PERTURBATIVE ESTIMATES FOR THE ONE-PHASE STEFAN PROBLEM

D. DE SILVA, N. FORCILLO, AND O. SAVIN

ABSTRACT. We provide perturbative estimates for the one-phase Stefan free boundary problem and obtain the regularity of flat free boundaries via a linearization technique in the spirit of the elliptic counterpart established in [D].

1. INTRODUCTION

In this paper we are concerned with perturbative estimates for the one-phase Stefan problem,

$$\begin{cases} u_t = \triangle u & \text{in } (\Omega \times (0,T]) \cap \{u > 0\}, \\ u_t = |\nabla u|^2 & \text{on } (\Omega \times (0,T]) \cap \partial\{u > 0\}, \end{cases} \tag{1.1}$$

with $\Omega \subset \mathbb{R}^n$, $u : \Omega \times [0,T] \to \mathbb{R}$, $u \geq 0$.

The classical one-phase Stefan problem describes the phase transition between solids and liquids, such as the melting of the ice (see for example [F], [R]). In this setting u represents the temperature of the liquid, and the region $\{u = 0\}$ the unmelted region of ice.

The main object of interest is the behavior of the free boundary $\partial\{u > 0\}$. In problems of this type free boundaries may not regularize instantaneously. A two dimensional example in which a Lipschitz free boundary preserves corners can be found for instance in [CS]. Athanasopoulos, Caffarelli, and Salsa studied the regularizing properties of the free boundary under reasonable assumptions in the more general setting of the two-phase Stefan problem. In [ACS1] they showed that Lipschitz free boundaries in space-time become smooth provided a nondegeneracy condition holds, while in [ACS2] the same conclusion was established for sufficiently "flat" free boundaries. The techniques are based on the original work of Caffarelli in the elliptic case [C1, C2].

A related result is due to S. Choi and I. Kim who showed in [CK] that solutions regularize instantaneously if the initial free boundary is locally Lipschitz with bounded Lipschitz constant and the initial data has subquadratic growth.

In this paper we study the regularity of flat free boundaries for (1.1) based on perturbation arguments leading to a linearization of the problem, which are in the spirit of the elliptic counterpart developed by the first author in [D]. Our result is basically equivalent to the previously mentioned flatness result in [ACS2]. The techniques in [D] are very flexible and have been widely generalized to a variety of free boundary problems, including two-phase inhomogeneous problems, "thin" free

N.F. is partially supported by INDAM-GNAMPA-2019 project: *Proprietà di regolarità delle soluzioni viscose con applicazioni a problemi di frontiera libera,* and by the project: *GHAIA Horizon 2020 MCSA RISE programme grant No 777822*.

boundary problems, minimization problems (see for example [DFS], [DR], [DSV]). The methods of the current paper are suitable to further extensions as well.

Our main theorem roughly states that a solution to the Stefan problem in a ball of size λ in space-time which is of size λ and has a "flat free boundary" in space, must have smooth free boundary in the interior provided that a necessary nondegeneracy condition holds. The nondegeneracy condition for u requires that u is bounded below by a small multiple of λ at some point in the domain at distance λ from the free boundary. Precisely, we assume that $u : \Omega\times[0,T] \to \mathbb{R}^+$ solves (1.1) in the viscosity sense. This means that u is continuous and its graph cannot be touched by above (resp. below) at a point (x_0,t_0) in a parabolic cylinder $B_r(x_0)\times(t_0-r^2,t_0]$, by the graph of a classical strict supersolution φ^+ (resp. subsolution). By a classical strict supersolution we mean that $\varphi(x,t) \in C^2$, $\nabla_x\varphi \neq 0$, and it solves

$$\begin{cases} \varphi_t > \triangle\varphi & \text{in } (\Omega\times(0,T])\cap\{\varphi>0\}, \\ \varphi_t > |\nabla\varphi|^2 & \text{on } (\Omega\times(0,T])\cap\partial\{\varphi>0\}. \end{cases} \tag{1.2}$$

Similarly we can define a strict classical subsolution.

Throughout the paper, given a space-time function, ∇, Δ, and D^2 are computed with respect to the space variable x.

