldsymbol{I})^{-1}(F_2(z)-\boldsymbol{I}). \tag{4-6}$$

Proof. The identity (4-5) is a direct application of Theorem 4.2. The identity (4-6) follows immediately from the trivial relation

$$F_2(z) = \int_{\mathbb{T}} M\,\mathrm{d}\mu + 2F_1(z) = \boldsymbol{I}_{\mathfrak{D}} + 2F_1(z);$$

the equality $\int_{\mathbb{T}} M\,\mathrm{d}\mu = \boldsymbol{I}_{\mathfrak{D}} = \boldsymbol{I}_{\mathbb{C}^d}$ is just a restatement of the fact that the functions $b_1, b_2, \dots, b_d$ form an orthonormal basis in $\mathcal{H}$. □

5. Relations between characteristic functions $\boldsymbol{\theta_\Gamma}$

5A. ***Characteristic functions and linear fractional transformations.*** When $d = 1$, it is known that the characteristic functions are related by a linear fractional transformation

$$\theta_\gamma(z) = \frac{\theta_0(z)-\gamma}{1-\bar{\gamma}\theta_0(z)};$$

see [Liaw and Treil 2016, equation (2.9)].

It turns out that a similar formula holds for finite-rank perturbations.

Theorem 5.1. *Let T be the operator given in* (1-3), *with Γ being a strict contraction. Then the characteristic functions $\theta_\Gamma := \theta_{T_\Gamma}$ and $\theta_{\mathbf{0}} = \theta_{T_{\mathbf{0}}}$ are related via linear fractional transformation*,

$$\theta_\Gamma = D_{\Gamma^*}^{-1}(\theta_{\mathbf{0}}-\Gamma)(\boldsymbol{I}_{\mathfrak{D}}-\Gamma^*\theta_{\mathbf{0}})^{-1}D_\Gamma = D_{\Gamma^*}(\boldsymbol{I}_{\mathfrak{D}}-\theta_{\mathbf{0}}\Gamma^*)^{-1}(\theta_{\mathbf{0}}-\Gamma)D_\Gamma^{-1}.$$

Remark. At first sight, this formula looks like a formula in [Nikolski and Vasyunin 1998, p. 234]. However, their result expresses the characteristic function in terms of a linear fractional transformation in T, whereas, here we have a linear fractional transformation in Γ.

Theorem 5.2. *Under assumptions of the above Theorem 5.1*

$$\theta_{\mathbf{0}} = D_{\Gamma^*}(\boldsymbol{I} + \theta_\Gamma\Gamma^*)^{-1}(\theta_\Gamma + \Gamma)D_\Gamma^{-1} = D_{\Gamma^*}^{-1}(\theta_\Gamma + \Gamma)(\boldsymbol{I} + \Gamma^*\theta_\Gamma)^{-1}D_\Gamma.$$

To prove Theorem 5.1 we start with the following simpler statement.

Proposition 5.3. *The matrix-valued characteristic functions θ_Γ and $\theta_{\mathbf{0}}$ are related via*

$$\theta_\Gamma = -\Gamma + D_{\Gamma^*}\theta_{\mathbf{0}}(\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}D_\Gamma = -\Gamma + D_{\Gamma^*}(\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}\Gamma^*)^{-1}\theta_{\mathbf{0}}D_\Gamma.$$

Proof. Solving (4-5) for F_1 we get

$$F_1(z) = \theta_{\mathbf{0}}(z)[\boldsymbol{I} - \theta_{\mathbf{0}}(z)]^{-1}.$$

Substituting this expression into the formula for the characteristic function from Theorem 4.2, we see that

$$\theta_\Gamma = -\Gamma + D_{\Gamma^*}\theta_{\mathbf{0}}[\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}]^{-1}\{\boldsymbol{I}_{\mathfrak{D}} - (\Gamma^* - \boldsymbol{I}_{\mathfrak{D}})\theta_{\mathbf{0}}[\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}]^{-1}\}^{-1}D_\Gamma. \tag{5-1}$$

We manipulate the term inside the curly brackets

$$\begin{aligned}\boldsymbol{I}_{\mathfrak{D}} - (\Gamma^* - \boldsymbol{I}_{\mathfrak{D}})\theta_{\mathbf{0}}[\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}]^{-1} &= (\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}} - (\Gamma^* - \boldsymbol{I}_{\mathfrak{D}})\theta_{\mathbf{0}})[\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}]^{-1}\\ &= (\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})[\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}]^{-1},\end{aligned}$$

so that

$$\{\boldsymbol{I}_{\mathfrak{D}} - (\Gamma^* - \boldsymbol{I}_{\mathfrak{D}})\theta_{\mathbf{0}}[\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}]^{-1}\}^{-1} = [\boldsymbol{I}_{\mathfrak{D}} - \theta_{\mathbf{0}}](\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}.$$

Substituting this back into (5-1), we get the first equation in the proposition.

The second equation is obtained similarly. □

Lemma 5.4. *For $\|\Gamma\| < 1$ we have for all $\alpha \in \mathbb{R}$*

$$D_{\Gamma^*}^\alpha\Gamma = \Gamma D_\Gamma^\alpha, \tag{5-2}$$

$$D_\Gamma^\alpha\Gamma^* = \Gamma^* D_{\Gamma^*}^\alpha, \tag{5-3}$$

where, recall $D_\Gamma := (\boldsymbol{I} - \Gamma^*\Gamma)^{1/2}$, $D_{\Gamma^*} := (\boldsymbol{I} - \Gamma\Gamma^*)^{1/2}$ are the defect operators.

Proof. Let us prove (5-2). It is trivially true for $\alpha = 2$, and by induction we get that it is true for $\alpha = 2n$, $n \in \mathbb{N}$. Since $\|\Gamma\| < 1$, the spectrum of D_Γ lies in the interval $[a, 1]$, $a = (1 - \|\Gamma\|^2)^{1/2} > 0$.

Approximating $\varphi(x) = x^\alpha$ uniformly on $[a, 1]$ by polynomials of x^2 we get (5-2).

Applying (5-2) to Γ^* we get (5-3). □

Proof of Theorem 5.1. From (5-2) we get $D_{\Gamma^*}^{-1}\Gamma D_\Gamma^{-1} = D_{\Gamma^*}^{-2}\Gamma$, so

$$\begin{aligned}\theta_\Gamma &= -\Gamma + D_{\Gamma^*}\theta_{\mathbf{0}}(\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}D_\Gamma\\ &= D_{\Gamma^*}[-D_{\Gamma^*}^{-2}\Gamma + \theta_{\mathbf{0}}(\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}]D_\Gamma\\ &= D_{\Gamma^*}^{-1}[-\Gamma + D_{\Gamma^*}^2\theta_{\mathbf{0}}(\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}]D_\Gamma\\ &= D_{\Gamma^*}^{-1}[-\Gamma(\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}}) + (\boldsymbol{I} - \Gamma\Gamma^*)\theta_{\mathbf{0}}](\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}D_\Gamma\\ &= D_{\Gamma^*}^{-1}[-\Gamma + \theta_{\mathbf{0}}](\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}D_\Gamma,\end{aligned}$$

which is exactly the first identity.

The second identity is obtained similarly, using the formula $D_{\Gamma^*}^{-1}\Gamma D_{\Gamma}^{-1} = \Gamma D_{\Gamma}^{-2}$ and taking the factor $(\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}$ out of brackets on the left. □

Proof of Theorem 5.2. Right-multiplying the first identity in Theorem 5.1 by $D_{\Gamma}^{-1}(\boldsymbol{I} - \Gamma^*\theta_{\mathbf{0}})$ we get

$$\theta_\Gamma D_{\Gamma}^{-1} - \theta_\Gamma D_{\Gamma}^{-1}\Gamma^*\theta_{\mathbf{0}} = D_{\Gamma^*}^{-1}\theta_{\mathbf{0}} - D_{\Gamma^*}^{-1}\Gamma.$$

Using identities $D_{\Gamma^*}^{-1}\Gamma = \Gamma D_{\Gamma}^{-1}$ and $D_{\Gamma}^{-1}\Gamma^* = \Gamma^* D_{\Gamma^*}^{-1}$, see Lemma 5.4, we rewrite the above equality as

$$\theta_\Gamma D_{\Gamma}^{-1} + \Gamma D_{\Gamma}^{-1} = \theta_\Gamma \Gamma^* D_{\Gamma^*}^{-1}\theta_{\mathbf{0}} + D_{\Gamma^*}^{-1}\theta_{\mathbf{0}}.$$

Right-multiplying both sides by $D_{\Gamma^*}(\theta_\Gamma\Gamma^* + \boldsymbol{I})^{-1}$ we get the first equality in the theorem.

The second one is proved similarly. □

5B. ***The defect functions $\boldsymbol{\Delta}_{\boldsymbol{\Gamma}}$ and relations between them.*** Recall that every strict contraction Γ yields a characteristic matrix-valued function θ_Γ through association with the c.n.u. contraction U_Γ. The definition of the Sz.-Nagy–Foiaş model space, see, e.g., formula (3-16), reveals immediately that the defect functions $\Delta_\Gamma = (\boldsymbol{I} - \theta_\Gamma^*\theta_\Gamma)^{1/2}$ are central objects in model theory. We express the defect function Δ_Γ in terms of $\Delta_{\mathbf{0}}$ (and Γ and $\theta_{\mathbf{0}}$).

Theorem 5.5. *The defect functions of θ_Γ and $\theta_{\mathbf{0}}$ are related by*

$$\Delta_\Gamma^2 = D_\Gamma(I - \theta_{\mathbf{0}}^*\Gamma)^{-1}\Delta_{\mathbf{0}}^2(I - \Gamma^*\theta_{\mathbf{0}})^{-1}D_\Gamma.$$

Proof. By Theorem 5.1

$$\theta_\Gamma = D_{\Gamma^*}^{-1}(\theta_{\mathbf{0}} - \Gamma)(\boldsymbol{I}_{\mathfrak{D}} - \Gamma^*\theta_{\mathbf{0}})^{-1}D_\Gamma,$$

so $\theta_\Gamma^*\theta_\Gamma = A^*BA$, where

$$A = (\boldsymbol{I} - \Gamma^*\theta_{\mathbf{0}})D_\Gamma, \quad B = (\theta_{\mathbf{0}}^* - \Gamma^*)D_{\Gamma^*}^{-2}(\theta_{\mathbf{0}} - \Gamma).$$

Then $\Delta_\Gamma = \boldsymbol{I} - \theta_\Gamma^*\theta_\Gamma = A^*XA$, where

$$\begin{aligned} X &= (A^*)^{-1}A^{-1} - B = (\boldsymbol{I} - \theta_{\mathbf{0}}^*\Gamma)D_\Gamma^{-2}(\boldsymbol{I} - \Gamma^*\theta_{\mathbf{0}}) - (\theta_{\mathbf{0}}^* - \Gamma^*)D_{\Gamma^*}^{-2}(\theta_{\mathbf{0}} - \Gamma) \\ &= D_\Gamma^{-2} - \theta_{\mathbf{0}}^*\Gamma D_\Gamma^{-2} - D_\Gamma^{-2}\Gamma^*\theta_{\mathbf{0}} + \theta_{\mathbf{0}}^*\Gamma D_\Gamma^{-2}\Gamma^*\theta_{\mathbf{0}} - \theta_{\mathbf{0}}^* D_{\Gamma^*}^{-2}\theta_{\mathbf{0}} + \Gamma^* D_{\Gamma^*}^{-2}\theta_{\mathbf{0}} + \theta_{\mathbf{0}}^* D_{\Gamma^*}^{-2}\Gamma - \Gamma^* D_{\Gamma^*}^{-2}\Gamma. \end{aligned}$$

It follows from Lemma 5.4 that $D_\Gamma^{-2}\Gamma^* = \Gamma^* D_{\Gamma^*}^{-2}$ and that $\Gamma^* D_\Gamma^{-2} = D_{\Gamma^*}^{-2}\Gamma$, so in the above identity we have cancellation of nonsymmetric terms,

$$-\theta_{\mathbf{0}}^*\Gamma D_\Gamma^{-2} - D_\Gamma^{-2}\Gamma^*\theta_{\mathbf{0}} + \Gamma^* D_{\Gamma^*}^{-2}\theta_{\mathbf{0}} + \theta_{\mathbf{0}}^* D_{\Gamma^*}^{-2}\Gamma = 0.$$

Therefore

$$\begin{aligned} X &= D_\Gamma^{-2} + \theta_{\mathbf{0}}^*\Gamma D_\Gamma^{-2}\Gamma^*\theta_{\mathbf{0}} - \theta_{\mathbf{0}}^* D_{\Gamma^*}^{-2}\theta_{\mathbf{0}} - \Gamma^* D_{\Gamma^*}^{-2}\Gamma \\ &= D_\Gamma^{-2} + \theta_{\mathbf{0}}^* D_{\Gamma^*}^{-2}\Gamma\Gamma^*\theta_{\mathbf{0}} - \theta_{\mathbf{0}}^* D_{\Gamma^*}^{-2}\theta_{\mathbf{0}} - D_\Gamma^{-2}\Gamma^*\Gamma \\ &= D_\Gamma^{-2}(\boldsymbol{I} - \Gamma^*\Gamma) + \theta_{\mathbf{0}}^* D_\Gamma^{-2}(\Gamma^*\Gamma - \boldsymbol{I})\theta_{\mathbf{0}} = \boldsymbol{I} - \theta_{\mathbf{0}}^*\theta_{\mathbf{0}} = \Delta_{\mathbf{0}}. \end{aligned}$$

Thus we get $\Delta_\Gamma = A^*\Delta_{\mathbf{0}}A$, which is exactly the conclusion of the theorem. □

5C. ***Multiplicity of the absolutely continuous spectrum.*** It is well known that the Sz.-Nagy–Foiaş model space reduces to the familiar one-story setting with $\mathcal{K}_\theta = H^2(\mathfrak{D}_*) \ominus \theta H^2(\mathfrak{D})$ when θ is inner. Indeed, for inner θ the nontangential boundary values of the defect $\Delta(\xi) = (\boldsymbol{I} - \theta^*(\xi)\theta(\xi))^{1/2} = 0$ for Lebesgue a.e. $\xi \in \mathbb{T}$. So, the second component of the Sz.-Nagy–Foiaş model space collapses completely.

Here we provide a finer result that reveals the matrix-valued weight function and the multiplicity of U's absolutely continuous part.

Before we formulate the statement, we recall some terminology. First, we Lebesgue decompose the (scalar) measure $\mathrm{d}\mu = \mathrm{d}\mu_{\mathrm{ac}} + \mathrm{d}\mu_{\mathrm{sing}}$. The absolutely continuous part of U is unitarily equivalent to the multiplication by the independent variable ξ on the von Neumann direct integral $\mathcal{H}_{\mathrm{ac}} = \int_{\mathbb{T}}^{\oplus} E(\xi)\,\mathrm{d}\mu_{\mathrm{ac}}(\xi)$. Note that the dimension of $E(\xi)$ is the multiplicity function of the spectrum.

Let w denote the density of the absolutely continuous part of μ, i.e., $\mathrm{d}\mu_{\mathrm{ac}}(\xi) = w(\xi)\,\mathrm{d}m(\xi)$. Then the matrix-valued function $\xi \mapsto B^*(\xi)B(\xi)w(\xi)$ is the absolutely continuous part of the matrix-valued measure $B^*B\mu$.

Theorem 5.6. *The defect function $\Delta_{\mathbf{0}}$ of $\theta_{\mathbf{0}}$ and the absolutely continuous part B^*Bw of the matrix-valued measure $B^*B\mu$ are related by*

$$(\boldsymbol{I} - \theta_{\mathbf{0}}^*(\xi))B^*(\xi)B(\xi)w(\xi)(\boldsymbol{I} - \theta_{\mathbf{0}}(\xi)) = (\Delta_{\mathbf{0}}(\xi))^2 \tag{5-4}$$

for Lebesgue a.e. $\xi \in \mathbb{T}$.

The function $\boldsymbol{I} - \theta_{\mathbf{0}}$ is invertible a.e. on $\mathbb{T}$, so the multiplicity of the absolutely continuous part of μ is given by

$$\dim E(\xi) = \operatorname{rank}(\boldsymbol{I} - \theta_{\mathbf{0}}^*(\xi)\theta_{\mathbf{0}}(\xi)) = \operatorname{rank} \Delta_{\mathbf{0}}(\xi), \tag{5-5}$$

of course, with respect to Lebesgue a.e. $\xi \in \mathbb{T}$.

Combining (5-5) with Theorem 5.5 we obtain:

Corollary 5.7. *For Lebesgue a.e. $\xi \in \mathbb{T}$ we have* $\dim E(\xi) = \operatorname{rank} \Delta_\Gamma(\xi)$ *for all strict contractions* Γ.

Another immediate consequence is the following:

Corollary 5.8. *The operator U has no absolutely continuous part on a Borel set $B \subset \mathbb{T}$ if and only if $\theta_{\mathbf{0}}(\xi)$ (or, equivalently, $\theta_\Gamma(\xi)$ for all strict contractions Γ) is unitary for Lebesgue a.e. $\xi \in B$.*

This corollary is closely related to the main result of [Douglas and Liaw 2013, Theorem 3.1]. Interestingly, it appears that their proof of that result cannot be refined to yield our current result (Theorem 5.6).

Corollary 5.9. *In particular, we confirm that the following are equivalent*:

(i) *U is purely singular.*

(ii) *$\theta_\Gamma(\xi)$ is inner for one (equivalently any) strict contraction Γ.*

(iii) *$\Delta_\Gamma \equiv \mathbf{0}$ for one (equivalently any) strict contraction Γ.*

(iv) *The second story of the Sz.-Nagy–Foiaş model space collapses (and we are dealing with the model space $\mathcal{K}_{\theta_\Gamma} = H^2(\mathbb{C}^d) \ominus \theta_\Gamma H^2(\mathbb{C}^d)$ for one (equivalently any) strict contraction Γ).*

Proof of Theorem 5.6. Take $\Gamma \equiv \mathbf{0}$. Solving (4-6) for F_2 we see

$$F_2(z) = [\boldsymbol{I} + \theta_{\mathbf{0}}(z)][\boldsymbol{I} - \theta_{\mathbf{0}}(z)]^{-1}.$$

Let $\mathcal{P}(B^* B\mu)$ denote the Poisson extension of the matrix-valued measure $B^* B\mu$ to the unit disc $\mathbb{D}$. Since $F_2 = \mathcal{C}_2 B^* B\mu$, we can see that $\mathcal{P}(B^* B\mu) = \operatorname{Re} F_2$ on $\mathbb{D}$, so

$$\mathcal{P}(B^* B\mu) = \operatorname{Re} F_2 = \operatorname{Re}[(\boldsymbol{I} + \theta_{\mathbf{0}})(\boldsymbol{I} - \theta_{\mathbf{0}})^{-1}].$$

Standard computations yield

$$\begin{aligned}
\mathcal{P}(B^* B\mu) &= \operatorname{Re}[(\boldsymbol{I} + \theta_{\mathbf{0}})(\boldsymbol{I} - \theta_{\mathbf{0}})^{-1}] = \tfrac{1}{2}[(\boldsymbol{I} + \theta_{\mathbf{0}})(\boldsymbol{I} - \theta_{\mathbf{0}})^{-1} + (\boldsymbol{I} - \theta_{\mathbf{0}}^*)^{-1}(\boldsymbol{I} + \theta_{\mathbf{0}}^*)] \\
&= \tfrac{1}{2}(\boldsymbol{I} - \theta_{\mathbf{0}}^*)^{-1}[(\boldsymbol{I} - \theta_{\mathbf{0}}^*)(\boldsymbol{I} + \theta_{\mathbf{0}}) + (\boldsymbol{I} + \theta_{\mathbf{0}}^*)(\boldsymbol{I} - \theta_{\mathbf{0}})](\boldsymbol{I} - \theta_{\mathbf{0}})^{-1} \\
&= \tfrac{1}{2}(\boldsymbol{I} - \theta_{\mathbf{0}}^*)^{-1}[\boldsymbol{I} - \theta_{\mathbf{0}}^*\theta_{\mathbf{0}}](\boldsymbol{I} - \theta_{\mathbf{0}})^{-1} = (\boldsymbol{I} - \theta_{\mathbf{0}}^*)^{-1} \operatorname{Re}[\boldsymbol{I} - \theta_{\mathbf{0}}^*\theta_{\mathbf{0}}](\boldsymbol{I} - \theta_{\mathbf{0}})^{-1} \\
&= (\boldsymbol{I} - \theta_{\mathbf{0}}^*)^{-1}[\boldsymbol{I} - \theta_{\mathbf{0}}^*\theta_{\mathbf{0}}](\boldsymbol{I} - \theta_{\mathbf{0}})^{-1}
\end{aligned}$$

on $\mathbb{D}$. Note that for any characteristic function θ and $z \in \mathbb{D}$ the matrix $\theta(z)$ is a strict contraction, so in our case $\boldsymbol{I} - \theta_{\mathbf{0}}$ is invertible on $\mathbb{D}$, and all computations are justified.

We can rewrite the above identity as

$$(\boldsymbol{I} - \theta_{\mathbf{0}})^* \mathcal{P}(B^* B\mu)(\boldsymbol{I} - \theta_{\mathbf{0}}) = \boldsymbol{I} - \theta_{\mathbf{0}}^*\theta_{\mathbf{0}},$$

and taking the nontangential boundary values we get (5-4). Here we used the Fatou lemma, see, e.g., [Nikolski 2002a, Theorem 3.11.7], which says that for a complex measure τ the nontangential boundary values of its Poisson extension $\mathcal{P}\tau$ coincide a.e. with the density of the absolutely continuous part of τ; applying this lemma entrywise we get what we need in the left-hand side.

To see that the boundary values of $\boldsymbol{I} - \theta_{\mathbf{0}}$ are invertible a.e. on $\mathbb{T}$ we notice that $z \mapsto \det(\boldsymbol{I} - \theta_{\mathbf{0}}(z))$ is a bounded analytic function on $\mathbb{D}$, so its boundary values are nonzero a.e. on $\mathbb{T}$. □

6. What is wrong with the universal representation formula and what to do about it?

There are several things that are not completely satisfactory with the universal representation formula given by Theorem 2.4.

First of all, it is defined only on functions of form hb, where $h \in C^1$ is a scalar function and $b \in \operatorname{Ran} \boldsymbol{B}$. Of course, one can then define it on a dense set, for example on the dense set of linear combinations $f = \sum_k h_k, b_k$, where b_k are columns of the matrix B, $b_k = \boldsymbol{B}e_k$, and $h_k \in C^1(\mathbb{T})$. But the use of functions b (or b_k) in the representation is a bit bothersome, especially taking into account that the representation $f = \sum_k h_k b_k$ is not always unique. So, it would be a good idea to get rid of the function b.

The second thing is that while the representation formula looks like a singular integral operator (Cauchy transform), it is not represented as a classical singular integral operator, so it is not especially clear if the (well-developed) theory of such operators applies in our case. So, we would like to represent the operator in more classical way.

Defining $C_1(z) := C_*(z) - zC(z)$ and using the formal Cauchy-type expression

$$(T^{B^*\mu} f)(z) = \int_{\mathbb{T}} \frac{1}{1 - z\bar{\xi}} B^*(\xi) f(\xi)\, \mathrm{d}\mu(\xi),$$

we can, performing formal algebraic manipulations, rewrite (2-10) as

$$(\Phi^* hb)(z) = C_1(z)(T^{B^*\mu} hb)(z) + h(z)[C_*(z)\boldsymbol{B}^* b - C_1(z)(T^{B^*\mu} b)(z)], \quad z \in \mathbb{T}. \tag{6-1}$$

So, is it possible to turn these formal manipulations into meaningful mathematics? And the answer is “yes”: the formula (6-1) gives the representation of Φ^* if one interprets $T^{B^*\mu} f$ as the boundary values of the Cauchy transform $\mathcal{C}[B^* f\mu](z)$, $z \notin \mathbb{T}$; see the definition in the next section.

In the next section (Section 7) we present necessary facts about the (vector-valued) Cauchy transform and its regularization that will allow us to interpret and justify the formal expression (6-1). We will complete this justification in Section 8; see (8-12). This representation is a universal one, meaning that it works in any transcription of the model, but still involves the function $b \in \operatorname{Ran} \boldsymbol{B}$.

The function b is kind of eliminated in Proposition 8.4 below, and as it is usually happens in the theory of singular integral operators, the operator Φ^* splits into the singular integral part (weighted boundary values of the Cauchy transform) and the multiplication part. The function b becomes hidden in the multiplication part, and at first glance it is not clear why this part is well-defined.

Thus the representation given by Proposition 8.4 is still not completely satisfactory (the price one pays for the universality), but it is a step towards obtaining a nice representation for a fixed transcription of a model. Thus we were able to obtain a precise and unambiguous representation of Φ^* in the Sz.-Nagy–Foiaş transcription; see Theorem 8.1 which is the main result of Section 8.

7. Singular integral operators

7A. ***Cauchy-type integrals.*** For a finite (signed or even complex-valued) measure ν on $\mathbb{T}$, its Cauchy transform $\mathcal{C}\nu$ is defined as

$$\mathcal{C}\nu(z) = \mathcal{C}[\nu](z) = \int_{\mathbb{T}} \frac{\mathrm{d}\nu(\xi)}{1 - \bar{\xi} z}, \quad z \in \mathbb{C} \setminus \mathbb{T}.$$

It is a classical fact that $\mathcal{C}\nu(z)$ has nontangential boundary values a.e. on $\mathbb{T}$ as $z \to z_0 \in \mathbb{T}$ from the inside and from the outside of the disc $\mathbb{D}$. So, given a finite positive Borel measure μ one can define operators $T^{\mu}_{\pm}$ from $L^1(\mu; E)$ to the space of measurable functions on $\mathbb{T}$ as the nontangential boundary values from inside and outside of the unit disc $\mathbb{D}$,

$$(T^{\mu}_{+} f)(z_0) = \text{n.t.-}\lim_{\substack{z \to z_0 \\ z \in \mathbb{D}}} \mathcal{C}[f\mu](z), \quad (T^{\mu}_{-} f)(z_0) = \text{n.t.-}\lim_{\substack{z \to z_0 \\ z \notin \overline{\mathbb{D}}}} \mathcal{C}[f\mu](z).$$

One can also define the regularized operators T^{μ}_{r}, $r \in (0, \infty) \setminus \{1\}$, and the restriction of $\mathcal{C}[f\mu]$ to the circle of radius r,

$$T^{\mu}_{r} f(z) = \mathcal{C}[f\mu](rz).$$

Everything can be extended to the case of vector- and matrix-valued measures; there are some technical details that should be taken care of in the infinite-dimensional case, but in our case everything is finite-dimensional ($\dim E \le d < \infty$), so the generalization is pretty straightforward.

So, given a (finite, positive) scalar measure μ and a matrix-valued function B^* (with entries in $L^2(\mu)$) and vector-valued function $f \in L^2(\mu; E)$ we can define $T_\pm^{B^*\mu} f$ and $T_r^{B^*\mu} f$ as the nontangential boundary values and the restriction to the circle of radius r respectively of the Cauchy transform $\mathcal{C}[B^* f\mu](z)$. Modulo slight abuse of notation this notation agrees with the accepted notation for the scalar case.

In what follows the function B^* will be the function B^* from Theorem 2.4.

7B. ***Uniform boundedness of the boundary Cauchy operator and its regularization.*** For a finite Borel measure ν on $\mathbb{T}$ and $n \in \mathbb{Z}$ define

$$P_n\nu(z) = \begin{cases} \sum_{k=0}^{n} \hat{\nu}(k) z^k, & n \ge 0, \\ \sum_{k=n}^{-1} \hat{\nu}(k) z^k, & n < 0; \end{cases}$$

here $\hat{\nu}(k)$ is the Fourier coefficient of ν, $\hat{\nu}(k) = \int_{\mathbb{T}} \xi^{-k}\, \mathrm{d}\nu(\xi)$.

Recall that $C_1(z) := C_*(z) - zC(z)$, where C_* and C are from Theorem 2.4.

Recall that if W is a matrix-valued weight (i.e., a function whose values $W(\xi)$ are positive semidefinite operators on a finite-dimensional space H), then the norm in the weighted space $L^2(W; H)$ is defined as

$$\|f\|^2_{L^2(W;H)} = \int_{\mathbb{T}} (W(\xi) f(\xi), f(\xi))_H \,\mathrm{d}m(\xi).$$

We are working with the model space $\mathcal{K}_\theta$ which is a subspace of a weighted space $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$ (the weight could be trivial, $W \equiv \boldsymbol{I}$, as in the case of Sz.-Nagy–Foiaş model).

Define $\widetilde{C}_1 := W^{1/2} C_1$. The function $\widetilde{C}_1^* \widetilde{C}_1$ is a matrix-valued weight, whose values are operators on $\mathfrak{D}_* \oplus \mathfrak{D}$, so we can define the weighted space $L^2(\widetilde{C}_1^* \widetilde{C}_1) = L^2(\widetilde{C}_1^* \widetilde{C}_1; \mathfrak{D}_* \oplus \mathfrak{D})$. Note that

$$\|f\|_{L^2(\widetilde{C}_1^* \widetilde{C}_1)} = \|\widetilde{C}_1 f\|_{L^2(\mathfrak{D}_* \oplus \mathfrak{D})} = \|C_1 f\|_{L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})}.$$

Lemma 7.1. *The operators $P_n^{B^*\mu} : \mathcal{H} \subset L^2(\mu; E) \to L^2(\widetilde{C}_1^* \widetilde{C}_1; \mathfrak{D}_* \oplus \mathfrak{D})$ defined by*

$$P_n^{B^*\mu} f := P_n(B^* f\mu), \quad n \in \mathbb{Z},$$

are bounded uniformly in n with norm at most 2; *i.e.*,

$$\|\widetilde{C}_1 P_n(B^*\mu f)\|_{L^2(\mathfrak{D}_* \oplus \mathfrak{D})} \le 2\|f\|_{L^2(\mu;E)}.$$

Proof. The columns b_k of B are in $\mathcal{H} \subset L^2(\mu; E)$, so $B^* f\mu \in L^1(\mu; \mathfrak{D})$, and therefore the operators $P_n^{B^*\mu}$ are bounded operators $\mathcal{H} \to L^2(\mathfrak{D})$. It follows from Lemma 3.4 that $\|\widetilde{C}_1\|_\infty \le 2$, so the operators $f \mapsto \widetilde{C}_1 P_n^{B^*\mu} f$ are bounded operators $\mathcal{H} \to L^2(\mathfrak{D}_* \oplus \mathfrak{D})$ (notice that we do not claim the uniform-in-n bound here). Therefore, it is sufficient to check the uniform boundedness on a dense set.

Take $f = hb$, where $b \in \operatorname{Ran} \boldsymbol{B}$ and $h \in C^1(\mathbb{T})$ is scalar-valued. Then for $n \in \mathbb{Z}$ we have by Theorem 2.4

$$\begin{aligned}\Phi^* f - z^n \Phi^*(\bar{\xi}^n f) &= C_1(z) \int_{\mathbb{T}} \frac{h(\xi) - h(z)}{1 - \bar{\xi} z} B^* b \,\mathrm{d}\mu(\xi) - z^n C_1(z) \int_{\mathbb{T}} \frac{\bar{\xi}^n h(\xi) - \bar{z}^n h(z)}{1 - \bar{\xi} z} B^* b \,\mathrm{d}\mu(\xi) \\ &= C_1(z) \int_{\mathbb{T}} \frac{1 - (\bar{\xi} z)^n}{1 - \bar{\xi} z} B^* h b \,\mathrm{d}\mu(\xi).\end{aligned}$$

Expressing $(1 - (\bar{\xi} z)^n)/(1 - \bar{\xi} z)$ as a sum of geometric series we get that for $f = hb,\ h \in \mathbb{C}^1(\mathbb{T})$,

$$\Phi^* f - z^n \Phi^*(\bar{\xi}^n f) = \begin{cases} C_1 P_{n-1}(B^* f\mu), & n \geq 1, \\ -C_1 P_n(B^* f\mu), & n < 0. \end{cases}$$

By linearity the above identity holds for a dense set of linear combinations $f = \sum_k h_k b_k,\ h_k \in C^1(\mathbb{T})$. The operators $\Phi^* : \mathcal{H} \to \mathcal{K}_\theta \subset L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$ are bounded (unitary) operators, so the desired estimate holds on the above dense set. □

For a measure ν on $\mathbb{T}$ let $T_r \nu$ be the restriction of the Cauchy transform of ν to the circle of radius $r \neq 1$,

$$T_r \nu(z) = \int_{\mathbb{T}} \frac{\mathrm{d}\nu(\xi)}{1 - r\bar{\xi} z}, \quad z \in \mathbb{T}.$$

Define operators $T_r^{B^*\mu}$ on $L^2(\mu; E)$ as

$$T_r^{B^*\mu} f = T_r(B^* f\mu).$$

The lemma below is an immediate corollary of the above Lemma 7.1.

Lemma 7.2. *The operators* $T_r^{B^*\mu} : \mathcal{H} \subset L^2(\mu; E) \to L^2(\widetilde{C}_1^* \widetilde{C}_1; \mathfrak{D}_* \oplus \mathfrak{D})$ *are bounded uniformly in* r *with norm at most* 2; *i.e.*,

$$\|\widetilde{C}_1 T_r^{B^*\mu} f\|_{L^2(\mathfrak{D}_* \oplus \mathfrak{D})} \leq 2 \|f\|_{L^2(\mu; E)}.$$

Proof. The result follows immediately from Lemma 7.1, since the operators $T_r^{B^*\mu}$ can be represented as averages of operators $P_n^{B^*\mu}$,

$$T_r^{B^*\mu} = \begin{cases} \sum_{n=0}^{\infty} (r^n - r^{n+1}) P_n^{B^*\mu}, & 0 < r < 1, \\ \sum_{n=1}^{\infty} (r^{-n+1} - r^{-n}) P_{-n}^{B^*\mu}, & r > 1. \end{cases}$$ □

Using uniform boundedness of the operators $\widetilde{C}_1 T_r^{B^*\mu}$ (Lemma 7.2) and existence of nontangential boundary values $T_\pm^{B^*\mu} f$ we can get the convergence of operators $\widetilde{C}_1 T_r^{B^*\mu}$ in the weak operator topology.

Proposition 7.3. *The operators* $\widetilde{C}_1 T_\pm^{B^*\mu} : \mathcal{H} \subset L^2(\mu; E) \to L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$ *are bounded and*

$$C_1 T_\pm^{B^*\mu} = \text{w.o.t.-} \lim_{r \to 1^\mp} C_1 T_r^{B^*\mu}.$$

Proof. We want to show that for any $f \in \mathcal{H} \subset L^2(\mu; E)$

$$C_1 T_\pm^{B^*\mu} f = \text{w-} \lim_{r \to 1^\mp} C_1 T_r^{B^*\mu} f,$$

where the limit is in the weak topology of $L^2(W;\mathfrak{D}_*\oplus\mathfrak{D})$. This is equivalent to

$$\widetilde{C}_1 T_{\pm}^{B^*\mu} f = \text{w-}\lim_{r\to 1^{\mp}} \widetilde{C}_1 T_r^{B^*\mu} f,$$

with the limit being in the weak topology of $L^2(\mathfrak{D}_*\oplus\mathfrak{D})$.

Let us prove this identity for $\widetilde{C}_1 T_+^{B^*\mu} f$. Assume that for some $f\in L^2(\mu;E)$

$$\widetilde{C}_1 T_{+}^{B^*\mu} f \neq \text{w-}\lim_{r\to 1^{-}} \widetilde{C}_1 T_r^{B^*\mu} f.$$

Then for some $h\in L^2(\mathfrak{D}_*\oplus\mathfrak{D})$

$$(\widetilde{C}_1 T_r^{B^*\mu} f, h)_{L^2(\mathfrak{D}_*\oplus\mathfrak{D})} \nrightarrow (\widetilde{C}_1 T_+^{B^*\mu} f, h)_{L^2(\mathfrak{D}_*\oplus\mathfrak{D})} \quad \text{as } r\to 1^-, \tag{7-1}$$

so there exists a sequence $r_k \nearrow 1$ such that

$$\lim_{k\to\infty} (\widetilde{C}_1 T_{r_k}^{B^*\mu} f, h)_{L^2(\mathfrak{D}_*\oplus\mathfrak{D})} \neq (\widetilde{C}_1 T_+^{B^*\mu} f, h)_{L^2(\mathfrak{D}_*\oplus\mathfrak{D})};$$

note that taking a subsequence we can assume without loss of generality that the limit in the left-hand side exists.

Taking a subsequence again, we can assume without loss of generality that $\widetilde{C}_1 T_{r_k}^{B^*\mu} f \to g$ in the weak topology, and (7-1) implies $g \neq \widetilde{C}_1 T_+^{B^*\mu} f$.

The existence of nontangential boundary values and the definition of $T_+^{B^*\mu}$ implies $\widetilde{C}_1 T_{r_k}^{B^*\mu} f \to \widetilde{C}_1 T_+^{B^*\mu} f$ a.e. on $\mathbb{T}$. But as [Liaw and Treil 2009, Lemma 3.3] asserts, if $f_n\to f$ a.e. and $f_n\to g$ in the weak topology of L^2, then $f=g$, so we arrive at a contradiction.

Note, that in [Liaw and Treil 2009, Lemma 3.3] everything was stated for scalar functions, but applying this scalar lemma componentwise we immediately get the same result for $L^2(\mu;E)$ with values in a separable Hilbert space. □

8. Adjoint Clark operator in Sz.-Nagy–Foiaş transcription

The main result of this section is Theorem 8.1 below, giving a formula for the adjoint Clark operator Φ^*.

Denote by F the Cauchy transform of the matrix-valued measure $B^*B\mu$,

$$F(z) = \mathcal{C}[B^*B\mu](z) = \int_{\mathbb{T}} \frac{1}{1-z\bar{\xi}} B^*(\xi)B(\xi)\,\mathrm{d}\mu(\xi), \quad z\in\mathbb{D}, \tag{8-1}$$

and let us use the same symbol for its nontangential boundary values, which exist a.e. on $\mathbb{T}$. Using the operator $T_+^{B^*\mu}$ introduced in the previous section, we give the following formula for Φ^*.

Theorem 8.1. *The adjoint Clark operator in Sz.-Nagy–Foiaş transcription reduces to*

$$\Phi^* f = \begin{pmatrix} 0 \\ \Psi_2 \end{pmatrix} f + \begin{pmatrix} (\boldsymbol{I}+\theta_\Gamma\Gamma^*)D_{\Gamma^*}^{-1}F^{-1} \\ \Delta_\Gamma D_\Gamma^{-1}(\Gamma^*-\boldsymbol{I}) \end{pmatrix} T_+^{B^*\mu} f, \quad f\in\mathcal{H}, \tag{8-2}$$

with $\Psi_2(z) = \widetilde{\Psi}_2(z)R(z)$, *where*

$$\begin{aligned}\widetilde{\Psi}_2(z) &= \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* + (\boldsymbol{I} - \Gamma^*)F(z)) \\ &= \Delta_\Gamma D_\Gamma^{-1}(\boldsymbol{I} - \Gamma^*\theta_{\mathbf{0}}(z))F(z) \quad \textit{a.e. on } \mathbb{T}, \end{aligned} \tag{8-3}$$

and R is a measurable right inverse for the matrix-valued function B.

Remark. When $d = 1$, this result reduces to [Liaw and Treil 2016, equation (4.5)].

Remark 8.2. As one should expect, the matrix-valued function Ψ_2 does not depend on the choice of the right inverse R. To prove this, it is sufficient to show that $\ker B(z) \subset \ker \widetilde{\Psi}_2(z)$ a.e., which follows from the proposition below.

Proposition 8.3. *For $\widetilde{\Psi}_2$ defined above in* (8-16) *and w being the density of μ_{ac} we have*

$$\widetilde{\Psi}_2(\xi)^*\widetilde{\Psi}_2(\xi) = F(\xi)^*\Delta_{\mathbf{0}}(\xi)^2F(\xi) = B(\xi)^*B(\xi)w(\xi) \quad \mu_{\mathrm{ac}}\text{-}a.e., \tag{8-4}$$

and so

$$\Psi_2(\xi)^*\Psi_2(\xi) = w(\xi)\boldsymbol{I}_{E(\xi)} \quad \mu_{\mathrm{ac}}\text{-}a.e. \tag{8-5}$$

Proof. Since $\Psi_2 = \widetilde{\Psi}_2 R$, (8-5) follows immediately from (8-4).

To prove (8-4), consider first the case, $\Gamma = \mathbf{0}$. In this case $\widetilde{\Psi} = \Delta_{\mathbf{0}}F$, so

$$\begin{aligned}\widetilde{\Psi}_2^*\widetilde{\Psi}_2 &= F^*\Delta_{\mathbf{0}}^2F = (\boldsymbol{I} - \theta_{\mathbf{0}}^*)^{-1}\Delta_{\mathbf{0}}^2(\boldsymbol{I} - \theta_{\mathbf{0}})^{-1} \\ &= B^*Bw \qquad \text{by (5-4).}\end{aligned} \tag{8-6}$$

Consider now the case of general Γ. We get

$$\begin{aligned}\widetilde{\Psi}_2^*\widetilde{\Psi}_2 &= F^*(\boldsymbol{I} - \theta_{\mathbf{0}}^*\Gamma)D_\Gamma^{-1}\Delta_\Gamma^2 D_\Gamma^{-1}(\boldsymbol{I} - \Gamma^*\theta_{\mathbf{0}})F \\ &= F^*\Delta_{\mathbf{0}}^2F \qquad \text{by Theorem 5.5} \\ &= B^*Bw \qquad \text{by (8-6).}\end{aligned}$$

□

8A. ***A preliminary formula.*** We start proving Theorem 8.1 by first proving this preliminary result, which holds for any transcription of the model. Below, the matrix-valued functions $\boldsymbol{C}_*$ and $\boldsymbol{C}$ are from Theorem 2.4, and $\boldsymbol{C}_1(z) := \boldsymbol{C}_*(z) - z\boldsymbol{C}(z)$.

Proposition 8.4. *The adjoint Clark operator is represented for $f \in \mathcal{H} \subset L^2(\mu; E)$ by*

$$(\Phi^* f)(z) = \boldsymbol{C}_1(z)(T_\pm^{\boldsymbol{B}^*\mu} f)(z) + \Psi_\pm(z)f(z), \quad z \in \mathbb{T}, \tag{8-7}$$

where the matrix-functions $\Psi_\pm$, $\Psi_\pm(z) : E(z) \to \mathbb{C}^{2d} = \mathfrak{D}_* \oplus \mathfrak{D}$ are defined via the identities

$$\Psi_\pm(z)b(z) := \boldsymbol{C}_*(z)\boldsymbol{B}^*b - \boldsymbol{C}_1(z)(T_\pm^{\boldsymbol{B}^*\mu}b)(z), \quad b \in \operatorname{Ran} \boldsymbol{B}; \tag{8-8}$$

here two choices of sign (*the same sign for all terms*) *give two different representation formulas.*

Remark. When $d = 1$ and $b \equiv 1$ this alternative representation formula reduces to a formula that occurs in the proof of [Liaw and Treil 2016, Theorem 4.7].

Remark. It is clear that relations (8-8) with $b = b_k,\ k = 1, 2, \dots, d$, completely define the matrix-valued function Ψ. However, it is not immediately clear that such a function Ψ exists; the existence of Ψ will be shown in the proof.

Recalling the definition (8-1) of the function F, we can see that $\Psi(z)b_k(z)$ can be given as the (nontangential) boundary values of the vector-valued function

$$C_*(z)e_k - C_1(z)F(z)e_k, \quad z \in \mathbb{D}, \tag{8-9}$$

where $e_1, e_2, \dots, e_d$ is the standard orthonormal basis in $\mathbb{C}^d$.

Proof of Proposition 8.4. Let us first show the result for functions of the form $f = hb \in L^2(\mu; E)$, where $b \in \operatorname{Ran} \boldsymbol{B}$ and h is a scalar function. We want to show that

$$(\Phi^* hb)(z) = C_1(z)(T_{\pm}^{\boldsymbol{B}^*\mu} hb)(z) + h(z)\psi_b^{\pm}(z), \quad z \in \mathbb{T}, \tag{8-10}$$

where

$$\psi_b^{\pm}(z) := C_*(z)\boldsymbol{B}^* b - C_1(z)(T_{\pm}^{\boldsymbol{B}^*\mu} b)(z).$$

First note that (2-10) implies that for $b \in \operatorname{Ran} \boldsymbol{B}$

$$\Phi^* b(z) = C_*(z)\boldsymbol{B}^* b.$$

Observe that for (scalar) $h \in C^1$ we have uniform-on-z convergence, $z \in \mathbb{T}$, as $r \to 1^{\mp}$:

$$\int_{\mathbb{T}} \frac{h(\xi) - h(z)}{1 - rz\bar{\xi}} B^*(\xi)b(\xi)\,\mathrm{d}\mu(\xi) \rightrightarrows \int_{\mathbb{T}} \frac{h(\xi) - h(z)}{1 - z\bar{\xi}} B^*(\xi)b(\xi)\,\mathrm{d}\mu(\xi). \tag{8-11}$$

Multiplying both sides by $C_1(z)$ we get in the left-hand side $C_1(z)(T_r^{\boldsymbol{B}^*\mu} hb)(z) - h(z)C_1(z)(T_r^{\boldsymbol{B}^*\mu} b)(z)$, and in the right-hand side the part with the integral in the representation (2-10).

Recall that the model space $\mathcal{K}_{\theta_\Gamma}$ is a subspace of a weighted space $L^2(W, \mathfrak{D}_* \oplus \mathfrak{D})$. Uniform convergence in (8-11) implies the convergence in $L^2(\mathfrak{D}_* \oplus \mathfrak{D})$, and by Lemma 3.4 the multiplications by C_* and C_1 are bounded operators from $L^2(\mathfrak{D})$ to $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$. Thus (because h is bounded)

$$hC_*\boldsymbol{B}^* b + C_1 T_r^{\boldsymbol{B}^*\mu} hb - hC_1 T_r^{\boldsymbol{B}^*\mu} b \to \Phi^* hb$$

as $r \to 1^{\mp}$ in the norm of $L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$. By Proposition 7.3, $C_1 T_r^{\boldsymbol{B}^*\mu} \to C_1 T_{\pm}^{\boldsymbol{B}^*\mu}$ in weak operator topology as $r \to 1^{\mp}$, so

$$\Phi^* hb = C_1 T_{\pm}^{\boldsymbol{B}^*\mu} hb + hC_*\boldsymbol{B}^* b - hC_1 T_{\pm}^{\boldsymbol{B}^*\mu} b, \tag{8-12}$$

which immediately implies (8-10). Thus, (8-10) is proved for $h \in C^1(\mathbb{T})$.

To get (8-12), and so (8-10) for general h such that $hb \in L^2(\mu; E)$ (recall that $b \in \operatorname{Ran} \boldsymbol{B}$), we use the standard approximation argument: the operators $\Phi^*, C_1 T_{\pm}^{\boldsymbol{B}^*\mu} : \mathcal{H} \to L^2(W; \mathfrak{D}_* \oplus \mathfrak{D})$ are bounded, and therefore for a fixed $b \in \operatorname{Ran} \boldsymbol{B}$ the operators $hb \mapsto h\psi_b^{\pm}$ (which are defined initially on a submanifold of $\mathcal{H}$ consisting of functions of the form hb, $h \in C^1(\mathbb{T})$) are bounded (as a difference of two bounded operators). Approximating in $L^2(\mu; E)$ the function hb by functions $h_n b$, $h_n \in C^1(\mathbb{T})$ we get (8-12) and (8-10) for general h.

Let us now prove the existence of Ψ. Consider the (bounded) linear operator $\Phi^* - C_1 T^{\boldsymbol{B}^*\mu}$. We know that for $f = hb \in L^2(\mu; E)$ with $b \in \operatorname{Ran} \boldsymbol{B}$ and scalar h

$$(\Phi^* - C_1 T_{\pm}^{\boldsymbol{B}^*\mu}) hb = h\psi_b^{\pm},$$

so on functions $f = hb$ the operators $\Phi^* - C_1 T_{\pm}^{\boldsymbol{B}^*\mu}$ intertwine the multiplication operators M_ξ and M_z. Since linear combinations of functions $h_k b_k$ are dense in $\mathcal{H}$, we conclude that the operators $\Phi^* - C_1 T_{\pm}^{\boldsymbol{B}^*\mu}$ intertwine M_ξ and M_z on all $\mathcal{H}$, and so these operators are the multiplications by some matrix functions $\Psi_\pm$.

Using (8-12) with $h = 1$ we can see that

$$\Psi_\pm b = \Phi^* b - C_1 T_{\pm}^{\boldsymbol{B}^*\mu} b = \boldsymbol{C}_* B^* b - C_1 T_{\pm}^{\boldsymbol{B}^*\mu} b,$$

so $\Psi_\pm$ are defined exactly as stated in the proposition. ☐

8B. ***Some calculations.*** Let us start with writing more-detailed formulas for the matrix functions $\boldsymbol{C}_*$ and $\boldsymbol{C}_1$ from Proposition 8.4.

Lemma 8.5. *We have*

$$C_*(z) = \begin{pmatrix} \boldsymbol{I} + \theta_\Gamma(z)\Gamma^* \\ \Delta_\Gamma(z)\Gamma^* \end{pmatrix} D_{\Gamma^*}^{-1}, \quad C_1(z) = \begin{pmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{pmatrix} D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \begin{pmatrix} \theta_\Gamma(z) \\ \Delta_\Gamma(z) \end{pmatrix} D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}).$$

Proof. The formula for $C_*(z)$ is just (3-21) and the identity $\theta_\Gamma(0) = -\Gamma$. Similarly, equation (3-22) gives us

$$C(z) = \begin{pmatrix} z^{-1}(\theta_\Gamma(z) + \Gamma) \\ z^{-1}\Delta_\Gamma(z) \end{pmatrix} D_\Gamma^{-1}.$$

Substituting these expressions into $C_1(z) = C_*(z) - zC(z)$ and applying the commutation relations from Lemma 5.4 we see

$$\begin{aligned} C_1(z) &= \begin{pmatrix} D_{\Gamma^*}^{-1} + \theta_\Gamma \Gamma^* D_{\Gamma^*}^{-1} - \theta_\Gamma D_\Gamma^{-1} - \Gamma D_\Gamma^{-1} \\ \Delta_\Gamma \Gamma^* D_{\Gamma^*}^{-1} - \Delta_\Gamma D_\Gamma^{-1} \end{pmatrix} \\ &= \begin{pmatrix} D_{\Gamma^*}^{-1} + \theta_\Gamma D_\Gamma^{-1}\Gamma^* - \theta_\Gamma D_\Gamma^{-1} - D_{\Gamma^*}^{-1}\Gamma \\ \Delta_\Gamma D_\Gamma^{-1}\Gamma^* - \Delta_\Gamma D_\Gamma^{-1} \end{pmatrix} \\ &= \begin{pmatrix} D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \theta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) \\ \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{pmatrix} D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \begin{pmatrix} \theta_\Gamma \\ \Delta_\Gamma \end{pmatrix} D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}), \end{aligned}$$

and the second statement in the lemma is verified. ☐

Recall that $F(z)$, $z \in \mathbb{D}$, is the matrix-valued Cauchy transform of the measure $B^* B\mu$, see (8-1), and that for $z \in \mathbb{T}$ the symbol $F(z)$ denotes the nontangential boundary values of F. We need the following simple relations between F and $\theta_{\boldsymbol{0}}$.

Lemma 8.6. *For all $z \in \mathbb{D}$ and a.e. on $\mathbb{T}$*

$$F(z) = (\boldsymbol{I} - \theta_{\mathbf{0}}(z))^{-1};$$

note that for all $z \in \mathbb{D}$ the matrix $\theta_{\mathbf{0}}(z)$ is a strict contraction, so $\boldsymbol{I} - \theta_{\mathbf{0}}(z)$ is invertible.

Proof. Recall that the function F_1 was defined by $F_1(z) = \mathcal{C}_1[B^* B\mu](z)$. Since $F(z) = \boldsymbol{I} + F_1(z)$, we get from (4-5) that

$$\theta_{\mathbf{0}}(z) = F_1(z)(\boldsymbol{I} + F_1(z))^{-1} = (F(z) - \boldsymbol{I})F(z)^{-1}.$$

Solving for F we get the conclusion of the lemma. □

8C. ***Proof of Theorem 8.1.*** Let us first prove the second identity in (8-3). Using the identity $F = (\boldsymbol{I} - \theta_{\mathbf{0}})^{-1}$ we compute

$$\Gamma^* + (\boldsymbol{I} - \Gamma^*)F = (\Gamma^*(\boldsymbol{I} - \theta_{\mathbf{0}}) + \boldsymbol{I} - \Gamma^*)F = (\boldsymbol{I} - \Gamma^*\theta_{\mathbf{0}})F,$$

which is exactly what we need.

We now prove that Ψ from Proposition 8.4 is given by $\Psi = \left(\begin{smallmatrix} \mathbf{0} \\ \Psi_2 \end{smallmatrix}\right)$, with Ψ_2 defined above in Theorem 8.1. Since $R(z)b_k(z) = e_k$, it is sufficient to show that $\Psi = \left(\begin{smallmatrix} \mathbf{0} \\ \Psi_2 \end{smallmatrix}\right)$ and that

$$\Psi_2(z)b_k(z) = \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* + (\boldsymbol{I} - \Gamma^*)F(z))e_k, \quad k = 1, 2, \dots, d. \tag{8-13}$$

Using the formulas for C_* and C_1 provided in Lemma 8.5 we get from (8-9)

$$\Psi(z)b_k(z) = C_*(z)e_k - C_1(z)F(z)e_k = \begin{pmatrix} (\boldsymbol{I} + \theta_\Gamma\Gamma^*)D_{\Gamma^*}^{-1} - [D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \theta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I})]F \\ \Delta_\Gamma\Gamma^* D_{\Gamma^*}^{-1} - \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I})F \end{pmatrix} e_k.$$

Note that it is clear from the representation (8-7) that the top entry of Ψ should disappear, i.e., that

$$(\boldsymbol{I} + \theta_\Gamma\Gamma^*)D_{\Gamma^*}^{-1} = [D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \theta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I})]F. \tag{8-14}$$

Indeed, by the definition of $\mathcal{K}_\theta$ in the Sz.-Nagy–Foiaş transcription the top entry of $\Phi^* f$ belongs to $H^2(\mathfrak{D}_*)$. One can see from Lemma 8.5, for example, that the top entry of C_1 belongs to the matrix-valued H^∞, so the top entry of $C_1 T_+^{B^*\mu} f$ is also in $H^2(\mathfrak{D}_*)$. Therefore the top entry of Ψf must be in $H^2(\mathfrak{D}_*)$ for all f. But that is impossible, because f can be any function in $L^2(\mu; E)$.

For a reader that is not comfortable with such "soft" reasoning, we present a "hard" computational proof of (8-14). This computation also helps to assure the reader that the previous computations were correct.

To do the computation, consider the term in the square brackets in the right-hand side of (8-14). Using the commutation relations from Lemma 5.4 in the second equality, we get

$$\begin{aligned} D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \theta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) &= D_{\Gamma^*}^{-1} + \theta D_\Gamma^{-1}\Gamma^* - \theta D_\Gamma^{-1} - D_{\Gamma^*}^{-1}\Gamma \\ &= D_{\Gamma^*}^{-1} + \theta\Gamma^* D_{\Gamma^*}^{-1} - \theta D_\Gamma^{-1} - \Gamma D_\Gamma^{-1} \\ &= (\boldsymbol{I} + \theta_\Gamma\Gamma^*)D_{\Gamma^*}^{-1}\{\boldsymbol{I} - D_{\Gamma^*}(\boldsymbol{I} + \theta_\Gamma\Gamma^*)^{-1}(\theta_\Gamma + \Gamma)D_\Gamma^{-1}\} \\ &= (\boldsymbol{I} + \theta_\Gamma\Gamma^*)D_{\Gamma^*}^{-1}\{\boldsymbol{I} - \theta_{\mathbf{0}}\}; \end{aligned}$$

the last equality holds by Theorem 5.2.

By Lemma 8.6 we have $\boldsymbol{I} - \theta_{\mathbf{0}} = F^{-1}$, so we have for the term in the square brackets

$$[D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \theta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I})] = (\boldsymbol{I} + \theta_\Gamma \Gamma^*) D_{\Gamma^*}^{-1} F^{-1},$$

which proves (8-14).

To deal with the bottom entry of Ψ we use the commutation relations from Lemma 5.4,

$$\begin{aligned}\Delta_\Gamma \Gamma^* D_{\Gamma^*}^{-1} - \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I})F &= \Delta_\Gamma D_\Gamma^{-1}\Gamma^* - \Delta_\Gamma D_\Gamma^{-1}\Gamma^* F + \Delta_\Gamma D_\Gamma^{-1} F \\ &= \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* + (\boldsymbol{I} - \Gamma^*)F),\end{aligned}$$

which gives the desired formula (8-13) for Ψ_2.

Finally, let us deal with the second term in the right-hand side of (8-2). We know from Proposition 8.4 that the term in front of $T_+^{B^*\mu} f$ is given by C_1. From Lemma 8.5 we get

$$C_1 = \begin{pmatrix} D_{\Gamma^*}^{-1}(\boldsymbol{I} - \Gamma) + \theta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) \\ \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) \end{pmatrix}.$$

But the top entry of C_1 here is the expression in brackets in the right-hand side of (8-14), so it is equal to $(\boldsymbol{I} + \theta_\Gamma \Gamma^*) D_{\Gamma^*}^{-1} F^{-1}$. Therefore

$$C_1 = \begin{pmatrix} (\boldsymbol{I} + \theta_\Gamma \Gamma^*) D_{\Gamma^*}^{-1} F^{-1} \\ \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) \end{pmatrix},$$

which is exactly what we have in (8-2). □

8D. ***Representation of Φ^* using matrix-valued measures.*** The above Theorem 8.1 is more transparent if we represent the direct integral $\mathcal{H}$ as the weighted L^2 space with a matrix-valued measure.

Namely, consider the weighted space $L^2(B^* B\mu)$,

$$\|f\|_{L^2(B^* B\mu)}^2 = \int_{\mathbb{T}} \big(B(\xi)^* B(\xi) f(\xi), f(\xi)\big)_{\mathbb{C}^d} \,\mathrm{d}\mu(\xi) = \int_{\mathbb{T}} \|B(\xi) f(\xi)\|_{\mathbb{C}^d}^2 \,\mathrm{d}\mu(\xi)$$

(of course one needs to take the quotient space over the set of functions with norm 0).

Then for all scalar functions φ_k we have

$$\bigg\| \sum_{k=1}^d \varphi_k e_k \bigg\|_{L^2(B^* B\mu)} = \bigg\| \sum_{k=1}^d \varphi_k b_k \bigg\|_{L^2};$$

recall that $e_1, e_2, \dots, e_d$ is the standard basis in $\mathbb{C}^d$ and $b_k(\xi) = B(\xi) e_k$. Then the map $\mathcal{U}$,

$$\mathcal{U}\bigg(\sum_{k=1}^d \varphi_k e_k \bigg) = \sum_{k=1}^d \varphi_k b_k, \quad \text{or, equivalently,} \quad \mathcal{U} f = Bf,$$

defines a unitary operator from $L^2(B^* B\mu)$ to $\mathcal{H}$.

The inverse operator $\mathcal{U}^*$ is given by $\mathcal{U}^* f(\xi) = R(\xi) f(\xi)$, where, recall, R is a measurable pointwise right inverse of B, $B(\xi) R(\xi) = \boldsymbol{I}_{E(\xi)}$ μ-a.e.

We define $\widetilde{\Phi} := \mathcal{U}^*\Phi$, so $\widetilde{\Phi}^* = \Phi^*\mathcal{U}$, and denote by $T_+^{B^*B\mu} f$ the nontangential boundary values of the Cauchy integral $\mathcal{C}[B^*Bf\mu](z)$, $z \in \mathbb{D}$. Substituting $f = Bg$ into (8-2) we can restate Theorem 8.1 as follows.