The rigorous statement of the main theorem is as follows.

Theorem 1.1. *Fix a constant K (large) and let u be a solution to the one-phase Stefan problem (1.1) in $B_\lambda\times[-K^{-1}\lambda,0]$ for some $\lambda\leq 1$. Assume that*

$$0\leq u\leq K\lambda, \qquad u(x_0,t)\geq K^{-1}\lambda \quad \text{for some } x_0\in B_{\frac{3}{4}\lambda} \text{ and all } t\in[-K^{-1}\lambda,0].$$

There exists ϵ_0 depending only on K and n such that if, for each t, $\partial_x\{u>0\}$ is ϵ_0-flat in B_λ, then the free boundary $\partial\{u>0\}$ (and u up to the free boundary) is smooth in $B_{\frac{\lambda}{2}}\times[-(2K)^{-1}\lambda,0]$.

Here we use the notation $\partial_x\{u>0\}$ to denote the boundary in $\mathbb{R}^n$ of $\{u(\cdot,t)>0)\}$, with t being fixed. By $\partial_x\{u>0\}$ is ϵ_0-flat in B_λ we understand that, for each t, $\partial_x\{u>0\}\cap B_\lambda$ is trapped in a strip of width $\epsilon_0\lambda$ (the region between two parallel hyperplanes at distance $\epsilon_0\lambda$ from each other), and $u=0$ on one side of this strip while $u>0$ on the other side.

The assumption that u is of size λ in a domain of size λ around the free boundary is natural, since this eventually holds for all classical solutions by choosing λ small. We point out that in Theorem 1.1 the behavior of the solution depends strongly on the value of λ. If we scale the domain to unit size and keep the function u of size 1, then the rescaled function

$$(x,t) \quad \mapsto \quad \frac{1}{\lambda}u(\lambda x,\lambda t), \qquad (x,t)\in B_1\times[-K^{-1},0],$$

solves a Stefan problem with possibly large diffusion coefficient λ^{-1}

$$\begin{cases} \lambda u_t = \triangle u & \text{in } (B_1\times(-K^{-1},0])\cap\{u>0\}, \\ u_t = |\nabla u|^2 & \text{on } (B_1\times(-K^{-1},0])\cap\partial\{u>0\}. \end{cases} \tag{1.3}$$

Our theorem states that nondegenerate solutions of size 1 of (1.3) which have ϵ_0-flat free boundaries in B_1 are smooth up to the free boundary. We remark that ϵ_0 is independent of λ, which means that we need to obtain uniform estimates in λ for the oscillation of the free boundaries of solutions of (1.3). Our results show that the free boundary has a uniform $C^{1,\alpha}$ bound in space. On the other hand, the estimates for u in the set where it is positive depend on the parameter λ. The strategy is to approximate u with a family of explicit functions $l_{a,b}$ which in the direction perpendicular to the free boundary depend on λ while on the tangential directions to the free boundary are independent of the parameter λ.

Formally as $\lambda \to 0^+$, a solution u to (1.3) solves the Hele-Shaw equation. Estimates for this problem by similar methods as ours were obtained by H. Chang-Lara and N. Guillen in [CG].

To prove our main theorem, we show that if a solution u satisfies the hypotheses of Theorem 1.1 then, after a convenient dilation, the flatness assumption can be extended to the whole function u instead of just the free boundary. Then Theorem 1.1 follows from the following result.

Theorem 1.2. *Fix a constant K (large) and let u be a solution to the one-phase Stefan problem* (1.1) *in $B_{2\lambda} \times [-2\lambda, 0]$ for some $\lambda \leq 1$. Assume that $0 \in \partial\{u > 0\}$, and*

$$a_n(t)\ (x_n - b(t) - \epsilon_1\lambda)^+ \ \leq u \ \leq a_n(t)\ (x_n - b(t) + \epsilon_1\lambda)^+ \,,$$

with

$$K^{-1} \leq a_n \leq K, \qquad |a_n'(t)| \leq \lambda^{-2}, \qquad b'(t) = -a_n(t),$$

for some small ϵ_1 depending only on K and n. Then in $B_\lambda \times [-\lambda, 0]$ the free boundary $\partial\{u > 0\}$ is a $C^{1,\alpha}$ graph in the x_n direction.

The assumption that $b' = -a_n(t)$ means that the approximating linear functions in x, $a_n(t)(x_n - b(t))^+$, satisfy the free boundary condition, while $|a_n'(t)| \leq \lambda^{-2}$ respects the parabolic scaling of the interior equation and represents that a_n can change at most $o(1)$ in a time interval of length $o(\lambda^2)$.

We remark that it suffices to prove Theorem 1.2 under the more relaxed hypotheses

$$\lambda \leq \lambda_0 \quad \text{and} \quad |a_n'(t)| \leq c_0\lambda^{-2}, \tag{1.4}$$

with λ_0, c_0 small depending on K, n. We end up in this setting by working in balls of size $\tau\lambda$ with τ sufficiently small, and then relabel $\tau\lambda$ by λ and $\epsilon_1\tau^{-1}$ by ϵ_1.

Theorem 1.2 applies, for example, when u is a perturbation of order $o(1)\lambda$ of a traveling wave solution

$$(e^{ax_n + a^2t} - 1)^+, \qquad K^{-1} \leq a \leq K.$$

In this case we choose $a_n(t) = a$, $b(t) = -at$, and consider $\lambda \leq \lambda_0$ small so that the difference between the approximating linear part $a_n(t)(x_n - b(t))$ and the exact solution above is less than $\frac{1}{2}\epsilon_1\lambda$ in B_λ.

The proof of Theorem 1.2 is based on linearization techniques. The linearized equation in our setting has the form of an oblique derivative parabolic problem

$$\begin{cases} \lambda v_t = tr(A(t)D^2v) & \text{in } \{x_n > 0\}, \\ v_t = \gamma(t) \cdot \nabla v & \text{on } \{x_n = 0\}, \end{cases} \tag{1.5}$$

with $A(t)$ uniformly elliptic and $\gamma_n > 0$. An important task in our analysis is to develop Schauder-type estimates for equation (1.5) with respect to an appropriate distance d_λ and to capture both features of the mixed parabolic/hyperbolic scaling.

The paper is organized as follows. In the next section we show that Theorem 1.1 can be deduced from Theorem 1.2. In Section 3, we use a Hodograph transform to obtain an equivalent quasilinear parabolic equation with oblique derivative boundary condition. In the following section, we state an improvement of flatness result

Proposition 4.1 for solutions of such nonlinear problem, then we show how this implies Theorem 1.2. The proof of Proposition 4.1 is presented in Section 5, and it relies on various Hölder estimates (with respect to the appropriate distance) for solutions to the linearized problem associated to the nonlinear problem. Sections 6 and 7 are devoted to the proofs of such Hölder estimates, while Section 8 focuses on the one dimensional linear problem, which plays an essential role. The last section contains some general technical results on solutions to the linear problem.

2. From flat free boundaries to flat solutions.

In this section, we show that Theorem 1.1 can be reduced to Theorem 1.2.

We assume that the function u satisfies the ϵ_0-flatness hypothesis of the free boundary from Theorem 1.1 for some $\lambda \leq 1$, and that $(0,0)$ is a free boundary point. Precisely, by $\partial_x\{u>0\}$ is ϵ_0-flat in B_λ we understand that, for each t, there exists a direction ν such that

$$\partial_x\{u(\cdot,t)>0\}\cap B_\lambda \subset \{|\langle x-x_0,\nu\rangle| \leq \epsilon_0\lambda\},$$

and

$$\begin{aligned} u=0 &\quad \text{in } \{\langle x-x_0,\nu\rangle \leq -\epsilon_0\lambda\},\\ u>0 &\quad \text{in } \{\langle x-x_0,\nu\rangle \geq \epsilon_0\lambda\}. \end{aligned}$$

First, we show that in a smaller domain $B_{\eta\lambda}\times[-\eta\lambda,0]$ the whole graph of u is η^β- flat, for some small β, provided that $\epsilon_0 \leq c(\eta,K)$. Then, in this domain the hypotheses of Theorem 1.2 are satisfied by choosing η sufficiently small.