Theorem 8.7. *The adjoint Clark operator $\widetilde{\Phi}^* : L^2(B^*B\mu) \to \mathcal{K}_\theta$ in the Sz.-Nagy–Foiaş transcription is given by*

$$\widetilde{\Phi}^* g = \begin{pmatrix} 0 \\ \widetilde{\Psi}_2 \end{pmatrix} g + \begin{pmatrix} (\boldsymbol{I} + \theta_\Gamma \Gamma^*) D_{\Gamma^*}^{-1} F^{-1} \\ \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) \end{pmatrix} T_+^{B^*B\mu} g, \quad g \in L^2(B^*B\mu), \tag{8-15}$$

where the matrix-valued function $\widetilde{\Psi}_2(z)$ is defined as

$$\widetilde{\Psi}_2(z) = \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* + (\boldsymbol{I} - \Gamma^*)F(z)). \tag{8-16}$$

8E. ***A generalization of the normalized Cauchy transform.*** Consider the case when the unitary operator U has purely singular spectrum. By virtue of Corollary 5.9, the second component of the Sz.-Nagy–Foiaş model space collapses, i.e., $\mathcal{K}_{\theta_\Gamma} = H^2(\mathbb{C}^d) \ominus \theta_\Gamma H^2(\mathbb{C}^d)$ for all strict contractions Γ.

The representation formula (8-2) then reduces to a generalization of the well-studied normalized Cauchy transform.

Corollary 8.8. *If $\theta = \theta_{\mathbf{0}}$ is inner, then*

$$(\Phi^* f)(z) = (\boldsymbol{I} - \theta(z))(T_+^{B^*\mu} f)(z) = (F(z))^{-1}(T_+^{B^*\mu} f)(z)$$

for $z \in \mathbb{D}$, $f \in L^2(\mu; E)$.

The first equation was also obtained in [Kapustin and Poltoratski 2006, Theorem 1].

Here we used $\Gamma = \mathbf{0}$ only for simplicity. With the linear fractional relation in Theorem 5.2, it is not hard to write the result in terms of θ_Γ for any strict contraction Γ.

Proof. Theorem 8.1 for inner θ and $\Gamma = \mathbf{0}$ immediately reduces to the first statement.

The equality of the second expression follows immediately from Lemma 8.6. □

9. The Clark operator

Let $f \in \mathcal{H} \subset L^2(\mu; E)$ and let

$$\Phi^* f = h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \in \mathcal{K}_\theta. \tag{9-1}$$

From the representation (8-15) we get, subtracting from the second component the first component multiplied by an appropriate matrix-valued function, that

$$\Psi_2 f = h_2 - \Delta_\Gamma D_\Gamma^{-1}(\Gamma^* - \boldsymbol{I}) F D_{\Gamma^*}(\boldsymbol{I} + \theta_\Gamma \Gamma^*)^{-1} h_1.$$

Right-multiplying this identity by Ψ_2^*, and using Proposition 8.3 and formulas for $\Psi_2, \widetilde{\Psi}_2$ from Theorem 8.1, we get an expression for the density of the absolutely continuous part of $\mu_{\rm ac}$. Namely, we find that a.e. (with respect to Lebesgue measure on $\mathbb{T}$)

$$
\begin{aligned}
wf &= R^*F^*(\boldsymbol{I}-\theta_{\mathbf{0}}^*\Gamma)D_\Gamma^{-1}\Delta_\Gamma h_2 R^*F^*(\boldsymbol{I}-\theta_{\mathbf{0}}^*\Gamma)D_\Gamma^{-1}\Delta_\Gamma^2 D_\Gamma^{-1}(\Gamma^*-\boldsymbol{I})FD_{\Gamma^*}(\boldsymbol{I}+\theta_\Gamma\Gamma^*)^{-1}h_1 \\
&= R^*F^*(\boldsymbol{I}-\theta_{\mathbf{0}}^*\Gamma)D_\Gamma^{-1}\Delta_\Gamma h_2 - R^*F^*\Delta_{\mathbf{0}}^2(\boldsymbol{I}-\Gamma^*\theta_{\mathbf{0}})^{-1}(\Gamma^*-\boldsymbol{I})FD_{\Gamma^*}(\boldsymbol{I}+\theta_\Gamma\Gamma^*)^{-1}h_1. \qquad (9\text{-}2)
\end{aligned}
$$

In the case $\Gamma = \mathbf{0}$ the above equation simplifies:

$$
\begin{aligned}
wf &= R^*F^*\Delta_{\mathbf{0}}h_2 + R^*F^*\Delta_{\mathbf{0}}^2 F h_1 \\
&= R^*F\Delta_{\mathbf{0}}h_2 + wBh_1; \qquad (9\text{-}3)
\end{aligned}
$$

in the second equality we used (8-4).

The above formulas (9-2), (9-3) determine the absolutely continuous part of f.

The singular part of f was in essence computed in [Kapustin and Poltoratski 2006]. Formally it was computed there only for inner functions θ, but using the ideas and results from that paper it is easy to get the general case from our Theorem 8.1.

For the convenience of the reader, we give a self-contained presentation.

Lemma 9.1. *Let $f \in L^2(\mathbb{T}, \mu; \mathbb{C}^d)$. Then μ_s-a.e. the nontangential boundary values of $\mathcal{C}[f\mu](z)/\mathcal{C}[\mu](z)$, $z \in \mathbb{D}$, exist and equal $f(\xi)$, $\xi \in \mathbb{T}$.*

This lemma was proved in [Kapustin and Poltoratski 2006] even for the more general case of $f \in L^2(\mu; E)$, where E is a separable Hilbert space. Note that our case $E = \mathbb{C}^d$ follows trivially by applying the corresponding scalar result ($E = \mathbb{C}$) proved in [Poltoratskiĭ 1993] to entries of the vector f.

Applying the above lemma to the representation giving by the first coordinate of (8-2) from Theorem 8.1, we get that for f and h related by (9-1) we have

$$
B^*f = \frac{1}{\mathcal{C}[\mu]}FD_{\Gamma^*}(\boldsymbol{I}+\theta_\Gamma\Gamma^*)^{-1}h_1 \quad \mu_s\text{-a.e.}
$$

Left-multiplying this identity by R^* we get

$$
\Phi h = f = \frac{1}{\mathcal{C}[\mu]}R^*FD_{\Gamma^*}(\boldsymbol{I}+\theta_\Gamma\Gamma^*)^{-1}h_1 \quad \mu_s\text{-a.e.} \qquad (9\text{-}4)
$$

Summarizing, we get the following theorem, describing the direct Clark operator Φ.

Theorem 9.2. *If $\Phi^*f = h$ as in* (9-1), *so $f = \Phi h$, then the absolutely continuous part of f is given by* (9-2) *and the singular part of f is given by* (9-4).

References

[Clark 1972] D. N. Clark, "One dimensional perturbations of restricted shifts", *J. Analyse Math.* **25** (1972), 169–191. MR Zbl

[Douglas and Liaw 2013] R. G. Douglas and C. Liaw, "A geometric approach to finite rank unitary perturbations", *Indiana Univ. Math. J.* **62**:1 (2013), 333–354. MR Zbl

[Kapustin and Poltoratski 2006] V. Kapustin and A. Poltoratski, "Boundary convergence of vector-valued pseudocontinuable functions", *J. Funct. Anal.* **238**:1 (2006), 313–326. MR Zbl

[Liaw and Treil 2009] C. Liaw and S. Treil, "Rank one perturbations and singular integral operators", *J. Funct. Anal.* **257**:6 (2009), 1947–1975. MR Zbl

[Liaw and Treil 2016] C. Liaw and S. Treil, “Clark model in the general situation”, *J. Anal. Math.* **130** (2016), 287–328. MR Zbl

[Nikolski 2002a] N. K. Nikolski, *Operators, functions, and systems: an easy reading, I*, Mathematical Surveys and Monographs **92**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Nikolski 2002b] N. K. Nikolski, *Operators, functions, and systems: an easy reading, II*, Mathematical Surveys and Monographs **93**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Nikolski and Vasyunin 1998] N. Nikolski and V. Vasyunin, “Elements of spectral theory in terms of the free function model, I: Basic constructions”, pp. 211–302 in *Holomorphic spaces* (Berkeley, CA, 1995), edited by S. Axler et al., Math. Sci. Res. Inst. Publ. **33**, Cambridge Univ. Press, 1998. MR Zbl

[Poltoratskiĭ 1993] A. G. Poltoratskiĭ, “Boundary behavior of pseudocontinuable functions”, *Algebra i Analiz* **5**:2 (1993), 189–210. In Russian; translated in *St. Petersburg Math. J.* **5**:2 (1994), 389–406. MR Zbl

[Sarason 1994] D. Sarason, *Sub-Hardy Hilbert spaces in the unit disk*, University of Arkansas Lecture Notes in the Mathematical Sciences **10**, John Wiley & Sons, New York, 1994. MR Zbl

[Sz.-Nagy et al. 2010] B. Sz.-Nagy, C. Foiaş, H. Bercovici, and L. Kérchy, *Harmonic analysis of operators on Hilbert space*, 2nd ed., Springer, 2010. MR Zbl

[Woodbury 1950] M. A. Woodbury, “Inverting modified matrices”, Statistical Research Group, memo. report no. 42, Princeton University, 1950. MR

Received 20 Nov 2017. Revised 17 Apr 2018. Accepted 30 May 2018.

CONSTANZE LIAW: liaw@udel.edu
Department of Mathematical Sciences, University of Delaware, Newark, DE, United States

and

CASPER, Baylor University, Waco, TX, United States

SERGEI TREIL: treil@math.brown.edu
Department of Mathematics, Brown University, Providence, RI, United States

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.493

ON THE MAXIMAL RANK PROBLEM FOR THE COMPLEX HOMOGENEOUS MONGE–AMPÈRE EQUATION

JULIUS ROSS AND DAVID WITT NYSTRÖM

We give examples of regular boundary data for the Dirichlet problem for the complex homogeneous Monge–Ampère equation over the unit disc, whose solution is completely degenerate on a nonempty open set and thus fails to have maximal rank.

1. Introduction

Let (X, ω) be a compact Kähler manifold of dimension n and B be a Riemann surface with boundary ∂B. Suppose $(\phi_\tau)_{\tau\in\partial B}$ is a smooth family of Kähler potentials on X; so each ϕ_τ is a smooth function on X, varying smoothly in τ, that satisfies

$$\omega + dd^c\phi_\tau > 0.$$

Then let Φ be the solution to the Dirichlet problem for the complex homogeneous Monge–Ampère equation (HMAE) with this boundary data, so Φ is a function on $X \times B$ that satisfies

$$\begin{aligned} \Phi(\,\cdot\,, \tau) &= \phi_\tau(\,\cdot\,) \quad \text{for } \tau \in \partial B, \\ \pi_X^*\omega + dd^c\Phi &\geq 0, \\ (\pi_X^*\omega + dd^c\Phi)^{n+1} &= 0, \end{aligned} \tag{1}$$

where $\pi_X : X \times B \to X$ is the projection. From standard pluripotential theory we know there exists a unique weak solution Φ to this equation. The *maximal rank problem* in this setting asks whether the current

$$\pi_X^*\omega + dd^c\Phi$$

has maximal rank in the fibre directions, that is, whether the current $\omega + dd^c\Phi(\,\cdot\,, \tau)$ on X is strictly positive for each $\tau \in B$. Said another way, this asks if the rank of $\pi_X^*\omega + dd^c\Phi$ is precisely n at every point in $X \times B$, which is the maximum possible since $(\pi_X^*\omega + dd^c\Phi)^{n+1} = 0$. Similarly one has the *constant-rank problem* in which one asks if the rank of $\pi_X^*\omega + dd^c\Phi$ is the same at every point. The purpose of this note is to answer this question negatively, giving an explicit example in which the rank fails to be maximal.

MSC2010: 32W20, 58J32.
Keywords: 32W20, 35J60, 31C10, 35J70.

Theorem 1.1. *Let $B = \overline{\mathbb{D}} \subset \mathbb{C}$ be the closed unit disc and $(X, \omega) = (\mathbb{P}^1, \omega_{\mathrm{FS}})$, where ω_{FS} denotes the Fubini–Study form. Then there exists a smooth family of Kähler potentials $(\phi_\tau)_{\tau\in\partial\mathbb{D}}$ on $\mathbb{P}^1$ such that the solution Φ to the HMAE* (1) *is completely degenerate on some nonempty open subset $S \subset \mathbb{P}^1 \times \mathbb{D}$, i.e.,*

$$\pi^*_{\mathbb{P}^1}\omega_{\mathrm{FS}} + dd^c\Phi|_S = 0.$$

A more precise version of this statement is provided in Theorem 2.1. The motivation and ideas build on previous work of the authors [Ross and Witt Nyström 2015a; 2015b; 2015c] in which we understand the solution to the HMAE of a certain kind through a free boundary problem in the plane called the Hele-Shaw flow. But rather than expecting the reader to be an expert in this topic we have chosen to give a direct proof, which can be found in Section 2, that is both self-contained and rather simple. Then in Section 3 we explain the motivation behind our construction, as well as give a second (but essentially equivalent) proof that relies on more machinery. We then end with some questions and possible extensions.

Of course in the above theorem, $\pi^*_{\mathbb{P}^1}\omega_{\mathrm{FS}} + dd^c\Phi$ is not identically zero, and so does not have constant rank. In fact we can say more, and it is possible to arrange so that there is a nonempty open set in $\mathbb{P}^1 \times \mathbb{D}$ on which $\pi^*_{\mathbb{P}^1}\omega_{\mathrm{FS}} + dd^c\Phi$ is regular (i.e., smooth and of maximal rank). It is worth commenting from the outset that we do not expect the solution we have here to be everywhere smooth, but it should be possible to describe precisely where it is regular and where it is degenerate. All of this will be discussed in more detail in Section 3.

1A. ***Comparison with other work.*** It is known that convex solutions to elliptic partial differential equations have a constant-rank property. Early works of this include [Caffarelli and Friedman 1985; Singer, Wong, Yau, and Yau 1985]. These have since been built upon by many others, and it is now known that the constant-rank property holds for a wide class of elliptic equations; see, for instance, [Korevaar and Lewis 1987; Bian and Guan 2009; 2010; Caffarelli, Guan and Ma 2007; Székelyhidi and Weinkove 2016]. In this paper we are interested in the complex degenerate situation, about which much less has been written. The most famous result along these lines, and in the positive direction, is that of Lempert [1981] who proved that on a convex domain in $\mathbb{C}^n$ the solution to the complex HMAE with prescribed singularity at an interior point (the pluricomplex Green function) is smooth and of maximal rank. The maximal rank problem for other partial differential equations in the complex case has also been taken up by Guan, Li and Zhang [2009] and by Li [2009].

The closest previous work to that of this paper is probably that of Guan and Phong [2012], who studied the problem of finding uniform lower bounds for the eigenvalues of the solution to the (nondegenerate) Monge–Ampère equation in the limit as the equation becomes degenerate. Moreover, they asked whether solutions to the complex HMAE have maximal rank [Guan and Phong 2012, discussion after Theorem 4]. The idea of maximal rank for the complex HMAE also appears in the ideas of Chen and Tian [2008] through the concept of an almost-regular solution to the HMAE, which fails to have maximal rank only on a set which is small in a precise sense. The kinds of envelopes that we use in our proof also can be defined more generally, and even in higher dimensions, which is the topic of previous work of the authors [Ross and Witt Nyström 2017b], in which we prove a constant-rank theorem, Theorem 6.2 of that paper, that we call “optimal regularity”.

Questions concerning the regularity of the solution to the Dirichlet problem for the kind of complex HMAE we consider here go back at least as far as [Semmes 1992; Donaldson 2002], and this HMAE has been the focus of much interest due to it being the geodesic equation in the space of Kähler metrics. By [Chen 2000] with complements by Błocki [2012] we know such a solution always has bounded Laplacian (so in particular is $C^{1,\alpha}$ for any $\alpha < 1$). In fact in our case, since we are working on $\mathbb{P}^1$, the results of [Błocki 2012] imply that Φ is $C^{1,1}$. (We observe that we do not actually need to know this regularity for the direct proof of our main theorem). Donaldson [2002] gives examples of boundary data for which the solution is not regular, but the nature of the irregularity there is left unknown (for instance Donaldson's example may have maximal rank but fail to be everywhere smooth).

2. Main theorem

2A. *Notation.* We let $\mathbb{D}_r$ be the open disc of radius r in the complex plane about the origin, $\mathbb{D} = \mathbb{D}_1$ and $\mathbb{D}^\times = \mathbb{D} \setminus \{0\}$. Throughout we consider the standard cover of $\mathbb{P}^1$ by two charts equal to the complex plane with coordinates z and $w = 1/z$. We shall denote these two charts by $\mathbb{C}_z$ and $\mathbb{C}_w$ respectively. We use the convention $d^c = \frac{1}{2\pi}(\bar{\partial} - \partial)$ so $dd^c \log |z|^2 = \delta_0$, and normalise the Fubini–Study form ω_{FS} so $\int_{\mathbb{P}^1} \omega_{\mathrm{FS}} = 1$. Thus $\omega_{\mathrm{FS}} = dd^c \log(1 + |z|^2)$ locally on $\mathbb{C}_z$.

2B. *Statement of the main theorem.* The following is a precise version of our main theorem. By an *arc* in $\mathbb{C}$ we mean the image γ of a smooth map $[0, 1] \to \mathbb{C}$ that does not intersect itself. From now on $B = \overline{\mathbb{D}}$ is the closed unit disc and $(X, \omega) = (\mathbb{P}^1, \omega_{\mathrm{FS}})$.

Theorem 2.1. *Suppose that $\phi \in C^\infty(\mathbb{P}^1)$ satisfies*:

(1) $\omega_{\mathrm{FS}} + dd^c\phi > 0$.

(2) *On $\mathbb{C}_w \subset \mathbb{P}^1$ it holds that*
$$\phi(w) \geq -\ln(1 + |w|^2)$$
with equality precisely on an arc in $\mathbb{C}_w$.

Then setting
$$\phi_\tau(z) := \phi(\tau z) \quad \text{for } \tau \in \partial\mathbb{D},$$
the solution Φ to the HMAE (1) *does not have maximal rank. In fact there is a nonempty open subset $S \subset \mathbb{P}^1 \times \mathbb{D}$ such that*
$$\pi^*_{\mathbb{P}^1}\omega_{\mathrm{FS}} + dd^c\Phi|_S = 0.$$

2C. *Envelopes.* For the proof we need some background concerning envelopes of subharmonic functions. Fix a potential $\phi \in C^\infty(\mathbb{P}^1)$ so $\omega_{\mathrm{FS}} + dd^c\phi > 0$. For a topological space X let
$$\mathrm{USC}(X) = \{\psi : X \to \mathbb{R} \cup \{-\infty\} \text{ such that } \psi \text{ is upper semicontinuous}\}.$$

Definition 2.2. For $t \in (0, 1]$ set
$$\psi_t := \sup\{\psi \in \mathrm{USC}(\mathbb{P}^1) : \psi \leq \phi \text{ and } \omega_{\mathrm{FS}} + dd^c\psi \geq 0 \text{ and } \nu_{z=0}(\psi) \geq t\}.$$

Here $\nu_{z=0}$ denotes the Lelong number at the point $z=0$, so $\nu_{z=0}(\psi) \geq t$ means $\psi(z) \leq t \ln |z|^2 + O(1)$ near $z=0$. As the upper-semicontinuous regularisation of ψ_t is itself a candidate for the envelope defining ψ_t, we see that ψ_t is itself upper-semicontinuous.

Definition 2.3. For $t \in (0, 1]$ set

$$\Omega_t := \Omega_t(\phi) := \{z \in \mathbb{P}^1 : \psi_t(z) < \phi(z)\}. \tag{2}$$

Clearly if $t \leq t'$ then $\psi_{t'} \leq \psi_t$ and so $\Omega_t \subset \Omega_{t'}$. Now, unless one assumes some additional symmetry of ϕ, it is generally quite hard to describe the sets Ω_t. However, as the next lemma shows, it is possible, under a suitable hypothesis, to describe the largest one Ω_1 by looking at the level set on which ϕ takes its minimum value.

Lemma 2.4. *Let $\phi \in C^\infty(\mathbb{P}^1)$ be such that $\omega_{\mathrm{FS}} + dd^c\phi > 0$ and $\phi(w) \geq -\ln(1+|w|^2)$ on $\mathbb{C}_w$ with equality precisely on some nonempty subset $\gamma \subset \mathbb{C}_w$ containing $w=0$. Then*

$$\psi_1(z) = \ln\left(\frac{|z|^2}{1+|z|^2}\right)$$

and

$$\Omega_1(\phi) = \mathbb{P}^1 \setminus \gamma.$$

Proof. Observe first that the only upper-semicontinuous $\psi : \mathbb{P}^1 \to \mathbb{R} \cup \{-\infty\}$ with $\omega_{\mathrm{FS}} + dd^c\psi \geq 0$ and $\nu_{z=0}(\psi) \geq 1$ is, up to an additive constant, equal to

$$\zeta(z) := \ln\left(\frac{|z|^2}{1+|z|^2}\right) \quad \text{on } \mathbb{C}_z.$$

To see this observe first that we certainly cannot have $\nu_{z=0}(\psi) > 1$ since we have normalised so $\int_{\mathbb{P}^1} \omega_{\mathrm{FS}} = 1$. Thus we may assume $\nu_{z=0}\psi = 1$. Then observe that ζ is ω_{FS}-harmonic on $\mathbb{C}_z \setminus \{0\}$, and that the difference $\psi - \zeta$ is bounded near 0. Thus $\psi - \zeta$ extends to a bounded subharmonic function on all of $\mathbb{C}_z$, and hence is constant by the Liouville property. Thus the envelope ψ_1 from Definition 2.2 must be

$$\psi_1 = \zeta + C,$$

where C is the largest constant one can choose so that $\psi_1 \leq \phi$. Now on $\mathbb{C}_w$ we have

$$\psi_1(w) = -\ln(1+|w|^2) + C$$

and so as γ is nontrivial our hypothesis forces $C=0$. Thus

$$\Omega_1 = \{-\ln(1+|w|^2) < \phi(w)\} = \mathbb{P}^1 \setminus \gamma. \qquad \square$$

2D. ***Weak solutions to the HMAE.*** We now discuss the weak solution to two versions of the Dirichlet problem for the complex HMAE, first over the disc and second over the punctured disc; this follows the discussion in [Ross and Witt Nyström 2015b]. Again we let $\phi \in C^\infty(\mathbb{P}^1)$ be such that $\omega_{\mathrm{FS}} + dd^c\phi > 0$.

Definition 2.5. Let

$$\Phi := \sup\big\{\psi \in \mathrm{USC}(\mathbb{P}^1 \times \overline{\mathbb{D}}) : \pi_{\mathbb{P}^1}^* \omega_{\mathrm{FS}} + dd^c\psi \geq 0 \text{ and } \psi(z, \tau) \leq \phi(\tau z) \text{ for } (z, \tau) \in \mathbb{P}^1 \times \partial\mathbb{D}\big\}.$$

and

$$\widetilde{\Phi} := \sup\big\{\psi \in \mathrm{USC}(\mathbb{P}^1 \times \overline{\mathbb{D}}) : \pi_{\mathbb{P}^1}^* \omega_{\mathrm{FS}} + dd^c\psi \geq 0 \\ \text{and } \psi(z, \tau) \leq \phi(z) \text{ for } (z, \tau) \in \mathbb{P}^1 \times \partial\mathbb{D} \text{ and } \nu_{(z=0,\tau=0)}(\psi) \geq 1\big\}. \quad (3)$$

The function Φ is the weak solution to the complex HMAE with boundary data $\phi(\tau z)$ for $\tau \in \partial\mathbb{D}$, that is, the solution to (1). Similarly $\widetilde{\Phi}$ is the weak solution to the Dirichlet problem with boundary data $\phi(z)$, but with the additional requirement of having a prescribed singularity at the point $p := (0, 0) \subset \mathbb{C}_z \times \mathbb{D} \subset \mathbb{P}^1 \times \mathbb{D}$. That is, $\widetilde{\Phi}$ is upper-semicontinuous, $\pi_{\mathbb{P}^1}^* \omega_{\mathrm{FS}} + dd^c\widetilde{\Phi} \geq 0$ and $(\pi_{\mathbb{P}^1}^* \omega_{\mathrm{FS}} + dd^c\widetilde{\Phi})^2 = 0$ away from p and $\widetilde{\Phi}(z, \tau) = \phi(z)$ for $\tau \in \partial\mathbb{D}$. Moreover it is not hard to show that $\widetilde{\Phi}$ is locally bounded away from p and $\nu_p\widetilde{\Phi} = 1$. These two quantities carry the same information, as given by:

Proposition 2.6. *We have that*

$$\Phi(z, \tau) + \ln|\tau|^2 + \ln(1 + |z|^2) = \widetilde{\Phi}(\tau z, \tau) + \ln(1 + |\tau z|^2) \quad \textit{for } (z, \tau) \in \mathbb{P}^1 \times \overline{\mathbb{D}}^\times.$$

Proof. It is easily seen from the definition that $\Phi(z, \tau) + \ln|\tau|^2 + \ln(1 + |z|^2) - \ln(1 + |\tau z|^2)$ is a candidate for the envelope defining $\widetilde{\Phi}(\tau z, \tau)$, giving one inequality and the other inequality is proved similarly. □

2E. *Proof of Theorem 2.1.* Without loss of generality we assume the arc γ goes through the point $w = 0$. By Lemma 2.4

$$\psi_1(z) = \ln\left(\frac{|z|^2}{1 + |z|^2}\right)$$

and

$$\Omega_1 = \mathbb{P}^1 \setminus \gamma.$$

Looking at the other coordinate patch $\mathbb{C}_z$, we have that γ is a curve passing through infinity, and so $\mathbb{C}_z \setminus \gamma$ is an open, simply connected proper subset of $\mathbb{C}_z$. Hence by the Riemann mapping theorem there is a biholomorphism

$$f : \mathbb{D} \to \mathbb{C}_z \setminus \gamma \quad \text{with } f(0) = 0.$$

For $\tau \in \mathbb{D}^\times$ set

$$A_\tau := f(\mathbb{D}_{|\tau|}) \subset \mathbb{C}_z \subset \mathbb{P}^1.$$

Clearly each A_τ is a proper subset of $\mathbb{C}_z$ containing the origin, whose complement has nonempty interior.

Proposition 2.7. *We have*

$$\widetilde{\Phi}(z, \tau) = \psi_1(z) \quad \textit{for all } \tau \in \mathbb{D}^\times \textit{ and } z \in \mathbb{P}^1 \setminus A_\tau.$$

Proof. By abuse of notation we write ψ_1 also for the pullback of ψ_1 to $\mathbb{P}^1 \times \overline{\mathbb{D}}$. Then

$$\widetilde{\Phi}(z, \tau) \geq \psi_1(z) \quad \text{for } (z, \tau) \in \mathbb{P}^1 \times \overline{\mathbb{D}} \quad (4)$$

since ψ_1 is a candidate for the envelope (3) defining $\widetilde{\Phi}$.

We next claim that

$$\widetilde{\Phi}(f(\tau),\tau)=\psi_1(f(\tau)) \quad \text{for all } \tau\in\mathbb{D}. \tag{5}$$

To see this, observe that $\tau\mapsto\widetilde{\Phi}(f(\tau),\tau)$ is $f^*\omega_{\mathrm{FS}}$-subharmonic and has Lelong number 1 at $\tau=0$. On the other hand $\psi_1(f(\tau))$ is $f^*\omega_{\mathrm{FS}}$-harmonic except at $\tau=0$ where it has Lelong number 1. But $\widetilde{\Phi}(f(\tau),\tau)$ tends to $\psi_1(f(\tau))$ as $|\tau|$ tends to 1, and hence from the maximum principle along with (4), we get (5).

Now fix some $\tau\in\mathbb{D}^\times$ and set

$$\phi_\tau(z):=\widetilde{\Phi}(z,\tau).$$

Then the above says that $\phi_\tau=\psi_1$ on ∂A_τ. On the other hand by (4) we have $\phi_\tau\ge\psi_1$ everywhere. Moreover ϕ_τ is ω_{FS}-subharmonic on A_τ^c, whereas ψ_1 is bounded and ω_{FS}-harmonic on A_τ^c. Thus by the maximum principle we deduce $\phi_\tau=\psi_1$ on A_τ^c as required. □

Proof of Theorem 2.1. Set

$$S:=\{(z,\tau)\in\mathbb{P}^1\times\mathbb{D}^\times : \tau z\in(A_\tau^c)^\circ\},$$

which is nonempty and open in $\mathbb{P}^1\times\mathbb{D}^\times$. Then by Proposition 2.6 and then Proposition 2.7 if $(z,\tau)\in S$ we have

$$\Phi(z,\tau)=\widetilde{\Phi}(\tau z,\tau)+\ln\left(\frac{1+|\tau z|^2}{|\tau|^2(1+|z|^2)}\right)=\psi_1(\tau z)+\ln\left(\frac{1+|\tau z|^2}{|\tau|^2(1+|z|^2)}\right).$$

Thus on S we have

$$\pi_{\mathbb{P}^1}\omega_{\mathrm{FS}}+dd^c\Phi=\pi_{\mathbb{P}^1}\omega_{\mathrm{FS}}+dd^c\psi_1(\tau z)=0$$

as ψ_1 is ω_{FS}-harmonic away from $z=0$. □

2F. *A specific example.* We now construct a specific potential ϕ that satisfies the hypotheses of Theorem 2.1. Fix γ to be the interval $[-1,1]\subset\mathbb{R}\subset\mathbb{C}_w$. Our goal is to find a $\phi\in C^\infty(\mathbb{P}^1)$ such that $\omega_{\mathrm{FS}}+dd^c\phi>0$ and $\phi\ge-\ln(1+|w|^2)$ with equality precisely on γ.

To do so, let $\alpha:\mathbb{R}\to\mathbb{R}$ be a nonnegative smooth nondecreasing convex function with $\alpha(t)=0$ for $t\le 1$ and $\alpha(t)>0$ for $t>1$. Let

$$u(w):=\alpha(|w|^2)+\mathrm{Im}(w)^2.$$

Thus u is a smooth strictly subharmonic function on $\mathbb{C}_w$ that vanishes precisely on γ. Then $\epsilon u-\ln(1+|w|^2)$ for some small constant $\epsilon>0$ is essentially the function that we want; we simply need to adjust it to have the correct behaviour far away from γ.

To do so we shall use a regularised version of the maximum function, which can be explicitly given as follows: Let $|\cdot|_{\mathrm{reg}}$ be a smooth convex function on $\mathbb{R}$ so that $|t|_{\mathrm{reg}}=|t|$ for $|t|\ge 1$. Set $\max_{\mathrm{reg}}(a,b):=\frac{1}{2}(|a-b|_{\mathrm{reg}}+a+b)$ and for $\delta>0$ put

$$\max_\delta(a,b):=\delta\max_{\mathrm{reg}}(\delta^{-1}a,\delta^{-1}b). \tag{6}$$

Then $\max_\delta(\cdot\,,\cdot)$ is smooth, and satisfies

$$\max_\delta(a,b)=\begin{cases}a & \text{if } a>b+\delta,\\ b & \text{if } b>a+\delta.\end{cases}$$

Returning to the construction of ϕ, fix a sufficiently large constant C and a sufficiently small positive constant ϵ so that

$$\begin{aligned}\epsilon u &\geq \ln(1+|w|^2)-C+1 \quad \text{on } \mathbb{D}_2,\\ \epsilon u &\leq \ln(1+|w|^2)-C-1 \quad \text{on } \mathbb{D}_4\setminus\mathbb{D}_3.\end{aligned}$$

Then for $0<\delta\ll 1$ set

$$v:=\max_\delta(\epsilon u,\ln(1+|w|^2)-C).$$

So v is smooth, nonnegative, strictly subharmonic, equal to $\ln(1+|w|^2)-C$ on $\mathbb{D}_4\setminus\mathbb{D}_3$ and vanishes precisely on γ. We then put

$$\phi:=v-\ln(1+|w|^2)$$

and extend ϕ to take the constant value C in $\mathbb{C}_w\setminus\mathbb{D}_4$. So ϕ extends to a smooth function over $\mathbb{P}^1$ with the desired properties.

3. Discussion

3A. ***Context.*** Fix a $\phi\in C^\infty(\mathbb{P}^1)$ such that $\omega_{\mathrm{FS}}+dd^c\phi>0$. Then associated to ϕ we have two constructions:

(1) The solution $\widetilde{\Phi}$ to the complex HMAE on $\mathbb{P}^1\times\overline{\mathbb{D}}$ with boundary data given by $\phi_\tau=\phi$ for all $\tau\in\partial\mathbb{D}$ and the requirement of having Lelong number 1 at the point $(z,\tau)=(0,0)\in\mathbb{C}_z\times\mathbb{D}\subset\mathbb{P}^1\times\overline{\mathbb{D}}$.

(2) The envelopes ψ_t for $t\in(0,1]$ and the associated sets $\Omega_t(\phi)=\{\psi_t<\phi\}$.

In previous work we showed that these sets of data are intimately connected. First, $\widetilde{\Phi}$ and ψ_t are Legendre dual to each other [Ross and Witt Nyström 2015b, Theorem 2.7] in that

$$\psi_t(z)=\inf_{|\tau|>0}\{\widetilde{\Phi}(z,\tau)-(1-t)\ln|\tau|^2\} \tag{7}$$

and

$$\widetilde{\Phi}(z,\tau)=\sup_t\{\psi_t(z)+(1-t)\ln|\tau|^2\}. \tag{8}$$

Second, the collection of sets $\Omega_t(\phi)$ that are biholomorphic to a disc describes the harmonic discs of $\widetilde{\Phi}$. That is, if t is such that $\Omega_t(\phi)$ is a proper simply connected subset of $\mathbb{C}_z$ and $f:\mathbb{D}\to\Omega_t$ is a Riemann map with $f(0)=0$ then the restriction of $\widetilde{\Phi}$ to the graph $\{(f(\tau),\tau)\in\mathbb{P}^1\times\mathbb{D}\}$ is ω_{FS}-harmonic. Furthermore it is shown in [Ross and Witt Nyström 2015b, Theorem 3.1] these are (essentially) the only harmonic discs that occur.

We can say more. For $\tau\in\mathbb{D}^\times$ set

$$\phi_\tau(z):=\widetilde{\Phi}(z,\tau).$$

If $\widetilde{\Phi}$ is regular then each ϕ_τ will be a smooth Kähler potential, but in general this will not be the case. Nevertheless, by [Błocki 2012] we know ϕ_τ is $C^{1,1}$ and since $\pi_{\mathbb{P}^1}^*\omega_{\mathrm{FS}}+dd^c\widetilde{\Phi}\geq 0$, we know $\omega_{\mathrm{FS}}+dd^c\phi_\tau$ is semipositive. We can then define the associated sets $\Omega_t(\phi_\tau)$ in exactly the same way as before.

Proposition 3.1. *Suppose t is such that $\Omega_t(\phi)\subset\mathbb{C}_z$ is proper and simply connected and let $f_t:\mathbb{D}\to\Omega_t(\phi)$ be a Riemann map with $f(0)=0$. Then for each $\tau\in\mathbb{D}^\times$ we have*

$$f_t(\mathbb{D}_{|\tau|})=\Omega_t(\phi_\tau).$$

We shall give a proof of this fact below, but assuming it for now we can give an alternative proof that, under the hypotheses of Theorem 2.1, for each $\tau\in\mathbb{D}^\times$ the current $\omega_{\mathrm{FS}}+dd^c\widetilde{\Phi}(\,\cdot\,,\tau)$ is degenerate on some nonempty open subset of $\mathbb{P}^1$. First Lemma 2.4 gives

$$\Omega_1(\phi)=\mathbb{P}^1\setminus\gamma,$$

which is a simply connected proper subset of $\mathbb{C}_z$. We then take our Riemann map $f:\mathbb{D}\to\Omega_1(\phi)$ and consider the image

$$A_\tau:=f(\mathbb{D}_{|\tau|})=\Omega_1(\phi_\tau)\quad\text{for }\tau\in\mathbb{D}^\times.$$

As observed before, A_τ is a proper subset of $\mathbb{C}_z$ whose complement has nonempty interior.

On the other hand, it is a general fact that for each t the set $\Omega_t(\phi_\tau)$ has measure t with respect to the current $\omega_{\mathrm{FS}}+dd^c\phi_\tau$. (If ϕ_τ is smooth and $\omega_{\mathrm{FS}}+dd^c\phi_\tau$ is strictly positive, this is a standard piece of potential theory and is discussed in [Ross and Witt Nyström 2015b, Proposition 1.1]; when ϕ_τ is merely C^2 and $t<1$ then this is proved in [Ross and Witt Nyström 2017a, Theorem 1.2] and the case $t=1$ follows from this by continuity as $\Omega_1(\phi_\tau)=\bigcup_{t<1}\Omega_t(\phi_\tau)$; finally when ϕ_τ is merely $C^{1,1}$ this is given in [Berman and Demailly 2012, Remark 1.19, Corollary 2.5].)

Therefore

$$\int_{A_\tau}(\omega_{\mathrm{FS}}+dd^c\phi_\tau)=\int_{\Omega_1(\phi_\tau)}(\omega_{\mathrm{FS}}+dd^c\phi_\tau)=1.$$

But our normalisation is that $\int_{\mathbb{P}^1}(\omega_{\mathrm{FS}}+dd^c\phi_\tau)=\int_{\mathbb{P}^1}\omega_{\mathrm{FS}}=1$ as well, and so the current $\omega_{\mathrm{FS}}+dd^c\phi_\tau$ gives zero measure to the complement of A_τ, which is precisely what we were aiming to prove.

Proof of Proposition 3.1. Fix $\sigma\in\mathbb{D}^\times$ and set $r:=|\sigma|$. Our aim is to show

$$f_t(\mathbb{D}_r)=\Omega_t(\phi_\sigma).$$

Consider the S^1-action on $\mathbb{P}^1\times\overline{\mathbb{D}}$ given by $e^{i\theta}\cdot(z,\tau)=(z,e^{i\theta}\tau)$, and observe that the boundary data used to define $\widetilde{\Phi}$ from (3) is S^1-invariant, which implies $\widetilde{\Phi}$ is S^1-invariant as well. Thus we may as well assume that σ is real, so $\phi_\sigma=\phi_r$.

For a function F on $\mathbb{P}^1\times\overline{\mathbb{D}}$ and $D\subset\overline{\mathbb{D}}$ we write $F|_D$ for the restriction of F to $\mathbb{P}^1\times D$. Then $\widetilde{\Phi}|_{\overline{\mathbb{D}}_r}$ is the solution to the Dirichlet problem for the HMAE with boundary data $(\phi_\tau)_{\tau\in\partial\mathbb{D}_r}=\phi_r$ and the requirement that $\widetilde{\Phi}|_{\overline{\mathbb{D}}_r}$ has Lelong number 1 at the point $(0,0)\in\mathbb{C}_z\times\mathbb{D}_r\subset\mathbb{P}^1\times\mathbb{D}_r$.

Letting $s:=-\ln|\tau|^2$, consider the function on $\mathbb{P}^1\times\overline{\mathbb{D}}^\times$ given by

$$H(z,\tau):=\frac{\partial}{\partial s}\widetilde{\Phi}(z,e^{-s/2})$$

(when $|\tau|=1$ and thus $s=0$, we take the right derivative). As $\widetilde{\Phi}$ is $C^{1,1}$ on $\mathbb{P}^1\times\overline{\mathbb{D}}^\times$, the function H is well-defined and Lipschitz. Clearly this is compatible with restriction; i.e.,

$$H|_{\overline{\mathbb{D}}_r^\times}(z,\tau)=\frac{\partial}{\partial s}\widetilde{\Phi}|_{\overline{\mathbb{D}}_r}(z,e^{-s/2}).$$

Now, as discussed above, and proved in [Ross and Witt Nyström 2015b, Theorem 3.1], the graph $\{(f(\tau),\tau):\tau\in\mathbb{D}\}$ of f is a harmonic disc for $\widetilde{\Phi}$. What is also proved is that H takes the constant value $t-1$ along this disc so

$$H(f(\tau),\tau)=t-1\quad\text{for }\tau\in\mathbb{D}^\times.$$

Now H is also S^1-invariant and so this in particular implies

$$H(f(re^{i\theta}),r)=H(f(re^{i\theta}),re^{i\theta})=t-1\quad\text{for all }\theta\in[0,2\pi].$$

In other words, the function $H(\cdot,r)$ takes the value $t-1$ on the boundary of $f(\mathbb{D}_r)$. On the other hand, we prove in [Ross and Witt Nyström 2015b, Proposition 2.9] that the function $H(\cdot,r)$ describes the set $\Omega_t(\phi_r)$, in that

$$H(z,r)+1=\sup\{s:z\notin\Omega_s(\phi_r)\}$$

(we remark that the proof of that proposition does not require any regularity or strict positivity assumptions on the potential ϕ_σ). Thus $\Omega_t(\phi_r)$ is the interior component of the curve $\theta\mapsto f(re^{i\theta})$ (that is, the component containing $z=0$), which gives $\Omega_t(\phi_r)=f(\mathbb{D}_r)$ as desired. □

3B. ***Extensions and questions.*** Under the hypotheses of Theorem 2.1 we have shown that the current $\omega_{\mathrm{FS}}+dd^c\Phi(\cdot,\tau)$ fails to be strictly positive on any interior fibre (that is, for any τ with $0<|\tau|<1$). Furthermore we have no reason to expect our solution to be smooth everywhere. Thus the following two questions are natural:

Question 3.2. Does there exist a smooth family of potentials $(\phi_\tau)_{\tau\in\partial B}$ for which the solution to the complex HMAE (1) is everywhere smooth but not of maximal rank?

Question 3.3. Does there exist a smooth family of potentials $(\phi_\tau)_{\tau\in\partial B}$ for which the solution to the complex HMAE (1) such that $\omega+dd^c\Phi(\cdot,\tau)$ is a Kähler form for some τ with $0<|\tau|<1$ but not for others.

We are not currently able to answer these questions. However, we believe that the degenerate solutions we describe in this paper are actually regular in the interior of the complement of the degenerate set S (that is, they are smooth there and of maximal rank). In fact from our previous work in [Ross and Witt Nyström 2015b] we can understand the set on which our solution is regular in terms of the collection of sets $\Omega_t(\phi)$ that are simply connected. Now, our specific potential ϕ (Section 2F) was constructed to have curvature equal to ω_{FS} far away from the arc $\gamma=[-1,1]\subset\mathbb{C}_w\subset\mathbb{P}^1$, from which one can see that $\Omega_t(\phi)$ is a disc for sufficiently small t. This gives an open set of $\mathbb{P}^1\times\mathbb{D}$ for which the solution Φ is regular. Furthermore, by construction, $\Omega_1(\phi)$ is simply connected. We think it likely that $\Omega_t(\phi)$ is actually simply connected for all t, which would give rather precise information about the set on which our solution is regular, but it does not seem easy to prove that this is the case.

We furthermore believe that the fibrewise Laplacian of such a solution is uniformly bounded from below on the complement of S, and so has a discontinuity on the boundary ∂S where it jumps to zero. A somewhat bold conjecture would be that any solution to the HMAE is regular away from the set where it fails to have maximal rank, and is smooth away from the boundary of this set.

Acknowledgements

The authors would like to thank Valentino Tosatti for discussions on this topic, and the organisers of the American Institute of Mathematics (AIM) workshop on The Complex Monge–Ampère Equation (http://aimath.org/pastworkshops/mongeampere.html) where this problem was brought up for discussion. During this work Ross was supported by an EPSRC Career Acceleration Fellowship (EP/J002062/1).

References

[Berman and Demailly 2012] R. Berman and J.-P. Demailly, "Regularity of plurisubharmonic upper envelopes in big cohomology classes", pp. 39–66 in *Perspectives in analysis, geometry, and topology*, edited by I. Itenberg et al., Progr. Math. **296**, Springer, 2012. MR Zbl

[Bian and Guan 2009] B. Bian and P. Guan, "A microscopic convexity principle for nonlinear partial differential equations", *Invent. Math.* **177**:2 (2009), 307–335. MR Zbl

[Bian and Guan 2010] B. Bian and P. Guan, "A structural condition for microscopic convexity principle", *Discrete Contin. Dyn. Syst.* **28**:2 (2010), 789–807. MR Zbl

[Błocki 2012] Z. Błocki, "On geodesics in the space of Kähler metrics", pp. 3–19 in *Advances in geometric analysis* (Warsaw, 2009), edited by S. Janeczko et al., Adv. Lect. Math. (ALM) **21**, International, Somerville, MA, 2012. MR Zbl

[Caffarelli and Friedman 1985] L. A. Caffarelli and A. Friedman, "Convexity of solutions of semilinear elliptic equations", *Duke Math. J.* **52**:2 (1985), 431–456. MR Zbl

[Caffarelli, Guan and Ma 2007] L. Caffarelli, P. Guan, and X.-N. Ma, "A constant rank theorem for solutions of fully nonlinear elliptic equations", *Comm. Pure Appl. Math.* **60**:12 (2007), 1769–1791. MR Zbl

[Chen 2000] X. Chen, "The space of Kähler metrics", *J. Differential Geom.* **56**:2 (2000), 189–234. MR Zbl

[Chen and Tian 2008] X. X. Chen and G. Tian, "Geometry of Kähler metrics and foliations by holomorphic discs", *Publ. Math. Inst. Hautes Études Sci.* 107 (2008), 1–107. MR Zbl

[Donaldson 2002] S. K. Donaldson, "Holomorphic discs and the complex Monge–Ampère equation", *J. Symplectic Geom.* **1**:2 (2002), 171–196. MR Zbl

[Guan and Phong 2012] P. Guan and D. H. Phong, "A maximum rank problem for degenerate elliptic fully nonlinear equations", *Math. Ann.* **354**:1 (2012), 147–169. MR Zbl

[Guan, Li, and Zhang 2009] P. Guan, Q. Li, and X. Zhang, "A uniqueness theorem in Kähler geometry", *Math. Ann.* **345**:2 (2009), 377–393. MR Zbl

[Korevaar and Lewis 1987] N. J. Korevaar and J. L. Lewis, "Convex solutions of certain elliptic equations have constant rank Hessians", *Arch. Rational Mech. Anal.* **97**:1 (1987), 19–32. MR Zbl

[Lempert 1981] L. Lempert, "La métrique de Kobayashi et la représentation des domaines sur la boule", *Bull. Soc. Math. France* **109**:4 (1981), 427–474. MR Zbl

[Li 2009] Q. Li, "Constant rank theorem in complex variables", *Indiana Univ. Math. J.* **58**:3 (2009), 1235–1256. MR Zbl

[Ross and Witt Nyström 2015a] J. Ross and D. Witt Nyström, "Applications of the duality between the complex Monge–Ampère equation and the Hele-Shaw flow", preprint, 2015. To appear in *Annales de l'Institut Fourier*. arXiv

[Ross and Witt Nyström 2015b] J. Ross and D. Witt Nyström, "Harmonic discs of solutions to the complex homogeneous Monge–Ampère equation", *Publ. Math. Inst. Hautes Études Sci.* **122** (2015), 315–335. MR Zbl

[Ross and Witt Nyström 2015c] J. Ross and D. Witt Nyström, "The Hele-Shaw flow and moduli of holomorphic discs", *Compos. Math.* **151**:12 (2015), 2301–2328. MR Zbl

[Ross and Witt Nyström 2017a] J. Ross and D. Witt Nyström, "Envelopes of positive metrics with prescribed singularities", *Ann. Fac. Sci. Toulouse Math.* (6) **26**:3 (2017), 687–728. MR Zbl

[Ross and Witt Nyström 2017b] J. Ross and D. Witt Nyström, "Homogeneous Monge–Ampère equations and canonical tubular neighbourhoods in Kähler geometry", *Int. Math. Res. Not.* **2017**:23 (2017), 7069–7108. MR

[Semmes 1992] S. Semmes, "Complex Monge–Ampère and symplectic manifolds", *Amer. J. Math.* **114**:3 (1992), 495–550. MR Zbl

[Singer, Wong, Yau, and Yau 1985] I. M. Singer, B. Wong, S.-T. Yau, and S. S.-T. Yau, "An estimate of the gap of the first two eigenvalues in the Schrödinger operator", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **12**:2 (1985), 319–333. MR Zbl

[Székelyhidi and Weinkove 2016] G. Székelyhidi and B. Weinkove, "On a constant rank theorem for nonlinear elliptic PDEs", *Discrete Contin. Dyn. Syst.* **36**:11 (2016), 6523–6532. MR Zbl

Received 8 Dec 2017. Revised 6 Apr 2018. Accepted 14 May 2018.

JULIUS ROSS: j.ross@dpmms.cam.ac.uk
Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, United Kingdom

DAVID WITT NYSTRÖM: wittnyst@chalmers.se
Department of Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, Gothenburg, Sweden

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.505

A VISCOSITY APPROACH TO THE DIRICHLET PROBLEM FOR DEGENERATE COMPLEX HESSIAN-TYPE EQUATIONS

SŁAWOMIR DINEW, HOANG-SON DO AND TAT DAT TÔ

A viscosity approach is introduced for the Dirichlet problem associated to complex Hessian-type equations on domains in $\mathbb{C}^n$. The arguments are modeled on the theory of viscosity solutions for real Hessian-type equations developed by Trudinger (1990). As a consequence we solve the Dirichlet problem for the Hessian quotient and special Lagrangian equations. We also establish basic regularity results for the solutions.

1. Introduction

Partial differential equations play a pivotal role in modern complex geometric analysis. Their applications typically involve a geometric problem which can be reduced to the solvability of an associated equation. This solvability can be deduced by various methods, yet most of the basic approaches exploit a priori estimates for suitably defined weak solutions. Thus although geometers work in the smooth category, the associated weak theory plays an important role.

One of the most successful such theories is the pluripotential theory associated to the complex Monge–Ampère equation developed by Bedford and Taylor [1976; 1982], Kołodziej [1998], Guedj and Zeriahi [2005] and many others. Roughly speaking, pluripotential theory allows one to define $(i\partial\bar{\partial}u)^k$ as a measure-valued positive closed differential form (i.e., a closed positive current) for any locally bounded plurisubharmonic function, which in turn allows one to deal with nonsmooth weak solutions of Monge–Ampère equations. Unfortunately the pluripotential approach is applicable only for a limited class of nonlinear operators, such as the m-Hessian equations — see [Dinew and Kołodziej 2014; Lu 2013].

Dinew was supported by the NCN grant 2013/08/A/ST1/00312. Do was funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.02-2017.306. Tô was supported by the CFM foundation.
MSC2010: primary 35J70; secondary 32U15.
Keywords: nonlinear elliptic PDEs, viscosity solutions, pluripotential theory.

Some of the most important examples of nonlinear operators for which pluripotential tools do not seem to apply directly are the complex Hessian quotient operators. These are not only interesting for themselves but also appear in interesting geometrical problems. One such example is the Donaldson equation, which we describe below.

Given a compact Kähler manifold (X, ω) equipped with another Kähler form χ, one seeks a Kähler form $\tilde{\chi}$ cohomologous to χ such that

$$\omega \wedge \tilde{\chi}^{n-1} = c\tilde{\chi}^n, \tag{1}$$

with the constant c dependent only on the cohomology classes of χ and ω.

Donaldson [1999] introduced this equation in order to study the properness of the Mabuchi functional. Its parabolic version, known as the J-flow, was introduced independently by Donaldson [1999] and Chen [2000] and investigated afterwards by Song and Weinkove [Weinkove 2004; 2006; Song and Weinkove 2008]. It is known that (1) is not always solvable. It was shown in [Song and Weinkove 2008] that a necessary and sufficient condition for the solvability of (1) is that there exists a metric χ' in $[\chi]$, the Kähler class of χ, satisfying

$$(nc\chi' - (n-1)\omega) \wedge \chi'^{n-2} > 0 \tag{2}$$

in the sense of $(n-1, n-1)$-forms. A conjecture of Lejmi and Székelyhidi [2015] predicts that the solvability is linked to positivity of certain integrals which can be viewed as geometric stability conditions. It was also proved that, in general, these positivity conditions are equivalent to the existence of C-subsolutions introduced by Székelyhidi [2018]. They are also equivalent to the existence of parabolic C-subsolutions for the corresponding flows; see [Phong and Tô 2017]. It would be helpful to study the boundary case when we only have nonnegativity conditions; see [Fang et al. 2014] for Donaldson equation on surfaces. It is expected that in this boundary case the equation admits suitably defined singular solutions which are smooth except on some analytic set. This has been confirmed in complex dimension two in [Fang et al. 2014] but the proof cannot be generalized to higher dimensions. In fact a major part of the problem is to develop the associated theory of weak solutions for the given Hessian quotient equation. An essential problem in applying some version of pluripotential theory for this equation is that one has to define the quotient of two measure-valued operators.

In order to circumvent this difficulty one can look for a possibly different theory of weak solutions. One such approach, known as the viscosity method, was invented long ago in the real setting [Crandall et al. 1992], but was only recently introduced for complex Monge–Ampère equations by Eyssidieux, Guedj and Zeriahi [Eyssidieux et al. 2011], Wang [2012] and Harvey and Lawson [2009].

In the current note we initiate the viscosity theory for general complex nonlinear elliptic PDEs. As the manifold case is much harder we focus only on the local theory; i.e., we deal with functions defined over domains in $\mathbb{C}^n$. Precisely, let $\Omega \subset \mathbb{C}^n$ be a bounded domain. We consider the equation

$$F[u] := f(\lambda(Hu)) = \psi(x, u), \tag{3}$$

where $\lambda(Hu)$ denotes the vector of the eigenvalues of the complex Hessian Hu of the real-valued function u and $\psi : \Omega \times \mathbb{R} \to \mathbb{R}_+$ is a given nonnegative function which is weakly increasing in the second variable. We wish to point out that nonlinear PDEs appear also in geometric problems which are defined over domains in $\mathbb{C}^n$ — see for example [Collins et al. 2017], where a Dirichlet problem for the *special Lagrangian-type equation* is studied. These are the equations defined for a given function h by

$$F[u] := \sum_{i=1}^{n} \arctan \lambda_i = h(z),$$

with λ_i denoting the eigenvalues of the Hessian of u at z. In the real case, the special Lagrangian equations were introduced by Harvey and Lawson [1982] in the study of *calibrated geometries*. More precisely the graphs of gradients of the solutions correspond to calibrated minimal submanifolds. We show in Section 6 that our method can be applied to solve the Dirichlet problem for the special degenerate Lagrangian-type equation.

In our investigations we heavily rely on the corresponding real theory developed by Trudinger [1990]. It is worth pointing out that the real theory of Hessian and Hessian quotient equations is much better understood thanks to the fundamental results of [Trudinger 1995] and [Chou and Wang 2001]. Some of our results can be seen as complex analogues of the real results that can be found there. In particular we have focused on various comparison principles in Section 3. Our first major result can be summarized as follows (we refer to the next section for the definitions of the objects involved):

Theorem 1 (comparison principle). *Let Γ be the ellipticity cone associated to* (3). *Assume that the operator $F[u] = f(\lambda(Hu))$ in* (3) *satisfies*

$$f \in C^0(\overline{\Gamma}), \qquad f > 0 \quad \text{on } \Gamma, \qquad f = 0 \quad \text{on } \partial\Gamma,$$
$$f(\lambda + \mu) \geq f(\lambda) \quad \text{for all } \lambda \in \Gamma,\ \mu \in \Gamma_n.$$

Assume moreover that either

$$\sum_{i=1}^{n} \frac{\partial f}{\partial \lambda_i}\lambda_i = \sum_{i=1}^{n} f_i \lambda_i \geq \nu(f) \quad \text{in } \Gamma \qquad \text{and} \qquad \inf_{z \in \Omega} \psi(z, \cdot\,) > 0$$

for some positive increasing function ν, or

$$f \text{ is concave and homogeneous.}$$

Then any bounded subsolution u and supersolution v in Ω to (3) *satisfy*

$$\sup_{\Omega}(u - v) \leq \max_{\partial\Omega}\{(u - v)^*, 0\}.$$

We use later on this seemingly technical result to study existence, uniqueness and regularity of the associated Dirichlet problems. One of our main results is the solvability and sharp regularity for viscosity solutions to the Dirichlet problem for a very general class of operators including Hessian quotient-type equations.

Theorem 2. *The Dirichlet problem*

$$\begin{cases} F[w] = f(\lambda(Hw)) = \psi(z, w(z)), \\ w = \varphi \quad \text{on } \partial\Omega \end{cases}$$

admits a continuous solution for any bounded Γ-pseudoconvex domain Ω. Under natural growth assumptions on ψ, the solution is Hölder continuous for any Hölder continuous boundary data φ.

Another interesting topic is the comparison between viscosity and pluripotential theory whenever the latter can be reasonably defined. A guiding principle for us is the basic observation made by Eyssidieux, Guedj and Zeriahi [Eyssidieux et al. 2011] that plurisubharmonic functions correspond to viscosity subsolutions to the complex Monge–Ampère equation. We prove several analogous results for general complex nonlinear operators. It has to be stressed that the notion of a supersolution, which does not appear in pluripotential theory, is a very subtle one for nonlinear elliptic PDEs, and several alternative definitions are possible. We in particular compare these and introduce a notion of supersolution that unifies the previously known approaches.

A large part of the note is devoted to complex Hessian quotient equations in domains in $\mathbb{C}^n$. One of our goals in this case was to initiate the construction of the undeveloped pluripotential theory associated to such equations. We rely on connections with the corresponding viscosity theory. Our findings yield in particular that the natural domain of definition of these operators is *strictly smaller* than what standard pluripotential theory would predict. We prove the following theorem:

Theorem 3. *Assume that* $0 < \psi \in C^0(\Omega)$ *and* $u \in \mathrm{PSH}(\Omega) \cap L^\infty_{\mathrm{loc}}(\Omega)$ *is a viscosity subsolution of*

$$\frac{(dd^c u)^n}{(dd^c u)^{n-k} \wedge \omega^k} = \psi(z) \quad \text{in } \Omega.$$

Then

$$(dd^c u)^n \geq \psi (dd^c u)^{n-k} \wedge \omega^k$$

and

$$(dd^c u)^k \geq \binom{n}{k}^{-1} \psi \omega^k$$

in the pluripotential sense.

We guess that this observation, rather obvious in the case of smooth functions, will play an important role in the resolution of the issue caused by the division of measures.

The note is organized as follows: in the next section we collect the basic notions from linear algebra, viscosity and pluripotential theory. Then we investigate the various notions of supersolutions in [Eyssidieux et al. 2011; Lu 2013] and compare them with the complex analogue of Trudinger's supersolutions. Section 3 is devoted to the proof of a very general comparison principle. Then in Section 4 we restrict our attention to operators depending on the eigenvalues of the complex Hessian matrix of the unknown function. We show existence and uniqueness of viscosity solutions under fairly mild conditions. One subsection is devoted to the regularity of these weak solutions. Using classical methods due to Walsh [1968], see also [Bedford and Taylor 1976], we show the optimal Hölder regularity for sufficiently regular data. Section 5 is devoted to comparisons between viscosity and pluripotential subsolutions and supersolutions. Finally in Section 6 we solve the Dirichlet problem for the Lagrangian phase operator.

2. Preliminaries

In this section we collect the notation and the basic results and definitions that will be used throughout the note.

2.1. ***Linear algebra toolkit.*** We begin by introducing the notion of an admissible cone that will be used throughout the note:

Definition 4. A cone Γ in $\mathbb{R}^n$ with vertex at the origin is called admissible if:

(1) Γ is open and convex, $\Gamma \neq \mathbb{R}^n$.

(2) Γ is symmetric; i.e., if $x = (x_1, \dots, x_n) \in \Gamma$ then for any permutation of indices $i = (i_1, \dots, i_n)$, the vector $(x_{i_1}, \dots, x_{i_n})$ also belongs to Γ.

(3) $\Gamma_n \subset \Gamma$, where $\Gamma_n := \{x \in \mathbb{R}^n \mid x_i > 0,\ i \in 1, \dots, n\}$.