We work with the parabolic rescaling of the function u which is defined in $B_1 \times [-(K\lambda)^{-1},0]$ and keeps the function u of unit size:

$$(x,t) \quad \mapsto \quad \frac{1}{\lambda}u(\lambda x,\lambda^2 t), \qquad (x,t)\in B_1\times[-(K\lambda)^{-1},0].$$

By abuse of notation we denote this rescaling by u, and then u solves a Stefan problem with possibly small speed coefficient λ,

$$\begin{cases} u_t = \triangle u & \text{in } (B_1\times(-(K\lambda)^{-1},0])\cap\{u>0\},\\ u_t=\lambda|\nabla u|^2 & \text{on } (B_1\times(-(K\lambda)^{-1},0])\cap\partial\{u>0\}. \end{cases} \tag{2.1}$$

We prove the following main lemma. Universal constants only depend on n, K. As usual, in the body of the proofs, constants denoted by C may change from line to line.

Lemma 2.1. *Assume that u solves* (2.1)*,*

$$0\leq u\leq K, \qquad u(x_0,t)\geq K^{-1} \quad \textit{for some} \quad x_0\in B_{3/4}, \quad \textit{and all} \quad t\in(-(K\lambda)^{-1},0],$$

$$0\in\partial_x\{u(\cdot,0)>0\}, \qquad \textit{and } \partial_x\{u(\cdot,t)>0\} \textit{ is } \epsilon_0\textit{-flat in } B_1.$$

Then for all small $\eta>0$ we have up to rotations:

$$a_n(t)\left(x_n-b(t)-\eta^{1+\beta}\right)^+ \;\leq\; u \;\leq\; a_n(t)\left(x_n-b(t)+\eta^{1+\beta}\right)^+ \quad \textit{in } B_\eta\times[-\lambda^{-1}\eta,0],$$

with $\beta=1/20$ and for $c,C>0$ universal,

$$c\leq a_n(t)\leq C, \quad |a_n'(t)|\leq \eta^{\beta-2}, \qquad b'(t)=-\lambda a_n(t), \quad b(0)=0,$$

provided that $\epsilon_0\leq c(\eta,K)$.

When we rescale the conclusion back to the original coordinates, we obtain that the hypotheses of Theorem 1.2 are satisfied in the cylinder $B_{\eta\lambda} \times [-\eta\lambda, 0]$ with $\epsilon_1 = \eta^\beta$.

We start by proving a result about the location of the free boundary in time.

Lemma 2.2. *Assume u solves* (2.1) *in $B_2 \times [-K^{-1}, 1]$ and that $0 \le u \le K$. If $u(x,0) = 0$ in B_1, then*

$$u(x,t) \le C(|x| - 1)^+, \quad \textit{if} \quad t \in [-(2K)^{-1}, 0], \tag{2.2}$$

and

$$u(x,t) = 0 \quad \textit{if } |x| < 1 - C\lambda, \ t \in [0,1], \tag{2.3}$$

with $C > 0$ universal.

Proof. Since the support of u is increasing with time we deduce that $u = 0$ in B_1 for all $t \in [-K^{-1}, 0]$. Then, in the annular domain $(B_2 \setminus B_1) \times [-K^{-1}, 0]$, by the comparison principle, u is less than a multiple of the solution to the heat equation which equals 0 on $\partial B_1 \times (-K^{-1}, 0]$, and 1 on the remaining part of the parabolic boundary. This, together with the boundary regularity of such solution, implies the estimate (2.2).

Now, for times $t \in [0,1]$ we compare u with

$$w(x,t) = C_0 \ g(|x| - r(t)), \quad r(t) := 1 - C_0\lambda t,$$

with g a 1D function such that $g(s) = 0$ if $s \le 0$, and for positive s is defined by the ODE

$$g''(s) + 2ng'(s) = 0, \quad g(0) = 0, \quad g'(0) = 1.$$

Notice that $g' \in [0,1]$.

We may assume that $r(t) \ge 1/2$, otherwise the conclusion (2.3) is trivial (say for $C > 2C_0$).

The constant C_0 is chosen large such that $w \ge u$ at time $t = 0$ (by (2.2)) and also on $\partial B_2 \times [0,1]$. We check that w is a supersolution to (2.1); indeed in $\{w > 0\}$ we have (recall $r(t) \ge 1/2$),

$$w_t = C_0^2 \lambda g' \ge 0, \qquad \triangle w = C_0 \left(g'' + \frac{n-1}{|x|} g' \right) < 0,$$

and on $\partial\{w > 0\}$

$$w_t = \lambda C_0^2 = \lambda |\nabla w|^2.$$

In conclusion, $u \le w$ which gives the desired conclusion (2.3). □

Now, we turn to the proof of Lemma 2.1.

Proof of Lemma 2.1. We assume that u satisfies (2.1) in $B_1 \times [-(K\lambda)^{-1}, 0]$, and $\partial_x\{u > 0\}$ is ϵ_0-flat in B_1. Suppose that $(0,0) \in \partial\{u > 0\}$ and then, after a rotation,

$$u(x,0) > 0 \text{ if } x_n > \epsilon_0, \text{ and } u(x,0) = 0 \text{ if } x_n < -\epsilon_0.$$

From (2.2) in Lemma 2.2 (applied to balls tangent to $\{x_n = -\epsilon_0\}$) we find that $u \le C(x_n + \epsilon_0)^+$ in $B_{1/2} \times [-(2K)^{-1}, 0]$.

We define

$$u_\tau := \frac{1}{\tau} u(\tau x, \tau^2 t), \qquad \text{with} \quad \tau \ge \epsilon_0^{1/2},$$

and, if $\tau \in [\epsilon_0^{1/2}, c]$, then

$$u_\tau \leq C(x_n+\tau)^+ \quad \text{in } B_1 \times [-2,0]. \tag{2.4}$$

Notice that u_τ satisfies (2.1) with $\tau\lambda$ instead of λ. We apply (2.3) of Lemma 2.2 for u_τ and obtain that (since $(0,0) \in \partial\{u_\tau > 0\}$),

$$\partial_x\{u_\tau > 0\} \cap B_{1/2} \text{ intersects } \{x_n \leq C\lambda\tau\}, \quad \text{for all } t \in [-1,0]. \tag{2.5}$$

Moreover, $\partial_x\{u_\tau > 0\}$ is $\tau^{-1}\epsilon_0$-flat in B_1, which combined with (2.5) implies that

$$\partial\{u_\tau > 0\} \cap (B_{1/2} \times [-1,0]) \text{ is included in } \{x_n \leq C(\lambda\tau + \tau^{-1}\epsilon_0)\}. \tag{2.6}$$

In $(B_{1/2} \cap \{x_n > C\tau\}) \times [-1,0]$ we compare u_τ with the solution w to the heat equation which equals 0 on $\{x_n = C\tau\}$, and equals u_τ on the remaining part of the parabolic boundary. Notice that by (2.6), since $\tau \geq \epsilon_0^{1/2}$, $u_\tau > 0$ on $\{x_n = C\tau\}$. From (2.4) we find $|u_\tau - w| \leq C\tau$, and the boundary regularity of w gives

$$|u_\tau - ax_n| \leq C\rho^{3/2} + C\tau \leq 2C\rho^{3/2} \qquad \text{in} \quad B^+_{2\rho} \times [-\rho^2, 0], \tag{2.7}$$

for some constant $a < C$, provided that we choose $\tau = \rho^{3/2}$ with ρ small, to be made precise later.