From the very definition it follows that Γ_n is an admissible cone. Other examples involve the Γ_k cones that we describe below:

Consider the m-th elementary symmetric polynomial defined by

$$\sigma_m(x) = \sum_{1 \le j_1 < \cdots < j_m \le n} x_{j_1} x_{j_2} \cdots x_{j_m}.$$

We shall use also the normalized version

$$S_m(x) := \binom{n}{m}^{-1} \sigma_m.$$

Definition 5. For any $m = 1, \dots n$, the positive cone Γ_m of vectors $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ is defined by

$$\Gamma_m = \{x \in \mathbb{R}^n \mid \sigma_1(x) > 0, \dots, \sigma_m(x) > 0\}. \tag{4}$$

It is obvious that these cones are open and symmetric with respect to a permutation of the x_i's. It is a nontrivial but classical fact that Γ_m is also convex.

Exploiting the symmetry of Γ, it is possible to discuss Γ-positivity for Hermitian matrices:

Definition 6. A Hermitian $n \times n$ matrix A is called Γ-positive (respectively, Γ-semipositive) if the vector of eigenvalues $\lambda(A) := (\lambda_1(A), \dots, \lambda_n(A))$ belongs to Γ (respectively, to the Euclidean closure $\overline{\Gamma}$ of Γ). The definition is independent of the ordering of the eigenvalues.

Finally one can define, following [Li 2004], the notion of Γ-admissible and Γ-subharmonic functions through the following definitions:

Definition 7. A C^2 function u defined on a domain $\Omega \subset \mathbb{C}$ is called Γ-admissible if for any $z \in \Omega$ the complex Hessian $Hu(z) := [\partial^2/(\partial z_j\, \partial \bar{z}_k)]_{j,k=1}^n$ is Γ-positive.

In particular, if Γ is an admissible cone, then $\Gamma \subset \Gamma_1$, see [Caffarelli et al. 1985], and hence we have the following corollary:

Corollary 8. *Any Γ-admissible function is subharmonic.*

Definition 9. An upper semicontinuous function v defined on a domain $\Omega \subset \mathbb{C}^n$ is called Γ-subharmonic if near any $z \in \Omega$ it can be written as a decreasing limit of local Γ-admissible functions.

We refer to [Harvey and Lawson 2009] for a detailed discussion and potential-theoretic properties of general Γ-subharmonic functions.

2.2. ***Viscosity sub- and supersolutions.*** Let Ω be a bounded domain in $\mathbb{C}^n$. Consider the equation

$$F[u] := F(x, u, Du, Hu) = 0 \quad \text{on } \Omega, \tag{5}$$

where $Du = (\partial_{z_1} u, \dots, \partial_{z_n} u)$, $Hu = (u_{j\bar{k}})$ is the Hessian matrix of u and F is continuous on $\Omega \times \mathbb{R} \times \mathbb{C}^n \times \mathcal{H}^n$. The operator F is called *degenerate elliptic* at a point (z, s, p, M) if

$$F(z, s, p, M + N) \geq F(z, s, p, M) \quad \text{for all } N \geq 0,\ N \in \mathcal{H}^n, \tag{6}$$

where $\mathcal{H}^n$ is the set of Hermitian matrices of size $n \times n$. We remark that in our case $F(z, s, p, M)$ is not necessarily degenerate elliptic everywhere on $\Omega \times \mathbb{R} \times \mathbb{C}^n \times \mathcal{H}^n$. Motivated by [Trudinger 1990], we pose the following definition:

Definition 10. A function $u \in L^\infty(\Omega)$ is a viscosity subsolution of (5) if it is upper semicontinuous in Ω and for any $z_0 \in \Omega$, and any C^2 smooth function q defined in some neighborhood of z_0 and satisfying $u \leq q,\ u(z_0) = q(z_0)$, the inequality

$$F[q](z_0) \geq 0 \tag{7}$$

holds. We also say that $F[u] \geq 0$ in the viscosity sense and q is an upper (differential) test for u at z_0.

A function $v \in L^\infty(\Omega)$ is a viscosity supersolution of (5) if it is lower semicontinuous and there are no points $z_0 \in \Omega$ and C^2 smooth functions defined locally around z_0 such that $v \geq q$ in Ω, $v(z_0) = q(z_0)$ and

$$\inf_{N \geq 0} F\big(z_0, q(z_0), Dq(z_0), N + Hq(z_0)\big) > 0. \tag{8}$$

We also say that $F[u] \leq 0$ in the viscosity sense and q is a lower (differential) test for u at z_0.

For fixed $(z, s, p) \in \Omega \times \mathbb{R} \times \mathbb{C}^n$, the set of all Hermitian matrices M such that F is degenerate elliptic at (z, s, p, M) is called the *ellipticity set* $\mathcal{A}(z, s, p)$ for the data (z, s, p). Note that the ellipticity set has the property that

$$\mathcal{A}(z, s, p) + \Gamma_n \subset \mathcal{A}(z, s, p),$$

but it may not be a cone. Throughout the note we shall however focus on the situation when the ellipticity set is a cone which is moreover constant for all the possible data sets. We then define the *ellipticity cone* associated to the operator F which is modeled on the notion of a subequation coined by Harvey and Lawson [2009]:

Definition 11. An operator $F(z, s, p, M)$ has an ellipticity cone Γ if for any M in the ellipticity set the vector $\lambda(M)$ of the eigenvalues of M belongs to the closure $\overline{\Gamma}$ of Γ. Furthermore Γ is the minimal cone with such properties.

Throughout the note we consider only the situation when Γ is an admissible cone in the sense of Definition 4. We shall make also the following additional assumption (compare with the condition (2) in Section 4.1):

$$\text{for all } \lambda \in \partial\Gamma,\ \text{ for all } (z,\ s,\ p) \in \Omega \times \mathbb{R} \times \mathbb{C}^n, \quad F(z, s, p, \lambda) \leq 0. \tag{9}$$

This condition arises naturally whenever one seeks solutions to

$$F\big(z, u(z), Du(z), Hu(z)\big) = 0$$

with pointwise Hessian eigenvalues in Γ (recall that F increases in the Γ_n-directions).

It is evident that in Definition 10 the notion of a supersolution is different and substantially more difficult than the notion of a subsolution. The reason for this is that there is no analogue for the role of the positive cone Γ_n from the case of subsolutions in the supersolutions' case. As an illustration we recall that while any plurisubharmonic function is a subsolution for $F(u) := \det(H(u)) = 0$, see [Eyssidieux et al. 2011], it is far from true that all supersolutions can be written as the negative of a plurisubharmonic function.

Below we also give another notion of a supersolution that was coined in [Eyssidieux et al. 2011] for the Monge–Ampère equation; see also [Lu 2013] for the case of the m-Hessian operator. It can be generalized for all operators admitting an elliptic admissible cone:

Definition 12. A lower semicontinuous function u is said to be a supersolution for the operator $F(z,s,p,M)$ with the associated ellipticity cone Γ if and only if for any $z_0 \in \Omega$ and every lower differential test q at z_0 for which $\lambda(Hq(z_0)) \in \overline{\Gamma}$ one has

$$F\big(z, q(z_0), Dq(z_0), Hq(z_0)\big) \leq 0.$$

Note that in the definition we limit the differential tests only to those for which $\lambda(Hq(z_0)) \in \overline{\Gamma}$.

The next proposition shows that under the assumption (9) the definition above coincides with the one from Definition 10.

Proposition 13. *Suppose that the operator $F(z, s, p, M)$ satisfies* (9). *Then a lower semicontinuous function u defined on a domain Ω is a supersolution for $F(z, s, p, M) = 0$ in the sense of Definition 12 if and only if it is a supersolution in the sense of Definition 10.*

Proof. Suppose first that u is a supersolution in the sense of Definition 12. Fix any z_0 in Ω and q a lower differential test for u at z_0. If $\lambda(Hq(z_0)) \in \Gamma$ then

$$F\big(z, q(z_0), Dq(z_0), Hq(z_0)\big) \leq 0;$$

hence taking $N = 0$ in Definition 10 we see that the condition is fulfilled. If $\lambda(Hq(z_0))$ fails to be in Γ then there is a positive definite matrix N and a positive number t such that $\lambda(Hq(z_0)+tN) \in \partial\Gamma$. But this implies that $F(z, q(z_0), Dq(z_0), Hq(z_0)+tN) \leq 0$, which fulfills the condition in Definition 10 again.

Suppose now that u is a supersolution in the sense of Definition 10. Again choose z_0 in Ω and q a lower differential test for u at z_0. We can assume that $\lambda(Hq(z_0))$ is in $\overline{\Gamma}$, for otherwise such a differential test cannot be applied in Definition 12. But then by ellipticity

$$F\big(z, q(z_0), Dq(z_0), Hq(z_0)\big) \leq F\big(z, q(z_0), Dq(z_0), Hq(z_0)+N\big) \quad \text{for all } N \geq 0,\ N \in \mathcal{H}^n.$$

The infimum over N for the right-hand side is nonpositive by definition, which implies

$$F\big(z, q(z_0), Dq(z_0), Hq(z_0)\big) \leq 0,$$

which was to be proved. □

2.3. ***Aleksandrov–Bakelman–Pucci maximum principle.*** We now recall a variant of the Aleksandrov–Bakelman–Pucci (ABP) maximum principle following [Jensen 1988]. We first recall the following definition; see [Jensen 1988]:

Definition 14. Let Ω be a bounded domain in $\mathbb{R}^n$ centered at the origin and $u \in C(\overline{\Omega})$. We define

$$E_\delta = \{x \in \Omega \mid \text{for some } p \in \overline{B(0,\delta)},\ u(z) \leq u(x) + p.(z-x) \text{ for all } z \in \Omega\}.$$

Then we have the following lemma due to Jensen [1988], which will be used in the proof of Lemma 21. Recall that a function u is said to be semiconvex if $u + k|z|^2$ is convex for a sufficiently large constant k.

Lemma 15. *Let $u \in C(\overline{\Omega})$ be semiconvex for some constant $k > 0$. If u has an interior maximum and $\sup_\Omega u - \sup_{\partial\Omega} u = \delta_0 d > 0$, where $d = \operatorname{diam}(\Omega)$, then there is a constant $C = C(n,k) > 0$ such that*

$$|E_\delta| \geq C\delta^n \quad \textit{for all } \delta \in (0, \delta_0). \tag{10}$$

Proof. As in [Jensen 1988], by regularization, we can reduce to the case when $u \in C^2(\Omega)$. Now, suppose that u has an interior maximum at x_0 and

$$\delta_0 = \frac{\sup_\Omega u - \sup_{\partial\Omega} u}{d} = \frac{u(x_0) - \sup_{\partial\Omega} u}{d},$$

where $d = \operatorname{diam}(\Omega)$.

We now prove that for $\delta < \delta_0$ we have $B(0,\delta) \subset Du(E_\delta)$. Indeed, for any $p \in B(0,\delta)$, consider the hyperplane $\ell_p(x) = h + \langle p, x\rangle$, where $h = \sup_{y \in \Omega}(u(y) - \langle p, y\rangle)$. Then we have $u(x) \leq \ell_p(x)$ on Ω and $u(x_1) = \ell_p(x_1)$ for some $x_1 \in \overline{\Omega}$. If we can prove that $x_1 \in \Omega$, then $Du(x_1) = p$, so $B(0,\delta) \subset Du(E_\delta)$. Suppose by contradiction that $x_1 \in \partial\Omega$, then

$$\begin{aligned}
\sup_\Omega u &= u(x_0) \\
&\leq \ell_p(x_1) + \langle p, x_0 - x_1\rangle \\
&= u(x_1) + \langle p, x_0 - x_1\rangle \leq \sup_{\partial\Omega} u + \delta d < \sup_{\partial\Omega} u + \delta_0 d = \sup_\Omega u,
\end{aligned}$$

and hence we get a contradiction.

Next, as we have proved that $B(0,\delta) \subset Du(E_\delta)$, by comparing volumes we infer that

$$c(n)\delta^n \leq \int_{E_\delta} |\det(D^2 u)|. \tag{11}$$

Since u is semiconvex with the constant $k > 0$ and $D^2 u \leq 0$ in E_δ, we have $|\det(D^2 u)| \leq k^n$. It follows that $|E_\delta| \geq c(n) k^{-n} \delta^n$. ☐

2.4. **Γ*-subharmonic functions.*** We have defined Γ-subharmonic functions as limits of admissible ones. Below we present the alternative viscosity and pluripotential points of view:

Let $\Omega \subset \mathbb{C}^n$ be a bounded domain. Define $\omega = dd^c|z|^2$, where $d := i(\bar\partial + \partial)$ and $d^c := \frac{i}{2\pi}(\bar\partial - \partial)$ so that $dd^c = \frac{i}{\pi}\partial\bar\partial$. Let $\Gamma \subsetneq \mathbb{R}^n$ be an admissible cone as in Definition 4. We first recall the definition of k-subharmonic function:

Definition 16. We call a function $u \in C^2(\Omega)$ k-subharmonic if for any $z \in \Omega$, the Hessian matrix $(u_{i\bar{j}})$ has eigenvalues forming a vector in the closure of the cone Γ_k.

Following the ideas of Bedford and Taylor [1982], Błocki [2005] introduced the pluripotential definition of the k-subharmonic function.

Definition 17. Let u be subharmonic function on a domain $\Omega \subset \mathbb{C}^n$. Then u is called k-subharmonic (k-sh for short) if for any collection of C^2 smooth k-sh functions $v_1, \dots, v_{k-1}$, the inequality

$$dd^c u \wedge dd^c v_1 \wedge \cdots \wedge dd^c v_{k-1} \wedge \omega^{n-k} \geq 0$$

holds in the weak sense of currents.

For a general cone Γ, we have the following definition in the spirit of viscosity theory:

Definition 18. An upper semicontinuous function u is called Γ-subharmonic (respectively, strictly Γ-subharmonic) if for any $z \in \Omega$, and any upper test function q of u at z, we have

$$\lambda(Hq(z)) \in \overline{\Gamma} \quad (\text{respectively, } \lambda(Hq(z)) \in \Gamma).$$

By definition, if u is a Γ-subharmonic function, it is a Γ-subsolution in the sense of [Székelyhidi 2018]. In particular, when $\Gamma = \Gamma_k$ for $k = 1, \dots, n$, we have u is a viscosity subsolution of the equation

$$S_k(\lambda(Hu)) = 0,$$

where

$$S_k(\lambda(Hu)) = \frac{(dd^c u)^k \wedge \omega^{n-k}}{\omega^n}.$$

Then it follows from [Eyssidieux et al. 2011; Lu 2013] that u is a k-subharmonic function on Ω; hence u is a subharmonic function if $k = 1$ and a plurisubharmonic function if $k = n$.

We also have the following definition generalizing the pseudoconvex domains; see also [Li 2004] for a similar definition for smooth domains:

Definition 19. Let Ω be a bounded domain in $\mathbb{C}^n$. We say that Ω is a Γ-pseudoconvex domain if there is a constant $C_\Omega > 0$ depending only on Ω so that $-d(z) + C_\Omega d^2(z)$ is Γ-subharmonic on $\partial\Omega$, where

$$d(z) := \operatorname{dist}(z, \partial\Omega).$$

We recall the following lemma, which was proved in [Li 2004, Theorem 3.1].

Lemma 20. *Let Ω be bounded domain in $\mathbb{C}^n$ with C^2 smooth boundary. Let $\rho \in C^2(\overline{\Omega})$ be a defining function of Ω so that $\lambda(H\rho) \in \Gamma$ on $\partial\Omega$. Then there exists a defining function $\tilde{\rho} \in C^2(\overline{\Omega})$ for Ω such that $\lambda(H\tilde{\rho}) \in \Gamma$ on $\overline{\Omega}$.*

Finally we wish to recall the survey article [Zeriahi 2013] where the reader may find a thorough discussion of the viscosity theory associated to complex Monge–Ampère-type equations.

3. Comparison principles

Comparison principles are basic tools in pluripotential theory — we refer to [Kołodziej 1998; Guedj and Zeriahi 2017] for a thorough discussion of these inequalities. In viscosity theory one compares sub- and supersolutions to the same equation. It is a crucial observation, see [Eyssidieux et al. 2011], that even though supersolutions may fail to have nice pluripotential properties, a version of the comparison principle holds for the complex Monge–Ampère equation. In this section we discuss under what assumptions such comparison principles hold for general operators.

3.1. ***A preliminary comparison principle.*** Let Ω be a bounded domain in $\mathbb{C}^n$. In this subsection we prove a comparison principle for viscosity solutions of the equation

$$F[u] := F(x, u, Du, Hu) = 0. \tag{12}$$

It is well known that mere ellipticity is insufficient to guarantee a comparison-type result. Hence we add some natural structural conditions for (12).

First of all we assume that F is decreasing in the s-variable, namely

$$\text{for all } r > 0, \quad F(z, s, p, M) - F(z, s+r, p, M) \geq 0. \tag{13}$$

This is a natural assumption in the theory, see [Zeriahi 2013], as it yields an inequality in the "right" direction for the maximum principle.

Next we assume a particular continuity property with respect to the z- and p-variables:

$$|F(z_1, s, p_1, M) - F(z_2, s, p_2, M)| \leq \alpha_z(|z_1 - z_2|) + \alpha_p(|p_1 - p_2|) \tag{14}$$

for all $z_1, z_2 \in \Omega,\ \sigma \in \mathbb{R},\ p_1, p_2 \in \mathbb{C}^n,\ M \in \mathcal{H}^n$. Here α_z and α_p are certain moduli of continuity, i.e., increasing functions defined for nonnegative reals which tend to zero as the parameter decreases to zero.

We can now state the following general comparison principle for (12).

Lemma 21. *Suppose $u \in L^\infty(\bar{\Omega})$ (respectively, $v \in L^\infty(\bar{\Omega})$) satisfies $F[u] \geq \delta$ (respectively, $F[v] \leq 0$) in Ω in the viscosity sense for some $\delta > 0$. Then*

$$\sup_{\Omega}(u - v) \leq \max_{\partial\Omega}\{(u - v)^*, 0\}, \tag{15}$$

with * denoting the standard upper semicontinuous regularization.

Proof. The idea comes from [Trudinger 1990]. We use Jensen's approximation [1988] for u, v, which is defined by

$$u^\varepsilon(z) = \sup_{z' \in \Omega}\left\{u(z') - \frac{C_0}{\varepsilon}|z' - z|^2\right\}, \quad v_\varepsilon(z) = \inf_{z' \in \Omega}\left\{v(z') + \frac{C_0}{\varepsilon}|z' - z|^2\right\}, \tag{16}$$

where $\varepsilon > 0$ and $C_0 = \max\{\operatorname{osc}_\Omega u, \operatorname{osc}_\Omega v\}$ with $\operatorname{osc}(u) = \sup u_\Omega - \inf_\Omega u$. Then the supremum and infimum in (16) are achieved at points $z^*, z_* \in \Omega$ with $|z - z^*|, |z - z_*| < \varepsilon$ provided that $z \in \Omega_\varepsilon = \{z \in \Omega \mid \operatorname{dist}(z, \partial\Omega) > \varepsilon\}$. It follows from [Caffarelli and Cabré 1995] (see also [Wang 2012] for an adaptation in the complex case) that u^ε (respectively, v_ε) is Lipschitz and semiconvex (respectively, semiconcave) in Ω_ε, with

$$|Du^\varepsilon|, |Dv_\varepsilon| \leq \frac{2C_0}{\varepsilon}, \quad Hu^\varepsilon, -Hv_\varepsilon \geq -\frac{2C_0}{\varepsilon^2}\,\mathrm{Id}, \tag{17}$$

whenever these derivatives are well-defined.

Exploiting the definition of viscosity subsolution one can show that u^ε satisfies

$$F\big(z^*, u^\varepsilon(z), Du^\varepsilon(z), Hu^\varepsilon(z)\big) \geq \delta \tag{18}$$

in the viscosity sense for all $z \in \Omega_\varepsilon$. Indeed, let q be an upper test of u^ε at z_0. Then the function

$$\tilde{q}(z) := q(z + z_0 - z_0^*) + \frac{1}{\varepsilon}|z_0 - z_0^*|^2$$

is an upper test for u at z_0^*. Therefore we get (18) as u is a viscosity subsolution. This also implies that

$$F\big(z^*, u^\varepsilon(z), Du^\varepsilon(z), N + Hu^\varepsilon(z)\big) \geq \delta \tag{19}$$

in the viscosity sense for any fixed matrix $N \geq 0$. Since any locally semiconvex (semiconcave) function is twice differentiable almost everywhere by Aleksandrov's theorem, we infer that for almost all $z \in \Omega_\varepsilon$, F is degenerate elliptic at $(z^*, u^\varepsilon(z), Du^\varepsilon(z), Hu^\varepsilon(z))$ and

$$F\big(z^*, u^\varepsilon(z), Du^\varepsilon(z), N + Hu^\varepsilon(z)\big) \geq \delta \tag{20}$$

for all $N \in \mathcal{H}^n$ such that $N \geq 0$.

We assume by contradiction that $\sup_\Omega(u - v) = u(z_0) - v(z_0) = a > 0$ for some $z_0 \in \Omega$. For any ε sufficiently small, the function $w_\varepsilon := u^\varepsilon - v_\varepsilon$ has a positive maximum on Ω_ε at some point $z_\varepsilon \in \Omega_\varepsilon$ such that $z_\varepsilon \to z_0$ as $\varepsilon \to 0$. So we can choose $\varepsilon_0 > 0$ such that for any $\varepsilon < \varepsilon_0$, we know $w_\varepsilon := u^\varepsilon - v_\varepsilon$ has a positive maximum on Ω_ε at some point $z_\varepsilon \in \Omega$ with $d(z_\varepsilon, \partial\Omega) > \varepsilon_0$. Applying the ABP maximum principle (Lemma 15), for the function w_ε on Ω_{ε_0} and for any $\lambda > 0$ sufficiently small, there exists a set $E_\lambda \subset \Omega_{\varepsilon_0}$ containing z_ε with $|E_\lambda| \geq c\lambda^n$, where c is $c(n)\varepsilon^{2n}$, such that $|Dw_\varepsilon| \leq \lambda$ and $Hw_\varepsilon \leq 0$ almost everywhere in E_λ. Since $w_\varepsilon(z_\varepsilon) > 0$, we can choose λ small enough such that $w_\varepsilon \geq 0$ in E_λ. The condition (13) and the fact that F is degenerate elliptic at $(z^*, u^\varepsilon(z), Du^\varepsilon(z), Hu^\varepsilon(z))$ for almost all $z \in E_\lambda$ imply that

$$F\big(z^*, u^\varepsilon(z), Du^\varepsilon(z), N + Hu^\varepsilon(z)\big) \leq F\big(z^*, v_\varepsilon(z), Du^\varepsilon(z), N + Hv^\varepsilon(z)\big). \tag{21}$$

Using (14) and the fact that $|D(u^\varepsilon - v_\varepsilon)| \leq \lambda$, we get

$$F\big(z^*, v_\varepsilon(z), Du^\varepsilon(z), N + Hv^\varepsilon(z)\big) \leq F\big(z^*, v_\varepsilon(z), Dv^\varepsilon(z), N + Hv^\varepsilon(z)\big) + \alpha_p(\lambda).$$

Combining with (14), (20), (21) and $|z^* - z_*| < \varepsilon$ for almost all $z \in E_\lambda$,

$$F\big(z_*, v_\varepsilon(z), Dv_\varepsilon(z), N + Hv_\varepsilon(z)\big) \geq \delta - \alpha_z(\varepsilon) - \alpha_p(\lambda). \tag{22}$$

By taking λ, and then ε sufficiently small and using the fact that v_ε is twice differentiable almost everywhere on Ω, we can find at a fixed point $z_1 \in E_\lambda$ a lower test q of v at z_1 such that

$$F\big(z_0, q(z_0), Dq(z_0), N + Hq(z_0)\big) \geq \tfrac{1}{2}\delta \tag{23}$$

for all $N \geq 0$. This contradicts the definition of viscosity supersolution. Therefore we get (15). □

Remark. By assuming more properties of F, it is possible to obtain $\delta = 0$ in the previous result. This is the case for the Monge–Ampère equation. Otherwise we need to adjust the function u to achieve a strict inequality in order to use Lemma 21.

3.2. ***Comparison principle for Hessian-type equations.*** We now consider the Hessian-type equation of the form

$$F[u] = \psi(z, u), \tag{24}$$

where $\psi \in C^0(\Omega \times \mathbb{R})$ and $F[u] = f(\lambda(Hu))$ such that

$$s \mapsto \psi(\,\cdot\,, s) \text{ is weakly increasing,} \tag{25}$$

$$f \in C^0(\overline{\Gamma}), \qquad f > 0 \quad \text{on } \Gamma, \qquad f = 0 \quad \text{on } \partial\Gamma, \tag{26}$$

$$f(\lambda + \mu) \geq f(\lambda) \quad \text{for all } \lambda \in \Gamma,\ \mu \in \Gamma_n. \tag{27}$$

First, *in order to use Lemma 21, we extend f continuously on $\mathbb{R}^n$ by taking $f(\lambda) = 0$ for all $\lambda \in \mathbb{R}^n \setminus \Gamma$.*

For a δ-independent comparison principle, we need more assumptions on F. Similarly to [Trudinger 1990], we can assume that the operator $F[u] = f(\lambda(Hu))$ satisfies

$$\sum_{i=1}^{n} \frac{\partial f}{\partial \lambda_i}\lambda_i = \sum_{i=1}^{n} f_i \lambda_i \geq \nu(f) \quad \text{in } \Gamma, \qquad \inf_{z \in \Omega} \psi(z, \cdot\,) > 0 \tag{28}$$

for some positive increasing function ν.

This condition is satisfied for example in the case of the complex Hessian equations $F[u] := \sigma_k(\lambda(Hu))$, $k \in \{1, \dots, n\}$.

We also study a *new condition*, namely

$$f \text{ is concave and homogeneous;} \tag{29}$$

i.e., $f(t\lambda) = tf(\lambda)$ for all $t \in \mathbb{R}^+$.

Theorem 22. *Let $u, v \in L^\infty(\overline{\Omega})$ be a viscosity subsolution and a viscosity supersolution of* (24) *in Ω. Assume that either f satisfies either* (28) *or* (29). *Then*

$$\sup_{\Omega}(u - v) \leq \max_{\partial\Omega}\{(u - v)^*, 0\}. \tag{30}$$

Proof. Assume first that f satisfies (28). Then following [Trudinger 1990], we set for any $t \in (1, 2)$,

$$u_t(z) = tu(z) - C(t - 1),$$

where $C = \sup_\Omega u$. Therefore we have $u_t(z) \le u(z)$ on Ω for all $t \in (1,2)$. Then for any $z_0 \in \Omega$ and an upper test function $q_t(z)$ of u_t at z_0, we have $q(z) := t^{-1}q_t(z) - C(t^{-1}-1)$ is also an upper test for u at z_0. Set $\lambda = \lambda[q](z_0)$; then $\lambda[q_t](z_0) = t\lambda$ and $q(z_0) \ge q_t(z_0)$. We also recall that the function $s \mapsto f(s\lambda)$ is increasing on $\mathbb{R}^+$, by (28), and $f(\lambda) \ge \psi(z, u(z_0))$ since q is an upper test for u at z_0. It follows that at z_0,

$$\begin{aligned} F[q_t] &= f(\lambda[q_t]) = f(t\lambda) \\ &\ge f(\lambda) + (t-1)\sum \lambda_i f_i(t^*\lambda) \\ &\ge \psi(z_0, q(z_0)) + (t-1)\sum \lambda_i f_i(t^*\lambda) \\ &\ge \psi(z_0, q_t(z_0)) + \tfrac{1}{2}(t-1)\nu(\inf_\Omega \psi(z, \inf_\Omega u)) \end{aligned}$$

for $1 \le t^* < t$, sufficiently close to 1. Therefore we have for some $\delta > 0$

$$F[u_t] \ge \psi(z, u_t) + \delta$$

in the viscosity sense in Ω. Thus the inequality (30) follows from Lemma 21.

Next, consider the second case when f is concave and homogeneous. Suppose, without loss of generality, that $0 \in \Omega$. We set

$$u_\tau(z) = u(z) + \tau(|z|^2 - R),$$

where $R = \operatorname{diam}(\Omega)$. Then for any $q_\tau \in C^2(\Omega)$ such that $q_\tau \ge u_\tau$ near z_0 and $q_\tau(z_0) = u_\tau(z_0)$, we have $q = q_\tau - \tau(|z|^2 - R) \ge q_\tau$, and q is also an upper test for u at z_0. Therefore, we have at z_0,

$$F[q_\tau] = 2^d f\left(\frac{\lambda(Hq) + \tau\mathbf{1}}{2}\right) \ge f(\lambda(Hq)) + f(\tau\mathbf{1}) \ge \psi(z_0, q_\tau) + \delta. \tag{31}$$

Therefore $F[u_\tau] \ge \psi + \delta$ in the viscosity sense. Applying Lemma 21 we get (30). □

By definition, we have the following properties of sub- and supersolutions. Their proofs follow in a straightforward way from [Crandall et al. 1992, Proposition 4.3].

Lemma 23. (a) *Let $\{u_j\}$ be viscosity subsolutions of* (24) *in Ω which are uniformly bounded from above. Then* $(\limsup_\Omega u_j)^*$ *is also a viscosity subsolution of* (24) *in* Ω.

(b) *Let $\{v_j\}$ be viscosity supersolutions of* (24) *in* Ω, *which are uniformly bounded from below. Then* $(\liminf_\Omega v_j)_*$ *is also a viscosity supersolution of* (24) *in* Ω.

Now using Perron's method, see for instance [Crandall et al. 1992], we obtain the next result:

Lemma 24. *Suppose $\underline{u}, \bar{u} \in L^\infty(\Omega)$ are a subsolution and a supersolution of* (24) *on* Ω. *Suppose that $\underline{u}_*(z) = \bar{u}^*(z)$ on the boundary of* Ω. *Then the function*

$$u := \sup\{v \in L^\infty(\Omega) \cap \mathrm{USC}(\Omega) \mid v \text{ is a subsolution of } (24),\ \underline{u} \le v \le \bar{u}\}$$

satisfies $u \in C^0(\Omega)$ and

$$F[u] = \psi(x, u) \quad \text{in } \Omega$$

is the viscosity sense.

Proof. It is straightforward that u^* is a viscosity subsolution of (24). We next prove that u_* is a supersolution of (24). Assume by contradiction that u_* is not a supersolution of (24). Then there exists a point $z_0 \in \Omega$ and a lower differential test q for u_* at z_0 such that

$$F[q](z_0) > \psi(z_0, q(z_0)). \tag{32}$$

Set $\tilde{q}(z) = q(z) + b - a|z - z_0|^2$, where $b = \frac{1}{6}ar^2$ with $a, r > 0$ small enough so that $F[\tilde{q}] \geq \psi(x, \tilde{q})$ for all $|z - z_0| \leq r$. Since $u_* \geq q$ for $|z - z_0| \leq r$, we get $u^* \geq u_* > \tilde{q}$ for $\frac{1}{2}r \leq |z - z_0| < r$. Then the function

$$w(z) = \begin{cases} \max\{u^*(z), \tilde{q}(z)\} & \text{if } |z - z_0| \leq r, \\ u^*(z) & \text{otherwise} \end{cases}$$

is a viscosity subsolution of (24). By choosing a sequence $z_n \to z_0$ so that $u(z_n) \to u_*(z_0)$, we have $\tilde{q}(z_n) \to u_*(z_0) + b$. Therefore, for n sufficiently large, we have $w(z_n) > u(z_n)$ and this contradicts the definition of u. Thus we have u_* is also a supersolution. Then it follows from Theorem 22 and $\underline{u}_*(z) = \bar{u}^*(z)$ for $z \in \partial\Omega$ that $u^* \leq u_*$ on Ω; hence $u = u_* = u^*$. □

4. Dirichlet problems

4.1. *Viscosity solutions in Γ-pseudoconvex domains.* Let $\Omega \subset \mathbb{C}^n$ be a C^2 bounded domain. In this section, we study the Dirichlet problem

$$\begin{cases} F[u] = f(\lambda(Hu)) = \psi(x, u) & \text{on } \Omega, \\ u = \varphi & \text{on } \partial\Omega, \end{cases} \tag{33}$$

where $\varphi \in C^0(\partial\Omega)$ and $\psi \in C^0(\overline{\Omega} \times \mathbb{R})$ such that $\psi > 0$ and

$$s \mapsto \psi(\,\cdot\,, s) \text{ is weakly increasing.}$$

Let $\Gamma \subsetneq \mathbb{R}^n$ be an admissible cone. We assume further that $f \in C^0(\overline{\Gamma})$ satisfies:

(1) f is concave and $f(\lambda + \mu) \geq f(\lambda)$ for all $\lambda \in \Gamma$, $\mu \in \Gamma_n$.

(2) $\sup_{\partial\Gamma} f = 0$, and $f > 0$ in Γ.

(3) f is homogeneous on Γ.

We remark that the conditions (2) and (3) imply that for any $\lambda \in \Gamma$ we have

$$\lim_{t \to \infty} f(t\lambda) = +\infty. \tag{34}$$

We now can solve (33) in the viscosity sense:

Theorem 25. *Let Ω be a C^2 bounded Γ-pseudoconvex domain in $\mathbb{C}^n$. Then the Dirichlet problem*

$$f(\lambda[u]) = \psi(x, u) \quad \textit{in } \Omega, \qquad u = \varphi \quad \textit{on } \partial\Omega$$

admits a unique admissible solution $u \in C^0(\overline{\Omega})$.

In particular, we have an L^∞ bound for u which only depends on $\|\varphi\|_{L^\infty}$ and $\|\psi(x, C)\|_{L^\infty}$ and Ω, where C is a constant depending on Ω.

Proof. By Lemma 20, there is a defining function $\rho \in C^2(\overline{\Omega})$ for Ω such that $\lambda(H\rho) \in \Gamma$ on $\overline{\Omega}$. The C^2-smoothness of the boundary implies the existence of a harmonic function h on Ω for arbitrary given continuous boundary data φ. Set

$$\underline{u} = (A_1\rho + h) + A_2\rho,$$

where $A_1 > 0$ is chosen so that $A_1\rho + h$ is admissible and A_2 will be chosen later.

By the concavity of f and (34), for A_2 sufficiently large we get

$$\begin{aligned} f(\lambda[\underline{u}]) &\geq \tfrac{1}{2} f(2\lambda[A_1\rho + h]) + \tfrac{1}{2} f(2A_2\lambda[\rho]) \\ &\geq \max_{\overline{\Omega}} \psi(x, h) \geq \psi(x, \underline{u}). \end{aligned}$$

Therefore $\underline{u}$ is a subsolution of (33).

Since h is harmonic, for each $z \in \Omega$ there is a Hermitian matrix $N \geq 0$ such that $\lambda(N + H(h)(z)) \in \partial\Gamma$. But then $f(\lambda(N + H(h)(z))) = 0$. Therefore, $\bar{v} := h$ is a supersolution of (33).

Finally, the existence of solution follows from Perron's method. We set

$$u := \sup\{w \text{ is subsolution of (33) on } \Omega,\ \underline{u} \leq w \leq \bar{v}\}.$$

As in the argument from Lemma 24 we have u^* (respectively, u_*) is a subsolution (respectively, supersolution) of (33). It follows from the comparison principle (Theorem 22) that

$$u^*(z) - u_*(z) \leq \limsup_{w \to \partial\Omega} (u^* - u_*)^+(w).$$

Since $\underline{u}$ and $\bar{v}$ are continuous and $\underline{u} = \bar{v} = \varphi$ on $\partial\Omega$ we infer that $u^* \leq u_*$ on Ω and $u^* = u_*$ on $\partial\Omega$. Therefore $u = u^* = u_*$ is a viscosity solution of (33). The uniqueness follows from the comparison principle (Theorem 22). □

As a corollary of Theorem 25, we solve the following Dirichlet problem for Hessian quotient equations:

$$\begin{cases} S_{k,\ell}(\lambda(Hu)) := (S_k/S_\ell)(\lambda(Hu)) = \psi(x, u) & \text{on } \Omega, \\ u = \varphi & \text{on } \partial\Omega, \end{cases} \tag{35}$$

where $\Omega \subset \mathbb{C}^n$ is a smooth bounded Γ_k-pseudoconvex domain, $1 \leq \ell < k \leq n$, and

$$S_k(\lambda(Hu)) = \frac{(dd^c u)^k \wedge \omega^{n-k}}{\omega^n}.$$

Note that the operator $S_{k,\ell}^{1/(k-l)}$ is concave and homogeneous; see [Spruck 2005].

Corollary 26. *The Dirichlet problem* (35) *admits a unique viscosity solution $u \in C^0(\overline{\Omega})$ for any continuous data φ.*

We also remark that a viscosity subsolution is always a Γ-subharmonic function.

Lemma 27. *Any viscosity subsolution of the equation $f(\lambda(Hu)) = \psi(z, u)$ is a Γ-subharmonic function. In particular, if u is a viscosity subsolution of the equation*

$$S_{k,\ell}(\lambda(Hu)) = \psi(z, u), \tag{36}$$

then u is k-subharmonic.

Proof. Let $z_0 \in \Omega$ and $q \in C^2_{\text{loc}}(\{z_0\})$ such that $u - q$ attains its maximum at z_0 and $u(z_0) = q(z_0)$. By definition we have

$$f(\lambda(Hq)(z_0)) > 0.$$

Observe that for any semipositive Hermitian matrix N, the function

$$\tilde{q}(z) := q(z) + \langle N(z - z_0), z - z_0\rangle$$

is also an upper test function for u at z_0. By the definition of viscosity subsolutions we have

$$f(\lambda(H\tilde{q})(z_0)) > 0. \tag{37}$$

Suppose that $\lambda(Hq)(z_0) \notin \overline{\Gamma}$. Then we can find $N \geq 0$ such that $\lambda(H\tilde{q})(z_0) \in \partial\Gamma$, so $f(\lambda(H\tilde{q})(z_0)) = 0$ by the condition (3) above, and this contradicts (37). Hence we always have $\lambda[q](z_0) \geq 0$, and so u is Γ-subharmonic. □

4.2. *Hölder continuity of Hessian-type equations.* In this subsection, we study the Hölder continuity of the viscosity solution obtained in Section 4.1 to the Dirichlet problem

$$\begin{cases} F[u] = f(\lambda(Hu)) = \psi(x, u) & \text{on } \Omega, \\ u = \varphi & \text{on } \partial\Omega, \end{cases} \tag{38}$$

where f, φ and ψ satisfy the conditions spelled out in the previous subsection. We prove the following result:

Theorem 28. *Let Ω be a strictly Γ-pseudoconvex domain. Let u be the viscosity solution of* (38). *Suppose that $\varphi \in C^{2\alpha}(\partial\Omega)$ for some $\alpha \in (0, 1)$. If additionally $\psi(z, s)$ satisfies*

(1) $|\psi(z, s)| \leq M_1(s)$ *for some L^∞_{loc} function M_1,*

(2) $|\psi(z, s) - \psi(w, s)| \leq M_2(s)|z - w|^\alpha$ *for some L^∞_{loc} function M_2,*

then $u \in C^\alpha(\overline{\Omega})$.

Remark. Classical examples, see [Bedford and Taylor 1976], show that the claimed regularity cannot be improved. Conditions (1) and (2) can be regarded as weak growth conditions and seem to be optimal. If ψ does not depend on the second variable then these conditions mean that ψ is globally bounded and contained in C^α.

Proof. The proof relies on the classical idea of Walsh [1968]. A similar argument was used by Bedford and Taylor [1976], who dealt with the complex Monge–Ampère operator. We shall apply a small adjustment in the construction of the local barriers which is due to Charabati [2016].

Suppose for definiteness that $0 \in \Omega$. Assume without loss of generality that the Γ-subharmonic function $\rho = -\operatorname{dist}(z, \partial\Omega) + C_\Omega \operatorname{dist}(z, \partial\Omega)^2$ satisfies $F(\rho) \geq 2$ (multiply ρ by a constant if necessary and exploit the homogeneity of F). Recall that ρ vanishes on $\partial\Omega$. As $\partial\Omega \in C^2$ we know that $\rho \in C^2$ near the boundary. Then it is easy to find a continuation of ρ in the interior of Ω (still denoted by ρ), so that ρ is Γ-subharmonic and satisfies $F(\rho) \geq 1$.

Fix $\xi \in \partial\Omega$. There is a uniform $C \gg 1$ (dependent on Ω, but independent of ξ) such that the function

$$g_\xi(z) := C\rho(z) - |z-\xi|^2$$

is Γ-sh. In particular $g_\xi \le 0$ in $\bar{\Omega}$.

By definition there is a constant $\widetilde{C}$ such that for any $z \in \partial\Omega$

$$\varphi(z) \ge \varphi(\xi) - \widetilde{C}|z-\xi|^{2\alpha}.$$

Consider the function $h_\xi(z) := -\widetilde{C}(-g_\xi(z))^\alpha$. Then

$$H(h_\xi(z)) \ge \widetilde{C}\alpha(1-\alpha)(-g_\xi(z))^{\alpha-2}H(g_\xi(z)), \tag{39}$$

where $\lambda(H(g_\xi(z))) \in \Gamma$; thus h_ξ is Γ-subharmonic.

Observe that

$$h_\xi(z) \le -\widetilde{C}|z-\xi|^{2\alpha} \le \varphi(z) - \varphi(\xi).$$

Thus $h_\xi(z) + \varphi(\xi)$ are local boundary barriers constructed following the method of [Charabati 2016] (in [Bedford and Taylor 1976], where the Monge–Ampère case was considered, h_ξ was simply chosen as $-(x_n)^\alpha$ in a suitable coordinate system, but this is not possible in the general case).

At this stage we recall that u is bounded a priori by Theorem 25. Hence we know that for some uniform constant A one has $F[u] \le A$ in the viscosity sense.

From the gathered information, one can produce a global barrier for u in a standard way; see [Bedford and Taylor 1976]. Indeed, consider the function $\tilde{h}(z) := \sup_\xi\{ah_\xi(z) + \varphi(\xi)\}$ for a large but uniform constant a. Using the balayage procedure it is easy to show that $F(\tilde{h}(z)) \ge A$ in the viscosity sense once a is taken large enough. Thus $\tilde{h}$ majorizes u by the comparison principle and so is a global barrier for u matching the boundary data given by φ. By construction $\tilde{h}$ is globally α-Hölder continuous.

Note on the other hand that u is subharmonic as $\Gamma \subset \Gamma_1$; thus the harmonic extension u_φ of φ in Ω majorizes u from above. Recall that u_φ is α-Hölder continuous by classical elliptic regularity.

Coupling the information for both the lower and the upper barrier one obtains

$$\text{for all } z \in \bar{\Omega},\ \text{ for all } \xi \in \partial\Omega, \quad |u(z) - u(\xi)| \le K|z-\xi|^\alpha. \tag{40}$$

Denote by K_1 the quantity $K_2 \operatorname{diam}^2(\Omega)\max\{1, f(\mathbf{1})\} + K$, where $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$ is the vector of the eigenvalues of the identity matrix, while $K_2 := \widetilde{C} f(\mathbf{1})^{-1}$ and finally $\widetilde{C}$ is the α-Lipschitz constant of ψ. Consider for a small vector $\tau \in \mathbb{C}^n$ the function

$$v(z) := u(z+\tau) + K_2|\tau|^\alpha|z|^2 - K_1|\tau|^\alpha$$

defined over $\Omega_\tau := \{z \in \Omega \mid z+\tau \in \Omega\}$.

It is easy to see by using the barriers that if $z+\tau \in \partial\Omega$ or $z \in \partial\Omega$ then

$$v(z) \le u(z) + K|\tau|^\alpha + K_2 \operatorname{diam}^2\Omega|\tau|^\alpha - K_1|\tau|^\alpha \le u(z).$$

We now claim that $v(z) \le u(z)$ in Ω_τ. By the previous inequality this holds on $\partial(\Omega_\tau)$. Suppose the claim is false and consider the open subdomain U of Ω_τ defined by $U_\tau = \{z \in \Omega_\tau \mid v(z) > u(z)\}$.

We will now prove that v is a subsolution to $F[u] = f(\lambda(Hu)) = \psi(z, u(z))$ in U. To this end pick a point z_0 and an upper differential test q for v at z_0. Observe then that $\tilde{q}(z) := q(z) - K_2|\tau|^\alpha|z|^2 - K_1|\tau|^\alpha$ is then an upper differential test for $u(\tau + \cdot)$ at the point z_0. Hence

$$\begin{aligned} F[q(z_0)] &= f\big(\lambda(H\tilde{q}(z_0)) + K_2|\tau|^\alpha \mathbf{1}\big) \\ &\geq f(\lambda(H\tilde{q}(z_0))) + K_2|\tau|^\alpha f(\mathbf{1}) \\ &\geq \psi(z_0 + \tau, u(z_0 + \tau)) + K_2|\tau|^\alpha f(\mathbf{1}), \end{aligned}$$

where we have used the concavity and homogeneity of f in the first inequality and the fact that $\tilde{q}$ is an upper differential test for $u(\tau + \cdot)$ for the second one.

Next

$$\begin{aligned} \psi(z_0 + \tau, u(z_0 + \tau)) + K_2|\tau|^\alpha f(\mathbf{1}) &\geq \psi\big(z_0 + \tau, u(z_0 + \tau) + K_2|\tau|^\alpha|z_0|^2 - K_1|\tau|^\alpha\big) + K_2|\tau|^\alpha f(\mathbf{1}) \\ &= \psi(z_0 + \tau, v(z_0)) + K_2|\tau|^\alpha f(\mathbf{1}) \\ &\geq \psi(z_0 + \tau, u(z_0)) + K_2|\tau|^\alpha f(\mathbf{1}), \end{aligned}$$

where we have exploited twice the monotonicity of ψ with respect to the second variable (and the fact that $z_0 \in U_\tau$).

Exploiting now the Hölder continuity of ψ with respect to the first variable we obtain

$$\psi(z_0 + \tau, u(z_0 + \tau)) + K_2|\tau|^\alpha f(\mathbf{1}) \geq \psi(z_0 + \tau, u(z_0)) + K_2|\tau|^\alpha f(\mathbf{1}) \geq \psi(z_0, u(z_0)).$$

This proves that $F[q(z_0)] \geq \psi(z_0, u(z_0))$ and hence $F[v(z)] \geq \psi(z, v(z))$ in the viscosity sense.

Thus over U_τ, we know v is subsolution and u is a solution, which implies by comparison principle that $v \leq u$ there, a contradiction unless the set U_τ is empty.

We have thus proven that

$$\text{for all } z \in \Omega_\tau, \quad u(z + \tau) + K_2|\tau|^\alpha|z|^2 - K_1|\tau|^\alpha \leq u(z),$$

which implies the claimed α-Hölder continuity. □

5. Viscosity vs. pluripotential solutions

Let Ω be a bounded smooth strictly pseudoconvex domain in $\mathbb{C}^n$. Let $0 < \psi \in C(\overline{\Omega} \times \mathbb{R})$ be a continuous function nondecreasing in the last variable. In this section, we study the relations between viscosity concepts with respect to the inverse σ_k-equations

$$\frac{(dd^c u)^n}{(dd^c u)^{n-k} \wedge \omega^k} = \psi(z, u) \quad \text{in } \Omega \tag{41}$$

and pluripotential concepts with respect to the equation

$$(dd^c u)^n = \psi(z, u)(dd^c u)^{n-k} \wedge \omega^k \quad \text{in } \Omega. \tag{42}$$

For the regular case, the following result was shown in [Guan and Sun 2015]:

Theorem 29 (Guan–Sun). *Let $0 < h \in C^\infty(\overline{\Omega})$ and $\varphi \in C^\infty(\partial\Omega)$. Then, there exists a smooth strictly plurisubharmonic function u in $\overline{\Omega}$ such that*

$$\frac{(dd^c u)^n}{(dd^c u)^{n-k} \wedge \omega^k} = h(z) \quad \text{in } \Omega, \qquad u = \varphi \quad \text{in } \partial\Omega. \tag{43}$$

Note that the function u in Theorem 29 is a viscosity solution of (41) in the case when $\psi(z,u) = h(z)$. Using Theorem 29, we obtain:

Proposition 30. *If $u \in C(\overline{\Omega}) \cap \mathrm{PSH}(\Omega)$ is a viscosity solution of* (41) *then there exists a sequence of smooth plurisubharmonic functions u_j in Ω such that u_j is decreasing to u and the function $(dd^c u_j)^n/((dd^c u_j)^{n-k} \wedge \omega^k)$ converges uniformly to $\psi(z,u)$ as $j \to \infty$. In particular, u is a solution of* (42) *in the pluripotential sense.*

Proof. Let $\varphi_j \in C^\infty(\partial\Omega)$ and $0 < \psi_j \in C^\infty(\overline{\Omega})$ be sequences of smooth functions such that $\varphi_j \searrow \varphi$ and $\psi_j \nearrow \psi(z,u)$ as $j \to \infty$. Then, by Theorem 29, for any $j = 1, 2, \dots$, there exists a smooth strictly plurisubharmonic function u_j in $\overline{\Omega}$ such that

$$\frac{(dd^c u_j)^n}{(dd^c u_j)^{n-k} \wedge \omega^k} = \psi_j(z) \quad \text{in } \Omega, \qquad u_j = \varphi_j \quad \text{in } \partial\Omega. \tag{44}$$

By the comparison principle, we have

$$u_1 \ge u_2 \ge \cdots \ge u_j \ge \cdots \ge u.$$

Let $C > \sup_\Omega |z|^2$. By the homogeneity and the concavity of $S_{n,n-k}^{1/k}$, we have

$$\frac{(dd^c(u_j + \varepsilon|z|^2))^n}{(dd^c(u_j + \varepsilon|z|^2))^{n-k} \wedge \omega^k} \ge \frac{(dd^c u_j)^n}{(dd^c u_j)^{n-k} \wedge \omega^k} + \varepsilon^k.$$

Then, by the comparison principle, for any $\varepsilon > 0$, there exists $N > 0$ such that

$$u_j + \varepsilon(|z|^2 - C) \le u$$

for any $j > N$. Hence, u_j is decreasing to u as $j \to \infty$. □

Observe that a continuous solution of (42) in the pluripotential sense may not be a viscosity solution of (41). For example, if a continuous plurisubharmonic function $u : \Omega \to \mathbb{R}$ depends only on $n-k-1$ variables then u is a solution of (42) in the pluripotential sense but u is not a viscosity solution of (41). Moreover, by Theorem 34, we know that a viscosity solution of (41) has to satisfy $(dd^c u)^k \ge a\omega^k$ for some $a > 0$. The following question is natural:

Question 31. *If $u \in \mathrm{PSH}(\Omega) \cap C(\overline{\Omega})$ satisfies* (42) *in the pluripotential sense and*

$$(dd^c u)^k \ge a\omega^k \tag{45}$$

for some $a > 0$, does u satisfy (41) *in the viscosity sense*?

At the end of this section, we will give the answer to a special case of this question. Now, we consider the relation between viscosity subsolutions of (41) and *pluripotential subsolutions* of (42). Recall that according to the definition in Section 2.1 for any $n \times n$ complex matrix A and $k \in \{1, \dots, n\}$, $S_k(A)$ denotes the coefficient with respect to t^{n-k} of the polynomial $\binom{n}{k}^{-1} \det(A + t\,\mathrm{Id}_n)$.

Next we prove the following technical result:

Lemma 32. *Assume that A, B are $n \times n$ complex matrices and $k \in \{1, \dots, n\}$. Then*

$$S_k(AA^*)S_k(BB^*) \geq |S_k(AB^*)|^2.$$

Proof. Denote by $a_1, \dots, a_n$ and $b_1, \dots, b_n$, respectively, the row vectors of A and B. Then

$$S_k(AA^*) = \left(\binom{n}{k}\right)^{-1} \sum_{\sharp J = k} \det(\langle a_p, a_q \rangle)_{p,q \in J},$$

$$S_k(BB^*) = \left(\binom{n}{k}\right)^{-1} \sum_{\sharp J = k} \det(\langle b_p, b_q \rangle)_{p,q \in J},$$

$$S_k(AB^*) = \left(\binom{n}{k}\right)^{-1} \sum_{\sharp J = k} \det(\langle a_p, b_q \rangle)_{p,q \in J}.$$

We will show that, for any $J = \{p_1, \dots, p_k\}$ with $1 \leq p_1 < \cdots < p_k \leq n$,

$$\det(\langle a_p, a_q \rangle)_{p,q \in J} \cdot \det(\langle b_p, b_q \rangle)_{p,q \in J} \geq |\det(\langle a_p, b_q \rangle)_{p,q \in J}|^2. \tag{46}$$

Indeed, if either $\{a_{p_1}, \dots, a_{p_k}\}$ or $\{b_{p_1}, \dots, b_{p_k}\}$ are linearly dependent then both sides of (46) are equal to 0. Otherwise, exploiting the Gram–Schmidt process, we can assume that $\{a_{p_1}, \dots, a_{p_k}\}$ and $\{b_{p_1}, \dots, b_{p_k}\}$ are orthogonal systems (observe that the quantities in question do not change during the orthogonalization process). Next normalizing the vectors a_{p_j} and b_{p_j}, $j = 1, \dots, n$, to unit length, both sides change by the same factor. Hence it suffices to prove the statement for two collections of orthonormal bases.

Under this assumption we have

$$(\langle a_p, a_q \rangle)_{p,q \in J} = (\langle b_p, b_q \rangle)_{p,q \in J} = \mathrm{Id}_k. \tag{47}$$

Let $M = (\langle a_p, b_q \rangle)_{p,q \in J}$. Then MM^* is a semipositive Hermitian matrix, and

$$\mathrm{Tr}(MM^*) = \sum_{l=1}^{k} \sum_{j=1}^{k} |\langle b_{p_j}, a_{p_l} \rangle|^2 = \sum_{j=1}^{k} \left\langle b_{p_j}, \sum_{l=1}^{k} \langle b_{p_j}, a_{p_l} \rangle a_{p_l} \right\rangle \leq \sum_{j=1}^{k} \|b_{p_j}\|^2 = k.$$

Therefore, $|\det(M)| = \sqrt{\det(MM^*)} \leq 1$; hence we obtain (46). Finally, using (46) and the Cauchy–Schwarz inequality, we infer that

$$S_k(AA^*)S_k(BB^*) \geq |S_k(AB^*)|^2,$$

as required. □

For any $n \times n$ Hermitian matrix $A = (a_{j\bar{\ell}})$, we define

$$\omega_A = \sum_{j,\ell=1}^{n} a_{j\bar{\ell}} \frac{i}{\pi} dz_j \wedge d\bar{z}_\ell$$

and

$$\mathcal{B}(A,k) := \left\{ B \in \mathcal{H}_+^n \;\middle|\; \frac{\omega_B^k \wedge \omega_A^{n-k}}{\omega^n} = 1 \right\},$$

where $k = 1, 2 \dots, n$.

Theorem 33. *Let $u \in \mathrm{PSH}(\Omega) \cap L^\infty_{\mathrm{loc}}(\Omega)$ and $0 < g \in C(\Omega)$. Then the following are equivalent*:

(i) *$(dd^c u)^n / ((dd^c u)^{n-k} \wedge \omega^k) \geq g^k(z)$ in the viscosity sense.*

(ii) *For all $B \in \mathcal{B}(\mathrm{Id}, n-k)$,*

$$(dd^c u)^k \wedge \omega_{B^2}^{n-k} \geq g^k(z)\omega^n$$

in the viscosity sense.

(iii) *For any open set $U \Subset \Omega$, there are smooth plurisubharmonic functions u_ε and functions $0 < g^\varepsilon \in C^\infty(U)$ such that u_ε are decreasing to u and g^ε converge uniformly to g as $\varepsilon \searrow 0$, and*

$$(dd^c u_\varepsilon) \wedge \omega_{A_1} \wedge \cdots \wedge \omega_{A_{k-1}} \wedge \omega_{B^2}^{n-k} \geq g^\varepsilon \omega^n \tag{48}$$

pointwise in U for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$ and $A_1, \dots, A_{k-1} \in \mathcal{B}(B^2, k)$.

(iv) *For any open set $U \Subset \Omega$, there are smooth strictly plurisubharmonic functions u_ε and functions $0 < g^\varepsilon \in C^\infty(U)$ such that the sequence u_ε is decreasing to u and the sequence g^ε converges uniformly to g as $\varepsilon \searrow 0$, and*

$$\frac{(dd^c u_\varepsilon)^n}{(dd^c u_\varepsilon)^{n-k} \wedge \omega^k} \geq (g^\varepsilon)^k \tag{49}$$

pointwise in U for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$.

Proof. (iv) $\Rightarrow$ (i) is obvious. It remains to show (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv).

(i) $\Rightarrow$ (ii): Assume that $q \in C^2$ is an upper test for u at $z_0 \in \Omega$. Then q is strictly plurisubharmonic in a neighborhood of z_0 and

$$\frac{(dd^c q)^n}{(dd^c q)^{n-k} \wedge \omega^k} \geq g^k$$

at z_0.

By using Lemma 32 for $\sqrt{Hq}$ and $(\sqrt{Hq})^{-1}B$, we have

$$\begin{aligned}\frac{(dd^c q)^{n-k} \wedge \omega^k}{(dd^c q)^n} \frac{(dd^c q)^k \wedge \omega_{B^2}^{n-k}}{\omega^n} &= \frac{(dd^c q)^{n-k} \wedge \omega^k}{\omega^n} \frac{(dd^c q)^k \wedge \omega_{B^2}^{n-k}}{(dd^c q)^n} \\ &\geq \left(\frac{\omega_B^{n-k} \wedge \omega^k}{\omega^n} \right)^2\end{aligned}$$

for any $B \in \mathcal{H}_+^n$ (observe that

$$S_{n-k}(CC^*) = \frac{(dd^c q)^k \wedge \omega_{B^2}^{n-k}}{(dd^c q)^n} \quad \text{and} \quad S_{n-k}(\sqrt{Hq}C^*) = \frac{\omega_B^{n-k} \wedge \omega^k}{\omega^n}$$

for $C = (\sqrt{Hq})^{-1} B$).

Then, for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$ we have

$$(dd^c q)^k \wedge \omega_{B^2}^{n-k} \geq g^k \omega^n$$

at z_0. Hence

$$(dd^c u)^k \wedge \omega_{B^2}^{n-k} \geq g^k \omega^n$$

in the viscosity sense.

(ii) $\Rightarrow$ (iii): Assume that $q \in C^2$ touches u from above at $z_0 \in \Omega$. Then, for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$,

$$(dd^c q)^k \wedge \omega_{B^2}^{n-k} \geq g^k \omega^n$$

at z_0. By the same arguments as in [Lu 2013], we have

$$(dd^c q) \wedge \omega_{A_1} \wedge \cdots \wedge \omega_{A_{k-1}} \wedge \omega_{B^2}^{n-k} \geq g \omega^n$$

for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$, $A_1, \dots, A_{k-1} \in \mathcal{B}(B^2, k)$. Hence

$$(dd^c u) \wedge \omega_{A_1} \wedge \cdots \wedge \omega_{A_{k-1}} \wedge \omega_{B^2}^{n-k} \geq g \omega^n \tag{50}$$

in the viscosity sense for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$, $A_1, \dots, A_{k-1} \in \mathcal{B}(B^2, k)$.

Let g_j be a sequence of smooth functions in Ω such that $g_j \nearrow g$ as $j \to \infty$. Then

$$(dd^c u) \wedge \omega_{A_1} \wedge \cdots \wedge \omega_{A_{k-1}} \wedge \omega_{B^2}^{n-k} \geq g_j \omega^n \tag{51}$$

in the viscosity sense for any $j \in \mathbb{N}$, $B \in \mathcal{B}(\mathrm{Id}, n-k)$ and $A_1, \dots, A_{k-1} \in \mathcal{B}(B^2, k)$. By the same arguments as in [Eyssidieux et al. 2011, the proof of Proposition 1.5], u satisfies (51) in the sense of positive Radon measures. Using convolution to regularize u and setting $u_\varepsilon = u * \rho_\varepsilon$, we see that u_ε is smooth strictly plurisubharmonic and

$$(dd^c u_\varepsilon) \wedge \omega_{A_1} \wedge \cdots \wedge \omega_{A_{k-1}} \wedge \omega_{B^2}^{n-k} \geq (g_j)_\varepsilon \omega^n$$

pointwise in Ω_ε. Choosing $g^\varepsilon := (g_{[1/\varepsilon]})_\varepsilon$, we obtain (48).

(iii) $\Rightarrow$ (iv): At $z_0 \in \Omega_\varepsilon$, choosing

$$B = \frac{Hu_\varepsilon(z_0)}{(S_{n-k}(Hu_\varepsilon(z_0)))^{1/(n-k)}}$$

and

$$A_1 = A_2 = \cdots = A_{k-1} = \left(\frac{(dd^c u_\varepsilon(z_0))^k \wedge \omega_{B^2}^{n-k}}{\omega^n} \right)^{-1/k} Hu_\varepsilon(z_0),$$

we get

$$
\begin{aligned}
g^{\varepsilon} &\le \left(\frac{(dd^c u_{\varepsilon}(z_0))^k \wedge \omega_{B^2}^{n-k}}{\omega^n}\right)^{1/k} \\
&= \left(\frac{(dd^c u_{\varepsilon}(z_0))^n}{\omega^n}\frac{1}{S_{n-k}(Hu_{\varepsilon}(z_0))}\right)^{1/k} \\
&= \left(\frac{(dd^c u_{\varepsilon}(z_0))^n}{\omega^n}\frac{\omega^n}{(dd^c u_{\varepsilon})^{n-k}\wedge\omega^k}\right)^{1/k} \\
&= \left(\frac{(dd^c u_{\varepsilon})^n}{(dd^c u_{\varepsilon})^{n-k}\wedge\omega^k}\right)^{1/k}
\end{aligned}
$$

pointwise in Ω_{ε}. Then

$$\frac{(dd^c u_{\varepsilon})^n}{(dd^c u_{\varepsilon})^{n-k}\wedge\omega^k} \ge (g^{\varepsilon})^k. \qquad \square$$

As a consequence, our result implies that a viscosity subsolution is a pluripotential subsolution.

Theorem 34. *Assume that $\psi(z,s)=\psi(z)$ with $\psi\in C^0(\Omega)$ and $u\in \mathrm{PSH}(\Omega)\cap L^{\infty}_{\mathrm{loc}}(\Omega)$ is a viscosity subsolution of* (41). *Then*

$$(dd^c u)^n \ge \psi(dd^c u)^{n-k}\wedge\omega^k \tag{52}$$

and

$$(dd^c u)^k \ge \binom{n}{k}^{-1}\psi\,\omega^k \tag{53}$$

in the pluripotential sense. If u is continuous then the conclusion still holds in the case where ψ depends on both variables.

Proof. By Theorem 33, for any open set $U\Subset\Omega$, there are strictly plurisubharmonic functions $u_{\varepsilon}\in C^{\infty}(U)$ and functions $0<h^{\varepsilon}\in C^{\infty}(U)$ such that u_{ε} is decreasing to u and h^{ε} converges uniformly to ψ as $\varepsilon\searrow 0$, and

$$\frac{(dd^c u_{\varepsilon})^n}{(dd^c u_{\varepsilon})^{n-k}\wedge\omega^k} \ge h^{\varepsilon} \tag{54}$$

pointwise in U. Choosing $B=\mathrm{Id}_n$ and letting $\varepsilon\to 0$, we obtain (52).

It also follows from Theorem 33 that we can choose u_{ε} and h^{ε} so that

$$(dd^c u_{\varepsilon})^k\wedge\omega_{B^2}^{n-k} \ge h^{\varepsilon}\omega^n \tag{55}$$

pointwise in U for any $B\in\mathcal{B}(\mathrm{Id},n-k)$. Fix $z_0\in U$ and $0<\varepsilon\ll 1$. We can choose complex coordinates so that $Hu_{\varepsilon}(z_0)=\mathrm{diag}(\lambda_1,\dots,\lambda_n)$, where $0\le\lambda_1\le\cdots\le\lambda_n$. Choosing

$$B=\binom{n}{k}^{1/(n-k)}\mathrm{diag}(0,\dots,\underbrace{0}_{k\text{-th}},1,\dots,1),$$

we get

$$\lambda_1\cdots\lambda_k \ge \binom{n}{k}^{-1}h^{\varepsilon}.$$

Then
$$(dd^c u_\varepsilon)^k \ge \binom{n}{k}^{-1} h^\varepsilon \omega^k$$
pointwise in U. Letting $\varepsilon \to 0$, we obtain (53). □

Remark. Note that for strictly positive ψ, (53) implies that the *natural* space of functions to consider for the Hessian quotient problem (41) is *not* the space of bounded plurisubharmonic functions but a considerably smaller one.

By assuming some additional conditions, we can also prove that a pluripotential subsolution is a viscosity one.

Proposition 35. *Assume that $\psi(z,s) = \psi(z) > 0$ with $\psi \in C^0(\Omega)$ and u is a local bounded plurisubharmonic function in Ω satisfying*
$$(dd^c u)^k \ge \psi \omega^k$$
in the pluripotential sense. Then
$$\frac{(dd^c u)^n}{(dd^c u)^{n-k} \wedge \omega^k} \ge \psi$$
in the viscosity sense.

Proof. By the assumption, for any $A \in \mathcal{H}_+^n$,
$$(dd^c u)^k \wedge \omega_A^{n-k} \ge \psi \omega^k \wedge \omega_A^{n-k} \tag{56}$$
in the pluripotential sense. By [Lu 2013], (56) also holds in the viscosity sense. If $A = B^2$ for some $B \in \mathcal{B}(\mathrm{Id}, n-k)$ then, by using Lemma 32, we have
$$\omega^k \wedge \omega_{B^2}^{n-k} \ge \left(\frac{\omega_B^{n-k} \wedge \omega^k}{\omega^n} \right)^2 \omega^n = \omega^n.$$
Then
$$(dd^c u)^k \wedge \omega_{B^2}^{n-k} \ge \psi \omega^n$$
in the viscosity sense, for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$. Applying Theorem 33, we obtain
$$\frac{(dd^c u)^n}{(dd^c u)^{n-k} \wedge \omega^k} \ge \psi$$
in the viscosity sense. □

We now discuss the notion of a *supersolution*. By the same argument as in [Guedj et al. 2017], relying on the idea from [Berman 2013], we obtain the following relation between viscosity supersolutions of (41) and pluripotential supersolutions of (42):

Proposition 36. *Let $u \in \mathrm{PSH}(\Omega) \cap C(\bar{\Omega})$ be a viscosity supersolution of* (41). *Then there exists an increasing sequence of strictly plurisubharmonic functions $u_j \in C^\infty(\bar{\Omega})$ such that u_j converges in capacity to u as $j \to \infty$, and*
$$\frac{(dd^c u_j)^n}{(dd^c u_j)^{n-k} \wedge \omega^k} \le \psi(z, u)$$

pointwise in Ω. *In particular*,

$$(dd^c u)^n \le \psi(z,u)(dd^c u)^{n-k}\wedge\omega^k$$

in the pluripotential sense.

If there exists $a>0$ *such that* $(dd^c u)^k \ge a\omega^k$ *then* u_j *can be chosen such that*

$$\frac{(dd^c u_j)^n}{(dd^c u_j)^{n-k}\wedge\omega^k}\ge b$$

pointwise in Ω *for some* $b>0$.

For the definition of convergence in capacity, we refer to [Guedj and Zeriahi 2017].

Proof. Define $\varphi = u|_{\partial\Omega}$ and $g(z)=\psi(z,u(z))$. Then, for any $j\ge 1$, there exists a unique viscosity solution v_j of

$$\begin{cases}(dd^c v_j)^n/((dd^c v_j)^{n-k}\wedge\omega^k)=e^{j(v_j-u)}g(z) & \text{in } \Omega,\\ v_j=\varphi & \text{in } \partial\Omega.\end{cases}\tag{57}$$

Applying the comparison principle to the equation

$$\frac{(dd^c v)^n}{(dd^c v)^{n-k}\wedge\omega^k}=e^{j(v-u)}g(z),$$

we get $u\ge v_j$ and $v_{j+1}\ge v_j$ for any $j\ge 1$.

Note that, by Proposition 30,

$$(dd^c v_j)^n = e^{j(v_j-u)}g(z)(dd^c v_j)^{n-k}\wedge\omega^k$$

in the pluripotential sense. For any $h\in \mathrm{PSH}(\Omega)$ such that $-1\le h\le 0$, we have

$$\begin{aligned}\varepsilon^n\int_{\{v_j<u-2\varepsilon\}}(dd^c h)^n &\le \int_{\{v_j<u+\varepsilon h-\varepsilon\}}(dd^c(u+\varepsilon h))^n\\ &\le \int_{\{v_j<u+\varepsilon h-\varepsilon\}}(dd^c v_j)^n\\ &\le \int_{\{v_j<u-\varepsilon\}} e^{j(v_j-u)}g(z)(dd^c v_j)^{n-k}\wedge\omega^k\\ &\le e^{-j\varepsilon}\int_{\{v_1<u-\varepsilon\}} g(z)(dd^c v_j)^{n-k}\wedge\omega^k\\ &\le Ce^{-j\varepsilon},\end{aligned}$$

where $C>0$ is independent of j. The last inequality holds by the Chern–Levine–Nirenberg inequalities; see [Guedj and Zeriahi 2017]. This implies that v_j converges to u in capacity.

If there exists $a>0$ such that $(dd^c u)^k\ge a\omega^k$ then, by Proposition 35,

$$\frac{(dd^c u)^n}{(dd^c u)^{n-k}\wedge\omega^k}\ge a$$

in the viscosity sense. Choosing $M \gg 1$ such that $e^{-M}\sup_\Omega g < a$, we get

$$\frac{(dd^c v_j)^n}{(dd^c v_j)^{n-k}\wedge\omega^k} \le a e^{j(v_j-u)+M}.$$

Applying the comparison principle to the equation

$$\frac{(dd^c v)^n}{(dd^c v)^{n-k}\wedge\omega^k} = a e^{j(v-u)},$$

we get $v_j + M/j \ge u$ for any $j \ge 1$. Then

$$\frac{(dd^c v_j)^n}{(dd^c v_j)^{n-k}\wedge\omega^k} = e^{j(v_j-u)}g(z) \ge e^{-M}g(z)$$

for any $j \ge 1$. Hence, by Theorem 34,

$$(dd^c v_j)^k \ge \binom{n}{k}^{-1} e^{-M}g(z) \ge \binom{n}{k}^{-1} e^{-M}\min_{\overline{\Omega}} g$$

for any $j \ge 1$.

Now, by Proposition 30, for any j we can choose a strictly plurisubharmonic function $u_j \in C^\infty(\overline{\Omega})$ such that

$$v_j - \frac{1}{2^j} \le u_j \le v_j - \frac{1}{2^{j+1}}$$

and

$$-\frac{1}{2^j} \le \frac{(dd^c u_j)^n}{(dd^c u_j)^{n-k}\wedge\omega^k} - e^{j(v_j-u)}g(z) \le 0.$$

It is easy to see that u_j satisfies the required properties. □

The next result gives the answer to a special case of Question 31:

Theorem 37. *Let $u \in \mathrm{PSH}(\Omega)\cap C(\Omega)$ such that*

$$\frac{(dd^c u)^n}{(dd^c u)^{n-k}\wedge\omega^k} \le \psi(z,u) \tag{58}$$

in the viscosity sense and

$$(dd^c u)^n \ge \psi(z,u)(dd^c u)^{n-k}\wedge\omega^k \tag{59}$$

in the pluripotential sense. If there exists $a > 0$ such that $(dd^c u)^k \ge a\omega^k$ then u is a viscosity solution of the equation

$$\frac{(dd^c u)^n}{(dd^c u)^{n-k}\wedge\omega^k} = \psi(z,u). \tag{60}$$

Proof. It remains to show that u is a viscosity subsolution of (60) in any smooth strictly pseudoconvex domain $U \Subset \Omega$.

Let V be a smooth strictly pseudoconvex domain such that $U \Subset V \Subset \Omega$. By Proposition 36, there exists an increasing sequence of strictly plurisubharmonic functions $u_j \in C^\infty(\bar{V})$ such that u_j converges in capacity to u as $j \to \infty$, and

$$b \le \frac{(dd^c u_j)^n}{(dd^c u_j)^{n-k} \wedge \omega^k} \le \psi(z, u)$$

pointwise in V, where $b > 0$. By Theorem 34, we have $(dd^c u_j)^k \ge \binom{n}{k}^{-1} b\omega^k$. Then, there exists $C > 0$ such that

$$(dd^c u_j)^{n-k} \wedge \omega^k \ge \frac{1}{\psi(z, u)}(dd^c u_j)^n \ge C\omega^n.$$

Define

$$f_j(z) := \frac{(dd^c u_j)^n}{(dd^c u_j)^{n-k} \wedge \omega^k}.$$

Then $f_j(z) \le \psi(z, u)$ for any $z \in V$, and $(\psi - f_j)(dd^c u_j)^{n-k} \wedge \omega^k \ge C(\psi - f_j)\omega^n$ converges weakly to 0. Hence f_j converges in Lebesgue measure to ψ in V as $j \to \infty$.

Now, by Theorem 33, we have

$$(dd^c u_j) \wedge \omega_{A_1} \wedge \cdots \wedge \omega_{A_{k-1}} \wedge \omega_{B^2}^{n-k} \ge (f_j)^{1/k} \omega^n$$

pointwise in V for any $B \in \mathcal{B}(\mathrm{Id}, n-k)$ and $A_1, \dots, A_{k-1} \in \mathcal{B}(B^2, k)$. Letting $j \to \infty$, we get

$$(dd^c u) \wedge \omega_{A_1} \wedge \cdots \wedge \omega_{A_{k-1}} \wedge \omega_{B^2}^{n-k} \ge \psi^{1/k} \omega^n$$

in the sense of Radon measures. It follows from [Lu 2013] that

$$(dd^c u)^k \wedge \omega_{B^2}^{n-k} \ge \psi^{1/k} \omega^n$$

in the viscosity sense. Using Theorem 33, we get that u is a viscosity subsolution of (60) in U. □

6. Dirichlet problem for the Lagrangian phase operator

In this section, we prove the existence of a unique viscosity solution to the Dirichlet problem for the Lagrangian phase operator. The existence and uniqueness of the smooth version was obtained recently by Collins, Picard and Wu [Collins et al. 2017]. Let $\Omega \subset \mathbb{C}^n$ be a bounded domain. Consider the Dirichlet problem

$$\begin{cases} F[u] := \sum_{i=1}^n \arctan \lambda_i = h(z) & \text{on } \Omega, \\ u = \varphi & \text{on } \partial\Omega, \end{cases} \tag{LA}$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of the complex Hessian Hu. We can also write $F[u] = f(\lambda(Hu))$. We assume that $\varphi \in C^0(\partial\Omega)$ and $h : \bar{\Omega} \to \big[(n-2)\frac{\pi}{2} + \delta, n\frac{\pi}{2}\big)$ is continuous for some $\delta > 0$.

The Lagrangian phase operator F in (LA) arises in geometry and mathematical physics. We refer to [Collins et al. 2015; 2017; Harvey and Lawson 1982; Jacob and Yau 2017; Yuan 2006; Wang and Yuan 2013; 2014] for the details.

Since $h \geq (n-2)\frac{\pi}{2}$, this case is called the *supercritical phase* following [Yuan 2006; Jacob and Yau 2017; Collins et al. 2015; 2017]. Recall first the following properties; see [Yuan 2006; Wang and Yuan 2014; Collins et al. 2017].

Lemma 38. *Suppose $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ satisfy $\sum_i \arctan\lambda_i \geq (n-2)\frac{\pi}{2} + \delta$ for some $\delta > 0$. Then we have*:

(1) $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1} > 0$ *and* $|\lambda_n| \leq \lambda_{n-1}$.

(2) $\sum_i \lambda_i \geq 0$ *and* $\lambda_n \geq -C(\delta)$.

(3) $\sum \lambda_i^{-1} \leq -\tan(\delta)$ *when* $\lambda_n < 0$.

(4) *For any $\sigma \in \big((n-2)\frac{\pi}{2}, n\frac{\pi}{2}\big)$, the set $\Gamma^\sigma := \big\{\lambda \in \mathbb{R}^n \;\big|\; \sum_i \arctan\lambda_i > \sigma\big\}$ is a convex set and $\partial\Gamma^\sigma$ is a smooth convex hypersurface.*

It follows from Lemma 38 that the function f can be defined on a cone Γ satisfying $\Gamma_n \subset \Gamma \subset \Gamma_1$. We also remark that if $h \geq (n-1)\frac{\pi}{2}$, then F is concave, while F has concave level sets if $(n-2)\frac{\pi}{2} h \leq (n-1)\frac{\pi}{2}$, but in general F may not be concave; see [Collins et al. 2017]. Therefore we cannot apply Theorem 22 directly. Fortunately, we still have a comparison principle for the Lagrangian operator using Lemma 21.

Lemma 39. *Let $u, v \in L^\infty(\overline{\Omega})$ be a viscosity subsolution and a viscosity supersolution of the equation $F[u] = f(\lambda(Hu)) = h$ on Ω. Then*

$$\sup_{\Omega}(u - v) \leq \max_{\partial\Omega}\{(u - v)^*, 0\}. \tag{61}$$

Proof. We first define $\varepsilon > 0$ by $\max_{\overline{\Omega}} h = n\frac{\pi}{2} - \varepsilon$. Now for any $0 < \tau \leq \frac{1}{2}\varepsilon$, set $u_\tau = u + \tau|z|^2$. Let q_τ be any upper test for u_τ at any point $z_0 \in \Omega$; then $q = q_\tau - \tau|z|^2$ is also an upper test for u at z_0. By the definition we have

$$F[q](z_0) = \sum_{i=1}^n \arctan\lambda_i(z_0) \geq h(z_0),$$

where $\lambda(z_0) = \lambda(Hq(z_0))$. We also have

$$F[q_\tau](z_0) = \sum_{i=1}^n \arctan(\lambda_i(z_0) + \tau). \tag{62}$$

Next, if $F[q](z_0) \geq n\frac{\pi}{2} - \frac{\varepsilon}{2}$, then $F[q](z_0) \geq h(z_0) + \frac{\varepsilon}{2}$; hence

$$F[q_\tau](z_0) \geq h(z_0) + \frac{\varepsilon}{2}. \tag{63}$$

Conversely, if $F[q](z_0) < n\frac{\pi}{2} - \frac{\varepsilon}{2}$, this implies that $\arctan(\lambda_n(z_0)) \leq \frac{\pi}{2} - \frac{\varepsilon}{2n}$. Combining with Lemma 38(2), we get $-C(\delta) \leq \lambda_n(z_0) \leq C(\varepsilon)$. Using the mean value theorem, there exists $\hat{\lambda}_n \in (\lambda_n(z_0), \lambda_n(z_0) + \tau)$ such that

$$\arctan(\lambda_n(z_0) + \tau) - \arctan\lambda_n(z_0) = \frac{1}{1 + \hat{\lambda}_n^2}\tau \geq C(\delta, \varepsilon, \tau) > 0.$$

It follows that

$$F[q_\tau](z_0) \geq F[q](z_0) + C(\delta, \varepsilon, \tau) \geq h(z_0) + C(\delta, \varepsilon, \tau). \tag{64}$$

Combing with (63) yields

$$F[q_\tau](z_0) \geq h(z_0) + C,$$

where $C > 0$ depends only on δ, ε, τ. We thus infer that u_τ satisfies $F[u_\tau] \geq h(z) + C$ in the viscosity sense. Therefore applying Lemma 21 to u_τ and v, then letting $\tau \to 0$, we obtain the desired inequality. □

Theorem 40. *Suppose Ω is a bounded C^2 domain. Let $\underline{u}$ be a bounded upper semicontinuous function on Ω satisfying $F[\underline{u}] \geq h(z)$ in Ω in the viscosity sense and $\underline{u} = \varphi$ on $\partial\Omega$. Then the Dirichlet equation* (LA) *admits a unique viscosity solution $u \in C^0(\Omega)$.*

Proof. It suffices to find a viscosity supersolution $\bar{u}$ for the equation $F[u] = h(z)$ satisfying $\bar{u} = \varphi$ on $\partial\Omega$. The C^2-boundary implies the existence of a harmonic function ϕ on Ω for arbitrary given continuous boundary data φ. Since $\sum_i \lambda_i(H\phi) = 0$, it follows from Lemma 38 that we have $F[\phi] < (n-2)\frac{\pi}{2} + \delta \leq h$; hence ϕ is a supersolution for (LA). The rest of the proof is similar to the one of Theorem 25, by using Lemma 39. □

Acknowledgements

We are grateful to Vincent Guedj, Duong H. Phong and Ahmed Zeriahi for useful discussions. This work was partially written during Dinew's stay at the Institut de Mathématiques de Toulouse granted by the Szolem Mandelbrojt prize and Tô's visits to the Department of Mathematics of Columbia University and the Institute of Mathematics of Jagiellonian University funded by ATUPS and EDMITT travel grants. We would like to thank all these institutions for their hospitality.

References

[Bedford and Taylor 1976] E. Bedford and B. A. Taylor, "The Dirichlet problem for a complex Monge–Ampère equation", *Invent. Math.* **37**:1 (1976), 1–44. MR Zbl

[Bedford and Taylor 1982] E. Bedford and B. A. Taylor, "A new capacity for plurisubharmonic functions", *Acta Math.* **149**:1-2 (1982), 1–40. MR Zbl

[Berman 2013] R. J. Berman, "From Monge–Ampère equations to envelopes and geodesic rays in the zero temperature limit", preprint, 2013. arXiv

[Błocki 2005] Z. Błocki, "Weak solutions to the complex Hessian equation", *Ann. Inst. Fourier* (*Grenoble*) **55**:5 (2005), 1735–1756. MR Zbl

[Caffarelli and Cabré 1995] L. A. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, American Mathematical Society Colloquium Publications **43**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Caffarelli et al. 1985] L. Caffarelli, L. Nirenberg, and J. Spruck, "The Dirichlet problem for nonlinear second-order elliptic equations, III: Functions of the eigenvalues of the Hessian", *Acta Math.* **155**:3-4 (1985), 261–301. MR Zbl

[Charabati 2016] M. Charabati, "Modulus of continuity of solutions to complex Hessian equations", *Internat. J. Math.* **27**:1 (2016), art. id. 1650003. MR Zbl

[Chen 2000] X. Chen, "On the lower bound of the Mabuchi energy and its application", *Internat. Math. Res. Notices* **2000**:12 (2000), 607–623. MR Zbl

[Chou and Wang 2001] K.-S. Chou and X.-J. Wang, "A variational theory of the Hessian equation", *Comm. Pure Appl. Math.* **54**:9 (2001), 1029–1064. MR Zbl

[Collins et al. 2015] T. C. Collins, A. Jacob, and S.-T. Yau, "(1,1) forms with specified Lagrangian phase: a priori estimates and algebraic obstructions", preprint, 2015. arXiv

[Collins et al. 2017] T. C. Collins, S. Picard, and X. Wu, "Concavity of the Lagrangian phase operator and applications", *Calc. Var. Partial Differential Equations* **56**:4 (2017), art. id. 89. MR Zbl

[Crandall et al. 1992] M. G. Crandall, H. Ishii, and P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations", *Bull. Amer. Math. Soc.* (*N.S.*) **27**:1 (1992), 1–67. MR Zbl

[Dinew and Kołodziej 2014] S. Dinew and S. Kołodziej, "A priori estimates for complex Hessian equations", *Anal. PDE* **7**:1 (2014), 227–244. MR Zbl

[Donaldson 1999] S. K. Donaldson, "Moment maps and diffeomorphisms", *Asian J. Math.* **3**:1 (1999), 1–15. MR Zbl

[Eyssidieux et al. 2011] P. Eyssidieux, V. Guedj, and A. Zeriahi, "Viscosity solutions to degenerate complex Monge–Ampère equations", *Comm. Pure Appl. Math.* **64**:8 (2011), 1059–1094. Correction in **70**:5 (2017), 815–821. MR Zbl

[Fang et al. 2014] H. Fang, M. Lai, J. Song, and B. Weinkove, "The J-flow on Kähler surfaces: a boundary case", *Anal. PDE* **7**:1 (2014), 215–226. MR Zbl

[Guan and Sun 2015] B. Guan and W. Sun, "On a class of fully nonlinear elliptic equations on Hermitian manifolds", *Calc. Var. Partial Differential Equations* **54**:1 (2015), 901–916. MR Zbl

[Guedj and Zeriahi 2005] V. Guedj and A. Zeriahi, "Intrinsic capacities on compact Kähler manifolds", *J. Geom. Anal.* **15**:4 (2005), 607–639. MR Zbl

[Guedj and Zeriahi 2017] V. Guedj and A. Zeriahi, *Degenerate complex Monge–Ampère equations*, EMS Tracts in Mathematics **26**, European Mathematical Society, Zürich, 2017. MR Zbl

[Guedj et al. 2017] V. Guedj, C. H. Lu, and A. Zeriahi, "Plurisubharmonic envelopes and supersolutions", preprint, 2017. arXiv

[Harvey and Lawson 1982] R. Harvey and H. B. Lawson, Jr., "Calibrated geometries", *Acta Math.* **148** (1982), 47–157. MR Zbl

[Harvey and Lawson 2009] F. R. Harvey and H. B. Lawson, Jr., "Dirichlet duality and the nonlinear Dirichlet problem", *Comm. Pure Appl. Math.* **62**:3 (2009), 396–443. MR Zbl

[Jacob and Yau 2017] A. Jacob and S.-T. Yau, "A special Lagrangian type equation for holomorphic line bundles", *Math. Ann.* **369**:1-2 (2017), 869–898. MR Zbl

[Jensen 1988] R. Jensen, "The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations", *Arch. Rational Mech. Anal.* **101**:1 (1988), 1–27. MR Zbl

[Kołodziej 1998] S. Kołodziej, "The complex Monge–Ampère equation", *Acta Math.* **180**:1 (1998), 69–117. MR Zbl

[Lejmi and Székelyhidi 2015] M. Lejmi and G. Székelyhidi, "The J-flow and stability", *Adv. Math.* **274** (2015), 404–431. MR Zbl

[Li 2004] S.-Y. Li, "On the Dirichlet problems for symmetric function equations of the eigenvalues of the complex Hessian", *Asian J. Math.* **8**:1 (2004), 87–106. MR Zbl

[Lu 2013] H. C. Lu, "Viscosity solutions to complex Hessian equations", *J. Funct. Anal.* **264**:6 (2013), 1355–1379. MR Zbl

[Phong and Tô 2017] D. H. Phong and D. T. Tô, "Fully non-linear parabolic equations on compact Hermitian manifolds", preprint, 2017. arXiv

[Song and Weinkove 2008] J. Song and B. Weinkove, "On the convergence and singularities of the J-flow with applications to the Mabuchi energy", *Comm. Pure Appl. Math.* **61**:2 (2008), 210–229. MR Zbl

[Spruck 2005] J. Spruck, "Geometric aspects of the theory of fully nonlinear elliptic equations", pp. 283–309 in *Global theory of minimal surfaces*, edited by D. Hoffman, Clay Math. Proc. **2**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Székelyhidi 2018] G. Székelyhidi, "Fully non-linear elliptic equations on compact Hermitian manifolds", *J. Differential Geom.* **109**:2 (2018), 337–378. MR Zbl

[Trudinger 1990] N. S. Trudinger, "The Dirichlet problem for the prescribed curvature equations", *Arch. Rational Mech. Anal.* **111**:2 (1990), 153–179. MR Zbl

[Trudinger 1995] N. S. Trudinger, "On the Dirichlet problem for Hessian equations", *Acta Math.* **175**:2 (1995), 151–164. MR Zbl

[Walsh 1968] J. B. Walsh, "Continuity of envelopes of plurisubharmonic functions", *J. Math. Mech.* **18** (1968), 143–148. MR Zbl

[Wang 2012] Y. Wang, “A viscosity approach to the Dirichlet problem for complex Monge–Ampère equations”, *Math. Z.* **272**:1-2 (2012), 497–513. MR Zbl

[Wang and Yuan 2013] D. Wang and Y. Yuan, “Singular solutions to special Lagrangian equations with subcritical phases and minimal surface systems”, *Amer. J. Math.* **135**:5 (2013), 1157–1177. MR Zbl

[Wang and Yuan 2014] D. Wang and Y. Yuan, “Hessian estimates for special Lagrangian equations with critical and supercritical phases in general dimensions”, *Amer. J. Math.* **136**:2 (2014), 481–499. MR Zbl

[Weinkove 2004] B. Weinkove, “Convergence of the J-flow on Kähler surfaces”, *Comm. Anal. Geom.* **12**:4 (2004), 949–965. MR Zbl

[Weinkove 2006] B. Weinkove, “On the J-flow in higher dimensions and the lower boundedness of the Mabuchi energy”, *J. Differential Geom.* **73**:2 (2006), 351–358. MR Zbl

[Yuan 2006] Y. Yuan, “Global solutions to special Lagrangian equations”, *Proc. Amer. Math. Soc.* **134**:5 (2006), 1355–1358. MR Zbl

[Zeriahi 2013] A. Zeriahi, “A viscosity approach to degenerate complex Monge–Ampère equations”, *Ann. Fac. Sci. Toulouse Math.* (6) **22**:4 (2013), 843–913. MR Zbl

Received 13 Jan 2018. Accepted 30 May 2018.

SŁAWOMIR DINEW: slawomir.dinew@im.uj.edu.pl
Institute of Mathematics, Jagiellonian University, Kraków, Poland

HOANG-SON DO: hoangson.do.vn@gmail.com
Institute of Mathematics, Vietnam Academy of Science and Technology, Hanoi, Vietnam

TAT DAT TÔ: tat-dat.to@math.univ-toulouse.fr
Institut Mathématiques de Toulouse, Université Paul Sabatier, Toulouse, France

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.537

RESOLVENT ESTIMATES FOR SPACETIMES BOUNDED BY KILLING HORIZONS

ORAN GANNOT

We show that the resolvent grows at most exponentially with frequency for the wave equation on a class of stationary spacetimes which are bounded by nondegenerate Killing horizons, without any assumptions on the trapped set. Correspondingly, there exists an exponentially small resonance-free region, and solutions of the Cauchy problem exhibit logarithmic energy decay.

1. Introduction

1A. ***Statement of results.*** Let (M, g) be a connected $(n+1)$-dimensional Lorentzian manifold of signature $(1, n)$ with connected boundary ∂M, satisfying the following assumptions:

(1) ∂M is a Killing horizon generated by a complete Killing vector field T, whose surface gravity is a positive constant $\kappa > 0$ (see Section 2C for details).

(2) M is stationary in the sense that there is a compact spacelike hypersurface X with boundary such that each integral curve of T intersects X exactly once.

(3) T is timelike in M°.

Consider a formally self-adjoint (with respect to the volume density) operator $L \in \mathrm{Diff}^2(M)$ commuting with T, such that $L - \Box_g \in \mathrm{Diff}^1(M)$. Thus we can write

$$L = \Box_g + \mathcal{W} + \mathcal{V},$$

where $\mathcal{W}$ is a smooth vector field and $\mathcal{V} \in \mathcal{C}^\infty(M)$. In addition, assume that $\mathcal{W}$ is tangent to ∂M.

Identify $M = \mathbb{R}_t \times X$ under the flow of T. Since T commutes with L, the composition

$$\boldsymbol{P}(\omega) = e^{i\omega t} L e^{-i\omega t} \tag{1-1}$$

descends to a differential operator on X depending on $\omega \in \mathbb{C}$. Fredholm properties of $\boldsymbol{P}(\omega)$ were first examined in a robust fashion by Vasy [2013] using methods of microlocal analysis, and subsequently by Warnick [2015] via physical space arguments; see also [Gannot 2018].

Here we summarize a simple version of these results, which applies in any strip of fixed width near the real axis. For $k \in \mathbb{N}$, let

$$\mathcal{X}^k = \{u \in H^{k+1}(X) : \boldsymbol{P}(0)u \in H^k(X)\}, \tag{1-2}$$

MSC2010: 35L05, 35P25, 35R01, 83C57.
Keywords: logarithmic decay, resolvent bounds, Killing horizons, Carleman estimates.

equipped with the graph norm. Since $\boldsymbol{P}(\omega)-\boldsymbol{P}(0)\in\mathrm{Diff}^1(X)$, the operator $\boldsymbol{P}(\omega)$ is bounded $\mathcal{X}^k\to H^k(X)$ for each $\omega\in\mathbb{C}$.

Proposition 1.1 [Vasy 2013; Warnick 2015]. *The operator $\boldsymbol{P}(\omega):\mathcal{X}^k\to H^k(X)$ is Fredholm of index 0 in the half-plane $\{\operatorname{Im}\omega>-\kappa(k+\frac{1}{2})\}$ and is invertible for $\operatorname{Im}\omega>0$ sufficiently large.*

The inverse $\boldsymbol{P}(\omega)^{-1}:H^k(X)\to\mathcal{X}^k$ forms a meromorphic family of operators in $\{\operatorname{Im}\omega>-\kappa(k+\frac{1}{2})\}$, called the resolvent family, which is independent of k in a suitable sense [Vasy 2013, Remark 2.9]. Its complex poles in $\{\operatorname{Im}\omega>-\kappa(k+\frac{1}{2})\}$ are known as resonances, and correspond to nontrivial mode solutions $v=e^{-i\omega t}u$ of the equation $\Box_g v=0$, where $u\in\mathcal{C}^\infty(M)$ satisfies $Tu=0$. Thus mode solutions with $\operatorname{Im}\omega>0$ grow exponentially in time, whereas those with $\operatorname{Im}\omega<0$ exhibit exponential decay.

Given $\omega_0, C_0>0$, define the region

$$\Omega=\{|\operatorname{Im}\omega|\le e^{-C_0|\operatorname{Re}\omega|}\}\cap\{|\omega|>\omega_0\}.$$

These parameters are fixed in the next theorem, which is the main result of this paper.

Theorem 1. *There exist $\omega_0, C_0>0$ such that $\boldsymbol{P}(\omega)$ has no resonances in Ω. Furthermore, there exists $C>0$ such that if $\omega\in\Omega$, then*

$$\|\boldsymbol{P}(\omega)^{-1}f\|_{H^{k+1}}\le e^{C|\operatorname{Re}\omega|}\|f\|_{H^k} \tag{1-3}$$

for each $k\in\mathbb{N}$ and $f\in H^k(X)$.

Theorem 1 is also true when ∂M consists of several Killing horizons generated by T, each of which has a positive, constant surface gravity. In particular, Theorem 1 applies to any stationary perturbation of the Schwarzschild–de Sitter spacetime (which is bounded by two nondegenerate Killing horizons [Vasy 2013, Section 6]) that preserves the timelike nature of T, and for which the horizons remain nondegenerate Killing horizons. Other examples are even asymptotically hyperbolic spaces in the sense of [Guillarmou 2005].

1B. *Energy decay.* Theorem 1 can be used to prove logarithmic decay to constants for solutions of the Cauchy problem

$$\Box_g v=0,\quad v|_X=v_0,\quad Tv|_X=v_1. \tag{1-4}$$

Given initial data $(v_0,v_1)\in H^{k+1}(X)\times H^k(X)$, the equation (1-4) admits a unique solution

$$v\in\mathcal{C}^0(\mathbb{R}_+;H^{k+1}(X))\cap\mathcal{C}^1(\mathbb{R}_+;H^k(X)).$$

If N denotes the future-pointing unit normal to the level sets of t and $Q[v]$ is the stress-energy tensor (see Section 4C) associated to v, define the energy

$$\mathcal{E}[v](s)=\int_{\{t=s\}}Q[v](N,N)\,dS_X.$$

Here dS_X is the induced volume density on $X=\{t=0\}$, which is isometric to each time slice $\{t=s\}$. Since N is timelike, it is well known that $\mathcal{E}[v](s)$ is positive definite in dv. One consequence of the positivity of κ is the energy-boundedness statement

$$\mathcal{E}[v](t) \le C\mathcal{E}[v](0); \tag{1-5}$$

see for instance [Warnick 2015, Corollary 3.9]. One can also define an energy $\mathcal{E}_k[v]$ controlling all derivatives up to order k, with $\mathcal{E}[v] = \mathcal{E}_1[v]$, which is similarly uniformly bounded. This can be improved to a logarithmic energy-decay statement uniformly up to the horizon, with a loss of derivatives.

Corollary 1.2. *Given $k \in \mathbb{N}$, there exists $C > 0$ such that*

$$\mathcal{E}_k[v](t)^{1/2} \le \frac{C}{\log(2+t)} \|(v_0, v_1)\|_{\mathcal{X}^k \times H^{k+1}}$$

for each $v \in \mathcal{C}^0(\mathbb{R}_+; H^{k+1}(X)) \cap \mathcal{C}^1(\mathbb{R}_+; H^k(X))$ solving the Cauchy problem (1-4) *with initial data $(v_0, v_1) \in \mathcal{X}^k \times H^{k+1}(X)$.*

We can also improve Corollary 1.2 by showing that v decays logarithmically to a constant as follows. Given $(v_0, v_1) \in \mathcal{X}^k \times H^{k+1}(X)$, define the constant

$$v_\infty = \operatorname{vol}(\partial X)^{-1} \int_X \bigl(A^{-2} v_1 - 2A^{-2} W v_0 - \operatorname{div}_g(A^{-2} W) v_0\bigr) A \, dS_X.$$

Here $A > 0$ is the lapse function and W is the shift vector as described in Section 2D.

Corollary 1.3. *Given $k \in \mathbb{N}$, there exists $C > 0$ such that*

$$\|v(t) - v_\infty\|_{H^{k+1}} + \|\partial_t v(t)\|_{H^k} \le \frac{C}{\log(2+t)} \|(v_0, v_1)\|_{\mathcal{X}^k \times H^{k+1}}$$

for each $v \in \mathcal{C}^0(\mathbb{R}_+; H^{k+1}(X)) \cap \mathcal{C}^1(\mathbb{R}_+; H^k(X))$ solving the Cauchy problem (1-4) *with initial data $(v_0, v_1) \in \mathcal{X}^k \times H^{k+1}(X)$.*

By Sobolev embedding, Corollary 1.3 can be used to deduce pointwise decay estimates as well.

1C. ***Relationship with previous work.*** The analogue of Theorem 1 was first established for compactly supported perturbations of the Euclidean Laplacian in the landmark paper [Burq 1998]. There have been subsequent improvements and simplifications in the asymptotically Euclidean setting [Burq 2002; Vodev 2000; Datchev 2014], while Rodnianski and Tao [2015] considered asymptotically conic spaces. In a different direction, Holzegel and Smulevici [2013] established logarithmic energy decay on slowly rotating Kerr–AdS spacetimes, which contain a Killing horizon of the type described here in addition to a conformally timelike boundary. However, their approach made heavy use of the symmetries of Kerr–AdS, and is not adaptable to our setting.

Most relevant to the setting considered here are [Moschidis 2016; Cardoso and Vodev 2002]. The former reference shows logarithmic energy decay on Lorentzian spacetimes which may contain Killing horizons, but importantly also contain at least one asymptotically flat end. There, the mechanism of decay is radiation into the asymptotically flat region. In contrast, asymptotically flat ends are not considered in the present paper, but we do allow spacetimes which contain Killing horizons as their only boundary components. We therefore stress that the results of [Moschidis 2016] are disjoint from those of this paper.

Meanwhile, [Cardoso and Vodev 2002] applies to a wide class of Riemannian metrics, including those with hyperbolic ends. There is a close connection between asymptotically hyperbolic manifolds and black holes spacetimes, first exploited in the study of resonances by Sá Barreto and Zworski [1997]. This relationship has attracted a great deal of interest, especially following the paper [Vasy 2013]; for a survey of recent developments, see [Zworski 2017].

Common to the works described above is the use of Carleman estimates in the interior of the geometry, which is then combined with some other (typically more complicated) analysis near infinity. Although the proof of Theorem 1 adopts techniques from [Burq 1998; Moschidis 2016; Rodnianski and Tao 2015], one novelty (and simplifying feature) is that the Carleman estimate employed here is valid up to and including the horizon. In particular, this avoids the use of separation of variables and special function methods [Burq 1998; Holzegel and Smulevici 2013; Vodev 2000], Mourre-type estimates [Burq 2002], and spherical energies [Cardoso and Vodev 2002; Datchev 2014; Moschidis 2016; Rodnianski and Tao 2015].

2. Preliminaries

2A. *Semiclassical rescaling.* It is conceptually convenient to rescale the operator by

$$P(z) = h^2 \boldsymbol{P}(h^{-1}z). \tag{2-1}$$

Thus $\omega = h^{-1}z$, and uniform bounds on $P(z)$ for $\pm z$ in a compact set $[a, b] \subset (0, \infty)$ give high-frequency bounds for $\boldsymbol{P}(\omega)$ as $|\omega| \to \infty$. Theorem 1 is easily seen to be equivalent to the following.

Theorem 1′. *Given* $[a, b] \subset (0, \infty)$, *there exist* $C, C_1 > 0$ *such that*

$$\|u\|_{H_h^{k+1}} \le e^{C/h} \|P(z)u\|_{H_h^k} \tag{2-2}$$

for each $u \in \mathcal{X}^k$ *and* $\pm z \in [a, b] + ie^{-C_1/h}[-1, 1]$.

The norms in (2-2) are semiclassically rescaled Sobolev norms. For detailed expositions on semiclassical analysis, the reader is referred to [Zworski 2012] and [Dyatlov and Zworski 2018, Appendix E].

2B. *Stationarity.* A tensor on M will be called stationary if it is annihilated by the Lie derivative $\mathcal{L}_T$. The definition of stationarity can be extended to T^*M by observing that T lifts to a vector field on T^*M via the identification

$$T^*M = T^*\mathbb{R} \oplus T^*X.$$

Any covector $\varpi \in T_q^*M$ at a point $q = (t, x)$ can be decomposed as $\varpi = \xi + \tau\, dt$, where $\xi \in T_x^*X$ and $\tau\, dt \in T_t^*\mathbb{R}$. Thus a function $F \in \mathcal{C}^\infty(T^*M)$ is stationary if it depends only on $\xi \in T_x^*X$ and $\tau \in \mathbb{R}$, which we sometimes denote by $F(x, \xi, \tau)$. Furthermore, if $\tau = \tau_0$ is fixed, then F induces a function $F(\,\cdot\,, \tau_0)$ on T^*X. This is compatible with the Poisson bracket in the sense that for stationary $F_1, F_2 \in \mathcal{C}^\infty(T^*M)$, there is the equality

$$\{F_1, F_2\}(x, \xi, \tau_0) = \{F_1(\,\cdot\,, \tau_0), F_2(\,\cdot\,, \tau_0)\}(x, \xi). \tag{2-3}$$

On the left is the Poisson bracket on T^*M, and on the right the Poisson bracket on T^*X.

In particular, this discussion applies to the dual metric function $G \in \mathcal{C}^\infty(T^*M)$, whose value at $\varpi \in T_q^*M$ is given by

$$G(x, \varpi) = g_x^{-1}(\varpi, \varpi) = g^{\alpha\beta}(x)\varpi_\alpha \varpi_\beta.$$

The semiclassical principal symbol $p = \sigma_h(P(z))$ is given by $p(x, \xi; z) = -G(x, \xi - z\,dt)$.

Lemma 2.1. *The quadratic form $(x, \xi) \mapsto G(x, \xi)$ is negative definite on T^*X°.*

Proof. The condition $\tau = 0$ implies that $\varpi = \xi + 0\,dt$ is orthogonal to $T^\flat$. But $T^\flat$ is timelike on M°, whence the result follows. □

If $\tau_0 \in \mathbb{R}$ is fixed and $K \subset X^\circ$ is compact, then by Lemma 2.1 there exist $c, R > 0$ such that if $G(x, \xi) \geq R$, then

$$G(x, \xi + \tau_0\,dt) \geq cG(x, \xi)$$

for each $\xi \in T_K^*X^\circ$, where the constants c, R are locally uniform in τ_0. In particular, given a compact interval $I \subset \mathbb{R}$, the set

$$\{\xi \in T_K^*X^\circ : G(\xi + \tau\,dt) = 0 \text{ for some } \tau \in I\}$$

is a compact subset of T^*X°. This also implies that if Q is a stationary quadratic form on T^*M, then there exists $C > 0$ such that

$$|Q(x, \xi + \tau\,dt)| \leq C(1 + |G(x, \xi + \tau\,dt)|)$$

for each $\xi \in T_K^*X^\circ$ and $\tau \in I$.

2C. *Killing horizons and surface gravity.* Recall the hypotheses on (M, g) described in Section 1A and set

$$\mu = g(T, T).$$

The key property of (M, g) is that ∂M is a Killing horizon generated by T. By definition, this means that ∂M is a null hypersurface which agrees with a connected component of the set $\{\mu = 0,\ T \neq 0\}$. Of course in this case T is nowhere-vanishing. Since orthogonal null vectors are collinear, there is a smooth function $\kappa : \partial M \to \mathbb{R}$, called the surface gravity, such that

$$\nabla_g \mu = -2\kappa T \tag{2-4}$$

on ∂M. The nondegeneracy assumption means that $\kappa > 0$, and for simplicity it is assumed that κ is in fact constant along ∂M.

2D. *Properties of the metric.* Let N denote the future-pointing unit normal to the level sets of t, and define the lapse function $A > 0$ by $A^{-2} = g^{-1}(dt, dt)$. The shift vector is given by the formula

$$W = T - AN,$$

which by construction is tangent to the level sets of t. Let k denote the induced (positive definite) metric on X. If (x^i) are local coordinates on X, then

$$g = (A^2 - k_{ij}W^iW^j)\,dt^2 - 2k_{ij}W^i dx^j\,dt - k_{ij}\,dx^i\,dx^j.$$

Inverting this form of the metric gives

$$g^{-1} = A^{-2}(\partial_t - W^i\,\partial_i)^2 - k^{ij}\,\partial_i\,\partial_j. \tag{2-5}$$

Note that $k(W, W) = A^2 - \mu$, and hence $W \neq 0$ near ∂M.

Now use the condition that ∂M is a Killing horizon generated by T. The covariant form of (2-4) reads

$$\partial_i \mu = 2\kappa W^j k_{ij}. \tag{2-6}$$

By assumption $\kappa > 0$, so W is a nonzero inward-pointing normal to X along ∂X whose length with respect to k is A.

Introduce geodesic normal coordinates (r, y^A) on X near ∂X, so r is the distance to ∂X (uppercase indices will always range over $A = 2, \dots, n$). By construction, ∂_r is an inward-pointing unit normal along ∂X, so

$$W^r = A, \quad W^A = 0 \tag{2-7}$$

along the boundary. Also by construction, the components of the induced metric in (r, y^A)-coordinates satisfy $k^{rr} = 1$ and $k^{rA} = 0$.

Lemma 2.2. *The function r satisfies $g^{-1}(dr, dr) = -2\kappa A^{-1} r + r^2 \mathcal{C}^\infty(M)$.*

Proof. First observe that $k_{AB} W^A W^B \in r^2 \mathcal{C}^\infty(M)$ by (2-7), and since $k(W, W) = A^2 - \mu$,

$$A^2 - \mu = (W^r)^2 + k_{AB} W^A W^B.$$

Now μ and r are both boundary-defining functions, so $\mu = fr$ for some $f \in C^\infty(M)$, and hence $d\mu = f\,dr$ on ∂X. But on the boundary $\langle d\mu \rangle = 2\kappa A^2$ from (2-6), while $\langle W, dr\rangle = W^r = A$ from (2-7). Thus

$$\mu = fr = 2\kappa Ar + r^2 \mathcal{C}^\infty(M).$$

Plugging this back into the equation for $k(W, W)$ yields

$$(W^r)^2 = A^2 - 2\kappa Ar + r^2 \mathcal{C}^\infty(M),$$

and therefore $g^{-1}(dr, dr) = -k^{rr} + A^{-2}(W^r)^2 = -2\kappa A^{-1} r + r^2 \mathcal{C}^\infty(M)$ as desired. $\square$

Observe that the surface gravity depends on the choice of null generator T. Consider the rescaled vector field

$$\widehat{T} = \frac{T}{2\kappa},$$

which changes the time coordinate by the transformation $\hat{t} = 2\kappa t$. If $\widehat{\boldsymbol{P}}(\hat{\omega})$ is now defined as in (1-1) but with t replacing $\hat{t}$, then

$$\boldsymbol{P}(\omega) = \widehat{\boldsymbol{P}}\left(\frac{\omega}{2\kappa}\right).$$

It suffices to prove Theorem 1 for $\widehat{\boldsymbol{P}}(\omega)$ then, since rescaling the frequency only changes the constants ω_0, C_0, C. Dropping the hat notation, it will henceforth be assumed that $\kappa = \frac{1}{2}$.

Next, consider a conformal change $g = f\tilde{g}$, where $f > 0$ is stationary. The operator L can then be written as

$$L = f^{-1}\Box_{\tilde{g}} + (n-1)f^{-2}\nabla_{\tilde{g}} f + \mathcal{W} + \mathcal{V}. \tag{2-8}$$

Thus we can write $L = f^{-1}\tilde{L}$, where $\tilde{L}$ has the same form as L but with $\tilde{g}$ replacing g, provided that the vector field $\nabla_{\tilde{g}} f$ is tangent to ∂M. But this follows from the stationarity of f, since

$$g(T, \nabla_g f) = 0$$

and T is normal to ∂M. Thus it suffices to prove Theorem 1 with $\tilde{L}$ replacing L. Observe that ∂M remains a Killing horizon generated by T with respect to $\tilde{g}$, and the surface gravity is unchanged.

By making a conformal change and dropping the tilde notation, it will also be assumed that

$$g^{-1}(dr, dr) = -r. \tag{2-9}$$

If (τ, ρ, η_A) are dual variables to (t, r, y^A), define a stationary quadratic form $G_0 \in \mathcal{C}^\infty(T^*M)$ by

$$G_0 = -r\rho^2 - 2\rho\tau - k_0^{AB}\eta_A\eta_B. \tag{2-10}$$

Here k_0 is the restriction of k to ∂M, which is then extended to a neighborhood of ∂M by requiring that $\mathcal{L}_{\partial_r} k_0 = 0$. In the next section, the difference $G - G_0$ will be analyzed.

2E. *Negligible tensors.* We now define a class of tensors which will arise as errors throughout the proof of Theorem 1′.

Definition 2.3. (1) A stationary 1-tensor $F^\alpha\,\partial_\alpha$ is said to be negligible if its components in a coordinate system (t, r, y^A) satisfy

$$F^t \in r\mathcal{C}^\infty(M), \quad F^r \in r^2\mathcal{C}^\infty(M), \quad F^A \in r\mathcal{C}^\infty(M).$$

(2) A stationary 2-tensor $H^{\alpha\beta}\,\partial_\alpha\,\partial_\beta$ is said to be negligible if its components in a coordinate system (t, r, y^A) satisfy

$$H^{tt} \in \mathcal{C}^\infty(M), \quad H^{rr} \in r^2\mathcal{C}^\infty(M), \quad H^{AB} \in r\mathcal{C}^\infty(M),$$
$$H^{tA} \in \mathcal{C}^\infty(M), \quad H^{tr} \in r\mathcal{C}^\infty(M), \quad H^{rA} \in r\mathcal{C}^\infty(M).$$

Observe that negligibility is invariant under those coordinate changes which leave (t, r) invariant. Denote by $\mathcal{N}_1$ and $\mathcal{N}_2$ all $\mathcal{C}^\infty(T^*M)$ functions of the forms $F^\alpha\varpi_\alpha$ and $H^{\alpha\beta}\varpi_\alpha\varpi_\beta$, respectively.

Recall the definition of G_0 in (2-10). The notion of negligibility is motivated by the fact that

$$G = G_0 + \mathcal{N}_2.$$

This follows directly from (2-5), (2-7), and (2-9). We will also repeatedly reference the auxiliary functions

$$Y = (r\rho)^2 + \tau^2, \quad Z = r\rho^2 + k^{AB}\eta_A\eta_B. \tag{2-11}$$

It follows immediately from the Cauchy–Schwarz inequality $2ab < \delta a^2 + b^2/\delta$ that there exists $C > 0$ satisfying

$$Z \leq C\Big(|G_0| + \frac{\tau^2}{r}\Big). \tag{2-12}$$

The next two lemmas also follow from judicious applications of the Cauchy–Schwarz inequality and the trivial observation that $(r\rho)^2 = r(r\rho^2)$ is small relative to $r\rho^2$ for small values of r.

Lemma 2.4. *Let $F \in \mathcal{N}_1$. Then, for each $\gamma > 0$ there exists C_γ such that*

$$r^{-1}|\tau||F| \leq C_\gamma \tau^2 + \gamma Z.$$

Furthermore, $\rho\mathcal{N}_1 \subset \mathcal{N}_2$ and $\mathcal{N}_1 \cdot \mathcal{N}_1 \subset r\mathcal{N}_2$.

Lemma 2.5. *Let $H \in \mathcal{N}_2$. Then, for each $\gamma > 0$ there exist $C_\gamma, r_\gamma > 0$ such that*

$$|H| \leq C_\gamma Y + \gamma k^{AB}\eta_A\eta_B, \quad |H| \leq C_\gamma \tau^2 + \gamma Z$$

for $r \in [0, r_\gamma]$.

Now combine Lemma 2.5 with the bound (2-12) and the relation $G = G_0 + \mathcal{N}_2$. Thus there exists $R > 0$ and $C > 0$ such that

$$Z \leq C\Big(|G| + \frac{\tau^2}{r}\Big) \tag{2-13}$$

for $r \in [0, R]$.

The next goal is to compute the Poisson brackets $\{G, r\}$ and $\{G, \{G, r\}\}$. To begin, observe that

$$\{G_0, r\} = -2(r\rho + \tau), \quad \{G_0, \{G_0, r\}\} = 2(r\rho^2 + 2\tau\rho). \tag{2-14}$$

In order to replace G_0 with G we also need to consider the Poisson brackets of functions in $\mathcal{N}_1$ and $\mathcal{N}_2$.

Lemma 2.6. *The Poisson bracket satisfies $\{\mathcal{N}_2, r\} \subset \mathcal{N}_1$ and $\{\mathcal{N}_2, \mathcal{N}_1\} \subset \mathcal{N}_2$, as well as $\{G_0, \mathcal{N}_1\} \subset \mathcal{N}_2$ and $\{\{G_0, r\}, \mathcal{N}_2\} \subset \mathcal{N}_2$. Therefore,*

$$\{G, r\} = -2(r\rho + \tau) + \mathcal{N}_1, \quad \{G, \{G, r\}\} = 2(r\rho^2 + 2\tau\rho) + \mathcal{N}_2. \tag{2-15}$$

Furthermore, $\{G, \{G, r\}\} = -2r\rho^2 + \mathcal{N}_2$ whenever $\{G, r\} = 0$.

Proof. The first part is a direct calculation, while (2-15) follows from the first part and (2-14). The last statement follows from the inclusion $\rho\mathcal{N}_1 \subset \mathcal{N}_2$. □

3. Carleman estimates in the interior

3A. ***Statement of result.*** We now prove a Carleman estimate valid in the interior X°, but with uniform control over the exponential weight near ∂X.

Recall that r denotes the distance on X to the boundary with respect to the induced metric. Although this function is only well-defined in a small neighborhood of ∂X, for notational convenience we will assume that $[0, 3]$ is contained in the range of r (otherwise it is just a matter of replacing 3 with 3ε for an appropriate $\varepsilon > 0$).

Proposition 3.1. *Given $[a, b] \subset (0, \infty)$, there exists $r_1 \in (0, 1)$ and $\varphi_1, \varphi_2 \in \mathcal{C}^\infty(X)$ such that*

- *on $\{r \leq 1\}$ the functions φ_1, φ_2 are equal and depend only on r,*
- *$\varphi_i'(r) < 0$ is constant on $\{r \leq r_1\}$ for $i = 1, 2$,*

with the following property: given a compact set $K \subset X^\circ$ there exists $C > 0$ such that

$$\|(e^{\varphi_1/h} + e^{\varphi_2/h})u\|_{H^2_h(X)} \leq Ch^{-1/2}\|(e^{\varphi_1/h} + e^{\varphi_2/h})P(z)u\|_{L^2(X)}$$

for each $u \in \mathcal{C}_c^\infty(K^\circ)$ and $\pm z \in [a, b]$.

It clearly suffices to prove Proposition 3.1 for the operator $L = \Box_g$, since the lower-order terms can be absorbed as errors. In order to prove Theorem 1′, an additional estimate is needed near the boundary; this is achieved in Section 4 below.

3B. *The conjugated operator.* Given $\varphi \in \mathcal{C}^\infty(X)$, define the conjugated operator

$$P_\varphi(z) = e^{\varphi/h}P(z)e^{-\varphi/h}.$$

Let $p_\varphi(z)$ denote its semiclassical principal symbol. Define $L^2(X)$ with respect to the density $A \cdot dS_X$, where recall dS_X is the induced volume density on X, and $A > 0$ is the lapse function as in Section 2C. Defining $\operatorname{Re} P_\varphi(z)$ and $\operatorname{Im} P_\varphi(z)$ with respect to this inner product, integrate by parts to find

$$\|P_\varphi(\omega)u\|^2_{L^2(X)} = \langle P_\varphi(\omega)P_\varphi(\omega)^* u, u\rangle_{L^2(X)} + i\langle[\operatorname{Re} P_\varphi(\omega), \operatorname{Im} P_\varphi(\omega)]u, u\rangle_{L^2(X)} \tag{3-1}$$

for $u \in \mathcal{C}_c^\infty(X^\circ)$. The idea is to find φ which satisfies Hörmander's hypoellipticity condition

$$\{\operatorname{Re} p_\varphi, \operatorname{Im} p_\varphi\} > 0 \tag{3-2}$$

on the characteristic set $\{p_\varphi = 0\}$.

In order to apply the results of Section 2E without introducing additional notation, it is convenient to work with the dual metric function G directly. Define

$$G_\varphi(x, \varpi) = G(x, \varpi + i\, d\varphi),$$

so since we are assuming that τ is real, $\operatorname{Re} G_\varphi(x, \varpi) = G(x, \varpi) - G(x, d\varphi)$, and $\operatorname{Im} G_\varphi(x, \varpi) = (H_G\varphi)(x, \varpi)$. We will then construct φ (viewed as a stationary function on M) such that

$$\{\operatorname{Re} G_\varphi, \operatorname{Im} G_\varphi\}(x, \varpi) = (H_G^2\varphi)(x, \varpi) + (H_G^2\varphi)(x, d\varphi) > 0 \tag{3-3}$$

on $\{G_\varphi = 0\} \cap \{a \leq \pm\tau \leq b\}$. This will imply the original hypoellipticity condition from the discussion surrounding (2-3) and the identifications

$$p_\varphi(x, \xi; z) = -G_\varphi(x, \xi - z\, dt), \quad z = -\tau.$$

Note that the dual variable τ is now playing the role of a *rescaled* time frequency.

3C. ***Constructing the phase in a compact set.*** To avoid any undue topological restrictions, we will actually construct two weights φ_1, φ_2 in the interior which agree outside a large compact set. This appears already in [Burq 1998], but we will follow the closely related presentation in [Moschidis 2016; Rodnianski and Tao 2015].

Lemma 3.2. *There exist positive functions $\psi_1, \psi_2 \in \mathcal{C}^\infty(X)$ with the following properties*:

(1) *ψ_1, ψ_2 have finitely many nondegenerate critical points, all of which are contained in $\{r > 2\}$.*

(2) *$\psi_2 > \psi_1$ on $\{d\psi_1 = 0\}$, and $\psi_1 > \psi_2$ on $\{d\psi_2 = 0\}$.*

(3) *The functions ψ_1, ψ_2 are equal and depend only on r in $\{r \le 2\}$. Furthermore $\partial_r\psi_1$ and $\partial_r\psi_2$ are negative in this region.*

Proof. Let $\zeta \in \mathcal{C}^\infty(\{r \ge 2\})$ solve the boundary value problem

$$\Delta_k \zeta = 1, \quad \zeta|_{\{r=2\}} = 1.$$

Here Δ_k is the nonpositive Laplacian with respect to the induced metric k. Since $\Delta_k\zeta > 0$, none of the critical points of ζ in $\{r > 2\}$ are local maxima. In addition, since ζ clearly achieves its maximum at each point of $\{r = 2\}$, its outward-pointing normal derivative is strictly positive by Hopf's lemma [Gilbarg and Trudinger 1983, Lemma 3.4]. By construction, the outward-pointing unit normal is $-\partial_r$; hence $\zeta' < 0$ near $\{r = 2\}$ (for the remainder of the proof, prime will denote differentiation with respect to r).