We claim that the nondegeneracy assumption $u(x_0,t) \geq K^{-1}$ for some $x_0 \in B_{3/4}$ implies that $a > c$. For this we use (2.6) which, in terms of the function u, implies that $\partial_x\{u(\cdot,t) > 0\}$, at all times $t = -\tau^2 \leq -\epsilon_0$, intersects the x_n axis at distance at most $C(\lambda|t| + \epsilon_0)$ from the origin. As for (2.6), using that $\partial_x\{u > 0\}$ is ϵ_0-flat in B_1, we obtain that $u(x,t) > 0$ if $x_n > C\epsilon_0 + C\lambda|t|$ in $B_{1/2}$. Now we can use the nondegeneracy condition with a Hopf-type lemma for the heat equation and obtain

$$u \geq c(x_n - C(\epsilon_0 + \lambda|t|))^+ \quad \text{in} \quad B_{1/4} \times [-(4K)^{-1}, 0],$$

for some $c > 0$ that depends only on n and K. We use this inequality at time $t = 0$ in (2.7) and conclude $a > c$ since $\tau\rho > 2\tau^2 \geq 2\epsilon_0$. We can restate (2.7) as

$$(ax_n - C\eta^{1+\frac{1}{5}})^+ \leq u \leq (ax_n + C\eta^{1+\frac{1}{5}})^+ \quad \text{in} \quad B_{2\eta} \times [-\eta^2, 0],$$

with $\eta := \tau\rho = \rho^{5/2}$.

Similarly, by looking at the points $(b(t)e_n, t)$ where the free boundary intersects the x_n axis, we obtain that

$$|b(t)| \leq C(\lambda|t| + \epsilon_0) \leq C_0\eta \quad \text{if} \quad t \in [-\lambda^{-1}\eta, 0],$$

and in the domain $B_{2C_0\eta} \times [t - \eta^2, t]$ we have

$$\left(a(t)\cdot(x - b(t)e_n) - C\eta^{\frac{6}{5}}\right)^+ \leq u(x,s) \leq \left(a(t)\cdot(x - b(t)e_n) + C\eta^{\frac{6}{5}}\right)^+$$

with $c \leq |a(t)| \leq C$. The flatness assumption of the free boundary in B_1 implies

$$|a(t) - a_n(t)e_n| \leq C\eta,$$

so we may replace $a(t)\cdot(x - b(t)e_n)$ above by $a_n(t)(x_n - b(t))$.

The bounds on u above imply that $a_n(t)$ can vary at most $C\eta^{1/5}$ in an interval of length η^2. We can regularize $a_n(t)$ by averaging over such intervals (convolving with a mollifier) and the bounds for u still hold after changing the value of the constant C. Hence for all $t \in [-\lambda^{-1}\eta, 0]$, we can find $a_n(t) \in \mathbb{R}$ such that

$$a_n(t)\ \left(x_n - b(t) - C\eta^{\frac{6}{5}}\right)^+ \leq u \leq a_n(t)\ \left(x_n - b(t) + C\eta^{\frac{6}{5}}\right)^+ \tag{2.8}$$

in $B_{2C_0\eta} \times [t-\eta^2, t]$ with

$$c \le a_n(t) \le C, \qquad |a_n'(t)| \le C\eta^{\frac{1}{5}-2}, \qquad |b(t)| \le C_0\eta. \tag{2.9}$$

It remains to show that we can modify b slightly so that it satisfies the ODE $b' = -\lambda a_n$. Precisely, we let

$$\tilde{b}'(t) = -\lambda a_n(t), \quad \tilde{b}(0) = 0,$$

and we show that

$$|b(t) - \tilde{b}(t)| \le C\eta^{1+\beta} \quad \text{if} \quad t \in [-\lambda^{-1}\eta, 0], \qquad \beta = 1/10. \tag{2.10}$$

For this we perturb the family of evolving planes $a_n(t)(x_n - \tilde{b}(t))^+$ into a subsolution/supersolution. Let

$$d(t) := \tilde{b}(t) + C_1\eta^\beta\lambda t,$$

with C_1 large, to be specified later. We claim that

$$b(t) \ge d(t) - 2\eta^{1+\beta}. \tag{2.11}$$

For this we define the function

$$v := (1 - C_2\eta^\beta)\, a_n(t)\, (h(x - d(t)e_n))^+,$$

with

$$h(x) := x_n - \eta^{\beta-1}(|x'|^2 - 2nx_n^2),$$

and check that it is a subsolution to our problem (2.1) in the domain

$$\Omega := \bigcup_{t\in[-\lambda^{-1}\eta,0]} B_{2\eta}(d(t)e_n) \times \{t\}.$$

Notice that in a ball of radius 2η,

$$h \le C\eta, \qquad |\nabla h| = 1 + O(\eta^\beta), \tag{2.12}$$

and the constant $C_2 = C_2(n)$ is chosen depending only on n such that

$$v \le a_n(t)(x_n - d(t))^+, \tag{2.13}$$

with equality at $d(t)e_n$ and moreover, when $x \in \partial B_{2\eta}(d(t)e_n) \cap \{v(x,t) > 0\}$, the difference between the two functions above is greater than $\eta^{1+\beta}$.