The first step is to replace ζ by a Morse function. We may for instance embed $\{r \ge 2\}$ into a compact manifold X_0 without boundary, and approximate an arbitrary smooth extension of ζ to X_0 by a Morse function in the $\mathcal{C}^\infty(X_0)$ topology. Restricting to $\{r \ge 2\}$ and again calling this replacement ζ, we still have that ζ has no local maximum in $\{r > 2\}$ and $\zeta' < 0$ near $\{r = 2\}$. In particular, all critical points of ζ are nondegenerate and lie in a compact subset of $\{r > 2\}$.

Now fix any function $\bar\zeta = \bar\zeta(r) \in \mathcal{C}^\infty(\{r < 3\})$ such that $\bar\zeta' < 0$ everywhere, and $\bar\zeta \ge \zeta$ on their common domain of definition $\{2 \le r < 3\}$. Choose a cutoff $H = H(r) \in \mathcal{C}^\infty(X; [0, 1])$ such that

$$H = 1 \quad \text{for } r < 2+\gamma, \qquad \operatorname{supp} H \subset \{r \le 2+2\gamma\},$$

and $H' \le 0$. Set $\psi_1 = H\bar\zeta + (1-H)\zeta$, and compute $\psi_1' = H'(\bar\zeta - \zeta) + H\bar\zeta' + (1-H)\zeta'$. If $\gamma > 0$ is sufficiently small, then $\psi_1' < 0$ in a neighborhood of $\operatorname{supp} H$, since the sum of the last two terms is strictly positive on $\operatorname{supp} H$. On the other hand, outside of such a neighborhood the only critical points of ψ_1 are those of ζ.

Let $p_1, \dots, p_n$ enumerate the necessarily finite number of critical points of ψ_1, and choose $\gamma > 0$ such that the closed geodesic balls $B(p_1, \gamma), \dots, B(p_n, \gamma)$ are mutually disjoint and $B(p_j, \gamma) \subset \{r > 2\}$ for each j. Since p_j is not a local maximum, for each j there is a point $q_j \in B(p_j, r)$ such that

$$\psi_1(q_j) > \psi_1(p_j).$$

Now choose a diffeomorphism $g : X \to X$ which is the identity outside the union of the $B(q_j, r)$ and exchanges p_j with q_j. Then, set $\psi_2 = \psi_1 \circ g$. By construction the only critical points of ψ_2 are $q_1, \dots, q_n$, and furthermore

$$\psi_2(p_j) > \psi_1(p_j), \quad \psi_1(q_j) > \psi_2(q_j)$$

for each j. Since outside of $\{r > 2\}$ the functions $\psi_1 = \psi_2$ depend on r only, the proof is complete, adding an appropriate constant if necessary to ensure that both functions are positive. □

Let $B_1 \subset \{r > 2\}$ be a closed neighborhood of $\{d\psi_1 = 0\}$ such that $\psi_2 > \psi_1$ on B_1, and likewise for B_2, exchanging the roles of ψ_1 and ψ_2. Also, let $U_i \subset B_i$ be additional neighborhoods of $\{d\psi_i = 0\}$. Now define

$$\varphi_i = \exp(\alpha\psi_i), \quad i = 1, 2, \tag{3-4}$$

where $\alpha > 0$ is a parameter. The following lemma is a standard computation which is included for the sake of completeness.

Lemma 3.3. *Given $\varepsilon > 0$ and $\tau_0 > 0$, there exists $\alpha_0 > 0$ such that if $\alpha \geq \alpha_0$, then*

$$\{\operatorname{Re} G_{\varphi_i}, \operatorname{Im} G_{\varphi_i}\} > 0$$

on $(\{G_{\varphi_i} = 0\} \cap \{r \geq \varepsilon\} \cap \{|\tau| \leq \tau_0\}) \setminus T^*_{U_i} M$ for $i = 1, 2$.

Proof. The subscript $i = 1, 2$ will be suppressed. Use the definition (3-4) to compute

$$H_G\varphi = \alpha e^{\alpha\psi} H_G\psi, \quad H_G^2\varphi = \alpha^2 e^{\alpha\psi}(H_G\psi)^2 + \alpha e^{\alpha\psi} H_G^2\psi.$$

Assume that $G_\varphi(x, \varpi) = 0$. It follows from $\operatorname{Im} G_\varphi(x, \varpi) = 0$ that $(H_G\varphi)(x, \varpi) = 0$, and hence $(H_G\psi)(x, \varpi) = 0$. Therefore by (3-3),

$$\{G - G(x, d\varphi), H_G\varphi\}(x, \varpi) = \alpha e^{\alpha\psi}(H_G^2\psi)(x, \varpi) + \alpha^3 e^{3\alpha\psi}(H_G^2\psi)(x, d\psi) + \alpha^4 e^{3\alpha\psi}|G(x, d\psi)|^2.$$

Next, use the condition $(\operatorname{Re} G_\varphi)(x, \varpi) = 0$, which implies $G(x, \varpi) = \alpha^2 e^{2\alpha\psi} G(x, d\psi)$. By the discussion following Lemma 2.1, there exists $C > 0$ such that

$$|(H_G^2\psi)(x, \varpi)| \leq C(1 + |G(x, \varpi)|)$$

on $\{r \geq \varepsilon\} \cap \{|\tau| \leq \tau_0\}$. Thus on the set $\{G_\varphi = 0\} \cap \{r \geq \varepsilon\} \cap \{|\tau| \leq \tau_0\}$,

$$|\alpha e^{\alpha\psi}(H_G^2\psi)(x, \varpi)| + |\alpha^3 e^{3\alpha\psi}(H_G^2\psi)(x, d\psi)| \leq C\alpha^3 e^{3\alpha\psi}.$$

On the other hand, as soon as $d\psi \neq 0$ the third term $\alpha^4 e^{3\alpha\psi}|G(x, d\psi)|^2$ is positive by Lemma 2.1, and dominates the previous two terms for large $\alpha > 0$. Since $d\psi \neq 0$ away from B, the proof is complete. □

3D. ***Constructing the phase outside of a compact set.*** The most delicate part of the argument is the construction of the phase outside of a compact set. Since $g^{-1}(dr, dr) = -r$ and φ is a function only of r in this region,

$$G_\varphi = G + r(\varphi')^2 + i\varphi' H_G r.$$

Now compute the Poisson bracket

$$\begin{aligned}\{\operatorname{Re} G_\varphi, \operatorname{Im} G_\varphi\} &= \{G + r(\varphi')^2, \varphi' H_G r\} \\ &= \varphi' H_G^2 r + \varphi''(H_G r)^2 - ((\varphi')^3 + 2r(\varphi')^2\varphi'')\,\partial_\rho H_G r.\end{aligned}$$

Assume that $\varphi' < 0$, in which case $\operatorname{Im} G_\varphi = 0$ is equivalent to $H_G r = 0$. The goal is then to arrange negativity of the term

$$H_G^2 r - ((\varphi')^2 + 2r\varphi'\varphi'')\,\partial_\rho H_G r \tag{3-5}$$

on the set $\{\operatorname{Re} G_\varphi = 0\}$. Recall the definition of Z from (2-11).

Lemma 3.4. *There exists $C > 0$ and $R > 0$ such that $Z \le C(r(\phi')^2 + \tau^2/r)$ on $\{\operatorname{Re} G_\varphi = 0\} \cap \{0 < r \le R\}$.*

Proof. Apply (2-13), using that $\operatorname{Re} G_\varphi = 0$ implies $G = -r(\varphi')^2$. □

Putting everything together, it is now easy compute $H_G^2 r$ on $\{G_\varphi = 0\}$ near the boundary.

Lemma 3.5. *For each $\delta > 0$ there exists $R_\delta > 0$ such that*

$$\left| H_G^2 r + \frac{2\tau^2}{r} \right| \le \delta\Big(r(\phi')^2 + \frac{\tau^2}{r} \Big)$$

on $\{G_\varphi = 0\} \cap \{0 < r \le R_\delta\}$.

Proof. From the expression (2-15) for $H_G^2 r$ and Lemma 2.5, find $C_\gamma > 0$ and $r_\gamma > 0$ such that

$$|H_G^2 r + 2r\rho^2| < C_\gamma |\tau|^2 + \gamma Z \tag{3-6}$$

for $r \in (0, r_\gamma)$. Now multiply $H_G r$ by ρ, and use that $\rho\mathcal{N}_1 \subset \mathcal{N}_2$. Therefore by Lemma 2.5, there exist $C'_\gamma > 0$ and $r'_\gamma > 0$ such that

$$|2r\rho^2 + 2\tau\rho| < C'_\gamma |\tau|^2 + \gamma Z \tag{3-7}$$

for $r \in (0, r'_\gamma)$. On the other hand, from $H_G r = 0$, deduce that $-\tau\rho = \tau^2/r + \tau r^{-1}\mathcal{N}_1$. By Lemma 2.4, there exists $C''_\gamma > 0$ such that

$$\left| 2\tau\rho + \frac{2\tau^2}{r} \right| < C''_\gamma |\tau|^2 + \gamma Z. \tag{3-8}$$

Combine (3-6), (3-7), and (3-8) via the triangle inequality with Lemma 3.4 to find that

$$\left| H_G^2 r + \frac{2\tau^2}{r} \right| < 3\gamma C\Big(r(\phi')^2 + \frac{\tau^2}{r} \Big) + (C_\gamma + C'_\gamma + C''_\gamma)\tau^2$$

for $r \in (0, \min\{r_\gamma, r'_\gamma, R\})$; here $C > 0$ and $R > 0$ are provided by Lemma 3.4. Finally, choose γ sufficiently small depending on δ and a corresponding $R_\delta > 0$ such that the conclusion of the lemma holds for $r \in (0, R_\delta)$. □

Next, observe that $-\partial_\rho H_G r = 2r + r^2\mathcal{C}^\infty(M)$. Given $a > 0$, it follows from (3-5) and Lemma 3.5 that there exists $R_1 > 0$ such that

$$(\varphi')^{-1}\{\operatorname{Re} G_\varphi, \operatorname{Im} G_\varphi\} < -\frac{3a^2}{2r} + 3r(\varphi')^2 + 3r^2\varphi'\varphi'' \tag{3-9}$$

on $\{G_\varphi = 0\} \cap \{0 < r \le R_1\} \cap \{|\tau| \ge a\}$, provided that $\varphi'' \ge 0$.

Shrinking R_1 if necessary, it may be assumed that $\psi = \psi_i$ as in Lemma 3.2 satisfies $\psi' < 0$ on $[0, R_1 + 1]$. Recalling that $\varphi_i = \exp(\alpha\psi_i)$, choose $\alpha > 0$ satisfying the conclusion of Lemma 3.3 with $\varepsilon = R_1$. By further increasing α, (but keeping $a > 0$ fixed), it may also be assumed that $\varphi = \varphi_i$ satisfies

$$3(\varphi'(R_1)R_1)^2 > a^2, \qquad \varphi''(r) \geq -\frac{\varphi'(r)}{r} \quad \text{for } r \in [R_1, R_1 + 1]. \tag{3-10}$$

Although φ is already defined on all of X, the following lemma allows one to redefine φ on $\{r < R_1 + 1\}$ in such a way that its derivative is controlled; this new extension will still be denoted by φ. The idea comes from [Burq 1998, Section 3.1.2], but of course the form of the operator there is quite different.

Lemma 3.6. *There exists an extension of $\varphi = \varphi_i$ from $\{r \geq R_1 + 1\}$ to $\{r < R_1 + 1\}$ such that*

$$\{\operatorname{Re} G_\varphi, \operatorname{Im} G_\varphi\} > 0$$

on $\{G_\varphi = 0\} \cap \{0 < r \leq R_1\} \cap \{|\tau| \geq a\}$. Furthermore, there exists $r_1 \in (0, R_1)$ such that $\varphi'(r) < 0$ is constant for $r \in [0, r_1]$.

Proof. Motivated by (3-9), consider the differential equation

$$-\frac{a^2}{r} + 3rk^2 + 3r^2kk' = 0, \quad k(R_1) = \varphi'(R_1) < 0.$$

This is a Bernoulli equation whose solution is given by

$$k(r) = -r^{-1}\left((\varphi'(R_1)R_1)^2 + \frac{2}{3}a^2 \log\left(\frac{r}{R_1}\right)\right)^{1/2}.$$

The solution is certainly meaningful for $r \in [R_0, R_1]$, where we define R_0 by

$$R_0 = R_1 \exp\left(\frac{1}{2} - \frac{3}{2}\left(\varphi'(R_1)\frac{R_1}{a}\right)^2\right).$$

Note that we indeed have $R_0 < R_1$ by the assumption (3-10). The value R_0 was chosen such that $k'(R_0) = 0$, and it is easy to see that $k'(r) > 0$ for $r \in (R_0, R_1]$. In addition, $k(R_0) < 0$. Let $\theta = \theta(r)$ be defined on $[0, R_1 + 1]$ by

$$\theta(r) = \begin{cases} \varphi'(r), & r \in [R_1, R_1 + 1], \\ k(r), & r \in [R_0, R_1], \\ k(R_0), & r \in [0, R_0]. \end{cases}$$

The function θ is strictly negative, and the piecewise continuous function θ' satisfies

$$-\frac{a^2}{r} + 3r\theta^2 + 3r^2\theta\theta' \leq 0$$

for $r \in (0, R_1 + 1]$. Indeed, by the construction of k and R_0, the inequality holds for $r \in (0, R_1)$, and it is also true for $r \in (R_1, R_1 + 1]$ by (3-10). Rearranging,

$$\theta' \geq \frac{a^2}{3r^3\theta} - \frac{\theta}{r} \tag{3-11}$$

for $r \in (0, R_1 + 1]$.

We now proceed to mollify θ in such a way that the hypotheses of the lemma hold. Let $\eta_\varepsilon(r) = (1/\varepsilon)\eta(r/\varepsilon)$ denote a standard mollifier, where $\eta \in \mathcal{C}_c^\infty((-1, 1))$ has integral one. In addition, choose a cutoff $H = H(r) \in \mathcal{C}^\infty(X; [0, 1])$ such that

$$H = 1 \quad \text{for } r < R_1 + \tfrac{1}{4}, \qquad H = 0 \quad \text{for } r > R_1 + \tfrac{1}{2},$$

and $H' \leq 0$. Now define

$$\theta_\varepsilon = (1 - H)\theta + \eta_\varepsilon * (H\theta).$$

Clearly θ_ε is smooth, and $\theta_\varepsilon \to \theta$ uniformly for $r \in [0, R_1 + 1]$. Furthermore, there exists $\varepsilon_0 > 0$ such that if $\varepsilon \in (0, \varepsilon_0)$, then the following properties are satisfied:

- $\theta_\varepsilon(r) < 0$ and $\theta_\varepsilon'(r) \geq 0$ for $r \in [0, R_1 + 1]$.
- $\theta_\varepsilon(r) = \varphi'(r)$ for $r \in \left[R_1 + \frac{3}{4}, R_1 + 1\right]$,
- There exists $r_1 \in (0, R_0]$ such that $\theta_\varepsilon(r) = k(R_0)$ for $r \in [0, r_1]$.

Since θ is continuous and piecewise smooth,

$$\theta_\varepsilon' = (1 - H)\theta' - H'\theta + \eta_\varepsilon * (H'\theta + H\theta'). \tag{3-12}$$

Therefore by (3-11),

$$\theta_\varepsilon' \geq -H'\theta + \eta_\varepsilon * (H'\theta) + (1 - H)\left(\frac{a^2}{3r^3\theta} - \frac{\theta}{r}\right) + \eta_\varepsilon * \left(H\left(\frac{a^2}{3r^3\theta} - \frac{\theta}{r}\right)\right)$$

for $r \in (0, R_1 + 1]$. The right-hand side converges uniformly to $a^2/(3r^3\theta) - \theta/r$ for $r \in [r_1, R_1 + 1]$ since the latter function is continuous there. Since $\theta_\varepsilon \to \theta$ uniformly for $r \in [r_1, R_1 + 1]$ as well, there exists $\varepsilon \in (0, \varepsilon_0)$ such that

$$-\frac{3a^2}{2r} + 3r\theta_\varepsilon^2 + 3r^2\theta_\varepsilon\theta_\varepsilon' \leq 0$$

for $r \in [r_1, R_1 + 1]$. This inequality is also true for $r \in (0, r_1)$, since $\theta_\varepsilon = k(R_0)$ on that interval. Now extend φ from $\{r \geq R_1 + 1\}$ to $\{r < R_1 + 1\}$ by the formula

$$\varphi(r) = \varphi(R_1 + 1) + \int_{R_1+1}^{r} \theta_\varepsilon(s)\, ds.$$

This completes the proof according to (3-9) by observing that the φ just constructed satisfies $\varphi''(r) \geq 0$. $\square$

As a remark, if $\tau \neq 0$, then the hypoellipticity condition also holds along $\{r = 0\}$, simply because $\operatorname{Im} G_\varphi \neq 0$ in that case. However, since $(x, \xi) \mapsto G(x, \xi)$ is not elliptic along $\{r = 0\}$, the hypoellipticity condition alone, stated here in the semiclassical setting, is not sufficient to prove a Carleman estimate; see [Hörmander 1963, Section 8.4].

Now that the phases φ_1, φ_2 have been constructed globally, we are ready to finish the proof of Proposition 3.1. Here we come back to the operator $P_\varphi(z)$ on X. Fix a norm $|\cdot|$ on the fibers of T^*X (for instance using the induced metric k) and let $\langle\xi\rangle = (1 + |\xi|^2)^{1/2}$.

Proof of Proposition 3.1. Recall that we are given $[a, b] \subset (0, \infty)$ and a compact set $K \subset X^\circ$. Without loss, we may assume that $K = \{r \geq \varepsilon\}$ for some $\varepsilon > 0$. Let B_i, U_i be as in the discussion preceding Lemma 3.3. In particular,

$$\{\operatorname{Re} p_{\varphi_i}, \operatorname{Im} p_{\varphi_i}\} > 0$$

on $(\{p_{\varphi_i} = 0\} \cap \{r \geq \varepsilon/2\}) \setminus T^*_{U_i} X$. Let $\chi_i \in \mathcal{C}^\infty_c(B_i^\circ)$ be such that $\chi_i = 1$ near U_i. If $\varphi = \varphi_i$, then

$$|p_\varphi|^2 + \chi^2 + h\{\operatorname{Re} p_\varphi, \operatorname{Im} p_\varphi\} \geq h(M|p_\varphi|^2 + M\chi^2 + \{\operatorname{Re} p_\varphi, \operatorname{Im} p_\varphi\})$$

for any $M > 0$, provided that $h > 0$ is sufficiently small. On the other hand, the set $\{\operatorname{Re} p_\varphi = 0\} \cap \{r \geq \varepsilon/2\}$ is compact by Lemma 2.1, uniformly for $\pm z \in [a, b]$. Therefore,

$$\langle \xi \rangle^{-4}(M|p_\varphi|^2 + M\chi^2 + \{\operatorname{Re} p_\varphi, \operatorname{Im} p_\varphi\}) > 0$$

near $T^*X \cap \{r \geq \varepsilon/2\}$ for $M > 0$ sufficiently large. By (3-1) and the semiclassical Gårding inequality applied to $e^{\varphi_i/h}u$,

$$h\|e^{\varphi_i/h}u\|^2_{H^2_h(X)} \leq C\|e^{\varphi_i/h}P(z)u\|^2_{L^2(X)} + C\|e^{\varphi_i/h}u\|^2_{L^2(B_i)} \tag{3-13}$$

for $u \in \mathcal{C}^\infty_c(K^\circ)$ and $i = 1, 2$. Since $\varphi_1 > \varphi_2$ on B_2 and $\varphi_2 > \varphi_1$ on B_1, there is $\gamma > 0$ such that

$$e^{\varphi_i/h} \leq e^{-\gamma/h}(e^{\varphi_1/h} + e^{\varphi_2/h})$$

on B_i. Now add (3-13) for $i = 1, 2$ to absorb the integral over $B_1 \cup B_2$ into the left-hand side. □

4. Degenerate Carleman estimates near the boundary

4A. ***Statement of result.*** We now complement Proposition 3.1 with a result valid up to the boundary. Recall that the phases φ_1, φ_2 are equal on $\{r \leq 1\}$. Since we are working near ∂X, we will thus drop the subscript and simply write φ.

Proposition 4.1. *Given $[a, b] \subset (0, \infty)$ there exists $r_0 > 0$ and $C > 0$ such that*

$$\|e^{\varphi/h}u\|_{H^1_{b,h}} \leq C\big(h^{-1/2}\|e^{\varphi/h}P(z)u\|_{L^2} + e^{\varphi(0)/h}\|u\|_{L^2(\partial X)}\big) \tag{4-1}$$

for $u \in \mathcal{C}^\infty_c(\{r < r_0\})$ and $\pm z \in [a, b]$.

The Sobolev space appearing on the left-hand side of (4-1) is modeled on the space of vector fields $\mathcal{V}_b(X)$ which are tangent to the boundary; see [Melrose 1993]. Thus $u \in H^1_b(X)$ if $u \in L^2(X)$ and $Ku \in L^2(X)$ for any $K \in \mathcal{V}_b(X)$. If $u \in H^1_b(X)$ and $\operatorname{supp} u \subset \{r < 1\}$, we can set

$$\|u\|^2_{H^1_{b,h}} = \int_X |u|^2 + h^2|r\,\partial_r u|^2 + h^2 k^{AB}(\partial_A u \cdot \partial_B \bar{u})\, dS_X.$$

Of course away from ∂X this is equivalent to the full H^1_h norm. Observe that it is enough to prove Proposition 4.1 for the operator $L = \Box_g$, since the estimate (4-1) is stable under perturbations $B \in h\operatorname{Diff}^1_h(X)$ provided that the vector field part of B is tangent to ∂X. The latter condition is satisfied by the hypothesis that $\mathcal{W}$ is tangent to ∂M made in the Introduction.

Proposition 4.1 is proved through integration by parts. A convenient way of carrying out this procedure is by constructing an appropriate multiplier for the wave operator and applying the divergence theorem. This approach to Carleman estimates for certain geometric operators is partly inspired by [Alexakis and Shao 2015; Ionescu and Klainerman 2009].

4B. ***The divergence theorem.*** We will use the divergence theorem in the time-differentiated form

$$\frac{d}{dt}\int_X g(K,N)\,dS_X+\int_{\partial X} g(K,T)\,dS_{\partial X}=\int_X(\operatorname{div}_g K)A\,dS_X, \tag{4-2}$$

valid for any vector field K, see [Warnick 2015, Lemma 3.1] for instance, where recall $X=\{t=0\}$. Thus the first term on the left-hand side of (4-2) is short-hand for

$$\frac{d}{ds}\int_{\{t=s\}} g(K,N)\,dS_X \text{ evaluated at } s=0.$$

Here $dS_{\partial X}$ is the volume density on ∂X induced by k (the latter is Riemannian, and hence the induced volume density is well-defined).

4C. ***Stress-energy tensor.*** Given $v\in\mathcal{C}^\infty(M)$, let $Q=Q[v]$ denote the usual stress-energy tensor associated to v with components

$$Q_{\alpha\beta}=\operatorname{Re}(\partial_\alpha v\cdot\partial_\beta\bar v)-\tfrac12 g^{-1}(dv,d\bar v)g_{\alpha\beta}.$$

This tensor has the property that $(\nabla^\beta Q_{\alpha\beta})S^\alpha=\operatorname{Re}(\Box v\cdot S\bar v)$ for any vector field S. Given such a vector field and a function w, define the modified vector field $J=J[v]$ with components

$$J^\alpha=Q^\alpha_\beta S^\beta+\tfrac12 w\cdot\partial^\alpha(|v|^2)-\tfrac12(\partial^\alpha w)|v|^2.$$

The relevant choices in this context are

$$S=\nabla_g r,\qquad w=\lambda+\tfrac12\Box_g r, \tag{4-3}$$

where $\lambda=\lambda(r)$ is an undetermined function to be chosen in Lemma 4.4 below. Also, introduce the tensor Π with components

$$\Pi^{\alpha\beta}=-\nabla^{\alpha\beta}r-\lambda g^{\alpha\beta}.$$

The divergence of J satisfies

$$\operatorname{Re}(\Box_g u\cdot(S\bar v+w\bar v))=\operatorname{div}_g J+\Pi(dv,d\bar v)+\tfrac12(\Box_g w)|v|^2, \tag{4-4}$$

which is verified by a direct calculation.

4D. ***The conjugated operator.*** Near ∂M, consider the conjugated operator $L_\Phi=e^\Phi\Box_g e^{-\Phi}$, where $\Phi=\Phi(r)$. Then, L_Φ has the expression

$$\begin{aligned}L_\Phi&=\Box_g-2\Phi' S+((\Phi')^2-\Phi'')g^{-1}(dr,dr)-\Phi'\Box_g r\\&=\Box_g-2\Phi' S+V_0.\end{aligned}$$

Now $g^{-1}(dr, dr) = -r$ by assumption, and consequently the potential term V_0 satisfies

$$V_0 = r(\Phi'' - (\Phi')^2) - \Phi' \Box_g r.$$

Set $V_1 = V_0 - 2\Phi' w$, multiply $L_\Phi v$ by $S\bar{v} + w\bar{v}$, and take the real part to find that

$$\operatorname{Re}(L_\Phi v \cdot (S\bar{v} + w\bar{v})) = \operatorname{Re}(\Box_g v \cdot (S\bar{v} + w\bar{v})) - 2\Phi' |Sv|^2 + \operatorname{Re} V_1 v \cdot S\bar{v} + V_0 w |v|^2. \tag{4-5}$$

It is also convenient to write $\operatorname{Re}(V_1 v \cdot S\bar{v})$ as a divergence,

$$\operatorname{Re}(V_1 v \cdot S\bar{v}) = \tfrac{1}{2} \operatorname{div}_g (V_1 |v|^2 S) - \tfrac{1}{2}(S(V_1) + V_1 \Box_g r)|v|^2.$$

In view of this expression, define the vector field $K = J + \frac{1}{2} V_1 |v|^2 S$. For future use, also define the modified potential V by

$$V = \tfrac{1}{2}(\Box_g w) + V_0 w - \tfrac{1}{2} S(V_1) - \tfrac{1}{2} V_1 \Box_g r + \Phi' w^2. \tag{4-6}$$

On one hand, integrating the divergence of K yields boundary integrals; the following special case of this will suffice.

Lemma 4.2. *Let $v \in \mathcal{C}^\infty(M)$ be given by $v = e^{-izt/h} u$, where u is stationary and $z \in \mathbb{R}$. Then,*

$$\int_X (\operatorname{div}_g K) A \, dS_X = -\left| \frac{z}{h} \right|^2 \int_{\partial X} |u|^2 \, dS_{\partial X}.$$

Proof. Apply the divergence theorem (4-2). Since $z \in \mathbb{R}$, the vector field K is stationary, and hence there is no contribution from the time derivative. As for the integral over ∂M, observe that T is null and $S = -T$ on the horizon. Since $Tv = -i(z/h)v$, it follows that $g(T, K) = -|Tv|^2 = -|z/h|^2 |u|^2$ on ∂M. □

Note that the boundary contribution from Lemma 4.2 has an unfavorable sign, which will account for the boundary term in Proposition 4.1. On the other hand, the divergence of K can also be expressed in terms of (4-5).

Lemma 4.3. *If $\Phi' < 0$, then the divergence of K satisfies*

$$(2|\Phi'|)^{-1} |L_\Phi v|^2 \geq \operatorname{div}_g K + \Pi(dv, d\bar{v}) - \Phi' |Sv|^2 + V|v|^2, \tag{4-7}$$

where V is given by (4-6).

Proof. Combine (4-5) with (4-4), and then use the Cauchy–Schwarz inequality to find

$$\operatorname{Re}(L_\Phi v \cdot (S\bar{v} + w\bar{v})) \leq (2|\Phi'|)^{-1} |L_\Phi v|^2 - \Phi'(|Sv|^2 + w^2 |v|^2),$$

recalling that $\Phi' < 0$. □

4E. ***Pseudoconvexity.*** To examine positivity properties of $\Pi(dv, d\bar{v}) - \Phi' |Sv|^2$, we establish a certain pseudoconvexity condition. A criterion of this type first appeared in work of Alinhac [1984] on unique continuation, and was also employed in [Ionescu and Klainerman 2009; Alexakis and Shao 2015]. Recall that the Poisson bracket is related to the Hessian via the formula

$$\{G, \{G, f\}\}(x, \varpi) = 4 \varpi_\alpha \varpi_\beta \nabla^{\alpha\beta} f, \tag{4-8}$$

valid for any $f \in \mathcal{C}^\infty(M)$.

Lemma 4.4. *There exists $M, c, R_0 > 0$, and a function $\lambda = \lambda(r)$ such that*

$$M\{G, r\}^2 - \{G, \{G, r\}\} - 4\lambda G \geq c((r\rho)^2 + \tau^2 + k^{AB}\eta_A\eta_B) \tag{4-9}$$

for $r \in [0, R_0]$.

Proof. Throughout, assume that $M \geq 1$. Let $r \leq (4M)^{-1}$, and define the function λ by

$$\lambda = \tfrac{1}{2} - (1-\delta)rM,$$

where $\delta > 0$ will be chosen sufficiently small. Observe that $\frac{1}{4} \leq \lambda \leq \frac{1}{2}$ uniformly in $M \geq 1$ for $r \leq (4M)^{-1}$. Denote the left-hand side of (4-9) by $4\mathcal{E}$, and the corresponding quantity by $4\mathcal{E}_0$ if G is replaced with G_0. Dividing through by 4,

$$\mathcal{E}_0 = M((r\rho)^2 + 2r\rho\tau + \tau^2) - \tfrac{1}{2}(r\rho^2 + 2\rho\tau) - \lambda G_0. \tag{4-10}$$

Use the expression for λG_0 and the lower bound $\lambda \geq \frac{1}{4}$ on $\{r \leq (4M)^{-1}\}$ to find that

$$\mathcal{E}_0 \geq M\delta((r\rho)^2 + 2r\rho\tau) + M\tau^2 + \tfrac{1}{4}k_0^{AB}\eta_A\eta_B.$$

Therefore $\mathcal{E}_0 \geq c(MY + k^{AB}\eta_A\eta_B)$ if $\delta > 0$ is sufficiently small, where recall $Y = (r\rho)^2 + \tau^2$.

Now consider the error $\mathcal{E} - \mathcal{E}_0$ incurred by replacing G with G_0. Replacing $M\{G, r\}^2$ with $M\{G_0, r\}^2$ produces an error

$$2M\{G_0, r\}\{G - G_0, r\} + M\{G - G_0, r\}^2.$$

Using Cauchy–Schwarz on the first term to absorb a small multiple of $M\{G_0, r\}^2$ into $\mathcal{E}_0$ (in other words, changing the constant $c > 0$ in the lower bound for $\mathcal{E}_0$ above) leaves an overall error of the form

$$M(\mathcal{N}_1 \cdot \mathcal{N}_1) \subset (rM)\mathcal{N}_2.$$

The factor of rM is harmless since $rM \leq \frac{1}{4}$; thus the right-hand side is certainly in $\mathcal{N}_2$ uniformly in $M \geq 1$. Using that λ is uniformly bounded in $M \geq 1$ on $\{r \leq (4M)^{-1}\}$, the remaining errors $\lambda(G - G_0)$ and

$$\{G - G_0, \{G - G_0, r\}\} + \{G - G_0, \{G_0, r\}\} + \{G_0, \{G - G_0, r\}\}$$

are also in $\mathcal{N}_2$ by Lemma 2.6, uniformly in $M \geq 1$. Now apply the first bound in Lemma 2.5, choosing $\gamma > 0$ sufficiently small but independent of M so that $\gamma k^{AB}\eta_A\eta_B$ can be absorbed by $ck^{AB}\eta_A\eta_B$ on the right-hand side for $r \in [0, r_\gamma]$. This leaves a large multiple of Y, which is then absorbed by MY on the right-hand side by taking M sufficiently large. It then suffices to take $R_0 = \min\{(4M)^{-1}, r_\gamma\}$. □

Fix $M > 0$ such that Lemma 4.4 is valid. This fixes the function λ, and therefore the function w in (4-3). Lemma 4.3 will be applied with the weight $\Phi = \varphi_i/h$, viewed as a stationary function on M. In particular, $\Phi' = -C/h$ on $\{r \leq r_1\}$ for some constant $C > 0$ (recall the statement of Proposition 3.1).

Before proceeding, consider the potential term V from Lemma 4.3. Instead of analyzing its sign, we more simply note that for $F' = -C/h$ one has

$$V = f_0 + h^{-1}f_1 + h^{-2}f_2, \tag{4-11}$$

where $f_0, f_1 \in \mathcal{C}^\infty(M)$ and $f_2 \in r\,\mathcal{C}^\infty(M)$. The small coefficient of f_2 means V can be treated as an error. To be precise, we have the following positivity result for the bulk terms.

Lemma 4.5. *Given $a > 0$, there exists $c, r_0 \geq 0$ such that if $|z| \geq a$, then*

$$\Pi(dv, d\bar{v}) - \Phi'|Sv|^2 + V|v|^2 \geq c(h^{-2}|u|^2 + |r\,\partial_r u|^2 + k^{AB}\,\partial_A u\,\partial_B \bar{u}) \tag{4-12}$$

on $\{r \leq r_0\}$ for each $v \in \mathcal{C}^\infty(M)$ of the form $v = e^{-izt/h}u$, where u is stationary.

Proof. Since $\Phi' = -C/h$, an inequality of the form (4-12) is true for sufficiently small $h > 0$ if the term $V|v|^2$ is dropped from the left-hand side; this follows from Lemma 4.4 and (4-8). On the other hand, for a potential V satisfying (4-11), there is clearly $r_0 > 0$ such that $V|v|^2$ can be absorbed by $ch^{-2}|v|^2$ for $r \in [0, r_0]$ and $h > 0$ sufficiently small. □

The proof of Proposition 4.1 is now immediate:

Proof of Proposition 4.1. Given $[a, b] \subset (0, \infty)$, apply Lemmas 4.2, 4.3, and 4.5 to functions of the form $v = e^{-izt/h}e^{\varphi/h}u$, where $\pm z \in [a, b]$ and $\operatorname{supp} u \subset \{r < r_0\}$. □

5. Proof of Theorem 1

We prove the equivalent Theorem 1′. Assume that $[a, b] \subset (0, \infty)$ has been fixed. Choose a cutoff function $\chi \in \mathcal{C}^\infty(X)$ such that

$$\operatorname{supp}\chi \subset \{r < r_0\}, \qquad \chi = 1 \quad \text{near } \{r \leq r_0/2\},$$

where r_0 is provided by Lemma 4.5. Then, apply Proposition 4.1 to χu and Proposition 3.1 to $(1-\chi)u$, where $u \in \mathcal{C}^\infty(X)$. Since the commutator $[P(z), \chi]$ is supported away from ∂X, the error terms can be absorbed even though the left-hand side is only estimated in the $H^1_{b,h}$ norm. Bounding $e^{\varphi_1/h} + e^{\varphi_2/h}$ from below on the left and from above on the right yields

$$\|u\|_{H^1_{b,h}} \leq e^{C/h}(\|P(z)u\|_{L^2} + \|u\|_{L^2(\partial X)}) \tag{5-1}$$

for $u \in \mathcal{C}^\infty(X)$ and $\pm z \in [a, b]$.

Next, we remove the boundary term on the right-hand side of (5-1). In order to estimate the boundary term, we use that L is formally self-adjoint and that $\mathcal{W}$ is tangent to ∂M. Apply the divergence theorem (4-2) to the vector field $\bar{v}\nabla_g v - v\nabla_g \bar{v} + |v|^2 \cdot \mathcal{W}$ with $v = e^{-izt/h}u$. Since L is formally self-adjoint, we obtain Green's formula

$$(hz)\int_{\partial X} |u|^2\,dS_{\partial X} = -\operatorname{Im}\int_X P(z)u \cdot \bar{u}\,A\,dS_X.$$

There is no boundary contribution coming from $\mathcal{W}$ since we assumed $g(T, \mathcal{W})$ vanishes on ∂M. Applying Cauchy–Schwarz to the right-hand side implies

$$e^{C/h}\|u\|_{L^2(\partial X)} \leq C_\varepsilon h^{-1}e^{2C/h}\|P(z)u\|_{L^2} + \varepsilon\|u\|_{L^2}$$

for some C_ε and every $\pm z \in [a, b]$. Therefore the boundary term on the right-hand side of (5-1) can be absorbed into the left-hand side by taking ε sufficiently small, at the expense of increasing the constant in the exponent $e^{C/h}$. We then have

$$\|u\|_{H^1_{b,h}} \leq e^{C/h} \|P(z)u\|_{L^2}.$$

The final step is to apply a bound of the form

$$\|u\|_{H_h^{k+1}} \leq Ch^{-1}(\|P(z)u\|_{H_h^k} + \|u\|_{L^2}) \tag{5-2}$$

for $u \in \mathcal{C}^\infty(X)$ and $\pm z \in [a, b]$. The most conceptual way of understanding this estimate is in terms of the semiclassical trapping present in the interior of M. For an appropriate pseudodifferential complex absorbing operator $Q \in \Psi_h^{-\infty}(X^\circ)$ with compact support in X°, the nontrapping framework of [Vasy 2013, Section 2.8] shows that $P(z) - iQ$ satisfies the nontrapping bound

$$\|u\|_{H_h^k} \leq Ch^{-1}\|(P(z) - iQ)u\|_{H_h^k}$$

for $z \in [a, b]$. Here Q is chosen to be elliptic (with the correct choice of sign) on the trapped set. In this case Q can be chosen to have compact microsupport in X°, and hence $Q : \mathcal{C}^{-\infty}(X) \to \mathcal{C}^\infty(X)$, and in particular

$$\|Qu\|_{H_h^k} \leq C\|u\|_{L^2}.$$

This clearly implies (5-2) for $z \in [a, b]$, with a similar argument when $-z \in [a, b]$.

This completes the proof of Theorem 1′ in the case when $u \in \mathcal{C}^\infty(X)$ and $\pm z \in [a, b]$. By perturbation, this extends to a region $\pm z \in [a, b] + ie^{-C_1/h}[-1, 1]$. Simply write

$$P(z) - P(\operatorname{Re} z) = \operatorname{Im} z \cdot B(z),$$

where $B(z) \in \operatorname{Diff}_h^1(X)$ is bounded $H_h^{k+1}(X) \to H_h^k(X)$ uniformly for $z \in [a, b]$ (although $B(z)$ is not holomorphic in z). Thus the difference can be absorbed into the left-hand side if $|\operatorname{Im} z| \leq e^{-C_1/h}$ for $C_1 > 0$ sufficiently large. Finally, $\mathcal{C}^\infty(X)$ is dense in $\mathcal{X}^k$, see [Dyatlov and Zworski 2018, Lemma E.47], so (2-2) is valid for $u \in \mathcal{X}^k$ as well, thus completing the proof of Theorem 1′.

6. Logarithmic energy decay

6A. ***A semigroup formulation.*** In this section we outline how Corollary 1.2 can be deduced from the resolvent estimate (1-3) via semigroup theory. The starting point is that the Cauchy problem (1-4) is associated with a $\mathcal{C}^0$ semigroup $U(t) = e^{-itB}$ on $\mathcal{H}^k = H^{k+1}(X) \times H^k(X)$ satisfying

$$\|U(t)\|_{\mathcal{H}^k \to \mathcal{H}^k} \leq Ce^{\nu t} \tag{6-1}$$

for some $C, \nu > 0$ [Warnick 2015, Corollary 3.14]. Recalling the lapse function $A = g^{-1}(dt, dt)^{-1/2}$, write

$$\Box_g = L_2 + L_1\,\partial_t + A^{-2}\,\partial_t^2,$$

where L_j is identified with a differential operator on X of order j. Thus $L_2 = \boldsymbol{P}(0)$ and $L_1 = i\,\partial_\omega \boldsymbol{P}(0)$. More explicitly,

$$L_1 = -2A^{-2}W - \operatorname{div}_g(A^{-2}W),$$

where W is the shift vector from Section 2D. The infinitesimal generator is then given by

$$-iB = \begin{pmatrix} 0 & 1 \\ -A^2L_2 & -A^2L_1 \end{pmatrix}. \tag{6-2}$$

Indeed, applying $U(t)$ to initial data in $\mathcal{C}_c^\infty(X^\circ)$ shows that $-iB$ is given by (6-2) in the sense of distributions. Now the resolvent set of B is nonempty, and indeed $\sigma(B) \subset \{\operatorname{Im}\omega \le \nu\}$ by (6-1). Therefore the domain $D(B)$ of B is characterized as those distributions $(v_0, v_1) \in \mathcal{H}^k$ such that

$$v_1 \in H^{k+1}(X), \quad L_2v_0 + L_1v_1 \in H^k(X).$$

Since $L_2 = \boldsymbol{P}(0)$ and $L_1 \in \operatorname{Diff}^1(X)$, this shows that the domain of B is

$$D(B) = \mathcal{X}^k \times H^{k+1}(X),$$

where $\mathcal{X}^k$ is defined by (1-2). It is also easy to see that the graph norm on $D(B)$ satisfies

$$\|B(v_0, v_1)\|_{\mathcal{H}^k} + \|(v_0, v_1)\|_{\mathcal{H}^k} \le C\|(v_0, v_1)\|_{\mathcal{X}^k\times H^{k+1}};$$

hence the two norms on $D(B)$ are equivalent by the open mapping theorem. Furthermore, the spectrum of B in $\left\{\operatorname{Im}\omega > -\kappa\left(k+\frac{1}{2}\right)\right\}$ coincides with poles of $\boldsymbol{P}(\omega)^{-1}$, and the resolvent estimate (1-3) translates into the bound $\|(B-\omega)^{-1}\|_{\mathcal{H}^k\to\mathcal{H}^k} \le e^{C|\operatorname{Re}\omega|}$ for $\omega \in \Omega$.

6B. ***Logarithmic stabilization of semigroups.*** The goal now is to apply a theorem on the logarithmic stabilization of certain bounded semigroups:

Theorem 2 [Burq 1998, Theorem 3, Batty and Duyckaerts 2008, Theorem 1.5]. *Let $U(t) = e^{-itB}$ be a bounded $\mathcal{C}^0$ semigroup on a Hilbert space $\mathcal{H}$. If $\sigma(B)\cap\mathbb{R} = \varnothing$ and $\|(B-\omega)^{-1}\|_{\mathcal{H}\to\mathcal{H}} \le e^{C|\omega|}$ for $\omega\in\mathbb{R}$, then there exists $C > 0$ such that*

$$\|U(t)v\|_{\mathcal{H}} \le \frac{C}{\log(2+t)}\|(B-i)v\|_{\mathcal{H}}$$

for each $v \in D(B)$.

A priori the semigroup $U(t)$ from Section 6A is not uniformly bounded in time on $\mathcal{H}^k$, since the energy $\mathcal{E}_k[v](t)$ does not control the L^2 norm of $v(t)$. Instead, observe that $\operatorname{span}\{(1,0)\} \subset \mathcal{H}^k$ is invariant under $U(t)$, which therefore descends to a semigroup $\widehat{U}(t)$ on the quotient space

$$\widehat{\mathcal{H}}^k = \mathcal{H}^k/\operatorname{span}\{(1,0)\}.$$

If $\pi : \mathcal{H}^k \to \widehat{\mathcal{H}}^k$ is the natural projection, then, the infinitesimal generator of $\widehat{U}(t)$ is simply the operator $\widehat{B}$ induced by B on $\pi(D(B))$. It follows from (1-5) and the Poincaré inequality that $\widehat{U}(t)$ is a bounded $\mathcal{C}^0$ semigroup.

Since $\operatorname{span}\{(0,1)\}$ is finite-dimensional, the spectrum of $\widehat{B}$ is contained in the spectrum of B, and furthermore the bound

$$\|(\widehat{B}-\omega)^{-1}\|_{\widehat{\mathcal{H}}\to\widehat{\mathcal{H}}} \le e^{C|\operatorname{Re}\omega|}$$

also holds for $\omega \in \Omega$. The final step is to show $\sigma(\widehat{B}) \cap \mathbb{R} = \varnothing$. If $\omega \in \mathbb{R} \setminus 0$, this follows from the fact that $\boldsymbol{P}(\omega)^{-1}$ has no nonzero real poles [Warnick 2015, Lemma A.1].

Finally, consider the spectrum at $\omega = 0$. If ω_0 is a pole of $(B-\omega)^{-1}$ acting on $\mathcal{H}^k$ with $\operatorname{Im}\omega_0 > -\kappa\big(k+\frac{1}{2}\big)$, then its Laurent coefficients all map into $\mathcal{C}^\infty(X) \times \mathcal{C}^\infty(X)$ [Vasy 2013, Section 2.6]. Thus $\ker B \subset \mathcal{X}^k \times H^{k+1}(X)$ is in one-to-one correspondence with smooth stationary solutions of $\Box_g v = 0$. If $\Box_g v = 0$ for v smooth and stationary, then (4-2) applied to the vector field $\bar{v}\nabla_g v + v\nabla_g \bar{v}$ shows that $g^{-1}(dv, d\bar{v}) = 0$ on X. Again using that v is stationary, Lemma 2.1 implies $dv = 0$, and hence v is constant. Thus $\ker B = \operatorname{span}\{(1,0)\}$, so $0 \notin \sigma(\widehat{B})$.

The hypotheses of Theorem 2 are therefore satisfied by $\widehat{U}(t)$, which yields the bound

$$\|\widehat{U}(t)\circ\pi(v_0, v_1)\|_{\widehat{\mathcal{H}}^k} \le \frac{C}{\log(2+t)}\|(B-i)(v_0,v_1)\|_{\mathcal{H}^k} \tag{6-3}$$

for each $(v_0, v_1) \in \mathcal{X}^k \times H^{k+1}(X)$. This establishes Corollary 1.2, since the norm on the left-hand side of (6-3) is equivalent to $\mathcal{E}_k[v](t)^{1/2}$, where v solves the Cauchy problem (1-4) with initial data (v_0, v_1).

6C. ***Decay to a constant.*** To prove Corollary 1.3, consider the Laurent expansion of $(B-\omega)^{-1}$ about $\omega = 0$. The range of the corresponding residue Π_0 consists of all generalized eigenvectors, and contains $\operatorname{span}\{(1,0)\}$.

If the algebraic multiplicity of $\omega = 0$ was greater than 1, then there would exist a solution of $\Box_g v = 0$ of the form

$$v(t,x) = u(x) + t,$$

where $u \in \mathcal{C}^\infty(M)$ is stationary. This is compatible with energy boundedness, but not with the logarithmic energy decay established above. Thus $\omega = 0$ is a simple pole with algebraic multiplicity 1.

By standard spectral theory, Π_0 is the projection onto $\operatorname{span}\{(1,0)\}$ along $\operatorname{range}(B)$, so

$$\Pi_0 = \langle\,\cdot\,,\psi\rangle(1,0)$$

for some $\psi \in (\ker B)'$, which we identify with $(\mathcal{H}^k)'/\operatorname{range}(B^*) = \ker(B^*)$. Furthermore, ψ is uniquely determined by requiring that $\langle(1,0),\psi\rangle = 1$. Here the duality between $\mathcal{H}^k$ and

$$(\mathcal{H}^k)' = \dot{H}^{-k-1}(X) \times \dot{H}^{-k}(X)$$

is induced by the $L^2(X)$ inner product described in Section 3B, where $\dot{H}^s(X)$ is the Sobolev space of supported distributions in the sense of [Hörmander 1985, Appendix B.2].

The domain of B^* consists of all $w \in \dot{H}^{-k-1}(X) \times \dot{H}^{-k}(X)$ for which there exists $v \in \dot{H}^{-k-1}(X) \times \dot{H}^{-k}(X)$ satisfying $(w, Bu) = (v, u)$ for every $u \in D(B) = \mathcal{X}^k \times H^{k+1}(X)$. Thus

$$D(B^*) = \dot{H}^{-k-1}(X) \times \dot{\mathcal{X}}^{-k},$$

where we define

$$\dot{\mathcal{X}}^{-k} = \{u \in \dot{H}^{-k}(X) : \boldsymbol{P}(0) \in \dot{H}^{-k-1}(X)\}.$$

The action of B^* is given by

$$iB^* = \begin{pmatrix} 0 & -L_2A^2 \\ 1 & L_1A^2 \end{pmatrix},$$

using that L_1 is skew-adjoint.

Now we compute the kernel of B^*, which again by abstract spectral theory is one-dimensional. Let $\psi_1 = \mathrm{vol}(\partial X)^{-1}A^{-2} \in L^2(X)$, viewed as an element of $\dot{H}^{-k}(X)$ via the $L^2(X)$ inner product, and then set

$$\psi_0 = -\,\mathrm{vol}(\partial X)^{-1}L_1(1) \in \dot{H}^{-k-1}(X)$$

in the sense of supported distributions. If we set $\psi = (\psi_0, \psi_1)$, then $B^*\psi = 0$. Furthermore,

$$\begin{aligned} \mathrm{vol}(\partial X)\langle 1, \psi_0\rangle = \langle L_1(1), 1\rangle &= -\int_X \mathrm{div}_g(A^{-2}W)\, A\, dS_X \\ &= -\int_{\partial X} A^{-2}g(W,T)\, dS_{\partial X} = \int_{\partial X} dS_{\partial X} = \mathrm{vol}(\partial X), \end{aligned}$$

since $g(W,T) = -g(AN,T) = -A^2$ on ∂X. Thus $\psi \in \ker B^*$ has the appropriate normalization.

Finally, let $E = \mathrm{range}(I - \Pi_0)$, which is thus invariant under $U(t)$, and $U(t)|_E = U(t)(I-\Pi_0)$. Since

$$\mathcal{H}^k = E \dotplus \mathrm{span}\{(1,0)\},$$

with $\dotplus$ denoting a topological direct sum, it follows that E is isomorphic to the quotient $\widehat{\mathcal{H}}^k$ as a Banach space. Given $(v_0, v_1) \in D(B)$, define the constant $v_\infty = \langle v_0, \psi_0\rangle + \langle v_1, \psi_1\rangle$. Then

$$\|U(t)(v_0 - v_\infty, v_1)\|_{\mathcal{H}^k} = \|U(t)(v_0, v_1) - (v_\infty, 0)\|_{\mathcal{H}^k} \le C\|\widehat{U}(t) \circ \pi(v_0, v_1)\|_{\widehat{\mathcal{H}}^k},$$

which completes the proof of Corollary 1.3.

Acknowledgments

This work was supported by the NSF grant DMS-1502632. The author would like to thank Georgios Moschidis for carefully explaining parts of his paper [Moschidis 2016], and Daniel Tataru for a discussion on Carleman estimates.

References

[Alexakis and Shao 2015] S. Alexakis and A. Shao, "Global uniqueness theorems for linear and nonlinear waves", *J. Funct. Anal.* **269**:11 (2015), 3458–3499. MR Zbl

[Alinhac 1984] S. Alinhac, "Unicité du problème de Cauchy pour des opérateurs du second ordre à symboles réels", *Ann. Inst. Fourier* (*Grenoble*) **34**:2 (1984), 89–109. MR Zbl

[Batty and Duyckaerts 2008] C. J. K. Batty and T. Duyckaerts, "Non-uniform stability for bounded semi-groups on Banach spaces", *J. Evol. Equ.* **8**:4 (2008), 765–780. MR Zbl

[Burq 1998] N. Burq, "Décroissance de l'énergie locale de l'équation des ondes pour le problème extérieur et absence de résonance au voisinage du réel", *Acta Math.* **180**:1 (1998), 1–29. MR Zbl

[Burq 2002] N. Burq, "Lower bounds for shape resonances widths of long range Schrödinger operators", *Amer. J. Math.* **124**:4 (2002), 677–735. MR Zbl

[Cardoso and Vodev 2002] F. Cardoso and G. Vodev, "Uniform estimates of the resolvent of the Laplace–Beltrami operator on infinite volume Riemannian manifolds, II", *Ann. Henri Poincaré* **3**:4 (2002), 673–691. MR Zbl

[Datchev 2014] K. Datchev, "Quantitative limiting absorption principle in the semiclassical limit", *Geom. Funct. Anal.* **24**:3 (2014), 740–747. MR Zbl

[Dyatlov and Zworski 2018] S. Dyatlov and M. Zworski, "Mathematical theory of scattering resonances", preprint, 2018, available at http://math.mit.edu/~dyatlov/res/res_20180406.pdf.

[Gannot 2018] O. Gannot, "The null-geodesic flow near horizons", *Trans. Amer. Math. Soc.* (online publication 2018).

[Gilbarg and Trudinger 1983] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, 2nd ed., Grundlehren der Math. Wissenschaften **224**, Springer, 1983. MR Zbl

[Guillarmou 2005] C. Guillarmou, "Meromorphic properties of the resolvent on asymptotically hyperbolic manifolds", *Duke Math. J.* **129**:1 (2005), 1–37. MR Zbl

[Holzegel and Smulevici 2013] G. Holzegel and J. Smulevici, "Decay properties of Klein–Gordon fields on Kerr-AdS spacetimes", *Comm. Pure Appl. Math.* **66**:11 (2013), 1751–1802. MR Zbl

[Hörmander 1963] L. Hörmander, *Linear partial differential operators*, Grundlehren der Math. Wissenschaften **116**, Springer, 1963. MR Zbl

[Hörmander 1985] L. Hörmander, *The analysis of linear partial differential operators, III: Pseudo-differential operators*, Grundlehren der Math. Wissenschaften **274**, Springer, 1985. MR Zbl

[Ionescu and Klainerman 2009] A. D. Ionescu and S. Klainerman, "On the uniqueness of smooth, stationary black holes in vacuum", *Invent. Math.* **175**:1 (2009), 35–102. MR Zbl

[Melrose 1993] R. B. Melrose, *The Atiyah–Patodi–Singer index theorem*, Research Notes in Math. **4**, A. K. Peters, Wellesley, MA, 1993. MR Zbl

[Moschidis 2016] G. Moschidis, "Logarithmic local energy decay for scalar waves on a general class of asymptotically flat spacetimes", *Ann. PDE* **2**:1 (2016), art. id. 5. MR

[Rodnianski and Tao 2015] I. Rodnianski and T. Tao, "Effective limiting absorption principles, and applications", *Comm. Math. Phys.* **333**:1 (2015), 1–95. MR Zbl

[Sá Barreto and Zworski 1997] A. Sá Barreto and M. Zworski, "Distribution of resonances for spherical black holes", *Math. Res. Lett.* **4**:1 (1997), 103–121. MR Zbl

[Vasy 2013] A. Vasy, "Microlocal analysis of asymptotically hyperbolic and Kerr–de Sitter spaces", *Invent. Math.* **194**:2 (2013), 381–513. MR Zbl

[Vodev 2000] G. Vodev, "Exponential bounds of the resolvent for a class of noncompactly supported perturbations of the Laplacian", *Math. Res. Lett.* **7**:3 (2000), 287–298. MR Zbl

[Warnick 2015] C. M. Warnick, "On quasinormal modes of asymptotically anti-de Sitter black holes", *Comm. Math. Phys.* **333**:2 (2015), 959–1035. MR Zbl

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Math. **138**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

[Zworski 2017] M. Zworski, "Mathematical study of scattering resonances", *Bull. Math. Sci.* **7**:1 (2017), 1–85. MR Zbl

Received 20 Feb 2018. Revised 27 Apr 2018. Accepted 30 May 2018.

ORAN GANNOT: gannot@northwestern.edu
Department of Mathematics, Northwestern University, Evanston, IL, United States

ANALYSIS AND PDE
Vol. 12, No. 2, 2019
dx.doi.org/10.2140/apde.2019.12.561

INTERPOLATION BY CONFORMAL MINIMAL SURFACES AND DIRECTED HOLOMORPHIC CURVES

ANTONIO ALARCÓN AND ILDEFONSO CASTRO-INFANTES

Dedicated to Franc Forstnerič on the occasion of his sixtieth birthday

Let M be an open Riemann surface and $n \geq 3$ be an integer. We prove that on any closed discrete subset of M one can prescribe the values of a conformal minimal immersion $M \to \mathbb{R}^n$. Our result also ensures jet-interpolation of given finite order, and hence, in particular, one may in addition prescribe the values of the generalized Gauss map. Furthermore, the interpolating immersions can be chosen to be complete, proper into $\mathbb{R}^n$ if the prescription of values is proper, and injective if $n \geq 5$ and the prescription of values is injective. We may also prescribe the flux map of the examples.

We also show analogous results for a large family of directed holomorphic immersions $M \to \mathbb{C}^n$, including null curves.

1. Introduction and main results

The theory of interpolation by holomorphic functions is a central topic in complex analysis which began in the 19th century with the celebrated Weierstrass interpolation theorem [1876]: *on a closed discrete subset of a domain $D \subset \mathbb{C}$, one can prescribe the values of a holomorphic function $D \to \mathbb{C}$.* Much later, Florack [1948] extended the Weierstrass theorem to arbitrary open Riemann surfaces. In this paper we prove an analogue of this classical result for conformal minimal surfaces in Euclidean spaces.

Theorem 1.1 (Weierstrass interpolation theorem for conformal minimal surfaces). *Let Λ be a closed discrete subset of an open Riemann surface M, and let $n \geq 3$ be an integer. Every map $\Lambda \to \mathbb{R}^n$ extends to a conformal minimal immersion $M \to \mathbb{R}^n$.*

Let M be an open Riemann surface and $n \geq 3$ be an integer. By the identity principle it is not possible to prescribe values of a conformal minimal immersion $M \to \mathbb{R}^n$ on a subset that is not closed and discrete; hence the assumptions on Λ in Theorem 1.1 are necessary.

Recall that a conformal immersion $X = (X_1, \dots, X_n) : M \to \mathbb{R}^n$ is minimal if, and only if, X is a harmonic map. If this is the case then, denoting by ∂ the $\mathbb{C}$-linear part of the exterior differential $d = \partial + \bar\partial$ on M (here $\bar\partial$ denotes the $\mathbb{C}$-antilinear part of d), the 1-form $\partial X = (\partial X_1, \dots, \partial X_n)$ with values in $\mathbb{C}^n$ is holomorphic, has no zeros, and satisfies $\sum_{j=1}^n (\partial X_j)^2 = 0$ everywhere on M. Therefore, ∂X determines the Kodaira-type holomorphic map

$$G_X : M \to \mathbb{CP}^{n-1}, \quad M \ni p \mapsto G_X(p) = [\partial X_1(p) : \cdots : \partial X_n(p)],$$

MSC2010: 53A10, 32E30, 32H02, 53A05.
Keywords: minimal surface, directed holomorphic curve, Weierstrass theorem, Riemann surface, Oka manifold.

which takes values in the complex hyperquadric

$$Q_{n-2} = \{[z_1 : \cdots : z_n] \in \mathbb{CP}^{n-1} : z_1^2 + \cdots + z_n^2 = 0\} \subset \mathbb{CP}^{n-1}$$

and is known as *the generalized Gauss map of* X. Conversely, every holomorphic map $M \to Q_{n-2} \subset \mathbb{CP}^{n-1}$ is the generalized Gauss map of a conformal minimal immersion $M \to \mathbb{R}^n$; see [Alarcón, Forstnerič and López 2017]. The real part $\Re(\partial X)$ is an exact 1-form on M; the *flux map* (or simply, the *flux*) of X is the group homomorphism $\mathrm{Flux}_X : H_1(M;\mathbb{Z}) \to \mathbb{R}^n$ of the first homology group of M with integer coefficients, given by

$$\mathrm{Flux}_X(\gamma) = \int_\gamma \Im(\partial X) = -\mathfrak{i}\int_\gamma \partial X, \qquad \gamma \in H_1(M;\mathbb{Z}),$$

where $\Im$ denotes the imaginary part and $\mathfrak{i} = \sqrt{-1}$.

Conversely, every holomorphic 1-form $\Phi = (\phi_1, \ldots, \phi_n)$ with values in $\mathbb{C}^n$, vanishing nowhere on M, satisfying the nullity condition

$$\sum_{j=1}^{n} (\phi_j)^2 = 0 \quad \text{everywhere on } M, \tag{1-1}$$

and whose real part $\Re(\Phi)$ is exact on M, determines a conformal minimal immersion $X : M \to \mathbb{R}^n$ with $\partial X = \Phi$ by the classical Weierstrass formula

$$X(p) = x_0 + 2\int_{p_0}^{p} \Re(\Phi), \qquad p \in M, \tag{1-2}$$

for any fixed base point $p_0 \in M$ and initial condition $X(p_0) = x_0 \in \mathbb{R}^n$. (We refer to [Osserman 1986] for a standard reference on minimal surface theory.) This representation formula has greatly influenced the study of minimal surfaces in $\mathbb{R}^n$ by providing powerful tools coming from complex analysis in one and several variables. In particular, Runge and Mergelyan theorems for open Riemann surfaces (see [Bishop 1958; Runge 1885; Mergelyan 1951]) and, more recently, the modern *Oka theory* (we refer to the monograph [Forstnerič 2017] and to the surveys [Lárusson 2010; Forstnerič and Lárusson 2011; Forstnerič 2013; Kutzschebauch 2014]) have been exploited in order to develop a uniform approximation theory for conformal minimal surfaces in Euclidean spaces which is analogous to the one of holomorphic functions in one complex variable and has found plenty of applications; see [Alarcón and López 2012; 2014; 2015; Alarcón and Forstnerič 2014; Drinovec Drnovšek and Forstnerič 2016; Alarcón, Forstnerič and López 2016a; 2016b; Forstnerič and Lárusson 2016]. In this paper we extend some of the methods invented for developing this approximation theory in order to provide also interpolation on closed discrete subsets of the underlying complex structure.

Theorem 1.1 is a consequence of the following much more general result ensuring not only interpolation but also *jet-interpolation of given finite order*, approximation on holomorphically convex compact subsets, control on the flux, and global properties such as completeness and, under natural assumptions, properness and injectivity. If A is a compact domain in an open Riemann surface, by a *conformal minimal immersion* $A \to \mathbb{R}^n$ *of class* $\mathscr{C}^m(A)$, $m \in \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$, we mean an immersion $A \to \mathbb{R}^n$ of class $\mathscr{C}^m(A)$ whose restriction to the interior $\mathring{A} = A \setminus bA$ is a conformal minimal immersion; we use the same notation if A is a union of pairwise-disjoint such domains.

Theorem 1.2 (Runge approximation with jet-interpolation for conformal minimal surfaces). *Let M be an open Riemann surface, $\Lambda \subset M$ be a closed discrete subset, and $K \subset M$ be a smoothly bounded compact domain such that $M \setminus K$ has no relatively compact connected components. For each $p \in \Lambda$ let $\Omega_p \subset M$ be a compact neighborhood of p in M, assume that $\Omega_p \cap \Omega_q = \varnothing$ for all $p \neq q \in \Lambda$, and set $\Omega := \bigcup_{p\in\Lambda} \Omega_p$. Also let $X : K \cup \Omega \to \mathbb{R}^n$ $(n \geq 3)$ be a conformal minimal immersion of class $\mathscr{C}^1(K \cup \Omega)$ and let $\mathfrak{p} : H_1(M; \mathbb{Z}) \to \mathbb{R}^n$ be a group homomorphism satisfying*

$$\mathrm{Flux}_X(\gamma) = \mathfrak{p}(\gamma) \quad \textit{for all closed curves } \gamma \subset K.$$

Then, given $k \in \mathbb{Z}_+$, X may be approximated uniformly on K by complete conformal minimal immersions $\widetilde{X} : M \to \mathbb{R}^n$ enjoying the following properties:

(I) *$\widetilde{X}$ and X have a contact of order k at every point in Λ.*

(II) $\mathrm{Flux}_{\widetilde{X}} = \mathfrak{p}$.

(III) *If the map $X|_\Lambda : \Lambda \to \mathbb{R}^n$ is proper then we can choose $\widetilde{X} : M \to \mathbb{R}^n$ to be proper.*

(IV) *If $n \geq 5$ and the map $X|_\Lambda : \Lambda \to \mathbb{R}^n$ is injective, then we can choose $\widetilde{X} : M \to \mathbb{R}^n$ to be injective.*

Condition (I) in the above theorem is equivalent to $\widetilde{X}|_\Lambda = X|_\Lambda$ and, if $k > 0$, the holomorphic 1-form $\partial(\widetilde{X} - X)$, assuming values in $\mathbb{C}^n$, has a zero of multiplicity (at least) k at all points in Λ; in other words, *the maps $\widetilde{X}$ and X have the same k-jet at every point in* Λ (see Section 2B). This is reminiscent of the generalization of the Weierstrass interpolation theorem provided by [Behnke and Stein 1949] and asserting that on an open Riemann surface one may prescribe values to arbitrary finite order for a holomorphic function at the points in a given closed discrete subset; see [Napier and Ramachandran 2011, Theorem 2.15.1]. In particular, choosing $k = 1$ in Theorem 1.2 we obtain that *on a closed discrete subset of an open Riemann surface M, one can prescribe the values of a conformal minimal immersion $M \to \mathbb{R}^n$ $(n \geq 3)$ and of its generalized Gauss map $M \to Q_{n-2} \subset \mathbb{CP}^{n-1}$* (see Corollary 7.1). The case $\Lambda = \varnothing$ in Theorem 1.2 (that is, when one does not take care of the interpolation) was recently proved by Alarcón, Forstnerič, and López [2016a, Theorem 1.2].

Note that the assumptions on $X|_\Lambda$ in assertions (III) and (IV) in Theorem 1.2 are necessary. We also point out that if Λ is infinite then there are injective maps $\Lambda \to \mathbb{R}^n$ which do not extend to a topological embedding $M \to \mathbb{R}^n$; hence, in general, one cannot choose the conformal minimal immersion $\widetilde{X}$ in (IV) to be an *embedding* (i.e., a homeomorphism onto $\widetilde{X}(M)$ endowed with the subspace topology inherited from $\mathbb{R}^n$). On the other hand, since proper injective immersions $M \to \mathbb{R}^n$ are embeddings, we can choose $\widetilde{X}$ in Theorem 1.2 to be a proper conformal minimal embedding provided that $n \geq 5$ and $X|_\Lambda : \Lambda \to \mathbb{R}^n$ is both proper and injective.

Let us now say a word about our methods of proof. Given a holomorphic 1-form θ on M with no zeros (such a θ exists by the Oka–Grauert principle, see [Grauert 1957; 1958; Forstnerič 2017, Theorem 5.3.1]), any holomorphic 1-form $\Phi = (\phi_1, \ldots, \phi_n)$ on M with values in $\mathbb{C}^n$ and satisfying the nullity condition

(1-1) can be written in the form $\Phi = f\theta$ where $f : M \to \mathbb{C}^n$ is a holomorphic function taking values in the *null quadric* (also called the *complex light cone*)

$$\mathfrak{A} := \{z = (z_1, \dots, z_n) \in \mathbb{C}^n : z_1^2 + \cdots + z_n^2 = 0\}. \tag{1-3}$$

Therefore, in order to prove Theorem 1.1 one needs to find a holomorphic map

$$f : M \to \mathfrak{A} \setminus \{0\} \subset \mathbb{C}^n$$

such that $\Re(f\theta)$ is an exact real 1-form on M and

$$2\int_{p_0}^{p} \Re(f\theta) = \mathfrak{Z}(p) \quad \text{for all } p \in \Lambda,$$

where $p_0 \in M \setminus \Lambda$ is a fixed base point and $\mathfrak{Z} : \Lambda \to \mathbb{R}^n$ is the given map. Then the formula (1-2) with $x_0 = 0$ and $\Phi = f\theta$ provides a conformal minimal immersion satisfying the conclusion of the theorem. The key in this approach is that the *punctured null quadric*

$$\mathfrak{A}_* := \mathfrak{A} \setminus \{0\} \tag{1-4}$$

is a *complex homogeneous manifold* and hence an *Oka manifold* [Alarcón and Forstnerič 2014, Example 4.4]; thus, there are *many* holomorphic maps $M \to \mathfrak{A}_*$ (see Section 2C for more information).

The proof of Theorem 1.2 is much more involved and elaborate. It requires, in addition to the above, a subtle use of the Runge–Mergelyan theorem with jet-interpolation for holomorphic maps from open Riemann surfaces into Oka manifolds (see Theorem 2.6) to achieve condition (I). Additionally, it requires a conceptually new intrinsic-extrinsic version of the technique from [Jorge and Xavier 1980] to ensure completeness of the interpolating immersions (see Lemma 5.5 and Section 6B), and, in order to guarantee assertion (III), we must extend the recently developed methods in [Alarcón and López 2012; Alarcón and Forstnerič 2014; Alarcón, Forstnerič and López 2016a] for constructing proper minimal surfaces in $\mathbb{R}^n$ with arbitrary complex structure (see Lemma 5.6 and Section 6C). Moreover, in order to prove (IV) we adapt the transversality approach by Abraham [1963] in Theorem 5.3; see [Alarcón and Forstnerič 2014; Alarcón, Drinovec Drnovšek, Forstnerič and López 2015; Alarcón, Forstnerič and López 2016a] for its implementation in minimal surface theory.

The above-described method for constructing conformal minimal surfaces in $\mathbb{R}^n$, based on Oka theory, was introduced in [Alarcón and Forstnerič 2014] and it also works in the more general framework of *directed holomorphic immersions* of open Riemann surfaces into complex Euclidean spaces. Directed immersions have been the focus of interest in a number of classical geometries such as symplectic, contact, Lagrangian, totally real, etc.; we refer for instance to the monograph [Gromov 1986], to [Eliashberg and Mishachev 2002, Chapter 19], and to the introduction of [Alarcón and Forstnerič 2014] for motivation on this subject. Given a (topologically) closed conical complex subvariety $\mathfrak{S}$ of $\mathbb{C}^n$ ($n \geq 3$), a holomorphic immersion $F : M \to \mathbb{C}^n$ of an open Riemann surface M into $\mathbb{C}^n$ is said to be *directed by* $\mathfrak{S}$, or an $\mathfrak{S}$-*immersion*, if its complex derivative F' with respect to any local holomorphic coordinate on M assumes values in

$$\mathfrak{S}_* := \mathfrak{S} \setminus \{0\};$$

see [Alarcón and Forstnerič 2014, Definition 2.1]. If A is a compact domain in an open Riemann surface, or a union of pairwise-disjoint such domains, by an $\mathfrak{S}$*-immersion* $A \to \mathbb{C}^n$ *of class* $\mathscr{A}^m(A)$ ($m \in \mathbb{Z}_+$) we mean an immersion $A \to \mathbb{C}^n$ of class $\mathscr{C}^m(A)$ whose restriction to the interior, $\mathring{A}$, is a (holomorphic) $\mathfrak{S}$-immersion. Among others, general existence, approximation, and desingularization results were proved in [Alarcón and Forstnerič 2014] for certain families of directed holomorphic immersions, including *null curves*, i.e., holomorphic curves in $\mathbb{C}^n$ which are directed by the null quadric $\mathfrak{A} \subset \mathbb{C}^n$; see (1-3). It is well known that the real and imaginary parts of a null curve $M \to \mathbb{C}^n$ are conformal minimal immersions $M \to \mathbb{R}^n$ whose flux map vanishes everywhere on $H_1(M; \mathbb{Z})$; conversely, every conformal minimal immersion $M \to \mathbb{R}^n$ is locally, on every simply connected domain of M, the real part of a null curve $M \to \mathbb{C}^n$; see [Osserman 1986, Chapter 4].

The second main theorem of this paper is an analogue of Theorem 1.2 for a wide family of directed holomorphic curves in $\mathbb{C}^n$ which includes null curves. Given integers $1 \le j \le n$ we denote by $\pi_j : \mathbb{C}^n \to \mathbb{C}$ the coordinate projection $\pi_j(z_1, \dots, z_n) = z_j$.

Theorem 1.3 (Runge approximation with jet-interpolation for directed holomorphic curves). *Let $\mathfrak{S}$ be an irreducible closed conical complex subvariety of $\mathbb{C}^n$ ($n \ge 3$) which is contained in no hyperplane and such that $\mathfrak{S}_* = \mathfrak{S} \setminus \{0\}$ is smooth and an Oka manifold. Let M, Λ, K, and Ω be as in Theorem 1.2 and let $F : K \cup \Omega \to \mathbb{C}^n$ be an $\mathfrak{S}$-immersion of class $\mathscr{A}^1(K \cup \Omega)$. Then, given $k \in \mathbb{N}$, F may be approximated uniformly on K by $\mathfrak{S}$-immersions $\widetilde{F} : M \to \mathbb{C}^n$ such that $\widetilde{F} - F$ has a zero of multiplicity* (*at least*) *k at every point in Λ. Moreover*, *if the map $F|_\Lambda : \Lambda \to \mathbb{C}^n$ is injective*, *then we can choose $\widetilde{F} : M \to \mathbb{C}^n$ to be injective.*

Furthermore:

(I) *If $\mathfrak{S} \cap \{z_1 = 1\}$ is an Oka manifold and $\pi_1 : \mathfrak{S} \to \mathbb{C}$ admits a local holomorphic section h near $\zeta = 0 \in \mathbb{C}$ with $h(0) \neq 0$, then we may choose $\widetilde{F}$ to be complete.*

(II) *If $\mathfrak{S} \cap \{z_j = 1\}$ is an Oka manifold and $\pi_j : \mathfrak{S} \to \mathbb{C}$ admits a local holomorphic section h_j near $\zeta = 0 \in \mathbb{C}$ with $h_j(0) \neq 0$ for all $j \in \{1, \dots, n\}$*, *and if the map $F|_\Lambda : \Lambda \to \mathbb{C}^n$ is proper*, *then we may choose $\widetilde{F} : M \to \mathbb{C}^n$ to be proper.*

In particular, if we are given $\mathfrak{S}$, M, and Λ as in Theorem 1.3 then *every map $\Lambda \to \mathbb{C}^n$ extends to an $\mathfrak{S}$-immersion $M \to \mathbb{C}^n$*. When the subset $\Lambda \subset M$ is empty, the above theorem except for assertion (I) is implied by [Alarcón and Forstnerič 2014, Theorems 7.2 and 8.1]. It is perhaps worth mentioning in this respect that, if $\mathfrak{S}$ is as in assertion (I) and $F|_\Lambda : \Lambda \to \mathbb{C}^n$ is not proper, Theorem 1.3 provides *complete* $\mathfrak{S}$-immersions $M \to \mathbb{C}^n$ which are not proper maps; these seem to be the first known examples of such apart from the case when $\mathfrak{S}$ is the null quadric. Let us emphasize that the particular geometry of $\mathfrak{A}$ allows for the construction of complete null holomorphic curves in $\mathbb{C}^n$ and minimal surfaces in $\mathbb{R}^n$ with a number of different asymptotic behaviors (other than proper in space); see [Alarcón and López 2013a; Alarcón and Forstnerič 2015; Alarcón, Drinovec Drnovšek, Forstnerič and López 2015; $\ge$ 2019; Alarcón and Castro-Infantes 2018].

Most of the technical parts in the proofs of Theorems 1.2 and 1.3 will be furnished by a general result concerning periods of holomorphic 1-forms with values in a closed conical complex subvariety of $\mathbb{C}^n$

(see Theorem 4.4 for a precise statement). With this at hand, the proofs of Theorems 1.2 and 1.3 are very similar; this is why, with brevity of exposition in mind, we shall spell out in detail the proof of Theorem 1.3 (which is, in some sense, more general) but only briefly sketch the one of Theorem 1.2.

This paper is, to the best of our knowledge, the first contribution to the theory of interpolation by conformal minimal surfaces and directed holomorphic curves in a Euclidean space.

Organization of the paper. In Section 2 we state some notation and the preliminaries which are needed throughout the paper; we also show an observation which is crucial to ensure the jet-interpolation conditions in Theorems 1.2 and 1.3 (see Lemma 2.2). Section 3 is devoted to the proof of several preliminary results on the existence of period-dominating sprays of maps into conical complex subvarieties $\mathfrak{S}_*$ of $\mathbb{C}^n$; we use them in Section 4 to prove the noncritical case of a Mergelyan theorem with jet-interpolation and control on the periods for holomorphic maps into such a $\mathfrak{S}_*$ being Oka (see Lemma 4.2), and the main technical result of the paper (Theorem 4.4). In Section 5 we prove a general position theorem, a completeness lemma, and a properness lemma for $\mathfrak{S}$-immersions, which enable us to complete the proof of Theorem 1.3 in Section 6. Finally, Section 7 is devoted to explaining how the methods in the proof of Theorem 1.3 can be adapted to prove Theorem 1.2.

2. Preliminaries

We define $\mathfrak{i} = \sqrt{-1}$, $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$, and $\mathbb{R}_+ = [0, +\infty)$. Given an integer $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$ and $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, we denote by $|\cdot|$, $\operatorname{dist}(\cdot,\cdot)$, and $\operatorname{length}(\cdot)$ the Euclidean norm, distance, and length in $\mathbb{K}^n$, respectively. If K is a compact topological space and $f : K \to \mathbb{K}^n$ is a continuous map, we denote by

$$\|f\|_{0,K} := \max\{|f(p)| : p \in K\}$$

the maximum norm of f on K. Likewise, given $x = (x_1, \ldots, x_n)$ in $\mathbb{K}^n$ we define

$$|x|_\infty := \max\{|x_1|, \ldots, |x_n|\} \quad \text{and} \quad \|f\|_{\infty,K} := \max\{|f(p)|_\infty : p \in K\}.$$

If K is a subset of a Riemann surface M, then for any $r \in \mathbb{Z}_+$ we shall denote by $\|f\|_{r,K}$ the standard $\mathscr{C}^r$ norm of a function $f : K \to \mathbb{K}^n$ of class $\mathscr{C}^r(K)$, where the derivatives are measured with respect to a Riemannian metric on M (the precise choice of the metric will not be important).

Given a smooth connected surface S (possibly with nonempty boundary) and a smooth immersion $X : S \to \mathbb{K}^n$, we denote by $\operatorname{dist}_X : S \times S \to \mathbb{R}_+$ the Riemannian distance induced on S by the Euclidean metric of $\mathbb{K}^n$ via X; i.e.,

$$\operatorname{dist}_X(p, q) := \inf\{\operatorname{length}(X(\gamma)) : \gamma \subset S \text{ an arc connecting } p \text{ and } q\}, \qquad p, q \in S.$$

Likewise, if $K \subset S$ is a relatively compact subset we define

$$\operatorname{dist}_X(p, K) := \inf\{\operatorname{dist}_X(p, q) : q \in K\}, \qquad p \in S.$$

An immersed open surface $X : S \to \mathbb{K}^n$ ($n \geq 3$) is said to be *complete* if the image by X of any proper path $\gamma : [0, 1) \to S$ has infinite Euclidean length; equivalently, if the Riemannian metric on S induced by

dist_X is complete in the classical sense. On the other hand, $X : S \to \mathbb{K}^n$ is said to be *proper* if the image by X of every proper path $\gamma : [0, 1) \to S$ is a divergent path in $\mathbb{K}^n$.

2A. ***Riemann surfaces and spaces of maps.*** Throughout the paper every Riemann surface will be considered connected if the contrary is not indicated.

Let M be an open Riemann surface. Given a subset $A \subset M$ we denote by $\mathcal{O}(A)$ the space of functions $A \to \mathbb{C}$ which are holomorphic on an unspecified open neighborhood of A in M. If A is a smoothly bounded compact domain, or a union of pairwise-disjoint such domains, and $r \in \mathbb{Z}_+$, we denote by $\mathscr{A}^r(A)$ the space of $\mathscr{C}^r$ functions $A \to \mathbb{C}$ which are holomorphic on the interior $\mathring{A} = A \setminus bA$; for simplicity we write $\mathscr{A}(A)$ for $\mathscr{A}^0(A)$. Likewise, we define the spaces $\mathcal{O}(A, Z)$ and $\mathscr{A}^r(A, Z)$ of maps $A \to Z$ to any complex manifold Z. Thus, if $\mathfrak{S}$ is a closed conical complex subvariety of $\mathbb{C}^n$ ($n \geq 3$), by an $\mathfrak{S}$-immersion $A \to \mathbb{C}^n$ of class $\mathscr{A}^r(A)$ we simply mean an immersion of class $\mathscr{A}^r(A)$ whose restriction to $\mathring{A}$ is an $\mathfrak{S}$-immersion. In the same way, a conformal minimal immersion $A \to \mathbb{R}^n$ of class $\mathscr{C}^r(A)$ will be nothing but an immersion of class $\mathscr{C}^r(A)$ whose restriction to $\mathring{A}$ is a conformal minimal immersion.

By a *compact bordered Riemann surface* we mean a compact Riemann surface M with nonempty boundary bM consisting of finitely many pairwise-disjoint smooth Jordan curves. The interior $\mathring{M} = M \setminus bM$ of M is called a *bordered Riemann surface*. It is well known that every compact bordered Riemann surface M is diffeomorphic to a smoothly bounded compact domain in an open Riemann surface $\widetilde{M}$. The spaces $\mathscr{A}^r(M)$ and $\mathscr{A}^r(M, Z)$, for an integer $r \in \mathbb{Z}_+$ and a complex manifold Z, are defined as above.

A compact subset K in an open Riemann surface M is said to be *Runge* (also called *holomorphically convex* or $\mathscr{O}(M)$*-convex*) if every continuous function $K \to \mathbb{C}$, holomorphic in the interior $\mathring{K}$, may be approximated uniformly on K by holomorphic functions on M; by the Runge–Mergelyan theorem [Runge 1885; Mergelyan 1951; Bishop 1958] this is equivalent to $M \setminus K$ having no relatively compact connected components in M. The following particular kind of Runge subsets will play a crucial role in our argumentation.

Definition 2.1. A nonempty compact subset S of an open Riemann surface M is called *admissible* if it is Runge in M and of the form $S = K \cup \Gamma$, where K is the union of finitely many pairwise-disjoint smoothly bounded compact domains in M and $\Gamma := \overline{S \setminus K}$ is a finite union of pairwise-disjoint smooth Jordan arcs and closed Jordan curves meeting K only in their endpoints (or not at all) and such that their intersections with the boundary bK of K are transverse.

If C and C' are oriented arcs in M, and the initial point of C' is the final one of C, we denote by $C * C'$ the product of C and C', i.e., the oriented arc $C \cup C' \subset M$ with initial point the initial point of C and final point the final point of C'.

Every open connected Riemann surface M contains a 1-dimensional embedded CW-complex $C \subset M$ such that there is a strong deformation retraction $\rho_t : M \to M$ ($t \in [0, 1]$); i.e., $\rho_0 = \mathrm{Id}_M$, $\rho_t|_C = \mathrm{Id}|_C$ for all $t \in [0, 1]$, and $\rho_1(M) = C$. It follows that the complement $M \setminus C$ has no relatively compact connected components in M and hence C is Runge. Such a CW-complex $C \subset M$ represents the topology of M and can be obtained, for instance, as the Morse complex of a Morse strongly subharmonic exhaustion function on M. Recall that the first homology group satisfies $H_1(M; \mathbb{Z}) = \mathbb{Z}^l$ for some $l \in \mathbb{Z}_+ \cup \{\infty\}$. It is not difficult to see that, if M is finitely connected (for instance, if it is a bordered Riemann surface), i.e., if $l \in \mathbb{Z}_+$, then, given a point $p_0 \in M$ there is a CW-complex $C \subset M$ as above which is a bouquet of l circles with base point p_0; i.e., $\{p_0\}$ is the only 0-cell of C, and C has l 1-cells $C_1, \dots, C_l$ which are closed Jordan curves on M that only meet at p_0.

2B. ***Jets.*** Let $\mathcal{M}$ and $\mathcal{N}$ be smooth manifolds without boundary, $x_0 \in \mathcal{M}$ be a point, and $f, g : \mathcal{M} \to \mathcal{N}$ be smooth maps. The maps f and g have, by definition, a *contact of order* $k \in \mathbb{Z}_+$ at the point x_0 if their Taylor series at this point coincide up to the order k. An equivalence class of maps $\mathcal{M} \to \mathcal{N}$ which have a contact of order k at the point x_0 is called a *k-jet*; see, e.g., [Michor 1980, §1] for a basic reference. Recall that the Taylor series at x_0 of a smooth map $f : \mathcal{M} \to \mathcal{N}$ does not depend on the choice of coordinate charts on $\mathcal{M}$ and $\mathcal{N}$ centered at x_0 and $f(x_0)$ respectively. Therefore, fixing such a pair of coordinates, we can identify the k-jet of f at x_0, which is usually denoted by $j^k_{x_0}(f)$, with the set of derivatives of f at x_0 of order up to and including k; under this identification of jets we have

$$j^0_{x_0}(f) = f(x_0), \quad j^1_{x_0}(f) = \left(f(x_0), \left.\frac{\partial f}{\partial x}\right|_{x_0} \right), \quad j^2_{x_0}(f) = \left(f(x_0), \left.\frac{\partial f}{\partial x}\right|_{x_0}, \left.\frac{\partial^2 f}{\partial x^2}\right|_{x_0} \right), \quad \dots$$

Analogously, if $\mathcal{M}$ and $\mathcal{N}$ are complex manifolds then we consider the complex (holomorphic) derivatives with respect to some local holomorphic coordinates. It is clear that the definition of the k-jet of a map at a point is local and hence it can be made for germs of maps at the point. Moreover, if a pair of maps have the same k-jet at a point then, obviously, they also have the same k'-jet at the point for all $k' \in \mathbb{Z}_+,\ k' \le k$.

In particular, if Ω is a neighborhood of a point p in an open Riemann surface M and $f, g : \Omega \to \mathbb{C}^n$ are holomorphic functions, then they have a contact of order $k \in \mathbb{Z}_+$, or the same k-jet, at the point p if, and only if, $f - g$ has a zero of multiplicity (at least) $k+1$ at p; if this is the case then for any distance function $\mathrm{d} : M \times M \to \mathbb{R}_+$ on M (not necessarily conformal) we have

$$|f - g|(q) = O(\mathrm{d}(q, p)^{k+1}) \quad \text{as } q \to p. \tag{2-1}$$

If $f, g : \Omega \to \mathbb{R}^n$ are harmonic maps (as, for instance, conformal minimal immersions), then we say that they have a contact of order $k \in \mathbb{Z}_+$, or the same k-jet, at the point p if, assuming that Ω is simply connected, there are harmonic conjugates $\tilde{f}$ of f and $\tilde{g}$ of g such that the holomorphic functions $f + \mathrm{i}\tilde{f}, g + \mathrm{i}\tilde{g} : \Omega \to \mathbb{C}^n$ have a contact of order k at p; this is equivalent to $f(p) = g(p)$ and, if $k > 0$, the holomorphic 1-form $\partial(f - g)$ has a zero of multiplicity (at least) k at p. Again, if such a pair of maps f and g have the same k-jet at the point $p \in \Omega$ then (2-1) formally holds.

The following observation will be crucial in order to ensure the jet-interpolation in the main results of this paper.

Lemma 2.2. *Let V be a holomorphic vector field in $\mathbb{C}^n$ ($n \in \mathbb{N}$), vanishing at $0 \in \mathbb{C}^n$, and let ϕ_s denote the flow of V for small values of time $s \in \mathbb{C}$. Given an open Riemann surface M, a point $p \in M$, and holomorphic functions $f : M \to \mathbb{C}^n$ and $h : M \to \mathbb{C}$ such that h has a zero of multiplicity $k+1$ at p for some $k \in \mathbb{Z}_+$, then the holomorphic map*

$$q \mapsto \tilde{f}(q) = \phi_{h(q)}(f(q)),$$

which is defined on a neighborhood of p in M, has a contact of order k with f at the point p; that is, f and $\tilde{f}$ have the same k-jet at p.

Proof. The flow ϕ_s of the vector field V at a point $z \in \mathbb{C}^n$ may be expressed as

$$\phi_s(z) = z + sV(z) + O(|s|^2);$$

see, e.g., [Abraham, Marsden and Ratiu 1988, §4.1]. Since h has a zero of multiplicity $k+1$ at p, the conclusion of the lemma follows. □

We shall use the following notation in several places throughout the paper.

Notation 2.3. Let $n \geq 3$ be an integer and $\mathfrak{S}$ be a (topologically) closed conical complex subvariety of $\mathbb{C}^n$; by *conical* we mean that $t\mathfrak{S} = \mathfrak{S}$ for all $t \in \mathbb{C}_* = \mathbb{C} \setminus \{0\}$. We also assume that $\mathfrak{S}$ is contained in no hyperplane of $\mathbb{C}^n$, and $\mathfrak{S}_* := \mathfrak{S} \setminus \{0\}$ is smooth and connected (hence irreducible). We also fix a large integer $N \geq n$ and holomorphic vector fields $V_1, \ldots, V_N$ on $\mathbb{C}^n$ which are tangential to $\mathfrak{S}$ along $\mathfrak{S}$, vanish at $0 \in \mathfrak{S}$, and satisfy

$$\operatorname{span}\{V_1(z), \ldots, V_N(z)\} = T_z\mathfrak{S} \quad \text{for all } z \in \mathfrak{S}_*. \tag{2-2}$$

(Such vector fields exist by Cartan's theorem A [1953].)

$$\text{Let } \phi_s^j \text{ denote the flow of the vector field } V_j \tag{2-3}$$

for $j = 1, \ldots, N$ and small values of the time $s \in \mathbb{C}$.

Remark 2.4. Throughout the paper we shall say that a holomorphic function has a zero of multiplicity $k \in \mathbb{N}$ at a point to mean that the function has a zero of multiplicity *at least* k at the point. When the multiplicity of the zero is exactly k, it will be explicitly mentioned. We will follow the same pattern when claiming that two functions have the same k-jet or a contact of order k at a point.

2C. *Oka manifolds.*

We recall the notion of Oka manifold and state some of the properties of such manifolds which will be exploited in our argumentation. A comprehensive treatment of Oka theory can be found in [Forstnerič 2017]; for a briefer introduction to the topic we refer to [Lárusson 2010; Forstnerič and Lárusson 2011; Forstnerič 2013; Kutzschebauch 2014].

Definition 2.5. A complex manifold Z is said to be an *Oka manifold* if every holomorphic map from a neighborhood of a compact convex set $K \subset \mathbb{C}^N$ ($N \in \mathbb{N}$) to Z can be approximated uniformly on K by entire maps $\mathbb{C}^N \to Z$.

The central result of Oka theory is that maps $M \to Z$ from a Stein manifold (as, for instance, an open Riemann surface) to an Oka manifold satisfy all forms of the Oka principle; see [Forstnerič 2006]. In this paper we shall use as a fundamental tool the following version of the *Mergelyan theorem with jet-interpolation* which trivially follows from [Forstnerič 2017, Theorems 3.8.1 and 5.4.4]; see also [Forstnerič 2004, Theorem 3.2; Hörmander and Wermer 1968, Theorem 4.1].

Theorem 2.6. *Let Z be an Oka manifold, let M be an open Riemann surface, and let $S = K \cup \Gamma \subset M$ be an admissible subset in the sense of Definition 2.1. Given a finite subset $\Lambda \subset \mathring{K}$ and an integer $k \in \mathbb{Z}_+$, every continuous map $f : S \to Z$ which is holomorphic on $\mathring{K}$ can be approximated uniformly on S by holomorphic maps $M \to Z$ having the same k-jet as f at all points in Λ.*

As we emphasized in the Introduction, the punctured null quadric $\mathfrak{A}_* \subset \mathbb{C}^n$, see (1-3) and (1-4), directing minimal surfaces in $\mathbb{R}^n$ and null curves in $\mathbb{C}^n$ is an Oka manifold for all $n \geq 3$; see [Alarcón and Forstnerič 2014, Example 4.4 Forstnerič 2017, Example 5.6.2]. Furthermore, for each $j \in \{1, \dots, n\}$ the complex manifold $\mathfrak{A} \cap \{z_j = 1\}$ is an embedded copy of the complex $(n-2)$-sphere

$$\mathbb{C}S^{n-2} = \{w = (w_1, \dots, w_{n-1}) \in \mathbb{C}^{n-1} : w_1^2 + \cdots + w_{n-1}^2 = 1\}.$$

Observe that $\mathbb{C}S^{n-2}$ is homogeneous relative to the complex Lie group $\mathrm{SO}(n-1, \mathbb{C})$, and hence it is an Oka manifold; see [Grauert 1957; Forstnerič 2017, Proposition 5.6.1]. For a more detailed discussion, see [Forstnerič 2017, Example 6.15.7; Alarcón and Forstnerič 2014, Example 7.8]. Moreover, choosing $k \in \{1, \dots, n\}$, $k \neq j$, the map $h = (h_1, \dots, h_n) : \mathbb{C} \to \mathfrak{A}$ given by

$$h_j(\zeta) = \zeta, \quad h_k(\zeta) = \sqrt{1 - \zeta^2}, \quad h_l(\zeta) = \frac{\mathrm{i}}{\sqrt{n-2}} \quad \text{for all } l \neq j, k, \quad \zeta \in \mathbb{C},$$

is a local holomorphic section near $\zeta = 0 \in \mathbb{C}$ of the coordinate projection $\pi_j : \mathfrak{A} \to \mathbb{C}$, $\pi_j(z_1, \dots, z_n) = z_j$, which satisfies $h(0) \neq 0$. Thus, the null quadric $\mathfrak{A} \subset \mathbb{C}^n$ meets the requirements in Theorem 1.3, including the ones in assertions (I) and (II), for all $n \geq 3$.

3. Paths in closed conical complex subvarieties of $\mathbb{C}^n$

We now use Notation 2.3; in particular, $\mathfrak{S} \subset \mathbb{C}^n$ $(n \geq 3)$ denotes a closed conical complex subvariety which is contained in no hyperplane of $\mathbb{C}^n$ and such that $\mathfrak{S}_* = \mathfrak{S} \setminus \{0\}$ is smooth and connected. We need the following:

Definition 3.1. Let Q be a topological space and $n \geq 3$ be an integer. A continuous map $f : Q \to \mathbb{C}^n$ is said to be *flat* if $f(Q) \subset \mathbb{C}z_0 = \{\zeta z_0 : \zeta \in \mathbb{C}\}$ for some $z_0 \in \mathbb{C}^n$, and *nonflat* otherwise. The map f is said to be *nowhere flat* if $f|_A : A \to \mathbb{C}^n$ is nonflat for all open subsets $\varnothing \neq A \subset Q$.

It is easily seen that a continuous map $f : [0, 1] \to \mathfrak{S}_* \subset \mathbb{C}^n$ is nonflat if, and only if,

$$\mathrm{span}\{T_{f(t)}\mathfrak{S} : t \in [0, 1]\} = \mathbb{C}^n.$$

3A. *Paths on $I := [0, 1]$.* We prove a couple of technical results for paths $[0, 1] \to \mathfrak{S}_*$ which pave the way to the construction of period-dominating sprays of holomorphic maps of an open Riemann surface into $\mathfrak{S}_*$ (see Lemma 3.4 in the next subsection).

Lemma 3.2. *Let $f : I \to \mathfrak{S}_*$ and $\vartheta : I \to \mathbb{C}_*$ be continuous maps. Let $\varnothing \neq I' \subset I$ be a closed subinterval and assume that f is nowhere flat on I'. There exist continuous functions $h_1, \dots, h_N : I \to \mathbb{C}$, with* *support on I', and a neighborhood U of* $0 \in \mathbb{C}^N$ *such that the* period map $\mathcal{P} : U \to \mathbb{C}^n$ *given by*

$$\mathcal{P}(\zeta) = \int_0^1 \phi^1_{\zeta_1 h_1(t)} \circ \cdots \circ \phi^N_{\zeta_N h_N(t)}(f(t))\,\vartheta(t)\,dt, \quad \zeta = (\zeta_1, \ldots, \zeta_N) \in U,$$

see (2-3), *is well-defined and has maximal rank equal to n at $\zeta = 0$.*

Proof. We choose continuous functions $h_1, \ldots, h_N : I \to \mathbb{C}$, with support on I', which will be specified later. Then we define for a small neighborhood U of $0 \in \mathbb{C}^N$ a map

$$\Phi : U \times I \to \mathfrak{S}$$

given by

$$\Phi(\zeta, t) := \phi^1_{\zeta_1 h_1(t)} \circ \cdots \circ \phi^N_{\zeta_N h_N(t)}(f(t)), \quad \zeta = (\zeta_1, \ldots, \zeta_N) \in U,\ t \in I.$$

Note that $\Phi(0, t) = f(t)$ for all $t \in I$; recall that each V_j vanishes at 0 for all $j \in \{1, \ldots, N\}$. Thus, since $f(I) \subset \mathfrak{S}_*$ is compact, we may assume that U is small enough so that Φ is well-defined and takes values in $\mathfrak{S}_*$. Furthermore, Φ is holomorphic in the variable ζ and its derivative with respect to ζ_j is

$$\left.\frac{\partial \Phi(\zeta, t)}{\partial \zeta_j}\right|_{\zeta=0} = h_j(t) V_j(f(t)), \quad j = 1, \ldots, N. \tag{3-1}$$

(See (2-2) and (2-3).) Thus, the period map $\mathcal{P} : U \to \mathbb{C}^n$ in the statement of the lemma can be written as

$$\mathcal{P}(\zeta) = \int_0^1 \Phi(\zeta, t)\,\vartheta(t)\,dt, \quad \zeta \in U.$$

Observe that $\mathcal{P}$ is holomorphic and, in view of (3-1),

$$\left.\frac{\partial \mathcal{P}(\zeta)}{\partial \zeta_j}\right|_{\zeta=0} = \int_0^1 h_j(t) V_j(f(t))\,\vartheta(t)\,dt, \quad j = 1, \ldots, N. \tag{3-2}$$

Since f is nowhere flat on I' (see Definition 3.1), (2-2) guarantees the existence of distinct points $t_1, \ldots, t_N \in I'$ such that

$$\operatorname{span}\{V_1(f(t_1)), \ldots, V_N(f(t_N))\} = \mathbb{C}^n. \tag{3-3}$$

Now we specify the values of the function h_j in I' ($j = 1, \ldots, N$); recall that $\operatorname{supp}(h_j) \subset I'$. We choose h_j with support in a small neighborhood $[t_j - \epsilon, t_j + \epsilon]$ of t_j in I', for some $\epsilon > 0$, and such that

$$\int_0^1 h_j(t)\,dt = \int_{t_j-\epsilon}^{t_j+\epsilon} h_j(t)\,dt = 1.$$

Then, for small $\epsilon > 0$, we have

$$\int_0^1 h_j(t) V_j(f(t))\,\vartheta(t)\,dt \approx V_j(f(t_j))\,\vartheta(t_j), \quad j = 1, \ldots, N.$$

Since $\vartheta(t) \neq 0$, (3-3) ensures that the vectors on the right side of the above display span $\mathbb{C}^n$, and hence the same is true for the vectors on the left side provided that $\epsilon > 0$ is chosen sufficiently small. This concludes the proof in view of (3-2). $\square$

Lemma 3.3. *Let $\vartheta : I \to \mathbb{C}_*$ be a continuous map. Given points $u_0, u_1 \in \mathfrak{S}_*$ and $x \in \mathbb{C}^n$, and a domain Ω in $\mathbb{C}^n$ containing 0 and x, there exists a continuous function $g : I \to \mathfrak{S}_*$ which is nowhere flat on a neighborhood of 0 in I and such that*:

(I) $g(0) = u_0$ *and* $g(1) = u_1$.

(II) $\int_0^s g(t)\vartheta(t)\,dt \in \Omega$ *for all* $s \in I$.

(III) $\int_0^1 g(t)\vartheta(t)\,dt = x$.

Proof. Set $I_0 := \left[0, \frac{1}{2}\right]$ and choose any continuous nowhere-flat map $g_0 : I_0 \to \mathfrak{S}_*$ such that

$$g_0(0) = u_0, \qquad \int_0^s g_0(t)\vartheta(t)\,dt \in \Omega \quad \text{for all } s \in I_0. \tag{3-4}$$

Such a map can be constructed as follows. For any $0 < \delta < \frac{1}{2}$ let $f_\delta : I_0 \to [\delta, 1]$ be the continuous map given by $f_\delta(s) = 1 - ((1-\delta)/\delta)s$ for $s \in [0, \delta]$ and $f_\delta(s) = \delta$ for $s \in \left[\delta, \frac{1}{2}\right]$. Choose any continuous nowhere-flat map $\tilde{g}_0 : I_0 \to \mathfrak{S}_*$ with $\tilde{g}_0(0) = u_0$. Then $g_0 := f_\delta \tilde{g}_0 : I_0 \to \mathfrak{S}_*$ satisfies the requirements for any $\delta > 0$ sufficiently small.

Let $\varnothing \neq I' \subset \mathring{I}_0$ be a closed subinterval. Thus, Lemma 3.2 applied to g_0 provides continuous functions $h_1, \dots, h_N : I \to \mathbb{C}$, with support on I', and a neighborhood U of the origin in $\mathbb{C}^N$, such that the period map

$$U \ni \zeta \mapsto \mathcal{P}(\zeta) = \int_0^{1/2} \phi^1_{\zeta_1 h_1(t)} \circ \cdots \circ \phi^N_{\zeta_N h_N(t)}(g_0(t))\vartheta(t)\,dt, \quad \zeta = (\zeta_1, \dots, \zeta_N) \in \mathbb{C}^N,$$

has maximal rank equal to n at $\zeta = 0$. (See (2-3).) Set

$$\Phi(\zeta, t) := \phi^1_{\zeta_1 h_1(t)} \circ \cdots \circ \phi^N_{\zeta_N h_N(t)}(g_0(t)) \in \mathfrak{S}, \quad \zeta \in U,\ t \in I_0,$$

and observe that $\Phi(0, t) = g_0(t) \in \mathfrak{S}_*$ for all $t \in I_0$. Then, up to shrinking U if necessary, we have:

(a) $\Phi(U \times I_0) \subset \mathfrak{S}_*$ and $\mathcal{P}(U)$ contains a ball in $\mathbb{C}^n$ with radius $\epsilon > 0$ centered at

$$\mathcal{P}(0) = \int_0^{1/2} g_0(t)\vartheta(t)\,dt \in \Omega;$$

see (3-4).

(b) $\Phi(\zeta, t) = g_0(t)$ for all $(\zeta, t) \in U \times \left\{0, \frac{1}{2}\right\}$; recall that $h_j(0) = h_j\left(\frac{1}{2}\right) = 0$ for all $j = 1, \dots, N$.

(c) $\int_0^s \Phi(\zeta, t)\vartheta(t)\,dt \in \Omega$ for all $\zeta \in U$ and $s \in I_0$; see (3-4).

To conclude the proof we adapt the argument in [Alarcón and Forstnerič 2014, Lemma 7.3]. Since the convex hull of $\mathfrak{S}$ is $\mathbb{C}^n$, see [Alarcón and Forstnerič 2014, Lemma 3.1], we may construct a polygonal path $\Gamma \subset \Omega$ connecting $\mathcal{P}(0)$ and x; to be more precise, $\Gamma = \bigcup_{j=1}^m \Gamma_j$ where each Γ_j is a segment of the form $\Gamma_j = w_j + [0, 1]z_j$ for some $w_j \in \mathbb{C}^n$ and $z_j \in \mathfrak{S}_*$, the initial point w_1 of Γ_1 is $\mathcal{P}(0)$, the final point $w_m + z_m$ of Γ_m is x, and the initial point w_j of Γ_j agrees with the final one $w_{j-1} + z_{j-1}$ of Γ_{j-1} for all $j = 2, \dots, m$. Set

$$I_j := \left[\frac{1}{2} + \frac{j-1}{2m}, \frac{1}{2} + \frac{j}{2m}\right], \quad j = 1, \dots, m,$$

and observe that $\bigcup_{j=1}^m I_j = \left[\frac{1}{2}, 1\right]$. For any number $0 < \lambda < 1/(4m)$, set

$$I_j^\lambda := \left[\frac{1}{2} + \frac{j-1}{2m} + \lambda, \frac{1}{2} + \frac{j}{2m} - \lambda\right] \subset I_j, \quad j = 1, \dots, m.$$

Without loss of generality we may assume that $m \in \mathbb{N}$ is large enough so that

$$b_j(\lambda) := \int_{I_j^\lambda} \vartheta(t)\, dt \neq 0 \quad \text{for all } 0 < \lambda < \frac{1}{4m}, \ j = 1, \dots, m; \tag{3-5}$$

recall that ϑ has no zeros. Fix a number $0 < \lambda < 1/(4m)$ and set $b_j := b_j(\lambda)$. Pick a constant $\kappa > \max\{|u_0|, |u_1|, |z_1/b_1|, \dots, |z_m/b_m|\}$. Also choose numbers $0 < \tau < \mu < \lambda$, which will be specified later, and consider a continuous map $g_1 : \left[\frac{1}{2}, 1\right] \to \mathfrak{S}_*$ satisfying the following conditions:

(d) $g_1\left(\frac{1}{2}\right) = g_0\left(\frac{1}{2}\right)$ and $g_1(1) = u_1$.

(e) $g_1(t) = z_j/b_j$ for all $t \in I_j^\lambda$.

(f) $|g_1(t)| \leq \kappa$ for all $t \in \left[\frac{1}{2}, 1\right]$.

(g) $|g_1(t)| \leq \tau$ for all $t \in I_j^\tau \setminus I_j^\mu$.

If $\tau > 0$ is chosen sufficiently small, and if μ is close enough to λ, then (e), (f), (g), and (3-5) ensure that:

(h) The image of the map $\left[\frac{1}{2}, 1\right] \ni s \mapsto \mathcal{P}(0) + \int_{1/2}^s g_1(t)\vartheta(t)\, dt$ is close enough to Γ in the Hausdorff distance so that it lies in Ω.

(i) $\left|\mathcal{P}(0) + \int_{1/2}^1 g_1(t)\vartheta(t)\, dt - x\right| < \epsilon$, where $\epsilon > 0$ is the number appearing in (a).

For $\zeta \in U$, let $g^\zeta : I \to \mathfrak{S}_*$ denote the function given by $g^\zeta(t) = \Phi(\zeta, t)$ for $t \in \left[0, \frac{1}{2}\right]$ and $g^\zeta(t) = g_1(t)$ for $t \in \left[\frac{1}{2}, 1\right]$. Properties (a) and (i) guarantee the existence of $\zeta_0 \in U$ such that

$$\int_0^{1/2} g^{\zeta_0}(t)\vartheta(t)\, dt = x - \int_{1/2}^1 g_1(t)\vartheta(t)\, dt,$$

and so $\int_0^1 g^{\zeta_0}(t)\vartheta(t)\, dt = x$. Thus $g := g^{\zeta_0}$ meets (III). By (3-4), (b), and (d), we have that g is continuous and satisfies (I), whereas (c) and (h) ensure (II). □

3B. *Paths on open Riemann surfaces.* Let us now state and prove the main result of this section; recall that we are using Notation 2.3.

Lemma 3.4. *Let M be an open Riemann surface and let θ be a holomorphic 1-form vanishing nowhere on M. Let $p_0 \in M$ be a point, $C_1, \dots, C_l$ ($l \in \mathbb{N}$) be a family of oriented Jordan arcs or closed curves in M that only meet at p_0 (i.e., $C_i \cap C_j = \{p_0\}$ for all $i \neq j \in \{1, \dots, l\}$) and such that $C := \bigcup_{i=1}^l C_i$ is Runge in M. Also let $f : C \to \mathfrak{S}_*$ be a continuous map and assume that for each $i \in \{1, \dots, l\}$ there exists a subarc $\widetilde{C}_i \subset C_i$ such that f is nowhere flat on $\widetilde{C}_i$. Then there exist continuous functions $h_{i,1}, \dots, h_{i,N} : C \to \mathbb{C}$, with support on $\widetilde{C}_i$, $i = 1, \dots, l$, and a neighborhood U of $0 \in (\mathbb{C}^N)^l$ such that the* period map *$U \to (\mathbb{C}^n)^l$ whose i-th component $U \to \mathbb{C}^n$ is given by*

$$U \ni \zeta \mapsto \int_{C_i} \phi^1_{\zeta_1^1 h_{1,1}(p)} \circ \dots \circ \phi^N_{\zeta_N^1 h_{1,N}(p)} \circ \dots \circ \phi^1_{\zeta_1^l h_{l,1}(p)} \circ \dots \circ \phi^N_{\zeta_N^l h_{l,N}(p)}(f(p))\theta,$$

see (2-2) *and* (2-3), *where*

$$\zeta = (\zeta^1, \dots, \zeta^l) \in (\mathbb{C}^N)^l, \quad \zeta^i = (\zeta^i_1, \dots, \zeta^i_N) \in \mathbb{C}^N,$$

are holomorphic coordinates, *is well-defined and has maximal rank equal to nl at* $\zeta = 0$.

Proof. Consider the period map $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_l) : \mathscr{C}(C, \mathbb{C}^n) \to (\mathbb{C}^n)^l$ whose i-th component is defined by

$$\mathscr{C}(C, \mathbb{C}^n) \ni g \mapsto \mathcal{P}_i(g) = \int_{C_i} g\theta, \quad i = 1, \dots, l. \tag{3-6}$$

For each $i = 1, \dots, l$, let $\gamma_i : I = [0, 1] \to C_i$ be a smooth parametrization of C_i such that $\gamma_i(0) = p_0$. If C_i is closed then we choose γ_i with $\gamma_i(1) = p_0$; further, up to changing the orientation of C_i if necessary, we assume that the parametrization γ_i is compatible with the orientation of C_i. Thus,

$$\mathcal{P}_i(g) = \int_0^1 g(\gamma_i(t))\theta(\gamma_i(t), \dot\gamma_i(t))\, dt, \quad g \in \mathscr{C}(C, \mathbb{C}^n). \tag{3-7}$$

Let $\varnothing \neq I_i \subset \mathring{I}$ be a closed interval such that $\gamma_i(I_i) \subset \widetilde{C}_i$. Lemma 3.2 applied to I_i, $f \circ \gamma_i$, and $\theta(\gamma_i(\cdot), \dot\gamma_i(\cdot))$ provides continuous functions $h^i_1, \dots, h^i_N : I \to \mathbb{C}$, supported on I_i, and a neighborhood U_i of $0 \in \mathbb{C}^N$ such that the period map $\mathsf{P}_i : U_i \to \mathbb{C}^n$ given, for any $\zeta^i = (\zeta^i_1, \dots, \zeta^i_{N_i}) \in U_i$, by

$$\mathsf{P}_i(\zeta^i) = \int_0^1 \phi^1_{\zeta^i_1 h^i_1(t)} \circ \cdots \circ \phi^N_{\zeta^i_N h^i_N(t)}(f(\gamma_i(t)))\theta(\gamma_i(t), \dot\gamma_i(t))\, dt, \tag{3-8}$$

see (2-2) and (2-3), is well-defined and has maximal rank equal to n at $\zeta^i = 0$. Let U be a ball centered at the origin of $(\mathbb{C}^N)^l$ and contained in $U_1 \times \cdots \times U_l$. For each $i \in \{1, \dots, l\}$ and $j = 1, \dots, N$, we define $h_{i,j} : C \to \mathbb{C}$ by $h_{i,j}(\gamma_i(t)) = h^i_j(t)$ for all $t \in I$, and $h_{i,j}(p) = 0$ for all $p \in C \setminus C_i$. Recall that $h^i_j(0) = 0$ and so $h_{i,j}$ is continuous and $h_{i,j}(p_0) = 0$. Define $\Phi : U \times C \to \mathfrak{S}$ by

$$\Phi(\zeta, p) = \phi^1_{\zeta^1_1 h_{1,1}(p)} \circ \cdots \circ \phi^N_{\zeta^1_N h_{1,N}(p)} \circ \cdots \circ \phi^1_{\zeta^l_1 h_{l,1}(p)} \circ \cdots \circ \phi^N_{\zeta^l_N h_{l,N}(p)}(f(p)),$$

and, up to shrinking U if necessary, assume that $\Phi(U \times C) \subset \mathfrak{S}_*$.

Let $\mathsf{P} : U \to (\mathbb{C}^n)^l$ be the period map whose i-th component $U \to \mathbb{C}^n$, $i = 1, \dots, l$, is given by

$$U \ni \zeta \mapsto \int_{C_i} \Phi(\zeta, \cdot)\theta = \mathsf{P}_i(\zeta^i), \quad \zeta = (\zeta^1, \dots, \zeta^l) \in U;$$

see (3-8) and recall that $h_{i,j}$ vanishes everywhere on $C \setminus C_i$. It follows that P has maximal rank equal to nl at $\zeta = 0$. □

4. Jet-interpolation with approximation

We begin this section with some preparations.

Definition 4.1. Let M be an open Riemann surface. An admissible subset $S = K \cup \Gamma \subset M$ (see Definition 2.1) will be called *simple* if $K \neq \varnothing$, every component of Γ meets K, Γ does not contain closed Jordan curves, and every closed Jordan curve in S meets only one component of K. Further, S will be

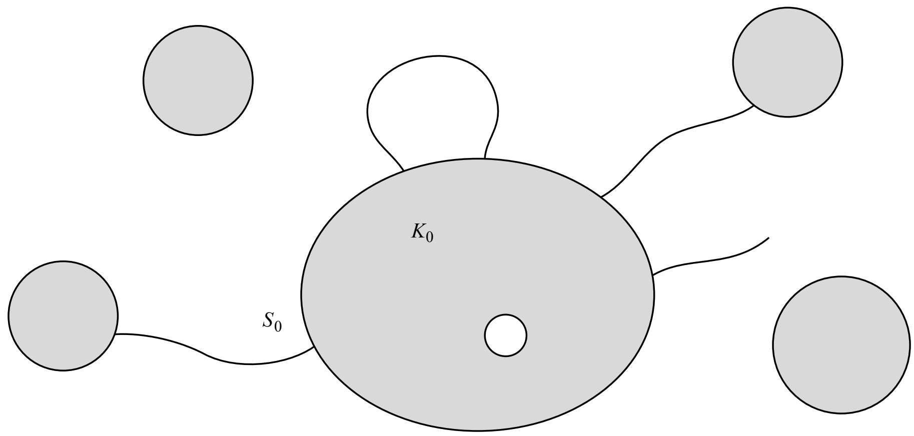

Figure 1. A very simple admissible set.

called *very simple* if it is simple, K has at most one nonsimply connected component K_0, which will be called the *kernel component* of K, and every component of Γ has at least one endpoint in K_0; in this case we denote by S_0 the component of S containing K_0 and call it the *kernel component* of S.

A connected admissible subset $S = K \cup \Gamma$ in an open Riemann surface M is very simple if, and only if, K has $m \in \mathbb{N}$ components $K_0, \ldots, K_{m-1}$, where K_i is simply connected for every $i > 0$, and $\Gamma = \Gamma' \cup \Gamma'' \cup \left(\bigcup_{i=1}^{m-1} \gamma_i\right)$ where Γ' consists of components of Γ with both endpoints in K_0, Γ'' consists of components of Γ with an endpoint in K_0 and the other one in $M \setminus K$, and γ_i is a component of Γ connecting K_0 and K_i for all $i = 1, \ldots, m-1$. Observe that, in this case, $K_0 \cup \Gamma'$ is a strong deformation retract of S. In general, a very simple admissible subset $S \subset M$ is of the form $S = (K \cup \Gamma) \cup K'$ where $K \cup \Gamma$ is a connected very simple admissible subset and $K' \subset M \setminus (K \cup \Gamma)$ is a (possibly empty) union of finitely many pairwise-disjoint smoothly bounded compact disks. (See Figure 1.)

If $S = K \cup \Gamma \subset M$ is admissible, we denote by $\mathscr{A}(S)$ the space of continuous functions $S \to \mathbb{C}$ which are of class $\mathscr{A}(K)$. Likewise, we define the space $\mathscr{A}(S, Z)$ for maps to any complex manifold Z.

In the remainder of this section we use Notation 2.3.

Lemma 4.2. *Let M be an open Riemann surface and θ be a holomorphic 1-form vanishing nowhere on M. Let $S = K \cup \Gamma \subset M$ be a very simple admissible subset and $L \subset M$ be a smoothly bounded compact domain such that $S \subset \mathring{L}$ and the kernel component S_0 of S is a strong deformation retract of L* (see Definition 4.1). *Denote by $l' \in \mathbb{Z}_+$ the dimension of the first homology group $H_1(S; \mathbb{Z}) = H_1(S_0; \mathbb{Z}) \cong H_1(L; \mathbb{Z})$. Let $K_0, \ldots, K_m$, $m \in \mathbb{Z}_+$, denote the components of K contained in S_0, where K_0 is the kernel component of K.*

Let $m' \in \mathbb{Z}_+$, $m' \geq m$, and let $p_0, \ldots, p_{m'}$ be distinct points in S such that $p_i \in \mathring{K}_i$ for all $i = 0, \ldots, m$ and $p_i \in \mathring{K}_0$ for all $i = m+1, \ldots, m'$, and let C_i, $i = 1, \ldots, m'$, be pairwise-disjoint oriented Jordan arcs in S with initial point p_0 and final point p_i. Set $l := l' + m'$. Also let C_i, $i = m'+1, \ldots, l$, be smooth Jordan curves in S determining a homology basis of S and such that $C_i \cap C_j = \{p_0\}$ for all $i \neq j \in \{1, \ldots, l\}$ and $C := \bigcup_{i=1}^{l} C_i$ is Runge in M. (See Figure 2.)

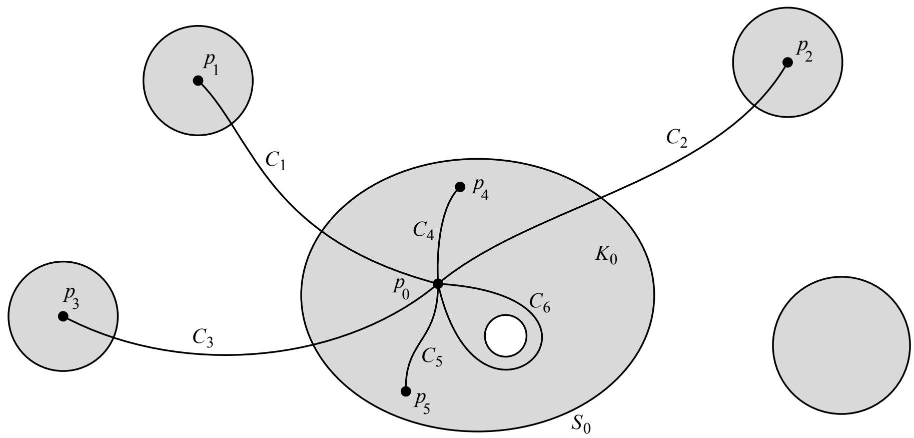

Figure 2. The sets in Lemma 4.2

Given $k \in \mathbb{N}$ and a map $f : S \to \mathfrak{S}_* \subset \mathbb{C}^n$ of class $\mathscr{A}(S)$ which is nonflat on $\mathring{K}_0$ (see Definition 3.1), *the following hold*:

(i) *There exist functions $h_{i,1}, \dots, h_{i,N} : L \to \mathbb{C}$, $i = 1, \dots, l$, of class $\mathscr{A}(L)$ and a neighborhood U of $0 \in (\mathbb{C}^N)^l$ such that*:

(i.1) *$h_{i,j}$ has a zero of multiplicity k at p_r for all $j = 1, \dots, N$ and $r = 1, \dots, m'$.*

(i.2) *Denoting by $\Phi_f : U \times S \to \mathfrak{S}$ the map*

$$\Phi_f(\zeta, p) = \phi^1_{\zeta^1_1 h_{1,1}(p)} \circ \cdots \circ \phi^N_{\zeta^1_N h_{1,N}(p)} \circ \cdots \circ \phi^1_{\zeta^l_1 h_{l,1}(p)} \circ \cdots \circ \phi^N_{\zeta^l_N h_{l,N}(p)}(f(p)),$$

see (2-2) *and* (2-3), *where $\zeta = (\zeta^1, \dots, \zeta^l) \in (\mathbb{C}^N)^l$ and $\zeta^i = (\zeta^i_1, \dots, \zeta^i_N) \in \mathbb{C}^N$, are holomorphic coordinates, the* period map *$U \to (\mathbb{C}^n)^l$ whose i-th component $U \to \mathbb{C}^n$ is given by*

$$U \ni \zeta \mapsto \int_{C_i} \Phi_f(\zeta, \cdot)\theta$$

has maximal rank equal to nl at $\zeta = 0$.

Furthermore, there is a neighborhood V of $g \in \mathscr{A}(S, \mathfrak{S}_*)$ such that the map $V \ni g \mapsto \Phi_g$ can be chosen to depend holomorphically on g.