Next, we check that v is a strict subsolution. In the interior $\{v > 0\}$, using (2.9), (2.12), the definition of $\tilde{b}$, we have (for η small)

$$|v_t| \le C|a_n'|\eta + C|d'| \le C\eta^{-4/5}, \qquad \triangle v \ge c\eta^{\beta-1} > v_t,$$

and on the free boundary (C' depending only on C_2, n),

$$v_t = (1 - C_2\eta^\beta)a_n(-d')h_n, \qquad |\nabla v|^2 \ge (1 - C'\eta^\beta)a_n^2.$$

Since

$$h_n = 1 + O(\eta^\beta), \qquad (-d')a_n = \lambda a_n^2 - C_1\lambda a_n\eta^\beta,$$

we can choose C_1 large such that $v_t < \lambda|\nabla v|^2$.

If

$$b(t_0) < d(t_0) - 2\eta^{1+\beta} \text{ for some } t_0 \in [-\lambda^{-1}\eta, 0],$$

then by (2.8) and (2.13) we find that $v < u$ at time $t = t_0$ in $B_{2\eta}(d(t_0)e_n) \cap \overline{\{v > 0\}}$. On the other hand $v = u$ at the origin $(0,0)$. This means that as we increase t from t_0 to 0, the graph of $v(\cdot, t)$ in $\overline{B_{2\eta}(d(t)e_n)} \cap \overline{\{v > 0\}}$ will touch by below the graph of u for a first time t, and the contact must be an interior point to $B_{2\eta}(d(t)e_n)$

due to the properties (2.8),(2.13) of u and v (in particular the difference between $a_n(t)\,(x_n - d(t))^+$ and v is greater than $\eta^{1+\beta}$ on $\partial B_{2\eta}(d(t)e_n)$). This contact point is either on the free boundary $\partial\{v > 0\}$ or on the positivity set $\{v > 0\}$, and we reach a contradiction since v is a strict subsolution. The claim (2.11) is proved, hence

$$b(t) \geq \tilde{b}(t) - C\eta^{1+\beta} \quad \text{if} \qquad t \in [-\lambda^{-1}\eta, 0].$$

The opposite inequality is obtained similarly and the claim (2.10) holds. Then from (2.8) we deduce that for all $\eta \leq c$ small

$$a_n(t)\left(x_n - \tilde{b}(t) - \eta^{1+\beta'}\right)^+ \;\leq\; u \;\leq\; a_n(t)\left(x_n - \tilde{b}(t) + \eta^{1+\beta'}\right)^+$$

in $B_\eta \times [-\lambda^{-1}\eta, 0]$ with $\beta' = 1/20$ and

$$c \leq a_n(t) \leq C, \quad |a_n'(t)| \leq \eta^{\beta'-2}, \qquad \tilde{b}'(t) = -\lambda a_n(t), \quad \tilde{b}(0) = 0.$$

□

3. The Nonlinear problem

In this section, we use a standard Hodograph transform to reduce our Stefan problem (1.1) to an equivalent nonlinear problem with fixed boundary and oblique derivative boundary condition (see (3.4)).