(ii) *If $\mathfrak{S}_*$ is an Oka manifold, then f may be approximated uniformly on S by maps $\tilde{f} : L \to \mathfrak{S}_*$ of class $\mathscr{A}(L)$ such that*:

(ii.1) *$(\tilde{f} - f)\theta$ is exact on S, equivalently, $\int_{C_r} (\tilde{f} - f)\theta = 0$ for all $r = m' + 1, \dots, l$.*

(ii.2) *$\int_{C_r} (\tilde{f} - f)\theta = 0$ for all $r = 1, \dots, m'$.*

(ii.3) *$\tilde{f} - f$ has a zero of multiplicity k at p_r for all $r = 1, \dots, m'$.*

(ii.4) *No component function of $\tilde{f}$ vanishes everywhere on M.*

Notice that conditions (ii.1) and (ii.2) in the above lemma may be written as a single one in the form

$$\int_{C_r} (\tilde{f} - f)\theta = 0 \quad \text{for all } r = 1, \dots, l.$$

However, we write them separately with the aim of emphasizing that they are useful for different purposes; indeed, (ii.1) concerns the period problem, whereas (ii.2) deals with the problem of interpolation.

Proof. Choose $k \in \mathbb{N}$ and let $f : S \to \mathfrak{S}_*$ be a map of class $\mathscr{A}(S)$ which is nonflat on K_0. Consider the period map $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_l) : \mathscr{C}(C, \mathbb{C}^n) \to (\mathbb{C}^n)^l$ whose i-th component $\mathcal{P}_i : \mathscr{C}(C, \mathbb{C}^n) \to \mathbb{C}^n$ is defined by

$$\mathscr{C}(C, \mathbb{C}^n) \ni g \mapsto \mathcal{P}_i(g) = \int_{C_i} g\theta, \quad i = 1, \dots, l. \tag{4-1}$$

Since S is very simple and f is holomorphic and nonflat on $\mathring{K}_0$, each C_i, $i = 1, \dots, l$, contains a subarc $\widetilde{C}_i \subset \mathring{K}_0 \setminus \{p_0\}$ such that f is nowhere flat on $\widetilde{C}_i$; if $i \in \{m+1, \dots, m'\}$ then we may choose $\widetilde{C}_i \subset C_i \setminus \{p_0, p_i\}$. Thus, Lemma 3.4 applied to the map $f|_C : C \to \mathfrak{S}_*$, the base point p_0, and the curves $C_1, \dots, C_l$ furnishes functions $g_{i,1}, \dots, g_{i,N} : C \to \mathbb{C}$, with support on $\widetilde{C}_i$, $i = 1, \dots, l$, and a neighborhood U of $0 \in (\mathbb{C}^N)^l$ such that the period map $\mathsf{P} : U \to (\mathbb{C}^n)^l$ whose i-th component $\mathsf{P}_i : U \to \mathbb{C}^n$ is given by

$$\mathsf{P}_i(\zeta) := \int_{C_i} \phi^1_{\zeta^1_1 g_{1,1}(p)} \circ \cdots \circ \phi^N_{\zeta^1_N g_{1,N}(p)} \circ \cdots \circ \phi^1_{\zeta^l_1 g_{l,1}(p)} \circ \cdots \circ \phi^N_{\zeta^l_N g_{l,N}(p)}(f(p))\,\theta,$$

see (2-2) and (2-3), is well-defined and has maximal rank equal to nl at $\zeta = 0$. Since $C \subset M$ is Runge, Theorem 2.6 enables us to approximate each $g_{i,j}$ by functions $h_{i,j} \in \mathcal{O}(M) \subset \mathscr{A}(L) \subset \mathscr{A}(S)$ satisfying condition (i.1); recall that every function g_{ij} vanishes on a neighborhood of p_r for all $r = 1, \dots, m'$. Furthermore, if the approximation of $g_{i,j}$ by $h_{i,j}$ is close enough then the period map defined in (i.2) also has maximal rank at $\zeta = 0$. Finally, by varying f locally (keeping the functions $h_{i,j}$ fixed) we obtain a holomorphic family of maps $f \mapsto \Phi_f$ with the desired properties. This proves (i).

Let us now prove assertion (ii), so assume that $\mathfrak{S}_*$ is an Oka manifold. Up to adding to S a smoothly bounded compact disk $D \subset M \setminus S$ and extending f to D as a function of class $\mathscr{A}(S)$ all of whose components are different from the constant 0 on D, we may assume that no component function of f vanishes everywhere on S. Consider the map $\Phi : U \times S \to \mathfrak{S}$ given in (i.2) and, up to shrinking U if necessary, assume that $\Phi(U \times S) \in \mathfrak{S}_*$. Note that the functions $h_{i,j}$ are defined on L but f only on S. By (i), the period map $\mathcal{Q} : U \to (\mathbb{C}^n)^l$ with i-th component

$$\mathcal{Q}_i(\zeta) = \int_{C_i} \Phi(\zeta, \cdot)\theta = \mathcal{P}_i(\Phi(\zeta, \cdot)), \quad \zeta \in U,$$

see (4-1), has maximal rank equal to nl at $\zeta = 0$. It follows that the image by $\mathcal{Q}$ of any open neighborhood of $0 \in U \subset (\mathbb{C}^N)^l$ contains an open ball in $(\mathbb{C}^n)^l$ centered at $\mathcal{Q}(0) = \mathcal{P}(f)$; see (4-1). Since $S \subset M$ is Runge and $\mathfrak{S}_*$ is Oka, Theorem 2.6 allows us to approximate f by holomorphic maps $\hat{f} : M \to \mathfrak{S}_*$ such that

$$\hat{f} - f \text{ has a zero of multiplicity } k \text{ at } p_r \text{ for all } r = 1, \dots, m'. \tag{4-2}$$

Define $\widehat{\Phi} : U \times L \to \mathfrak{S}$ by

$$\widehat{\Phi}(\zeta, p) = \phi^1_{\zeta^1_1 h_{1,1}(p)} \circ \cdots \circ \phi^N_{\zeta^1_N h_{1,N}(p)} \circ \cdots \circ \phi^1_{\zeta^l_1 h_{l,1}(p)} \circ \cdots \circ \phi^N_{\zeta^l_N h_{l,N}(p)}(\hat{f}(p)) \tag{4-3}$$

and, up to shrinking U once again if necessary, assume that $\widehat{\Phi}(U \times L) \subset \mathfrak{S}_*$. Consider now the period map $\widehat{\mathcal{Q}} : U \to (\mathbb{C}^n)^l$ whose i-th component $U \to \mathbb{C}^n$ is given by

$$U \ni \zeta \mapsto \int_{C_i} \widehat{\Phi}(\zeta, \cdot)\,\theta, \quad i = 1, \dots, l.$$

Thus, for any open ball $0 \in W \subset U$, if the approximation of f by $\hat{f}$ is close enough, the range of $\widehat{\mathcal{Q}}(W)$ also contains $\mathcal{P}(f)$. Therefore, there is $\zeta_0 \in W \subset U$ close to $0 \in (\mathbb{C}^N)^l$ such that

$$\tilde{f} := \widehat{\Phi}(\zeta_0, \cdot) : L \to \mathfrak{S}_* \tag{4-4}$$

lies in $\mathscr{A}(L)$ and satisfies (ii.1) and (ii.2); recall that S_0 is a strong deformation retract of L and so the curves C_i, $i = m' + 1, \dots, l$, determine a basis of $H_1(L; \mathbb{Z})$. To finish the proof, Lemma 2.2, (i.1), (4-3), and (4-4) guarantee that $\tilde{f} - \hat{f}$ has a zero of multiplicity (at least) k at p_r for all $r = 1, \dots, m'$. This and (4-2) ensure (ii.3). Finally, if the approximation of f by $\tilde{f}$ on S is close enough, since no component function of f vanishes everywhere on S, no component function of $\tilde{f}$ vanishes everywhere on M, which proves (ii.4) and concludes the proof. □

We now show the following technical result which will considerably simplify the subsequent proofs.

Proposition 4.3. *Let $n \geq 3$ be an integer and $\mathfrak{S}$ be an irreducible closed conical complex subvariety of $\mathbb{C}^n$ which is not contained in any hyperplane. Let $M = \mathring{M} \cup bM$ be a compact bordered Riemann surface, θ be a holomorphic 1-form vanishing nowhere on M, and $\Lambda \subset \mathring{M}$ be a finite subset. Choose $p_0 \in M \setminus \Lambda$ and, for each $p \in \Lambda$, let $C_p \subset \mathring{M}$ be a smooth Jordan arc with initial point p_0 and final point p such that $C_p \cap C_q = \{p_0\}$ for all $p \neq q \in \Lambda$.*

Let $f : M \to \mathfrak{S}_*$ be a map of class $\mathscr{A}(M)$ which is flat (see Definition 3.1) *and $k \in \mathbb{N}$ be an integer. Then f may be approximated uniformly on M by nonflat maps $\tilde{f} : M \to \mathfrak{S}_*$ of class $\mathscr{A}(M)$ satisfying the following properties*:

(i) *$(\tilde{f} - f)\theta$ is exact on M.*

(ii) *$\int_{C_p} (\tilde{f} - f)\theta = 0$ for all $p \in \Lambda$.*

(iii) *$\tilde{f} - f$ has a zero of multiplicity k at all points $p \in \Lambda$.*

Proof. Without loss of generality we assume that $\Lambda \neq \varnothing$, write $\Lambda = \{p_1, \dots, p_{l'}\}$, and set $C_i := C_{p_i}$, $i = 1, \dots, l'$. Choose $C_{l'+1}, \dots, C_l$ closed Jordan loops in $\mathring{M}$ forming a basis of $H_1(M, \mathbb{Z}) \cong \mathbb{Z}^{l-l'}$ such that $C_i \cap C_j = \{p_0\}$ for all $i, j \in \{1, \dots, l\}$, $i \neq j$, and $C := \bigcup_{j=1}^l C_j$ is a Runge subset of M; existence of such loops is ensured by basic topological arguments. Consider smooth parametrizations $\gamma_j : [0, 1] \to C_j$ of the respective curves verifying $\gamma_j(0) = p_0$ and $\gamma_j(1) = p_j$ for $j = 1, \dots, l'$, and $\gamma_j(0) = \gamma_j(1) = p_0$ for $j = l' + 1, \dots, l$.

Since f is flat there exists $z_0 \in \mathfrak{S}_*$ such that $f(M) \subset \mathbb{C}_* z_0$. Observe that $\mathbb{C}z_0$ is a proper complex subvariety of $\mathfrak{S}$. We consider the period map $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_l) : \mathscr{A}(M) \to (\mathbb{C}^n)^l$ defined by

$$\mathscr{A}(M) \ni g \mapsto \mathcal{P}_j(g) := \int_{C_j} g\theta = \int_0^1 g(\gamma_j(t))\theta(\gamma_j(t), \dot{\gamma}_j(t))\, dt, \quad j = 1, \dots, l. \tag{4-5}$$

Note that a map $g \in \mathscr{A}(M)$ meets (i) and (ii) if, and only if, $\mathcal{P}(g) = \mathcal{P}(f)$. So, to finish the proof it suffices to approximate f uniformly on M by nonflat maps $\tilde{f} \in \mathscr{A}(M)$ satisfying the latter condition and also (iii).

Choose a holomorphic vector field V on $\mathbb{C}^n$ which is tangential to $\mathfrak{S}$ along $\mathfrak{S}$, vanishes at 0, and is not everywhere tangential to $\mathbb{C}_* z_0$ along $f(M)$. Let $\phi_s(z)$ denote the flow of V for small values of time $s \in \mathbb{C}$. Choose a nonconstant function $h_1 : M \to \mathbb{C}$ of class $\mathscr{A}(M)$ such that $h_1(p_0) = 0$. Denote by $\mho$ the space of all functions $h : M \to \mathbb{C}$ of class $\mathscr{A}(M)$ having a zero of multiplicity $k \in \mathbb{N}$ at all points $p \in \Lambda$. The following map is well-defined and holomorphic on a small open neighborhood $\mho^*$ of the zero function in $\mho$:

$$\mho^* \ni h \mapsto \mathcal{P}(\phi_{h_1(\cdot)h(\cdot)}(f(\cdot))) \in (\mathbb{C}^n)^l.$$

Each component $\mathcal{P}_j$, $j = 1, \dots, l$, of this map at the point $h = 0$ equals

$$\mathcal{P}_j(\phi_0(f)) = \mathcal{P}_j(f)$$

(recall that V vanishes at $0 \in \mathbb{C}^n$). Since $\mho$ is infinite-dimensional, there is a function $h \in \mho$ arbitrarily close to the function 0 (in particular, we may take $h \in \mho^*$) and nonconstant on M, such that

$$\mathcal{P}(\phi_{h_1(\cdot)h(\cdot)}(f(\cdot))) = \mathcal{P}(f).$$

Set $\tilde{f}(p) = \phi_{h_1(p)h(p)}(f(p))$, $p \in M$. Assume that $\|h\|_{0,M}$ is sufficiently small so that $\tilde{f}$ is well-defined and of class $\mathscr{A}(M)$, $\tilde{f}$ approximates f on M, and $\tilde{f}(p) \in \mathfrak{S}_*$ for all $p \in M$. By the discussion below equation (4-5), $\tilde{f}$ satisfies (i) and (ii). On the other hand, since h has a zero of multiplicity k at every point of Λ and h_1 is not constant, we infer that hh_1 also has a zero of multiplicity (at least) k at all points of Λ. Thus, Lemma 2.2 ensures that $\tilde{f} - f$ satisfies (iii). Finally, since $h_1(p_0) = 0$ and V vanishes at 0, we have $\tilde{f}(p_0) = f(p_0) \in \mathbb{C}_* z_0$, whereas since hh_1 is nonconstant on M and V is not everywhere tangential to $\mathbb{C}_* z_0$ along $f(M)$, there is a point $q \in M$ such that $\tilde{f}(q) \notin \mathbb{C}_* z_0$. This proves that $\tilde{f}$ is nonflat, which concludes the proof. □

The following is the main technical result of this paper.

Theorem 4.4. *Let $n \geq 3$ be an integer and $\mathfrak{S}$ be an irreducible closed conical complex subvariety of $\mathbb{C}^n$ which is not contained in any hyperplane and such that $\mathfrak{S}_* = \mathfrak{S} \setminus \{0\}$ is smooth and an Oka manifold. Let M be an open Riemann surface, θ be a holomorphic 1-form vanishing nowhere on M, $\mathcal{K} \subset M$ be a smoothly bounded Runge compact domain, and $\Lambda \subset M$ be a closed discrete subset. Choose $p_0 \in \mathring{\mathcal{K}} \setminus \Lambda$ and, for each $p \in \Lambda$, let $C_p \subset M$ be an oriented Jordan arc with initial point p_0 and final point p such that $C_p \cap C_q = \{p_0\}$ for all $q \neq p \in \Lambda$ and $C_p \subset \mathcal{K}$ for all $p \in \Lambda \cap \mathcal{K}$. Also, for each $p \in \Lambda$, let $\Omega_p \subset M$ be a compact neighborhood of p in M such that $\Omega_p \cap (\Omega_q \cup C_q) = \varnothing$ for all $q \neq p \in \Lambda$. Set $\Omega := \bigcup_{p \in \Lambda} \Omega_p$.*

Let $f:\mathcal{K}\cup\Omega\to\mathfrak{S}_*$ *be a map of class* $\mathscr{A}(\mathcal{K}\cup\Omega)$, *let* $\mathfrak{q}:H_1(M;\mathbb{Z})\to\mathbb{C}^n$ *be a group homomorphism, and let* $\mathfrak{Z}:\Lambda\to\mathbb{C}^n$ *be a map, such that*:

(a) $\int_\gamma f\theta=\mathfrak{q}(\gamma)$ *for all closed curves* $\gamma\subset\mathcal{K}$.

(b) $\int_{C_p} f\theta=\mathfrak{Z}(p)$ *for all* $p\in\Lambda\cap\mathcal{K}$.

Then, *for any integer* $k\in\mathbb{N}$, f *may be approximated uniformly on* $\mathcal{K}$ *by holomorphic maps* $\tilde f:M\to\mathfrak{S}_*$ *satisfying the following conditions*:

(A) $\int_\gamma \tilde f\theta=\mathfrak{q}(\gamma)$ *for all closed curves* $\gamma\subset M$.

(B) $\tilde f-f$ *has a zero of multiplicity* k *at* p *for all* $p\in\Lambda$; *equivalently*, $\tilde f$ *and* f *have the same* $(k-1)$*-jet at every point* $p\in\Lambda$.

(C) $\int_{C_p}\tilde f\theta=\mathfrak{Z}(p)$ *for all* $p\in\Lambda$.

(D) *No component function of* $\tilde f$ *vanishes everywhere on* M.

Proof. Up to slightly enlarging $\mathcal{K}$ if necessary, we may assume without loss of generality that $\Lambda\cap b\mathcal{K}=\varnothing$. Further, up to shrinking the sets Ω_p, we may also assume that, for each $p\in\Lambda$, either $\Omega_p\subset\mathring{\mathcal{K}}$ or $\Omega_p\cap\mathcal{K}=\varnothing$. Finally, by Proposition 4.3 we may assume that $f:\mathcal{K}\to\mathfrak{S}_*$ is nonflat.

Set $M_0:=\mathcal{K}$ and let $\{M_j\}_{j\in\mathbb{N}}$ be a sequence of smoothly bounded Runge compact domains in M such that

$$M_0\Subset M_1\Subset M_2\Subset\cdots\Subset\bigcup_{j\in\mathbb{N}}M_j=M.$$

Assume also that the Euler characteristic $\chi(M_j\setminus\mathring{M}_{j-1})$ of $M_j\setminus\mathring{M}_{j-1}$ is either 0 or -1, and that $\Lambda\cap bM_j=\varnothing$ for all $j\in\mathbb{N}$. Such a sequence can be constructed by basic topological arguments; see, e.g., [Alarcón and López 2013b, Lemma 4.2]. Since Λ is closed and discrete, M_j is compact, and $\Lambda\cap bM_j=\varnothing$ for all $j\in\mathbb{Z}_+$, we know $\Lambda_j:=\Lambda\cap M_j=\Lambda\cap\mathring{M}_j$ is either empty or finite. Without loss of generality we assume that $\Lambda_0\neq\varnothing$ and $\Lambda_j\setminus\Lambda_{j-1}=\Lambda\cap(\mathring{M}_j\setminus M_{j-1})\neq\varnothing$ for all $j\in\mathbb{N}$, and hence Λ is infinite.

Set $f_0:=f|_{\mathcal{K}}$ and, for each $p\in\Lambda_0\neq\varnothing$, choose an oriented Jordan arc $C^p\subset\mathring{M}_0$ with initial point p and final point p_0, such that

$$C^p\cap C^q=\{p_0\}\quad\text{for all } p\neq q\in\Lambda_0. \tag{4-6}$$

Such curves trivially exist.

To prove the theorem we shall inductively construct a sequence of maps $f_j:M_j\to\mathfrak{S}_*\subset\mathbb{C}^n$ and a family of oriented Jordan arcs $C^p\subset\mathring{M}_j$, $p\in\Lambda_j\setminus\Lambda_{j-1}\neq\varnothing$, $j\in\mathbb{N}$, with initial point p and final point p_0, satisfying the following properties:

(i_j) $\|f_j-f_{j-1}\|_{0,M_{j-1}}<\epsilon_j$ for a certain constant $\epsilon_j>0$ which will be specified later.

(ii_j) $\int_\gamma f_j\theta=\mathfrak{q}(\gamma)$ for all closed curves $\gamma\subset M_j$.

(iii_j) $\int_{C^p}f_j\theta=\mathfrak{q}(C_p*C^p)-\mathfrak{Z}(p)$ for all $p\in\Lambda_j$. (Recall that $*$ denotes the product of oriented arcs; see Section 2A.)

(iv_j) f_j-f has a zero of multiplicity k at p for all $p\in\Lambda_j$.

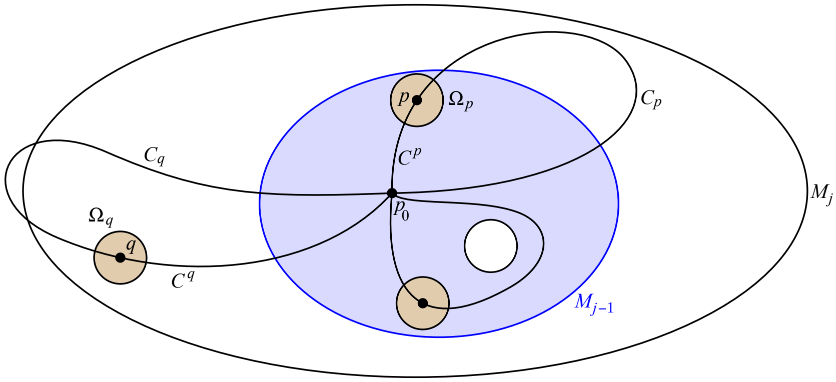

Figure 3. The set $\mathcal{K} \subset M$, the arcs C_p, and the domains Ω_p in Theorem 4.4.

(v$_j$) $C^p \cap C^q = \{p_0\}$ for all $p \neq q \in \Lambda_j$.

(vi$_j$) No component function of f_j vanishes everywhere on M_j.

(See Figure 3.) Assume for a moment that we have already constructed such sequence. Then choosing the sequence $\{\epsilon_j\}_{j\in\mathbb{N}}$ decreasing to zero fast enough, (i$_j$) ensures that there is a limit holomorphic map

$$\tilde{f} := \lim_{j\to\infty} f_j : M \to \mathfrak{S}_*$$

which is as close as desired to f uniformly on $\mathcal{K}$, whereas properties (ii$_j$), (iii$_j$), (iv$_j$), (v$_j$), and (vi$_j$) guarantee (A), (B), (C), and (D). This would conclude the proof.

The basis of the induction is given by the nonflat map $f_0 = f|_{\mathcal{K}}$ and the already fixed oriented arcs C^p, $p \in \Lambda_0$. Condition (i$_0$) is vacuous, (ii$_0$) = (a), (iii$_0$) is implied by (a) and (b), (iv$_0$) is trivial, and (v$_0$) = (4-6). For the inductive step, we assume that we already have a map $f_{j-1} : M_{j-1} \to \mathfrak{S}_*$ and arcs $C^p \subset \mathring{M}_{j-1}$, $p \in \Lambda_{j-1}$, satisfying properties (ii$_{j-1}$)–(v$_{j-1}$) for some $j \in \mathbb{N}$, and let us construct a map f_j and arcs C^p for $p \in \Lambda_j \setminus \Lambda_{j-1} = \Lambda \cap (\mathring{M}_j \setminus M_{j-1})$, enjoying conditions (i$_j$)–(vi$_j$). We distinguish cases depending on the Euler characteristic $\chi(M_j \setminus \mathring{M}_{j-1})$.

Case 1: the noncritical case. Assume that $\chi(M_j \setminus \mathring{M}_{j-1}) = 0$. In this case M_{j-1} is a strong deformation retract of M_j. Recall that $\Lambda_j \setminus \Lambda_{j-1}$ is a nonempty finite set. Choose, for each $p \in \Lambda_j \setminus \Lambda_{j-1}$, an oriented Jordan arc $C^p \subset \mathring{M}_j$ with initial point p and final point p_0, so that condition (v$_j$) holds; such arcs trivially exist. Up to shrinking Ω_p if necessary, we assume without loss of generality that $\Omega_p \subset \mathring{M}_j \setminus M_{j-1}$ for all $p \in \Lambda_j \setminus \Lambda_{j-1}$ and $\Omega_p \cap C^q = \varnothing$ for all $q \in \Lambda_j \setminus \Lambda_{j-1}$, $q \neq p$.

Set

$$K := M_{j-1} \cup \Bigl(\bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} \Omega_p \Bigr), \quad \Gamma := \Bigl(\bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} C^p \Bigr) \setminus \mathring{K},$$

and, up to slightly modifying the arcs C^p, $p \in \Lambda_j \setminus \Lambda_{j-1}$, assume that $S := K \cup \Gamma \subset \mathring{M}_j$ is an admissible subset of M (see Definition 2.1). Notice that S is connected and a strong deformation retract of M_j;

moreover, as admissible set, S is very simple and the kernel component of K is M_{j-1} (see Definition 4.1). Thus, Lemma 3.3 furnishes a map $\varphi : S \to \mathfrak{S}_*$ of class $\mathscr{A}(S)$ such that:

(I) $\varphi = f_{j-1}$ on M_{j-1}.

(II) $\varphi = f$ on $\bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} \Omega_p$.

(III) $\int_{C^p} \varphi\theta = \mathfrak{q}(C_p * C^p) - \mathfrak{Z}(p)$ for all $p \in \Lambda_j \setminus \Lambda_{j-1}$.

Now, given $\epsilon_j > 0$, Lemma 4.2(ii) applied to S, M_j, the arcs C^p, $p \in \Lambda_j$, the integer $k \in \mathbb{N}$, and the map φ, provides a map $f_j : M_j \to \mathfrak{S}_*$ of class $\mathscr{A}(M_j)$ satisfying the following conditions:

(IV) $\|f_j - \varphi\|_{0,S} < \epsilon_j$.

(V) $(f_j - \varphi)\theta$ is exact on S.

(VI) $\int_{C^p}(f_j - \varphi)\theta = 0$ for all $p \in \Lambda_j$.

(VII) $f_j - \varphi$ has a zero of multiplicity k at p for all $p \in \Lambda_j$.

(VIII) No component function of f_j vanishes everywhere on M_j.

We claim the map f_j satisfies properties (i_j)–(iv_j); recall that (v_j) is already guaranteed. Indeed, (i_j) follows from (I) and (IV); (ii_j) from (ii_{j-1}), (I), (V), and the fact that M_{j-1} is a strong deformation retract of M_j; (iii_j) from (iii_{j-1}), (I), (III), and (VI); (iv_j) from (iv_{j-1}), (I), (II), and (VII); and (vi_j) = (VIII).

Case 2: *the critical case. Assume that* $\chi(M_j \setminus \mathring{M}_{j-1}) = -1$. Now, the change of topology is described by attaching to M_{j-1} a smooth arc α in $\mathring{M}_j \setminus \mathring{M}_{j-1}$ satisfying M_{j-1} only at its endpoints. Thus, $M_{j-1} \cup \alpha$ is a strong deformation retract of M_j. Further, we may choose α such that $\alpha \cap \Lambda = \varnothing$ and $S := M_{j-1} \cup \alpha$ is an admissible subset of M, which is very simple (see Definition 4.1). Since both endpoints of α lie in bM_{j-1}, there is a closed curve $\beta \subset S$ which contains α as a subarc and is not in the homology of M_{j-1}. Now, Lemma 3.3 furnishes a map $\varphi : S \to \mathfrak{S}_*$ of class $\mathscr{A}(S)$ such that $\varphi = f_{j-1}$ on M_{j-1} and

$$\int_\beta \varphi\theta = \mathfrak{q}(\beta).$$

Choose a smoothly bounded compact domain $L \subset \mathring{M}_j$ such that $S \subset \mathring{L}$, S is a strong deformation retract of L, and $L \cap (\Lambda_j \setminus \Lambda_{j-1}) = \varnothing$. Given $\epsilon_j > 0$, Lemma 4.2(ii) applied to S, L, the arcs C^p, $p \in \Lambda_{j-1}$, the integer $k \in \mathbb{N}$, and the map φ, provides a map $\hat{f} : L \to \mathfrak{S}_*$ of class $\mathscr{A}(L)$ satisfying the following conditions:

(i) $\|\hat{f} - \varphi\|_{0,S} < \frac{1}{2}\epsilon_j$.

(ii) $(\hat{f} - \varphi)\theta$ is exact on S.

(iii) $\int_{C^p}(\hat{f} - \varphi)\theta = 0$ for all $p \in \Lambda_{j-1}$.

(iv) $\hat{f} - \varphi$ has a zero of multiplicity k at p for all $p \in \Lambda_{j-1}$.

Since the Euler characteristic satisfies $\chi(M_j \setminus \mathring{L}) = 0$, this reduces the construction to the noncritical case. This finishes the inductive process and concludes the proof of the theorem. □

To finish this section we prove a Runge–Mergelyan-type theorem with jet-interpolation for holomorphic maps into Oka subvarieties of $\mathbb{C}^n$ in which a component function is preserved provided that it holomorphically extends to the whole source Riemann surface. This will be an important tool to ensure conditions (III) and (IV) in Theorem 1.2 and (I) and (II) in Theorem 1.3.

Lemma 4.5. *Let $n \geq 3$ be an integer and $\mathfrak{S}$ be an irreducible closed conical complex subvariety of $\mathbb{C}^n$ which is not contained in any hyperplane. Assume that $\mathfrak{S}_* = \mathfrak{S} \setminus \{0\}$ is smooth and an Oka manifold, and that $\mathfrak{S} \cap \{z_1 = 1\}$ is also an Oka manifold and the coordinate projection $\pi_1 : \mathfrak{S} \to \mathbb{C}$ onto the z_1-axis admits a local holomorphic section h near $z_1 = 0$ with $h(0) \neq 0$. Let M be an open Riemann surface of finite topology, θ be a holomorphic 1-form vanishing nowhere on M, $S = K \cup \Gamma \subset M$ be a connected very simple admissible Runge subset* (see Definition 4.1) *which is a strong deformation retract of M. Let $\Lambda \subset \mathring{K}_0$ be a finite subset where K_0 is the kernel component of S. Choose $p_0 \in \mathring{K}_0 \setminus \Lambda$ and, for each $p \in \Lambda$, let $C_p \subset \mathring{K}_0$ be an oriented Jordan arc with initial point p_0 and final point p such that $C_p \cap C_q = \{p_0\}$ for all $q \neq p \in \Lambda$.*

Let $f = (f_1, \dots, f_n) : S \to \mathfrak{S}_*$ be a continuous map, holomorphic on K, such that f_1 extends to a holomorphic map $M \to \mathbb{C}$ which does not vanish everywhere on M. Assume also that $f|_K : K \to \mathfrak{S}_*$ is nonflat. Then, for any integer $k \in \mathbb{Z}_+$, f may be approximated in the $\mathscr{C}^0(S)$-topology by holomorphic maps $\tilde{f} = (\tilde{f}_1, \tilde{f}_2, \dots, \tilde{f}_n) : M \to \mathfrak{S}_*$ such that:

(i) *$\tilde{f}_1 = f_1$ everywhere on M.*

(ii) *$\tilde{f} - f$ has a zero of multiplicity k at p for all $p \in \Lambda$.*

(iii) *$\int_{C_p} (\tilde{f} - f)\theta = 0$ for all $p \in \Lambda$.*

(iv) *$\int_\gamma (\tilde{f} - f)\theta = 0$ for all closed curves $\gamma \subset S$.*

Proof. We adapt the ideas in [Alarcón and Forstnerič 2014, proof of Theorem 7.7]. Set $\mathfrak{S}' := \mathfrak{S} \cap \{z_1 = 1\}$. By dilations we see that $\mathfrak{S} \setminus \{z_1 = 0\}$ is biholomorphic to $\mathfrak{S}' \times \mathbb{C}_*$ (and hence is Oka), and the projection $\pi_1 : \mathfrak{S}' \to \mathbb{C}$ is a trivial fiber bundle with Oka fiber $\mathfrak{S}'$ except over $0 \in \mathbb{C}$. Write $(f_1, \hat{f}) = (f_1, f_2, \dots, f_n)$; that is, $\hat{f} := (f_2, \dots, f_n) : S \to \mathbb{C}^{n-1}$. Since f_1 is holomorphic and nonconstant on M, its zero set $f_1^{-1}(0) = \{a_1, a_2, \dots\}$ is a closed discrete subset of M. The pullback $f_1^*\pi_1 : E = f^*\mathfrak{S} \to M$ of the projection $\pi_1 : \mathfrak{S} \to \mathbb{C}$ is a trivial holomorphic fiber bundle with fiber $\mathfrak{S}'$ over $M \setminus f_1^{-1}(0)$, but it may be singular over the points $a_j \in f_1^{-1}(0)$. The map $\hat{f} : S \to \mathbb{C}^{n-1}$ satisfies $\hat{f}(x) \in \pi_1^{-1}(f_1(x))$ for all $x \in S$, so $\hat{f}$ corresponds to a section of $E \to M$ over the set S.

Now we need to approximate $\hat{f}$ uniformly by a section $E \to M$ solving the problem of periods and interpolation. (Except for the period and interpolation conditions, a solution is provided by the Oka principle for sections of ramified holomorphic maps with Oka fibers; see [Forstnerič 2003; 2017, §6.14]. We begin by choosing a local holomorphic solution on a small neighborhood of any point $a_j \in M \setminus S$ so that $\hat{f}(a_j) \neq 0$, and we add these neighborhoods to the domain of holomorphicity of $\hat{f}$. Then we need to approximate a holomorphic solution $\hat{f}$ on a smoothly bounded compact set $K \subset M$ by a holomorphic one on a larger domain $L \subset M$ assuming that K is a strong deformation retract of L and $L \setminus K$ does not contain any point a_j. This can be done by applying the Oka principle for maps to the Oka fiber G' of

$\pi : G \to \mathbb{C}$ over $\mathbb{C}_*$. In the critical case we add a smooth Jordan arc α to the domain $K \subset M$ disjoint from the points a_j and such that $K \cup \alpha$ is a strong deformation retract of the next domain. Next, we extend $\hat{f}$ smoothly over α so that the integral $\int_\alpha \hat{f}\theta$ takes the correct value by applying an analogous result of Lemma 3.3 but keeping the first coordinate fixed; this reduces the proof to the noncritical case and concludes the proof of the lemma. □

5. General position, completeness, and properness results

In this section we prove several results that pave the way to the proof of Theorem 1.3 in Section 6. Thus, all the results in this section concern directed holomorphic immersions of open Riemann surfaces into $\mathbb{C}^n$; we point out that the methods of proof easily adapt to give analogous results for conformal minimal immersions into $\mathbb{R}^n$ (see Section 7).

We begin with the following:

Definition 5.1. Let $\mathfrak{S}$ be a closed conical complex subvariety of $\mathbb{C}^n$ $(n \geq 3)$, M be an open Riemann surface, and $S = K \cup \Gamma \subset M$ be an admissible subset (see Definition 2.1). By a *generalized $\mathfrak{S}$-immersion* $S \to \mathbb{C}^n$ we mean a map $F : S \to \mathbb{C}^n$ of class $\mathscr{C}^1(S)$ whose restriction to K is an $\mathfrak{S}$-immersion of class $\mathscr{A}^1(S)$ and the derivative $F'(t)$ with respect to any local real parameter t on Γ belongs to $\mathfrak{S}_*$.

We now prove a Mergelyan-type theorem for generalized $\mathfrak{S}$-immersions which follows from Lemmas 4.2 and 4.5; it will be very useful in the subsequent results.

Proposition 5.2. *Let $\mathfrak{S} \subset \mathbb{C}^n$ be as in Theorem 4.4. Let M be a compact bordered Riemann surface and let $S = K \cup \Gamma \subset \mathring{M}$ be a very simple admissible Runge compact subset such that the kernel component S_0 of S* (see Definition 4.1) *is a strong deformation retract of M. Let $\Lambda \subset \mathring{K}$ be a finite subset and assume that $\Lambda \cap K'$ consists of at most a single point for each component K' of K, $K' \neq K_0$, where K_0 is the kernel component of K. Given an integer $k \in \mathbb{N}$, every generalized $\mathfrak{S}$-immersion $F = (F_1, \ldots, F_n) : S \to \mathbb{C}^n$ which is nonflat on $\mathring{K}_0$ may be approximated in the $\mathscr{C}^1(S)$-topology by $\mathfrak{S}$-immersions $\widetilde{F} = (\widetilde{F}_1, \ldots, \widetilde{F}_n) : M \to \mathbb{C}^n$ such that $\widetilde{F} - F$ has a zero of multiplicity $k \in \mathbb{N}$ at all points $p \in \Lambda$ and $\widetilde{F}$ has no constant component function.*

Furthermore, if $\mathfrak{S} \cap \{z_1 = 1\}$ is an Oka manifold, the coordinate projection $\pi_1 : \mathfrak{S} \to \mathbb{C}$ onto the z_1-axis admits a local holomorphic section h near $z_1 = 0$ with $h(0) \neq 0$, $\Lambda \subset \mathring{K}_0$, and F_1 extends to a nonconstant holomorphic function $M \to \mathbb{C}$, then $\widetilde{F}$ may be chosen with $\widetilde{F}_1 = F_1$.

We point out that an analogous result of the above proposition remains true for arbitrary admissible subsets; we shall not prove the most general statement for simplicity of exposition. Anyway, Proposition 5.2 will suffice for the aim of this paper.

Proof. Let θ be a holomorphic 1-form vanishing nowhere on M. Set $f = dF/\theta : S \to \mathfrak{S}_*$ and observe that f is nonflat on $\mathring{K}_0$ and of class $\mathscr{A}(S)$, and that $f\theta$ is exact on S. Fix a point $p_0 \in \mathring{K}_0 \setminus \Lambda$. If S is not connected then $S \setminus S_0$ consists of finitely many pairwise-disjoint, smoothly bounded compact disks $K_1, \ldots, K_m$. For each $i \in \{1, \ldots, m\}$ choose a smooth Jordan arc $\gamma_i \subset \mathring{M}$ with an endpoint in $(bK_0) \setminus \Gamma$, the other endpoint in bK_i, and otherwise disjoint from S. Choose these arcs so that $S' := S \cup \big(\bigcup_{i=1}^m \gamma_i\big)$ is an admissible subset of M. It follows that S' is connected, very simple, and a strong deformation retract of M. By Lemma 3.3 we may extend f to a map $f' : S' \to \mathfrak{S}_*$ of class $\mathscr{A}(S')$ such that $F(p_0) + \int_{p_0}^{p} f\theta = F(p)$ for all $p \in S'$. From now on we remove the primes and assume without loss of generality that S is connected.

For each $p \in \Lambda$ choose a smooth Jordan arc $C_p \subset S$ joining p_0 with p such that $C_p \cap C_q = \{p_0\}$ for all $p \neq q \in \Lambda$. By Lemma 4.2(ii) applied to the set $S \subset M$, the map f, the integer $k-1 \geq 0$, and the arcs $C_p,\ p \in \Lambda$, we may approximate f uniformly on S by a holomorphic map $\tilde{f} : M \to \mathfrak{S}_*$ such that:

(a) $\tilde{f}\theta$ is exact; recall that $f\theta$ is exact on (S and hence on) S_0 and that S_0 is a strong deformation retract of M.

(b) $F(p_0) + \int_{C_p} \tilde{f}\theta = F(p_0) + \int_{C_p} f\theta = F(p)$ for all $p \in \Lambda$.

(c) $\tilde{f} - f$ has a zero of multiplicity $k-1$ at all points $p \in \Lambda$.

(d) No component function of $\tilde{f}$ vanishes everywhere on M.

Then, property (a) ensures that the map $\widetilde{F} : M \to \mathbb{C}^n$ defined by

$$\widetilde{F}(p) := F(p_0) + \int_{p_0}^{p} \tilde{f}\theta, \quad p \in M,$$

is a well-defined $\mathfrak{S}$-immersion and is as close as desired to F in the $\mathscr{C}^1(S)$-topology. Moreover, properties (b) and (c) guarantee that $\widetilde{F} - F$ has a zero of multiplicity k at all points of Λ, whereas (d) ensures that $\widetilde{F}$ has no constant component function. This concludes the first part of the proof.

The second part of the lemma is proved in an analogous way but using Lemma 4.5 instead of Lemma 4.2(ii). Moreover, in order to reduce the proof to the case when S is connected, we need to extend f to a map f' on S' as above such that the first component of f' equals dF_1/θ; this is accomplished by a suitable analogue of Lemma 3.3, we leave the obvious details to the interested reader. □

5A. ***A general position theorem.*** We prove a desingularization result with jet-interpolation for directed immersions of class $\mathscr{A}^1$ on a compact bordered Riemann surface. We use Notation 2.3.

Theorem 5.3. *Let M be a compact bordered Riemann surface and $\Lambda \subset \mathring{M}$ be a finite set. Let $F : M \to \mathbb{C}^n$ ($n \geq 3$) be an $\mathfrak{S}$-immersion of class $\mathscr{A}^1(M)$ such that $F|_\Lambda$ is injective. Then, given $k \in \mathbb{N}$, F may be approximated uniformly on M by a $\mathfrak{S}$-embedding $\widetilde{F} : M \to \mathbb{C}^n$ of class $\mathscr{A}^1(M)$ such that $\widetilde{F} - F$ has a zero of multiplicity k at p for all $p \in \Lambda$.*

Proof. Proposition 4.3 allows us to assume without loss of generality that $F : M \to \mathbb{C}^n$ is nonflat. We assume that M is a smoothly bounded compact domain in an open Riemann surface R. We associate to F the difference map

$$\delta F : M \times M \to \mathbb{C}^n, \quad \delta F(x, y) = F(y) - F(x).$$

Obviously, F is injective if and only if $(\delta F)^{-1}(0) = D_M = \{(x, x) : x \in M\}$.

Since F is an immersion and $F|_\Lambda : \Lambda \to \mathbb{C}^n$ is injective, there is an open neighborhood $U \subset M \times M$ of $D_M \cup (\Lambda \times \Lambda)$ such that $\delta F \neq 0$ everywhere on $\overline{U} \setminus D_M$. To prove the theorem it suffices to find arbitrarily close to F another $\mathfrak{S}$-immersion $\widetilde{F}: M \to \mathbb{C}^n$ of class $\mathscr{A}^1(M)$ such that $\widetilde{F} - F$ has a zero of multiplicity k at all points of Λ whose difference map $\delta\widetilde{F}|_{M\times M\setminus U}$ is transverse to the origin. Indeed, since $\dim_{\mathbb{C}} M \times M = 2 < n$, this will imply that $\delta\widetilde{F}$ does not assume the value 0 on $M \times M \setminus U$, so $\widetilde{F}(y) \neq \widetilde{F}(x)$ when $(x, y) \in M \times M \setminus U$. On the other hand, if $(x, y) \in \bar{U} \setminus D_M$ then $\widetilde{F}(y) \neq \widetilde{F}(x)$ provided that $\widetilde{F}$ is sufficiently close to F.

To construct such an $\mathfrak{S}$-immersion we will use the standard transversality argument by Abraham [1963]. We need to find a neighborhood $\mathcal{V} \subset \mathbb{C}^N$ of the origin in a complex Euclidean space and a map $H: \mathcal{V} \times M \to \mathbb{C}^n$ of class $\mathscr{A}^1(\mathcal{V} \times M)$ such that:

(a) $H(0, \cdot\,) = F$.

(b) $H - F$ has a zero of multiplicity k at p for all $p \in \Lambda$.

(c) The difference map $\delta H: \mathcal{V} \times M \times M \to \mathbb{C}^n$, defined by

$$\delta H(\zeta, x, y) = H(\zeta, y) - H(\zeta, x), \quad \zeta \in \mathcal{V}, \quad x, y \in M,$$

is a submersive family of maps in the sense that the partial differential

$$d_\zeta \delta H(\zeta, x, y)|_{\zeta=0} : T_0\mathbb{C}^N \cong \mathbb{C}^N \to \mathbb{C}^n$$

is surjective for any $(x, y) \in M \times M \setminus U$.

By openness of the latter condition and compactness of $M \times M \setminus U$ it follows that the partial differential $d_\zeta \delta H$ is surjective for all ζ in a neighborhood $\mathcal{V}' \subset \mathcal{V}$ of the origin in $\mathbb{C}^N$. Hence, the map $\delta H : M \times M \setminus U \to \mathbb{C}^n$ is transverse to any submanifold of $\mathbb{C}^n$, in particular, to the origin $\{0\} \subset \mathbb{C}^n$. The standard argument then shows that for a generic member $H(\zeta, \cdot\,): M \to \mathbb{C}^n$ of this family, the difference map $\delta H(\zeta, \cdot\,)$ is also transverse to $0 \in \mathbb{C}^n$ on $M \times M \setminus U$. Choosing such a ζ sufficiently close to 0 we then obtain the desired $\mathfrak{S}$-embedding $\widetilde{F} := H(\zeta, \cdot\,)$.

To construct a map H as above we fix a nowhere-vanishing holomorphic 1-form θ on R and write $dF = f\theta$, where $f: M \to \mathfrak{S}_*$ is a map of class $\mathscr{A}^1(M)$. We begin with the following.

Lemma 5.4. *For any point $(p, q) \in M \times M \setminus (D_M \cup (\Lambda \times \Lambda))$ there is a deformation family $H = H^{(p,q)}(\zeta, \cdot\,)$ satisfying conditions* (a) *and* (b) *above, with $\zeta \in \mathbb{C}^n$, such that the differential $d_\zeta \delta H(\zeta, p, q)|_{\zeta=0} : \mathbb{C}^n \to \mathbb{C}^n$ is an isomorphism.*

For the proof we adapt the arguments in [Alarcón and Forstnerič 2014, Lemma 6.1] in order to guarantee also the jet-interpolation, i.e., condition (b) of the map H.

Proof. Pick $(p, q) \in M \times M \setminus (D_M \cup (\Lambda \times \Lambda))$. We distinguish cases.

Case 1: assume that $\{p, q\} \cap \Lambda \neq \varnothing$. Assume that $p \in \Lambda$ and hence $q \notin \Lambda$; otherwise we reason in a symmetric way. Write $\Lambda = \{p = p_1, \ldots, p_{l'}\}$. Pick a point $p_0 \in M \setminus (\Lambda \cup \{q\})$ and choose closed loops $C_j \subset M \setminus \Lambda$, $j = 1, \ldots, l''$, forming a basis of $H_1(M, \mathbb{Z}) = \mathbb{Z}^{l''}$, and smooth Jordan arcs $C_{l''+j}$ joining p_0 with p_j, $j = 1, \ldots, l'$, such that setting $l := l' + l''$, we have that $C_i \cap C_j = \{p_0\}$ for any $i, j \in \{1, \ldots, l\}$ and that $C := \bigcup_{j=1}^{l} C_j$ is a Runge set in M. Also choose another smooth Jordan arc C_q joining p_0 with q and verifying $C \cap C_q = \{p_0\}$. Finally let $\gamma_j : [0, 1] \to C_j$, $j = 1, \ldots, l$, and $\gamma : [0, 1] \to C_q$ be smooth parametrizations of the respective curves verifying $\gamma_j(0)=\gamma_j(1)=p_0$ for $j=1,\dots,l''$, $\gamma_j(0)=p_0$ and $\gamma_j(1)=p_j$ for $j=l''+1,\dots,l$, and $\gamma(0)=p_0$ and $\gamma(1)=q$.

Since F is nonflat, there exist tangential fields $V_1,\dots,V_n$ on $\mathfrak{S}$, vanishing at 0, and points $x_1,\dots,x_n\in C_q\setminus\{p_0,q\}$ such that, setting $z_i=f(x_i)\in\mathfrak{S}_*$, the vectors $V_1(z_1),\dots,V_n(z_n)$ span $\mathbb{C}^n$. Let $t_i\in(0,1)$ be such that $\gamma(t_i)=x_i$ and ϕ_t^i be the flow of the vector field V_i for small values of $t\in\mathbb{C}$ in the sense of Notation 2.3. Consider for any $i=1,\dots,n$ a smooth function $h_i:C\cup C_q\to\mathbb{R}_+\subset\mathbb{C}$ vanishing on $C\cup\{q\}$; its values on the relative interior of C_q will be specified later. As in the proof of Lemma 4.2, set $\zeta=(\zeta_1,\dots,\zeta_n)\in\mathbb{C}^n$ and consider the map

$$\psi(\zeta,x)=\phi^1_{\zeta_1 h_1(x)}\circ\cdots\circ\phi^n_{\zeta_n h_n(x)}(f(x))\in\mathfrak{S},\quad x\in C\cup C_q,$$

which is holomorphic in $\zeta\in\mathbb{C}^n$. Note that $\psi(0,\cdot)=f:M\to\mathfrak{S}_*$ (hence $\psi(\zeta,\cdot)$ does not vanish for ζ in a small neighborhood of the origin) and $\psi(\zeta,x)=f(x)$ for all $x\in C\cup\{q\}$. It follows that

$$\left.\frac{\partial\psi(\zeta,x)}{\partial\zeta_i}\right|_{\zeta=0}=h_i(x)V_i(f(x)),\quad i=1,\dots,n.$$

We choose h_i with support on a small compact neighborhood of $t_i\in(0,1)$ in such a way that

$$\int_0^1 h_i(\gamma(t))V_i(f(\gamma(t)))\theta(\gamma(t),\dot\gamma(t))\,dt\approx V_i(z_i)\,\theta(\gamma(t_i),\dot\gamma(t_i)).\tag{5-1}$$

Assuming that the neighborhoods are sufficiently small, the approximation in (5-1) is close enough so that, since the vectors on the right side above form a basis of $\mathbb{C}^n$, the ones in the left side also do.

Fix a number $\epsilon>0$. Theorem 2.6 furnishes holomorphic functions $g_i:M\to\mathbb{C}$ such that

$$g_i\text{ has a zero of multiplicity }k-1\text{ at all points of }\Lambda\tag{5-2}$$

and

$$\sup_{C\cup C_q}|g_i-h_i|<\epsilon,\quad i=1,\dots,n.$$

Following the arguments in the proof of Lemma 4.2, we define holomorphic maps

$$\Psi(\zeta,x,z)=\phi^1_{\zeta_1 g_1(x)}\circ\cdots\circ\phi^n_{\zeta_n g_n(x)}(z)\in\mathfrak{S},$$
$$\Psi_f(\zeta,x)=\Psi(\zeta,x,f(x))\in\mathfrak{S}_*,$$

where $x\in M$, $z\in\mathfrak{S}$, and ζ belongs to a sufficiently small neighborhood of the origin in $\mathbb{C}^n$. Observe that $\Psi_f(0,\cdot)=f$. In view of (5-1), if $\epsilon>0$ is small enough then we have that the vectors

$$\left.\frac{\partial}{\partial\zeta_i}\right|_{\zeta=0}\int_0^1\Psi_f(\zeta,\gamma(t))\theta(\gamma(t),\dot\gamma(t))\,dt=\int_0^1 g_i(\gamma(t))V_i(f(\gamma(t)))\theta(\gamma(t),\dot\gamma(t))\,dt,\tag{5-3}$$

$i=1,\dots,n$, are close enough to $V_i(z_i)\theta(\gamma(t_i),\dot\gamma(t_i))$ so that they also form a basis of $\mathbb{C}^n$.

To finish the proof it remains to perturb Ψ_f in order to solve the period problem and ensure the jet-interpolation at the points of Λ. From the Taylor expansion of the flow of a vector field it follows that

$$\Psi_f(\zeta,x)=f(x)+\sum_{i=1}^n\zeta_i g_i(x)V_i(f(x))+O(|\zeta|^2).$$

Since $|g_i| < \epsilon$ on C (recall that $h_i = 0$ on C), the integral of Ψ_f over the curves $C_1, \dots, C_l$ can be estimated by

$$\left|\int_{C_j} (\Psi_f(\zeta,\cdot) - f)\theta\right| = \left|\int_{C_j} \Psi_f(\zeta,\cdot)\theta\right| \le \eta_0\epsilon|\zeta|, \quad j = 1, \dots, l'' \tag{5-4}$$

(recall that $\int_{C_j} f\theta = \int_{C_j} dF = 0$ for all $j = 1, \dots, l''$, since these curves are closed),

$$\left|\int_{C_j} (\Psi_f(\zeta,\cdot) - f)\theta\right| = \left|F(p_0) - \left(F(p_j) - \int_{C_j} \Psi_f(\zeta,\cdot)\theta\right)\right| \le \eta_0\epsilon|\zeta|, \quad j = l''+1, \dots, l, \tag{5-5}$$

for some constant $\eta_0 > 0$ and sufficiently small $\zeta \in \mathbb{C}^n$. Furthermore, (5-2) guarantees that

$$\Psi_f(\zeta,\cdot) - f \text{ has a zero of multiplicity } k-1 \text{ at all points of } \Lambda \tag{5-6}$$

for ζ in a small neighborhood of the origin (see Lemma 2.2).

Now, Lemma 4.2(i) furnishes holomorphic maps $\Phi(\tilde{\zeta}, x, z)$ and $\Phi_f(\tilde{\zeta}, x) = \Phi(\tilde{\zeta}, x, f(x))$ with the parameter $\tilde{\zeta}$ in a small neighborhood of $0 \in \mathbb{C}^{\widetilde{N}}$ for some large $\widetilde{N} \in \mathbb{N}$ and $x \in M$ such that $\Phi(0, x, z) = z \in \mathfrak{S}$ for all $x \in M$ and

$$\Phi_f(0,\cdot) = \Phi_{\Psi_f(0,\cdot)}(0,\cdot) = f, \tag{5-7}$$

and the differential of the associated period map $\tilde{\zeta} \mapsto \mathcal{P}(\Phi_f(\tilde{\zeta},\cdot)) \in \mathbb{C}^{ln}$, see (4-1), at the point $\tilde{\zeta} = 0$ has maximal rank equal to ln. The same is true if we allow that f vary locally near the given initial map. Thus, replacing f by $\Psi_f(\zeta,\cdot)$ and considering the map

$$\mathbb{C}^{\widetilde{N}} \times \mathbb{C}^n \times M \ni (\tilde{\zeta}, \zeta, x) \mapsto \Phi(\tilde{\zeta}, x, \Psi_f(\zeta, x)) \in \mathfrak{S}_*$$

defined for $x \in M$ and $(\tilde{\zeta}, \zeta)$ in some sufficiently small neighborhood of $0 \in \mathbb{C}^{\widetilde{N}} \times \mathbb{C}^n$, the implicit function theorem provides a holomorphic map $\tilde{\zeta} = \rho(\zeta)$ near $\zeta = 0 \in \mathbb{C}^n$ with $\rho(0) = 0 \in \mathbb{C}^{\widetilde{N}}$ such that the map defined by $\Phi(\rho(\zeta), x, \Psi_f(\zeta, x))$ satisfies:

(i) $\mathcal{P}(\Phi(\rho(\zeta),\cdot,\Psi_f(\zeta,\cdot))) = \mathcal{P}(\Phi(0,\cdot,\Psi_f(0,\cdot))) = \mathcal{P}(\Psi_f(0,\cdot)) = \mathcal{P}(f)$.

(ii) $\Phi(\rho(\zeta),\cdot,\Psi_f(\zeta,\cdot)) - \Psi_f(\zeta,\cdot)$ has a zero of multiplicity $k-1$ at all points of Λ.

Condition (ii) together with (5-6) ensure that

$$\Phi(\rho(\zeta),\cdot,\Psi_f(\zeta,\cdot)) - f \text{ has a zero of multiplicity } k-1 \text{ at all points of } \Lambda \tag{5-8}$$

for all ζ in a small neighborhood of $0 \in \mathbb{C}^n$. Obviously the map $\rho = (\rho_1, \dots, \rho_n)$ also depends on f. It follows that the integral

$$H_F(\zeta, x) = F(p_0) + \int_{p_0}^{x} \Phi(\rho(\zeta),\cdot,\Psi_f(\zeta,\cdot))\theta \tag{5-9}$$

is independent of the choice of the arc from p_0 to $x \in M$. Moreover,

$$H_F(0,\cdot) = F, \tag{5-10}$$

see (5-7), and $H_F(\zeta,\cdot)$ is an $\mathfrak{S}$-immersion of class $\mathscr{A}^1(M)$ for every $\zeta\in\mathbb{C}^n$ sufficiently close to zero such that

$$H_F(\zeta,\cdot)=F \quad \text{on } \Lambda; \tag{5-11}$$

see (i). In addition, from (5-4) and (5-5) we have

$$|\rho(\zeta)|<\eta_1\epsilon|\zeta|$$

for some $\eta_1>0$. If we call $\widetilde{V}_j$ the vector fields and $\tilde{g}_j$ the functions involved in the construction of the map Φ (see Lemma 4.2), the above estimate gives

$$|\Phi(\rho(\zeta),x,\Psi_f(\zeta,x))-\Psi_f(\zeta,x)|=\Big|\sum\rho_j(\zeta)\tilde{g}_j(x)\widetilde{V}_j(\Psi_f(\zeta,x))+O(|\zeta|^2)\Big|<\eta_2\epsilon|\zeta|$$

for some $\eta_2>0$ and all $x\in M$ and all ζ near the origin in $\mathbb{C}^n$. Clearly, applying this estimate to the arc C_q we have

$$\left|\int_0^1\Phi(\rho(\zeta),\gamma(t),\Psi_f(\zeta,\gamma(t)))\theta(\gamma(t),\dot{\gamma}(t))-\int_0^1\Psi_f(\zeta,\gamma(t))\theta(\gamma(t),\dot{\gamma}(t))\right|<\eta_3\epsilon|\zeta|$$

for some $\eta_3>0$. Finally, choosing $\epsilon>0$ small enough, the derivatives

$$\frac{\partial}{\partial\zeta_i}\bigg|_{\zeta=0}\int_0^1\Phi(\rho(\zeta),\gamma(t),\Psi_f(\zeta,\gamma(t)))\,\theta(\gamma(t),\dot{\gamma}(t))\in\mathbb{C}^n,\quad i=1,\dots,n,$$

are so close to the vectors (5-3) that they also form a basis of $\mathbb{C}^n$. From the definition of H_F, (5-9), and (5-11), we have

$$\begin{aligned}\delta H_F(\zeta,p,q)&=H_F(\zeta,q)-H_F(\zeta,p)\\&=H_F(\zeta,q)-F(p)\\&=F(p_0)-F(p)+\int_0^1\Phi(\rho(\zeta),\gamma(t),\Psi_f(\zeta,\gamma(t)))\theta(\gamma(t),\dot{\gamma}(t)),\end{aligned}$$

and hence the partial differential

$$\frac{\partial}{\partial\zeta}\bigg|_{\zeta=0}\delta H(\zeta,p,q):\mathbb{C}^n\to\mathbb{C}^n$$

is an isomorphism. This, (5-10), (ii), and (5-11) show that H satisfies the conclusion of the lemma.

Case 2: *assume that* $\{p,q\}\cap\Lambda=\varnothing$. In this case setting $\Lambda':=\Lambda\cup\{p\}$ reduces the proof to Case 1. This proves the lemma. □

The family H_F depending on F given in (5-9) is holomorphically dependent also on F on a neighborhood of a given initial $\mathfrak{S}$-immersion F_0. In particular, if $F(\xi,\cdot):M\to\mathbb{C}^n$ is a family of holomorphic $\mathfrak{S}$-immersions depending holomorphically on $\xi\in\mathbb{C}$ such that $F(\xi,\cdot)-F$ has a zero of multiplicity k at all points $p\in\Lambda$ for any ξ, then $H_{F(\xi,\cdot)}(\zeta,\cdot)$ depends holomorphically on (ζ,ξ). This allows us to compose any finite number of such deformation families by an inductive process. For the case of two families suppose that $H=H_F(\zeta,\cdot)$ and $G=G_F(\xi,\cdot)$ are deformation families with $H_F(0,\cdot)=G_F(0,\cdot)=F$

and such that $H_F(\zeta,\cdot)-F$ and $G_F(\xi,\cdot)-F$ have a zero of multiplicity $k\in\mathbb{N}$ at all points of Λ for all ζ and ξ respectively. Then, we define the composed deformation family by

$$(H\,\#\,G)_F(\zeta,\xi,x):=G_{H_F(\zeta,\cdot)}(\xi,x),\quad x\in M.$$

Obviously, we have $(H\,\#\,G)_F(0,\xi,\cdot)=G_F(\xi,\cdot)$ and $(H\,\#\,G)_F(\zeta,0,\cdot)=H_F(\zeta,\cdot)$, and $H\,\#\,G-F$ has a zero of multiplicity k at p for all $p\in\Lambda$. The operation # is associative but not commutative.

To finish the proof of Theorem 5.3, Lemma 5.4 gives a finite open covering $\{U_i\}_{i=1}^m$ of the compact set $M\times M\setminus U$ and deformation families $H^i=H^i(\zeta^i,\cdot):M\to\mathbb{C}^n$, with $H^i(0,\cdot)=F$, where $\zeta^i=(\zeta_1^i,\dots,\zeta_{\eta_i}^i)\in\Omega_i\subset\mathbb{C}^{\eta_i}$ for positive integers $\eta_i\in\mathbb{N}$ and $i=1,\dots,m$. It follows that the difference map $\delta H^i(\zeta^i,p,q)$ is submersive at $\zeta^i=0$ for all $p,q\in U_i$. By taking $\zeta=(\zeta^1,\dots,\zeta^m)\in\mathbb{C}^N$, with $N=\sum_{i=1}^m\eta_i$, and setting

$$H(\zeta,x):=(H^1\,\#\,H^2\,\#\cdots\#\,H^m)(\zeta^1,\dots,\zeta^m,x),$$

we obtain a deformation family such that $H(0,\cdot)=F$, $H(q,\cdot)-F$ has a zero of multiplicity k at p for all $p\in\Lambda$, and δH is submersive everywhere on $M\times M\setminus U$ for all $\zeta\in\mathbb{C}^N$ sufficiently close to the origin. $\square$

5B. ***A completeness lemma.*** We develop an intrinsic-extrinsic version of the arguments from [Jorge and Xavier 1980] in order to prove the following

Lemma 5.5. *Let $\mathfrak{S}\subset\mathbb{C}^n$ $(n\geq 3)$ be as in Lemma 4.5. Let M be a compact bordered Riemann surface and $K\subset\mathring{M}$ be a smoothly bounded compact domain which is Runge and a strong deformation retract of M. Also let $\Lambda\subset\mathring{K}$ be a finite subset and $p_0\in\mathring{K}\setminus\Lambda$ be a point. Then, given an integer $k\in\mathbb{N}$ and a positive number $\tau>0$, every $\mathfrak{S}$-immersion $F:K\to\mathbb{C}^n$ of class $\mathscr{A}^1(K)$ may be approximated in the $\mathscr{C}^1(K)$-topology by $\mathfrak{S}$-immersions $\widetilde{F}:M\to\mathbb{C}^n$ of class $\mathscr{A}(M)$ satisfying the following conditions*:

(I) *$\widetilde{F}-F$ has a zero of multiplicity k at all points $p\in\Lambda$.*

(II) $\operatorname{dist}_{\widetilde{F}}(p_0,bM)>\tau$.

Proof. Without loss of generality we assume that M is a smoothly bounded compact domain in an open Riemann surface $\widetilde{M}$. By Proposition 5.2 we may assume that F is holomorphic on M and that its first component has no critical points on M. Fix a holomorphic 1-form θ vanishing nowhere on $\widetilde{M}$ and set $dF=f\theta$, where $f=(f_1,\dots,f_n):M\to\mathfrak{S}_*$ is a holomorphic map.

Since K is a strong deformation retract of M, we know $\mathring{M}\setminus K$ consists of a finite family of pairwise-disjoint open annuli. Thus, there exists a finite family of pairwise-disjoint, smoothly bounded, compact disks $L_1,\dots,L_m$ in $\mathring{M}\setminus K$ satisfying the following property: if $\alpha\subset M\setminus\bigcup_{j=1}^m L_j$ is a arc connecting p_0 with bM then

$$\int_\alpha|f_1\theta|>\tau. \tag{5-12}$$

Recall that $f_1\neq 0$. (Such disks can be found as pieces of labyrinths of Jorge–Xavier-type [1980] on the annuli forming $\mathring{M}\setminus K$; see [Alarcón, Fernández and López 2013, proof of Lemma 4.1] for a detailed explanation). Set $L:=\bigcup_{j=1}^m L_j$.

For each $j = 1, \dots, m$, choose a Jordan arc $\gamma_j \subset \mathring{M}$ with an endpoint in K, the other endpoint in L_j, and otherwise disjoint from $K \cup L$, such that $\gamma_i \cap \gamma_j = \varnothing$ for all $i \neq j \in \{1, \dots, m\}$ and the set

$$S := K \cup L \cup \Gamma,$$

where $\Gamma := \bigcup_{j=1}^m \gamma_j$, is an admissible subset of M. It follows that S is a connected very simple admissible subset of M with kernel component K (see Definition 4.1) such that K is a deformation retract of S (hence of M). Take a map $h = (h_1, \dots, h_n) : S \to \mathfrak{S}_*$ of class $\mathscr{A}(S)$ satisfying the following conditions:

(a) $h_1 = f_1|_S$.

(b) $h|_K = f|_K$.

(c) $\left|\int_\alpha h\theta\right| > \tau$ for all arcs $\alpha \subset S$ with initial point p_0 and final point in L.

Existence of such a map is clear; we may for instance choose h close to $0 \in \mathbb{C}^n$ on each component of L and such that $\left|\int_{\gamma_j} h\theta\right|$ is very large for every component γ_j of Γ. Also choose for each $p \in \Lambda$ a smooth Jordan arc $C_p \subset \mathring{K}$ with initial point p_0 and final one p, and assume that $C_p \cap C_q = \{p_0\}$ for all $p \neq q \in \Lambda$. Then, Lemma 4.5 provides a holomorphic map $\tilde{f} = (\tilde{f}_1, \tilde{f}_2, \dots, \tilde{f}_n) : M \to \mathfrak{S}_*$ such that:

(i) $\tilde{f}$ approximates h on S.

(ii) $\tilde{f}_1 = f_1$ everywhere on M.

(iii) $\tilde{f} - h$ has a zero of multiplicity $k - 1$ at p for all $p \in \Lambda$.

(iv) $\int_{C_p} (\tilde{f} - h)\theta = 0$ for all $p \in \Lambda$.

(v) $\int_\gamma (\tilde{f} - h)\theta = 0$ for all closed curves $\gamma \subset S$.

Since $f\theta = dF$ is exact, properties (b) and (v) and the fact that K is a strong deformation retract of M guarantee that $\tilde{f}\theta$ is exact on M as well. Therefore, the map $\widetilde{F} = (\widetilde{F}_1, \widetilde{F}_2, \dots, \widetilde{F}_n) : M \to \mathbb{C}^n$ defined by

$$\widetilde{F}(p) := F(p_0) + \int_{p_0}^p \tilde{f}\theta, \quad p \in M,$$

is well-defined and an $\mathfrak{S}$-immersion of class $\mathscr{A}^1(M)$. We claim that if the approximation in (i) is close enough then $\widetilde{F}$ satisfies the conclusion of the lemma. Indeed, properties (i) and (b) guarantee that $\widetilde{F}$ approximates F as close as desired in the $\mathscr{C}^1(K)$-topology. On the other hand, (iii), (iv), and (b) ensure that $\widetilde{F} - F$ has a zero of multiplicity k at all points of Λ, which proves (I). Finally, in order to check condition (II), let $\alpha \subset M$ be an arc with initial point p_0 and final one in bM. Assume first that $\alpha \cap L \neq \varnothing$ and let $\tilde{\alpha} \subset \alpha$ be a subarc with initial point p_0 and final point q for some $q \in L$. Then we have

$$\operatorname{length}(\widetilde{F}(\alpha)) > \operatorname{length}(\widetilde{F}(\tilde{\alpha})) \geq |\widetilde{F}(q) - \widetilde{F}(p_0)| = \left|\int_{p_0}^q \tilde{f}\theta\right| \overset{\text{(i)}}{\approx} \left|\int_{p_0}^q h\theta\right| \overset{\text{(c)}}{>} \tau.$$

Assume that, on the contrary, $\alpha \cap L = \varnothing$. In this case,

$$\operatorname{length}(\widetilde{F}(\alpha)) = \int_\alpha |\tilde{f}\theta| \geq \int_\alpha |\tilde{f}_1\theta| \overset{\text{(ii)}}{=} \int_\alpha |f_1\theta| \overset{\text{(5-12)}}{>} \tau.$$

This proves (II) and completes the proof. $\square$

5C. ***A properness lemma.*** Recall that given a vector $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$ or $\mathbb{C}^n$ we define $|x|_\infty = \max\{|x_1|, \dots, |x_n|\}$; see Section 2 for notation.

Lemma 5.6. *Let $n \geq 3$ be an integer and $\mathfrak{S}$ be an irreducible closed conical complex subvariety of $\mathbb{C}^n$ which is not contained in any hyperplane. Assume that $\mathfrak{S}_* = \mathfrak{S} \setminus \{0\}$ is smooth and an Oka manifold, and that $\mathfrak{S} \cap \{z_j = 1\}$ is an Oka manifold and the coordinate projection $\pi_j : \mathfrak{S} \to \mathbb{C}$ onto the z_j-axis admits a local holomorphic section h_j near $z_j = 0$ with $h_j(0) \neq 0$ for all $j = 1, \dots, n$. Let M be a compact bordered Riemann surface and $K \subset \mathring{M}$ be a smoothly bounded compact domain which is Runge and a strong deformation retract of M. Also let $\Lambda \subset \mathring{K}$ be a finite subset, $F : K \to \mathbb{C}^n$ be an $\mathfrak{S}$-immersion of class $\mathscr{A}^1(K)$, let $\tau > \rho > 0$ be numbers, and assume that*

$$|F(p)|_\infty > \rho \quad \text{for all } p \in bK. \tag{5-13}$$

Then, given an integer $k \in \mathbb{N}$, F may be approximated in the $\mathscr{C}^1(K)$-topology by $\mathfrak{S}$-immersions $\widetilde{F} : M \to \mathbb{C}^n$ of class $\mathscr{A}^1(M)$ satisfying the following conditions:

(I) *$\widetilde{F} - F$ has a zero of multiplicity k at all points $p \in \Lambda$.*

(II) *$|\widetilde{F}(p)|_\infty > \rho$ for all $p \in M \setminus \mathring{K}$.*

(III) *$|\widetilde{F}(p)|_\infty > \tau$ for all $p \in bM$.*

Proof. Without loss of generality we assume that M is a smoothly bounded compact domain in an open Riemann surface $\widetilde{M}$. By Proposition 5.2 we may assume that $F = (F_1, \dots, F_n)$ is holomorphic on $\widetilde{M}$. Since K is a strong deformation retract of M, we have that $M \setminus \mathring{K}$ consists of finitely many pairwise-disjoint compact annuli. For simplicity of exposition we assume that $\mathcal{A} := M \setminus \mathring{K}$ is connected (and hence a single annulus); the same proof applies in general by working separately on each connected component of $M \setminus \mathring{K}$. We denote by α the boundary component of $\mathcal{A}$ contained in bK and by β the one contained in bM; both α and β are smooth Jordan curves.

From inequality (5-13) there exist an integer $l \geq 3$, subsets $I_1, \dots, I_n$ of $\mathbb{Z}_l$ (where $\mathbb{Z}_l = \{0, 1, \dots, l-1\}$ denotes the additive cyclic group of integers modulus l), and a family of compact connected subarcs $\{\alpha_j : j \in \mathbb{Z}_l\}$ of bK, satisfying the following properties:

(a1) $\bigcup_{j \in \mathbb{Z}_l} \alpha_j = \alpha$.

(a2) α_j and α_{j+1} have a common endpoint p_j and are otherwise disjoint.

(a3) $\bigcup_{a=1}^n I_a = \mathbb{Z}_l$ and $I_a \cap I_b = \varnothing$ for all $a \neq b \in \{1, \dots, n\}$.

(a4) If $j \in I_a$ then $|F_a(p)| > \rho$ for all $p \in \alpha_j$, $a = 1, \dots, n$.

(Possibly $I_a = \varnothing$ for some $a \in \{1, \dots, n\}$.)

Consider for each $j \in \mathbb{Z}_l$ a smooth embedded arc $\gamma_j \subset \mathcal{A}$ with the following properties:

- γ_j joins $\alpha \subset bK$ with $\beta \subset bM$ and intersects them transversely.
- $\gamma_j \cap \alpha = \{p_j\}$.
- $\gamma_j \cap \beta$ consists of a single point, namely, q_j.
- The arcs γ_j, $j \in \mathbb{Z}_l$, are pairwise disjoint.

Consider the admissible set

$$S := K \cup \Big(\bigcup_{j\in\mathbb{Z}_l} \gamma_j\Big) \subset M$$

and fix a point $x_0 \in \mathring{K} \setminus \Lambda$. Let $z = (z_1, \dots, z_n)$ be the coordinates on $\mathbb{C}^n$ and recall that $\pi_a : \mathbb{C}^n \to \mathbb{C}$ is the a-th coordinate projection $\pi_a(z) = z_a$ for all $a = 1, \dots, n$. Let θ be a holomorphic 1-form vanishing nowhere on $\widetilde{M}$, and let $f : S \to \mathfrak{S}_*$ be a map of class $\mathscr{A}(S)$ such that $f = dF/\theta$ on K and the map $\widetilde{G} : S \to \mathbb{C}^n$ given by

$$\widetilde{G}(p) = F(p_0) + \int_{x_0}^{p} f\theta,$$

which is well-defined since K is a deformation retract of S, satisfies the following conditions:

(b1) $\widetilde{G} = F$ on K and on a neighborhood of p_j for all $j \in \mathbb{Z}_l$.

(b2) If $j \in I_a$ then $|\pi_a(\widetilde{G}(z))| > \rho$ for all $z \in \gamma_{j-1} \cup \gamma_j$, $a = 1, \dots, n$.

(b3) If $j \in I_a$ then $|\pi_a(\widetilde{G}(q_{j-1}))| > \tau$ and $|\pi_a(\widetilde{G}(q_j))| > \tau$, $a = 1, \dots, n$.

Existence of such an f is guaranteed by (a4). Theorem 4.4 provides a map $g : M \to \mathfrak{S}_*$ of class $\mathscr{A}(M)$ such that $g\theta$ is exact on M, and the $\mathfrak{S}$-immersion $G = (G_1, \dots, G_n) : M \to \mathbb{C}^n$ of class $\mathscr{A}^1(M)$ given by $G(p) = F(x_0) + \int_{x_0}^{p} g\theta$ enjoys the following properties:

(c1) G approximates $\widetilde{G}$ in the $\mathscr{C}^1(K)$-topology.

(c2) $G - \widetilde{G}$ has a zero of multiplicity k at all points of Λ.

(c3) If $j \in I_a$ then $|G_a(p)| > \rho$ for all $p \in \gamma_{j-1} \cup \alpha_j \cup \gamma_j$, $a = 1, \dots, n$.

(c4) If $j \in I_a$ then $|G_a(p)| > \tau$ for $p \in \{q_{j-1}, q_j\}$, $a = 1, \dots, n$.

Property (c3) follows from (a4) and (b2), whereas (c4) follows from (b3), provided that the approximation of f by g is close enough.

For each $j \in \mathbb{Z}_l$ let $\beta_j \subset \beta$ denote the subarc of β with endpoints q_{j-1} and q_j which does not contain q_i for any $i \in \mathbb{Z}_l \setminus \{j-1, j\}$. It is clear that

$$\bigcup_{j\in\mathbb{Z}_l} \beta_j = \beta. \tag{5-14}$$

Also denote by $D_j \subset \mathcal{A}$ the closed disk bounded by the arcs γ_{j-1}, α_j, γ_j, and β_j; see Figure 4. It follows that

$$\mathcal{A} = \bigcup_{j\in\mathbb{Z}_l} D_j. \tag{5-15}$$

Call $H_0 := G = (H_{0,1}, \dots, H_{0,n})$ and $I_0 := \varnothing$. We shall construct a sequence of $\mathfrak{S}$-immersions $H_b = (H_{b,1}, \dots, H_{b,n}) : M \to \mathbb{C}^n$, $b = 1, \dots, n$, of class $\mathscr{A}^1(M)$ satisfying the following requirements for all $b \in \{1, \dots, n\}$:

(d1$_b$) H_b approximates H_{b-1} in the $\mathscr{C}^1$-topology on $M \setminus \big(\bigcup_{j\in I_b} \mathring{D}_j\big)$.

(d2$_b$) $H_b - H_{b-1}$ has a zero of multiplicity k at all points of Λ.

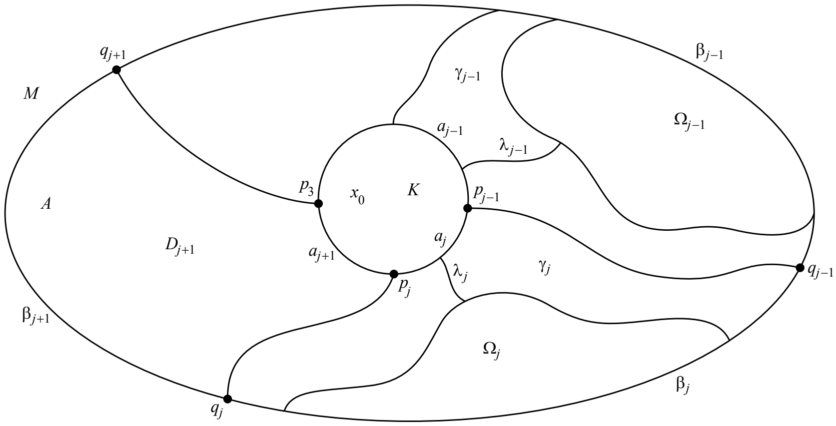

Figure 4. The annulus $\mathcal{A}$.

(d3_b) If $j \in \bigcup_{i=1}^{b} I_i$ then $|H_b(p)|_\infty > \rho$ for all $p \in D_j$.

(d4_b) If $j \in \bigcup_{i=1}^{b} I_i$ then $|H_b(p)|_\infty > \tau$ for all $p \in \beta_j$.

(d5_b) If $j \in I_a$ then $|H_{b,a}(p)| > \rho$ for all $p \in \gamma_{j-1} \cup \alpha_j \cup \gamma_j$, $a = 1, \ldots, n$.

(d6_b) If $j \in I_a$ then $|H_{b,a}(p)| > \tau$ for $p \in \{q_{j-1}, q_j\}$, $a = 1, \ldots, n$.

We claim that the $\mathfrak{S}$-immersion $\widetilde{F} := H_n : M \to \mathbb{C}^n$ satisfies the conclusion of the lemma. Indeed, $\widetilde{F}$ approximates F in the $\mathscr{C}^1(K)$-topology by properties (b1), (c1) and (d1_1)–(d1_n); condition (I) is guaranteed by (d2), (c2), and (b1); condition (II) by (d3_n), (a3), and (5-15); and condition (III) by (d4_n), (a3), and (5-14). So, to conclude the proof it suffices to construct the sequence $H_1, \ldots, H_n$ satisfying the above properties. We proceed by induction. Assume that we already have $H_0, \ldots, H_{b-1}$ for some $b \in \{1, \ldots, n\}$ with the desired properties and let us construct H_b. Notice that (d5_0) = (c3) and (d6_0) = (c4) formally hold. By the continuity of H_{b-1} and conditions (d5_{b-1}) and (d6_{b-1}), for each $j \in I_b$ there exists a closed disk $\Omega_j \subset D_j \setminus (\gamma_{j-1} \cup \alpha_j \cup \gamma_j)$ such that the following hold:

(i) $\Omega_j \cap \beta_j$ is a compact connected Jordan arc.

(ii) $|H_{b-1,b}(p)| > \rho$ for all $p \in \Upsilon_j := \overline{D_j \setminus \Omega_j}$.

(iii) $|H_{b-1,b}(p)| > \tau$ for all $p \in \overline{\beta_j \setminus \Omega_j}$.

Next, for each $j \in I_b$ choose a smooth embedded arc $\lambda_j \subset \Upsilon_j \setminus (\gamma_{j-1} \cup \gamma_j)$ with an endpoint in α_j and the other one in Ω_j and otherwise disjoint from $b\Upsilon_j$ (see Figure 4). Moreover, choose each λ_j so that the set

$$S_b := \Big(M \setminus \bigcup_{j \in I_b} \mathring{\Upsilon}_j\Big) \cup \Big(\bigcup_{j \in I_b} \lambda_j\Big)$$

is admissible. Notice that S_b is connected and very simple in the sense of Definition 4.1.

Set $h = (h_1, \ldots, h_n) = dH_{b-1}/\theta$ and let $\tilde{h} = (\tilde{h}_1, \ldots, \tilde{h}_n) : S_b \to \mathfrak{S}_*$ be a map of class $\mathscr{A}(S_b)$ such that:

(iv) $\tilde{h} = h$ on $M \setminus \big(\bigcup_{j \in I_b} \mathring{D}_j\big)$.

(v) $\tilde{h}_b = h_b$ on S_b.

(vi) The map $\widetilde{H} : S_b \to \mathbb{C}^n$ given by

$$\widetilde{H}(p) = H_{b-1}(x_0) + \int_{x_0}^{p} \tilde{h}\theta, \quad p \in S_b,$$

satisfies $|\widetilde{H}(p)|_\infty > \tau$ for all $p \in \bigcup_{j \in I_b} \Omega_j$.

To construct such a map $\tilde{h}$ we may for instance choose $\tilde{h} = h$ on $M \setminus \bigcup_{j \in I_b} \mathring{\Upsilon}_j$ and suitably define it on $\bigcup_{j \in I_b} \lambda_j$. Now, Lemma 4.5 furnishes a map $\phi : M \to \mathfrak{S}_*$ of class $\mathscr{A}(M)$ such that $\phi\theta$ is exact on M (take into account that $\tilde{h}\theta = h\theta = dH_{b-1}$ on K and that K is a deformation retract of M) and the map $H_b : M \to \mathbb{C}^n$ given by

$$H_b(p) = H_{b-1}(x_0) + \int_{x_0}^{p} \phi\theta, \quad p \in M,$$

is an $\mathfrak{S}$-immersion of class $\mathscr{A}^1(M)$ enjoying the following properties:

(vii) H_b is as close as desired to $\widetilde{H}$ in the $\mathscr{C}^1$-topology on $M \setminus \big(\bigcup_{j \in I_b} \mathring{D}_j\big)$.

(viii) $H_{b,b} = H_{b-1,b}$ (take into account (v)).

(ix) $H_b - \widetilde{H}$ has a zero of multiplicity k at all points of Λ.

Since $\widetilde{H} = H_{b-1}$ on $M \setminus \big(\bigcup_{j \in I_b} \mathring{D}_j\big) \supset \Lambda \cup \big(\bigcup_{j \in \mathbb{Z}_l} \gamma_{j-1} \cup \alpha_j \cup \gamma_j\big)$, we have (d1$_b$) = (vii), (d2$_b$) = (ix), and, if the approximation in (vii) is close enough, (d5$_b$) and (d6$_b$) follow from (d5$_{b-1}$) and (d6$_{b-1}$), respectively.

Pick $j \in \bigcup_{i=1}^{b} I_i$ and $p \in D_j$. If $j \notin I_b$ then (d3$_{b-1}$) and (vii) ensure that $|H_b(p)|_\infty > \rho$. On the other hand, if $j \in I_b$ then (ii) and (viii) ensure that $|H_b(p)| \geq |H_{b,b}(p)| > \rho$ provided that $p \in \Upsilon_j$, whereas (vi) and (vii) guarantee that $|H_b(p)| > \tau > \rho$ provided that $p \in \Omega_j$. This proves (d3$_b$).

Finally, choose $j \in \bigcup_{i=1}^{b} I_i$ and $p \in \beta_j$. As above, if $j \notin I_b$ then (d4$_{b-1}$) and (vii) give that $|H_b(p)|_\infty > \tau$. Likewise, if $j \in I_b$ then (iii) and (viii) ensure that $|H_b(p)| \geq |H_{b,b}(p)| > \tau$ provided that $p \in \beta_j \setminus \Omega_j$, whereas (vi) and (vii) imply that $|H_b(p)| > \tau$ provided that $p \in \beta_j \cap \Omega_j$. This proves (d4$_b$) and concludes the proof of the lemma. □

6. Proof of Theorem 1.3

As in the proof of Theorem 4.4 we can assume that $\Lambda \cap bK = \varnothing$ and also that for each $p \in \Lambda$ we have either $\Omega_p \subset \mathring{K}$ or $\Omega_p \cap K = \varnothing$.

Set $M_0 := K$ and let $\{M_j\}_{j \in \mathbb{N}}$ be an exhaustion of M by smoothly bounded Runge compact domains in M such that:

- $M_0 \Subset M_1 \Subset \cdots \Subset \bigcup_{j \in \mathbb{N}} M_j = M$.
- $\chi(M_j \setminus \mathring{M}_{j-1}) \in \{-1, 0\}$ for all $j \in \mathbb{N}$.
- $bM_j \cap \Lambda = \varnothing$ for all $j \in \mathbb{N}$ and so, up to shrinking the sets Ω_p if necessary, we may assume that $\Omega_p \subset \mathring{M}_j$ or $\Omega_p \cap M_j = \varnothing$ for all $p \in \Lambda$.

The existence of such a sequence is guaranteed as in the proof of Theorem 4.4. The set $\Lambda_j := \Lambda \cap M_j = \Lambda \cap \mathring{M}_j$, $j \in \mathbb{Z}_+$, is empty or finite; without loss of generality we may assume that $\Lambda_0 \neq \varnothing$ and $\Lambda_j \setminus \Lambda_{j-1} \neq \varnothing$ for all $j \in \mathbb{N}$, and hence Λ is infinite. Observe that $\Lambda_{j-1} \subsetneq \Lambda_j$ for all $j \in \mathbb{N}$. Fix a sequence $\{\epsilon_j\}_{j\in\mathbb{N}} \searrow 0$ which will be specified later, set

$$F_0 := F|_{M_0} : M_0 \to \mathbb{C}^n$$

and, by Proposition 4.3 and Theorem 5.3, assume without loss of generality that F_0 is nonflat and, if $F|_\Lambda$ is injective, an embedding.

6A. *Proof of the first part of the theorem.* For the first part of the theorem we shall construct a sequence $\{F_j\}_{j\in\mathbb{N}}$ of nonflat $\mathfrak{S}$-immersions $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$ satisfying:

(i_j) $\|F_j - F_{j-1}\|_{1,M_{j-1}} < \epsilon_j$.

(ii_j) $F_j - F$ has a zero of multiplicity $k \in \mathbb{N}$ at every point $p \in \Lambda_j$.

(iii_j) If $F|_\Lambda$ is injective then F_j is an embedding.

We proceed by induction. The basis is given by the $\mathfrak{S}$-immersion F_0, which clearly satisfies (ii_0) and (iii_0); condition (i_0) is vacuous. For the inductive step assume that we have an $\mathfrak{S}$-immersion $F_{j-1} : M_{j-1} \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_{j-1})$ satisfying (i_{j-1}), (ii_{j-1}), and (iii_{j-1}) for some $j \in \mathbb{N}$, and let us furnish $F_j : M_j \to \mathbb{C}^n$ enjoying the corresponding properties. We distinguish two different cases depending on the Euler characteristic of $M_j \setminus \mathring{M}_{j-1}$.

Noncritical case: *assume that* $\chi(M_j \setminus \mathring{M}_{j-1}) = 0$. It follows that M_{j-1} is a strong deformation retract of M_j, and then Proposition 5.2 applied to the data

$$S = M_{j-1} \cup \Bigl(\bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} \Omega_p \Bigr), \quad S_0 = M_{j-1}, \quad \Lambda = \Lambda_j, \quad k,$$

and the generalized $\mathfrak{S}$-immersion $S \to \mathbb{C}^n$ agreeing with F_{j-1} on M_{j-1} and with F on $\bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} \Omega_p$, provides an $\mathfrak{S}$-immersion $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$ that satisfies (i_j) and (ii_j). Finally, if $F|_\Lambda$ is injective then Theorem 5.3 enables us to choose F_j being an embedding; this ensures (iii_j).

Critical case: *assume that* $\chi(M_j \setminus \mathring{M}_{j-1}) = -1$. We then have that the change of topology is described by attaching to M_{j-1} a smooth arc α in $\mathring{M}_j \setminus \mathring{M}_{j-1}$ satisfying M_{j-1} only at its endpoints. Thus, $M_{j-1} \cup \alpha$ is a strong deformation retract of M_j. Further, we may choose α such that $\alpha \cap \Lambda = \varnothing$; and $S := M_{j-1} \cup \alpha$ is an admissible subset of M, which is clearly very simple (see Definition 4.1). We use Lemma 3.3 to extend F_{j-1} to S as a generalized $\mathfrak{S}$-immersion. By Proposition 5.2, we may approximate F_{j-1} in the $\mathscr{C}^1(M_{j-1} \cup \alpha)$-topology by nonflat $\mathfrak{S}$-immersions on a small compact tubular neighborhood $M_j' \Subset M_j$ of $M_{j-1} \cup \alpha$ having a contact of order k with F at all points of Λ_j. Since $\chi(M_j \setminus \mathring{M}_j') = 0$, this reduces the proof to the previous case and hence concludes the recursive construction of the sequence $\{F_j\}_{j\in\mathbb{N}}$.

Finally, if the number $\epsilon_j > 0$ is chosen sufficiently small at each step in the recursive construction, properties (i_j), (ii_j), and (iii_j) ensure that the sequence $\{F_j\}_{j\in\mathbb{N}}$ converges uniformly on compacta in M to an $\mathfrak{S}$-immersion

$$\widetilde{F} := \lim_{j\to\infty} F_j : M \to \mathbb{C}^n,$$

which is as close as desired to F uniformly on K, is injective if $F|_\Lambda$ is injective, and is such that $\widetilde{F} - F$ has a zero of multiplicity k at all points of Λ.

6B. ***Proof of assertion* (I).** Suppose that the assumptions in assertion (I) hold. Fix a point $p_0 \in \mathring{K} \setminus \Lambda$. We shall now construct a sequence of $\mathfrak{S}$-immersions $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$, $j \in \mathbb{N}$, satisfying conditions (i_j)–(iii_j) above and also:

(iv_j) $\mathrm{dist}_{F_j}(p_0, bM_j) > j$ for all $j \in \mathbb{N}$.

Observe that $F_0 = F|_{M_0}$ satisfies (iv_0) since it is an immersion and $p_0 \in \mathring{K}$. For the inductive step assume that we already have F_{j-1} satisfying (i_j)–(iv_j) for some $j \in \mathbb{N}$ and, reasoning as above, construct an $\mathfrak{S}$-immersion $F_j' : M_j \to \mathbb{C}^n$ satisfying (i_j), (ii_j), and (iii_j). Let $M_j' \subset \mathring{M}_j$ be a smoothly bounded compact domain which is Runge and a strong deformation retract of M_j and contains $M_{j-1} \cup \Lambda_j$ in its relative interior. Then, Lemma 5.5 applied to the data

$$M = M_j, \quad K = M_j', \quad \Lambda = \Lambda_j, \quad k, \quad \tau = j, \quad \text{and} \quad F = F_j'|_{M_j'},$$

gives an $\mathfrak{S}$-immersion $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$ satisfying (ii_j), (iv_j), and also (i_j) provided that the approximation of F_j' by F_j on M_j' is close enough; Theorem 5.3 enables us to assume that F_j also satisfies (iii_j). This closes the induction and concludes the construction of the sequence $\{F_j\}_{j\in\mathbb{N}}$ with the desired properties.

As above, if the number $\epsilon_j > 0$ is chosen sufficiently small at each step in the recursive construction, properties (i_j)–(iii_j) ensure that the sequence $\{F_j\}_{j\in\mathbb{N}}$ converges uniformly on compacta in M to an $\mathfrak{S}$-immersion $\widetilde{F} := \lim_{j\to\infty} F_j : M \to \mathbb{C}^n$ which is as close as desired to F uniformly on K, is injective if F is injective, and is such that $\widetilde{F} - F$ has a zero of multiplicity k at all points of Λ. In addition, property (iv_j) ensures that

$$\lim_{j\to\infty} \mathrm{dist}_{\widetilde{F}}(p_0, bM_j) = +\infty$$

whenever the number $\epsilon_j > 0$ is chosen small enough at each step in the recursive process. This implies that $\widetilde{F}$ is complete and concludes the proof of assertion (I).

6C. ***Proof of assertion* (II).** Suppose the assumptions in assertion (II) hold. Observe that $F|_\Lambda : \Lambda \to \mathbb{C}^n$ is a proper map if, and only if, $(F|_\Lambda)^{-1}(C)$ is finite for any compact set $C \subset \mathbb{C}^n$, or, equivalently, if either the closed discrete set Λ is finite or for some (and hence for any) ordering $\Lambda = \{p_1, p_2, p_3, \ldots\}$ of Λ, the sequence $\{F(p_1), F(p_2), F(p_3), \ldots\}$ is divergent in $\mathbb{C}^n$. Since we are assuming that Λ is infinite, there is $j_0 \in \mathbb{N}$ such that

$$F(p) \neq 0 \quad \text{for all } p \in \Lambda \setminus \Lambda_{j_0}. \tag{6-1}$$

In a first step we construct for each $j \in \{0, \ldots, j_0\}$ an $\mathfrak{S}$-immersion $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$ satisfying conditions (i_j)–(iii_j) above; we reason as in Section 6A. Now, up to a small deformation of M_{j_0}

if necessary, we may assume without loss of generality that F_{j_0} does not vanish anywhere on bM_{j_0}, and hence there exists $\rho_{j_0} > 0$ such that

$$|F_{j_0}(p)|_\infty > \rho_{j_0} > 0 \quad \text{for all } p \in bM_{j_0}. \tag{6-2}$$

Set

$$\rho_j := \min\{|F(p)|_\infty : p \in \Lambda_j \setminus \Lambda_{j-1}\} \quad \text{for all } j \geq j_0 + 1. \tag{6-3}$$

Recall that $\Lambda_j \setminus \Lambda_{j-1} \neq \varnothing$ for all $j \in \mathbb{N}$. In view of (6-1) and (6-2) we have $\rho_j > 0$ for all $j \geq j_0$. Moreover, since $F|_\Lambda$ is proper,

$$\lim_{j\to+\infty} \rho_j = +\infty. \tag{6-4}$$

In a second step, we shall construct a sequence of $\mathfrak{S}$-immersions $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$, for $j \geq j_0 + 1$, enjoying conditions (i_j)–(iii_j) and also:

(v_j.1) $|F_j(p)|_\infty > \frac{1}{2}\min\{\rho_{j-1}, \rho_j\}$ for all $p \in M_j \setminus \mathring{M}_{j-1}$.

(v_j.2) $|F_j(p)|_\infty > \rho_j$ for all $p \in bM_j$.

We proceed in an inductive way. The basis of the induction is accomplished by F_{j_0}; recall that it satisfies (i_{j_0})–(iii_{j_0}), whereas property (v_{j_0}.1) is vacuous and property (v_{j_0}.2) follows from (6-2). For the inductive step, assume that we already have $F_{j-1} : M_{j-1} \to \mathbb{C}^n$ for some $j \geq j_0 + 1$ satisfying (i_{j-1})–(iii_{j-1}), (v_{j-1}.1), and (v_{j-1}.2) and let us construct an $\mathfrak{S}$-immersion $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$ with the corresponding requirements.

By (6-3) and up to a shrinking of the set Ω_p if necessary, we may assume that

$$|F(q)|_\infty > \tfrac{1}{2}\rho_j \quad \text{for all points } q \text{ in } \Omega^j := \bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} \Omega_p \neq \varnothing. \tag{6-5}$$

Next, choose a smooth Jordan arc C_p for each $p \in \Lambda_j \setminus \Lambda_{j-1}$ with the initial point in bM_{j-1}, the final point in $b\Omega_p$, and otherwise disjoint from $M_{j-1} \cup \Omega^j$, and such that

$$S' := M_{j-1} \cup \Omega^j \cup \Bigg(\bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} C_p \Bigg)$$

is a very simple admissible subset of M_j; in particular $C_p \cap C_q = \varnothing$ if $p \neq q$. If $\chi(M_j \setminus \mathring{M}_{j-1}) = -1$ we then also choose another smooth Jordan arc $\alpha \subset \mathring{M}_j$ with its two endpoints in bM_{j-1} and otherwise disjoint from S' such that $S' \cup \alpha$ is admissible and a strong deformation retract of M_j. If $\chi(M_j \setminus \mathring{M}_{j-1}) = 0$ we set $\alpha := \varnothing$. In any case, the set

$$S := S' \cup \alpha \subset \mathring{M}_j$$

is admissible in M and a strong deformation retract of M_j. Set

$$C := \alpha \cup \Bigg(\bigcup_{p\in\Lambda_j\setminus\Lambda_{j-1}} C_p \Bigg),$$

and observe that $S = (M_{j-1} \cup \Omega^j) \cup C$. Consider a generalized $\mathfrak{S}$-immersion $\widetilde{F}_j : S \to \mathbb{C}^n$ of class $\mathscr{A}^1(S)$ such that:

(A.1) $\widetilde{F}_j|_{M_{j-1}} = F_{j-1}$.

(A.2) $\widetilde{F}_j|_{\Omega^j} = F|_{\Omega^j}$.

(A.3) $|\widetilde{F}_j(q)|_\infty > \frac{1}{2}\min\{\rho_{j-1}, \rho_j\}$ for all $q \in C$.

To ensure (A.3) we use Lemma 3.3; take into account (v_{j-1}.2) and (6-5). Thus, (v_{j-1}.2), (6-5), and (A.3) guarantee that:

(A.4) $|\widetilde{F}_j(p)|_\infty > \frac{1}{2}\min\{\rho_{j-1}, \rho_j\} > 0$ for all $p \in S \setminus \mathring{M}_{j-1} = (bM_{j-1}) \cup \Omega^j \cup C$.

Since $S \subset \mathring{M}_j$ is Runge and a strong deformation retract of M_j, Proposition 5.2 applied to the data

$$M = M_j, \quad S, \quad \Lambda = \Lambda_j, \quad k, \quad \text{and} \quad F = \widetilde{F}_j$$

gives a nonflat $\mathfrak{S}$-immersion $\widehat{F}_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$ such that:

(B.1) $\widehat{F}_j$ is as close as desired to $\widetilde{F}_j$ in the $\mathscr{C}^1(S)$-topology.

(B.2) $\widehat{F}_j - \widetilde{F}_j$ has a zero of multiplicity $k \in \mathbb{N}$ at every point $p \in \Lambda_j$.

If the approximation in (B.1) is close enough then, in view of (A.4), there exists a small compact neighborhood N of S in $\mathring{M}_j$, being a smoothly bounded compact domain and a strong deformation retract of M_j, and such that:

(B.3) $|\widehat{F}_j(p)|_\infty > \frac{1}{2}\min\{\rho_{j-1}, \rho_j\} > 0 \quad$ for all $p \in N \setminus \mathring{M}_{j-1}$.

Notice that $\Lambda \cap (M_j \setminus \mathring{N}) = \varnothing$ and hence we may apply Lemma 5.6 to the data

$$M = M_j, \quad K = N, \quad \Lambda = \Lambda_j, \quad F = \widehat{F}_j, \quad \rho = \tfrac{1}{2}\min\{\rho_{j-1}, \rho_j\}, \quad \tau = \rho_j, \quad k,$$

obtaining an $\mathfrak{S}$-immersion $F_j : M_j \to \mathbb{C}^n$ of class $\mathscr{A}^1(M_j)$ such that:

(C.1) F_j is as close as desired to $\widehat{F}_j$ in the $\mathscr{C}^1(N)$-topology.

(C.2) $F_j - \widehat{F}_j$ has a zero of multiplicity k at every point $p \in \Lambda_j \subset \mathring{N}$.

(C.3) $|F_j(p)|_\infty > \frac{1}{2}\min\{\rho_{j-1}, \rho_j\}$ for all $p \in M_j \setminus \mathring{N}$.

(C.4) $|F_j(p)|_\infty > \rho_j$ for all $p \in bM_j$.

We claim that, if the approximations in (B.1) and (C.1) are close enough, the $\mathfrak{S}$-immersion $F_j : M_j \to \mathbb{C}^n$ satisfies properties (i_j)–(iii_j), (v_j.1), and (v_j.2). Indeed, (A.1) ensures (i_j); properties (A.2), (B.2), and (C.2) guarantee (ii_j); condition (v_j.1) follows from (A.4), (B.3), and (C.3); and condition (v_j.2) is implied by (C.4). Finally, if $F|_\Lambda$ is injective then, by Theorem 5.3, we may assume without loss of generality that F_j is an embedding. This closes the inductive step and concludes the recursive construction of the sequence $\{F_j\}_{j \geq j_0+1}$ meeting the desired requirements.

As above, choosing the number $\epsilon_j > 0$ ($j \in \mathbb{N}$) small enough at each step in the construction, properties (i_j)–(iii_j) ensure that the sequence $\{F_j\}_{j\in\mathbb{N}}$ converges uniformly on compact subsets of M to an $\mathfrak{S}$-immersion $\widetilde{F} := \lim_{j\to\infty} F_j : M \to \mathbb{C}^n$ which is as close as desired to F uniformly on K, is injective if $F|_\Lambda$ is injective, and is such that $\widetilde{F} - F$ has a zero of multiplicity k at all points of Λ. Furthermore, properties (v_j.1) and (6-4) imply that $\widetilde{F}$ is a proper map. Indeed, take a number $R > 0$ and a sequence $\{q_m\}_{m\in\mathbb{N}}$ that diverges on M, and let us check that there is $m_0 \in \mathbb{N}$ such that $|\widetilde{F}(q_m)|_\infty > R$ for all $m \geq m_0$. Indeed, set

$$\varepsilon := \sum_{j\geq 1} \epsilon_j < +\infty$$

and observe that, by properties (i_j),

$$\|\widetilde{F} - F_j\|_{1,M_j} < \varepsilon \quad \text{for all } j \in \mathbb{Z}_+. \tag{6-6}$$

On the other hand, in view of (6-4) there is an integer $j_1 \geq j_0 + 1$ such that

$$\rho_{j-1} > 2(R+\varepsilon) \quad \text{for all } j \geq j_1. \tag{6-7}$$

Now, since the sequence $\{p_m\}_{m\in\mathbb{N}}$ diverges on M and $\{M_j\}_{j\in\mathbb{N}}$ is an exhaustion of M, there is $m_0 \in \mathbb{N}$ such that

$$p_m \in M \setminus M_{j_1} \quad \text{for all } m \geq m_0.$$

Thus, for any $m \geq m_0$ there is an integer $j_m \geq j_1$ such that $q_m \in M_{j_m} \setminus \mathring{M}_{j_m-1}$, and so

$$|\widetilde{F}(q_m)|_\infty \geq |F_{j_m}(q_m)|_\infty - |F_{j_m}(q_m) - \widetilde{F}(q_m)|_\infty \overset{(\mathrm{v}_{j_m}.1),\,(6\text{-}6)}{>} \tfrac{1}{2}\min\{\rho_{j_m-1}, \rho_{j_m}\} - \varepsilon \overset{(6\text{-}7)}{>} R.$$

This proves that $\widetilde{F} : M \to \mathbb{C}^n$ is a proper map and concludes the proof of Theorem 1.3.

7. Sketch of the proof of Theorem 1.2 and the proof of Theorem 1.1

We briefly explain how the arguments in Sections 5 and 6 which enabled us to prove Theorem 1.3 may be adapted in order to guarantee Theorem 1.2; we shall leave the obvious details of the proof to the interested reader. Afterward, we will use Theorem 1.2 to prove Theorem 1.1.

First of all recall that, as pointed out in Section 2C, for any integer $n \geq 3$ the punctured null quadric $\mathfrak{A}_* \subset \mathbb{C}^n$, see (1-3) and (1-4), directing minimal surfaces in $\mathbb{R}^n$ is an Oka manifold and satisfies the assumptions in assertions (I) and (II) in Theorem 1.3. Thus, Theorem 4.4 and Lemma 4.5 hold for $\mathfrak{S} = \mathfrak{A}$.

The first step in the proof of Theorem 1.2 consists of providing an analogue of Proposition 5.2 for *generalized conformal minimal immersions* in the sense of [Alarcón, Forstnerič and López 2016a, Definition 5.2]. In particular we need to show that if we are given M, S, and Λ as in Proposition 5.2 then, for any integer $k \in \mathbb{Z}_+$, every generalized conformal minimal immersion $X : S \to \mathbb{R}^n$ ($n \geq 3$) may be approximated in the $\mathscr{C}^1(S)$-topology by conformal minimal immersions $\widetilde{X} : M \to \mathbb{R}^n$ of class $\mathscr{C}^1(M)$ such that $\widetilde{X}$ and X have a contact of order k at every point in Λ and the flux map $\mathrm{Flux}_{\widetilde{X}}$ equals Flux_X everywhere in the first homology group $H_1(M;\mathbb{Z})$. To do that we reason as in the proof of Proposition 5.2 but working with the map $f := \partial X/\theta : S \to \mathfrak{A}_*$. Since $f\theta$ does not need to be exact (only its real part does) we replace conditions (a) and (b) in the proof of the proposition by the following ones:

- $(\tilde{f} - f)\theta$ is exact on S.
- $X(p_0) + 2\int_{C_p} \Re(\tilde{f}\theta) = X(p_0) + 2\int_{C_p} \Re(f\theta) = X(p)$ for all $p \in \Lambda$.

It can then be easily seen that the conformal minimal immersion $\widetilde{X} : M \to \mathbb{R}^n$ of class $\mathscr{C}^1(M)$ given by

$$\widetilde{X}(p) := X(p) + 2\int_{C_p} \Re(\tilde{f}\theta), \quad p \in M,$$

is well-defined and enjoys the desired properties.

In a second step and following the same spirit, we need to furnish a general position theorem, a completeness lemma, and a properness lemma for conformal minimal immersions of class $\mathscr{C}^1$ on a compact bordered Riemann surface, which are analogues of Theorem 5.3, Lemma 5.5, and Lemma 5.6, respectively. In this case the general position of conformal minimal surfaces is embedded in $\mathbb{R}^n$ for all integers $n \geq 5$; to adapt the proof of Theorem 5.3 to the minimal surfaces framework we combine the argument in [Alarcón, Forstnerič and López 2016a, proof of Theorem 1.1] with the new ideas in Section 5A which allow us to ensure the interpolation condition. Likewise, the analogues of Lemmas 5.5 and 5.6 for conformal minimal surfaces can be proved by adapting the proofs of the cited lemmas in Sections 5B and 5C, respectively; the required modifications follow the pattern described in the previous paragraph: at each step in the proofs we ensure that the real part of the 1-forms is exact and that the periods of the imaginary part agree with the flux map of the initial conformal minimal immersion. Furthermore, obviously, we are allowed to use only the real part in order to ensure the increasing of the intrinsic diameter of the surface to achieve completeness, see Lemma 5.5 (II), and the increasing of the $|\cdot|_\infty$-norm near the boundary to guarantee properness, see Lemma 5.6 (II) and (III). For the former we just replace condition (c) in the proof of Lemma 5.5 (which determines an extrinsic bound) by the following one:

- $|2\int_\alpha \Re(h\theta)| > \tau$ for all arcs $\alpha \subset S$ with initial point p_0 and final point in L.

For the latter, the adaptation is done straightforwardly since all the bounds are of the same nature, namely, extrinsic.

Finally, granted the analogues for conformal minimal surfaces in $\mathbb{R}^n$ of Proposition 5.2, Theorem 5.3, Lemma 5.5, and Lemma 5.6, the proof of Theorem 1.2 follows word for word, up to trivial modifications similar to the ones discussed in the previous paragraphs, the one of Theorem 1.3 in Section 6. It is perhaps worth pointing out that in the noncritical case in the recursive construction (see Section 6A) we now have to extend a conformal minimal immersion of class $\mathscr{C}^1(M_{j-1})$ to a generalized conformal minimal immersion on the admissible set $S = M_{j-1} \cup \alpha \subset \mathring{M}_j$ whose flux map equals $\mathfrak{p}$ for every closed curve in S (here $\mathfrak{p} : H_1(M; \mathbb{Z}) \to \mathbb{R}^n$ denotes the group homomorphism given in the statement of Theorem 1.2, whereas M_{j-1}, α, and M_j are as in Section 6A); this can be easily done as in [Alarcón, Forstnerič and López 2016a, proof of Theorem 1.2]. This concludes the sketch of the proof of Theorem 1.2; as we announced at the very beginning of this section, we leave the details to the interested reader.

To finish the paper we show how Theorem 1.2 can be used in order to prove the following extension to Theorem 1.1 in the Introduction.

Corollary 7.1. *Let M be an open Riemann surface and $\Lambda \subset M$ be a closed discrete subset. Consider also an integer $n \geq 3$ and maps $\mathfrak{Z} : \Lambda \to \mathbb{R}^n$ and*

$$G : \Lambda \to Q_{n-2} = \{[z_1 : \cdots : z_n] \in \mathbb{CP}^{n-1} : z_1^2 + \cdots + z_n^2 = 0\} \subset \mathbb{CP}^{n-1}.$$

Then there is a conformal minimal immersion $\widetilde{X} : M \to \mathbb{R}^n$ satisfying $\widetilde{X}|_\Lambda = \mathfrak{Z}$ and whose generalized Gauss map $G_{\widetilde{X}} : M \to \mathbb{CP}^{n-1}$ equals G on Λ.

Proof. For each $p \in \Lambda$ let Ω_p be a smoothly bounded, simply connected compact neighborhood of p in M, and assume that $\Omega_p \cap \Omega_q = \varnothing$ whenever $p \neq q \in \Lambda$. Set $\Omega := \bigcup_{p\in\Lambda} \Omega_p$ and let $X : \Omega \to \mathbb{R}^n$ be any conformal minimal immersion of class $\mathscr{C}^1(\Omega)$ such that $X|_\Lambda = \mathfrak{Z}$ and the generalized Gauss map satisfies $G_X|_\Lambda = G$. (Such an X always exists; we may for instance choose $X|_{\Omega_p}$ to be a suitable planar disk for each $p \in \Lambda$). Also fix a smoothly bounded simply connected compact domain $K \subset M \setminus \Lambda$, up to shrinking the sets Ω_p if necessary, assume that $K \subset M \setminus \Omega$, and extend X to $\Omega \cup K \to \mathbb{R}^n$ such that $X|_K : K \to \mathbb{R}^n$ is any conformal minimal immersion of class $\mathscr{C}^1(K \cup \Omega)$. Applying Theorem 1.2 to these data, any group homomorphism $H_1(M; \mathbb{Z}) \to \mathbb{R}^n$, and the integer $k = 1$, we obtain a conformal minimal immersion $\widetilde{X} : M \to \mathbb{R}^n$ which has a contact of order 1 with $X|_\Omega$ at every point in Λ. Thus, $\widetilde{X}|_\Lambda = X|_\Lambda = \mathfrak{Z}$ and the generalized Gauss map satisfies $G_{\widetilde{X}}|_\Lambda = [\partial \widetilde{X}]|_\Lambda = [\partial X]|_\Lambda = G_X|_\Lambda = G$. □

Acknowledgements

Research was partially supported by the MINECO/FEDER grant nos. MTM2014-52368-P and MTM2017-89677-P, Spain. Castro-Infantes was supported by the grant no. BES-2015-071993 from MINECO, Spain.

The authors wish to express their gratitude to Franc Forstnerič and Francisco J. López for helpful discussions which led to improvement of the paper.

References

[Abraham 1963] R. Abraham, "Transversality in manifolds of mappings", *Bull. Amer. Math. Soc.* **69** (1963), 470–474. MR Zbl

[Abraham, Marsden and Ratiu 1988] R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, tensor analysis, and applications*, 2nd ed., Applied Mathematical Sciences **75**, Springer, 1988. MR Zbl

[Alarcón and Castro-Infantes 2018] A. Alarcón and I. Castro-Infantes, "Complete minimal surfaces densely lying in arbitrary domains of $\mathbb{R}^n$", *Geom. Topol.* **22**:1 (2018), 571–590. MR Zbl

[Alarcón and Forstnerič 2014] A. Alarcón and F. Forstnerič, "Null curves and directed immersions of open Riemann surfaces", *Invent. Math.* **196**:3 (2014), 733–771. MR Zbl

[Alarcón and Forstnerič 2015] A. Alarcón and F. Forstnerič, "The Calabi–Yau problem, null curves, and Bryant surfaces", *Math. Ann.* **363**:3-4 (2015), 913–951. MR Zbl

[Alarcón and López 2012] A. Alarcón and F. J. López, "Minimal surfaces in $\mathbb{R}^3$ properly projecting into $\mathbb{R}^2$", *J. Differential Geom.* **90**:3 (2012), 351–381. MR Zbl

[Alarcón and López 2013a] A. Alarcón and F. J. López, "Null curves in $\mathbb{C}^3$ and Calabi–Yau conjectures", *Math. Ann.* **355**:2 (2013), 429–455. MR Zbl

[Alarcón and López 2013b] A. Alarcón and F. J. López, "Proper holomorphic embeddings of Riemann surfaces with arbitrary topology into $\mathbb{C}^2$", *J. Geom. Anal.* **23**:4 (2013), 1794–1805. MR Zbl

[Alarcón and López 2014] A. Alarcón and F. J. López, "Properness of associated minimal surfaces", *Trans. Amer. Math. Soc.* **366**:10 (2014), 5139–5154. MR Zbl

[Alarcón and López 2015] A. Alarcón and F. J. López, "Approximation theory for nonorientable minimal surfaces and applications", *Geom. Topol.* **19**:2 (2015), 1015–1062. MR Zbl

[Alarcón, Drinovec Drnovšek, Forstnerič and López 2015] A. Alarcón, B. Drinovec Drnovšek, F. Forstnerič, and F. J. López, "Every bordered Riemann surface is a complete conformal minimal surface bounded by Jordan curves", *Proc. Lond. Math. Soc.* (3) **111**:4 (2015), 851–886. MR Zbl

[Alarcón, Drinovec Drnovšek, Forstnerič and López ≥ 2019] A. Alarcón, B. Drinovec Drnovšek, F. Forstnerič, and F. J. López, "Minimal surfaces in minimally convex domains", *Trans. Amer. Math. Soc.*

[Alarcón, Fernández and López 2013] A. Alarcón, I. Fernández, and F. J. López, "Harmonic mappings and conformal minimal immersions of Riemann surfaces into $\mathbb{R}^{\mathrm{N}}$", *Calc. Var. Partial Differential Equations* **47**:1-2 (2013), 227–242. MR Zbl

[Alarcón, Forstnerič and López 2016a] A. Alarcón, F. Forstnerič, and F. J. López, "Embedded minimal surfaces in $\mathbb{R}^n$", *Math. Z.* **283**:1-2 (2016), 1–24. MR Zbl

[Alarcón, Forstnerič and López 2016b] A. Alarcón, F. Forstnerič, and F. J. López, "New complex analytic methods in the study of non-orientable minimal surfaces in $\mathbb{R}^n$", preprint, 2016. To appear in *Mem. Amer. Math. Soc.* arXiv

[Alarcón, Forstnerič and López 2017] A. Alarcón, F. Forstnerič, and F. J. López, "Every meromorphic function is the Gauss map of a conformal minimal surface", *J. Geom. Anal.* (online publication October 2017).

[Behnke and Stein 1949] H. Behnke and K. Stein, "Entwicklung analytischer Funktionen auf Riemannschen Flächen", *Math. Ann.* **120** (1949), 430–461. MR Zbl

[Bishop 1958] E. Bishop, "Subalgebras of functions on a Riemann surface", *Pacific J. Math.* **8** (1958), 29–50. MR Zbl

[Cartan 1953] H. Cartan, "Variétés analytiques complexes et cohomologie", pp. 41–55 in *Colloque sur les fonctions de plusieurs variables* (Bruxelles, 1953), Georges Thone, Liège, 1953. MR Zbl

[Drinovec Drnovšek and Forstnerič 2016] B. Drinovec Drnovšek and F. Forstnerič, "Minimal hulls of compact sets in $\mathbb{R}^3$", *Trans. Amer. Math. Soc.* **368**:10 (2016), 7477–7506. MR Zbl

[Eliashberg and Mishachev 2002] Y. Eliashberg and N. Mishachev, *Introduction to the h-principle*, Graduate Studies in Mathematics **48**, American Mathematical Society, Providence, RI, 2002. MR Zbl

[Florack 1948] H. Florack, "Reguläre und meromorphe Funktionen auf nicht geschlossenen Riemannschen Flächen", *Schr. Math. Inst. Univ. Münster,* **1948**:1 (1948), 34. MR Zbl

[Forstnerič 2003] F. Forstnerič, "The Oka principle for multivalued sections of ramified mappings", *Forum Math.* **15**:2 (2003), 309–328. MR Zbl

[Forstnerič 2004] F. Forstnerič, "Holomorphic submersions from Stein manifolds", *Ann. Inst. Fourier* (*Grenoble*) **54**:6 (2004), 1913–1942. MR Zbl

[Forstnerič 2006] F. Forstnerič, "Runge approximation on convex sets implies the Oka property", *Ann. of Math.* (2) **163**:2 (2006), 689–707. MR Zbl

[Forstnerič 2013] F. Forstnerič, "Oka manifolds: from Oka to Stein and back", *Ann. Fac. Sci. Toulouse Math.* (6) **22**:4 (2013), 747–809. MR Zbl

[Forstnerič 2017] F. Forstnerič, *Stein manifolds and holomorphic mappings: the homotopy principle in complex analysis*, 2nd ed., Ergebnisse der Mathematik (3) **56**, Springer, 2017. MR Zbl

[Forstnerič and Lárusson 2011] F. Forstnerič and F. Lárusson, "Survey of Oka theory", *New York J. Math.* **17A** (2011), 11–38. MR Zbl

[Forstnerič and Lárusson 2016] F. Forstnerič and F. Lárusson, "The parametric h-principle for minimal surfaces in $\mathbb{R}^n$ and null curves in $\mathbb{C}^n$", preprint, 2016. To appear in *Comm. Anal. Geom.* arXiv

[Grauert 1957] H. Grauert, "Approximationssätze für holomorphe Funktionen mit Werten in komplexen Rämen", *Math. Ann.* **133** (1957), 139–159. MR Zbl

[Grauert 1958] H. Grauert, "Analytische Faserungen über holomorph-vollständigen Räumen", *Math. Ann.* **135** (1958), 263–273. MR Zbl

[Gromov 1986] M. Gromov, *Partial differential relations*, Ergebnisse der Mathematik (3) **9**, Springer, 1986. MR Zbl

[Hörmander and Wermer 1968] L. Hörmander and J. Wermer, "Uniform approximation on compact sets in C^n", *Math. Scand.* **23** (1968), 5–21. MR Zbl

[Jorge and Xavier 1980] L. P. d. M. Jorge and F. Xavier, "A complete minimal surface in $\mathbb{R}^3$ between two parallel planes", *Ann. of Math.* (2) **112**:1 (1980), 203–206. MR Zbl

[Kutzschebauch 2014] F. Kutzschebauch, "Flexibility properties in complex analysis and affine algebraic geometry", pp. 387–405 in *Automorphisms in birational and affine geometry*, edited by I. Cheltsov et al., Springer Proc. Math. Stat. **79**, Springer, 2014. MR Zbl

[Lárusson 2010] F. Lárusson, "What is . . . an Oka manifold?", *Notices Amer. Math. Soc.* **57**:1 (2010), 50–52. MR Zbl

[Mergelyan 1951] S. N. Mergelyan, "On the representation of functions by series of polynomials on closed sets", *Doklady Akad. Nauk SSSR* (*N.S.*) **78** (1951), 405–408. In Russian. MR

[Michor 1980] P. W. Michor, *Manifolds of differentiable mappings*, Shiva Mathematics Series **3**, Shiva, Nantwich, 1980. MR Zbl

[Napier and Ramachandran 2011] T. Napier and M. Ramachandran, *An introduction to Riemann surfaces*, Springer, 2011. MR Zbl

[Osserman 1986] R. Osserman, *A survey of minimal surfaces*, 2nd ed., Dover, New York, 1986. MR Zbl

[Runge 1885] C. Runge, "Zur Theorie der Analytischen Functionen", *Acta Math.* **6**:1 (1885), 245–248. MR Zbl

[Weierstrauss 1876] K. Weierstrauss, "Zur Theorie der eindeutigen analytischen Functionen", *Berl. Abh.* **1876** (1876), 11–60.

Received 1 Mar 2018. Accepted 1 May 2018.

ANTONIO ALARCÓN: alarcon@ugr.es
Departamento de Geometría y Topología, Instituto de Matemáticas (IEMath-GR), Universidad de Granada, Campus de Fuentenueva, Granada, Spain

ILDEFONSO CASTRO-INFANTES: icastroinfantes@ugr.es
Departamento de Geometría y Topología, Instituto de Matemáticas (IEMath-GR), Universidad de Granada, Campus de Fuentenueva, Granada, Spain

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 12 No. 2 2019

2157-5045(2019)12:2;1-O