Here and henceforth, for $n \geq 2$, given $r > 0$ we set

$$Q_r := (-r, r)^n, \qquad Q_r^+ := Q_r \cap \{x_n \geq 0\}, \quad Q_r(x_0) := x_0 + Q_r,$$

$$\mathcal{C}_r := (Q_r \cap \{x_n > 0\}) \times (-r, 0], \quad \mathcal{F}_r := \{(x,t)|\;\; x \in Q_r \cap \{x_n = 0\},\;\; t \in (-r, 0]\}.$$

Also, by parabolic cylinders we mean

$$\mathcal{P}_r(x_0, t_0) := Q_r(x_0) \times (t_0 - r^2, t_0].$$

3.1. The Hodograph transform.

As mentioned above, we use a Hodograph transform to reduce the Stefan problem (1.1) to one with fixed boundary. Precisely, we view the graph of u in $\mathbb{R}^{n+2}$

$$\Gamma := \{(x, x_{n+1}, t)| \quad x_{n+1} = u(x_1, x_2, \ldots, x_n, t)\}$$

as the graph of a possibly multi-valued function $\bar{u}$ with respect to the x_n direction

$$\Gamma := \{(x, x_{n+1}, t)| \quad x_n = \bar{u}(x_1, x_2, \ldots, x_{n-1}, x_{n+1}, t)\}.$$

We use $(y_1, \ldots, y_n)$ to denote the coordinates $(x_1, x_2, \ldots, x_{n-1}, x_{n+1})$. Then, if Du and $D\bar{u}$ denote at some point on the graph Γ the gradients with respect to the first n entries of u and $\bar{u}$, we find

$$Du = -\frac{1}{\bar{u}_n}(\bar{u}_1, \ldots, \bar{u}_{n-1}, -1), \qquad u_t = -\frac{\bar{u}_t}{\bar{u}_n}$$

$$D^2 u = -\frac{1}{\bar{u}_n}\,(A(D\bar{u}))^T \;\, D^2\bar{u}\;A(D\bar{u}),$$

where $A(D\bar{u})$ is a square matrix which agrees with the identity matrix except on the nth row where the entries are given by the right hand side of Du above.

The Stefan problem (1.1) in terms of $\bar{u}$ can be written abstractly as the following quasilinear parabolic equation with oblique derivative boundary condition:

$$\begin{cases} \bar{u}_t = tr(\bar{A}(\nabla\bar{u})\, D^2\bar{u}) & \text{in } \{y_n > 0\}, \\ \bar{u}_t = g(\nabla\bar{u}) & \text{on } \{y_n = 0\}, \end{cases} \tag{3.1}$$

with $\bar{A}(p)$ symmetric, positive definite as long as $p_n \neq 0$, and $g_n(p) > 0$.

The free boundary of u is given by the graph of the trace of $\bar{u}$ on $\{y_n = 0\}$. Our goal becomes to show that $\bar{u}$ is $C^{1,\alpha}$ with respect to the y', t variables. Let us assume that u satisfies the hypotheses of Theorem 1.2 (it is now more convenient to work in cubes rather than in balls). Below we denote by c, C various constants depending on K and n. From the flatness assumption

$$|u - a_n(t)(x_n - b(t))^+| \leq C\epsilon_1\lambda \quad \text{in} \quad Q_\lambda \times [-\lambda, 0], \tag{3.2}$$

and $0 \in \partial\{u > 0\}$ implies $|b(0)| \leq C\epsilon_1\lambda$ which together with $|b'| \leq C\lambda$ gives

$$|b(t)| \leq C(\epsilon_1 + |t|)\lambda.$$

Thus, if $(x,t) \in Q_\lambda \times [-c\lambda, 0]$, then (for ϵ_1 possibly smaller), $|b(t)| \leq \lambda/2$ and by (3.2) the domain of definition of $\bar{u}$ at time t contains $Q^+_{\bar{c}\lambda}$ for $\bar{c}$ small enough. We conclude that $\bar{u}$ is well-defined in $Q^+_{\bar{\lambda}} \times [-\bar{\lambda}, 0]$, with $\bar{\lambda} := c_1\lambda$, c_1 sufficiently small.

Moreover, the graph of $\bar{u}$ in this set is closed in $\mathbb{R}^{n+2}$ (since it is obtained as a rigid motion from the graph of u) and it satisfies equation (3.1) in the viscosity sense, see Definition 3.2 below.

Remark 3.1. We observe that $\bar{u}$ is single-valued in the region $y_n \geq C\epsilon_1\lambda$, and possibly multi-valued near $y_n = 0$. Indeed, similarly as above, if $t \in [t_0 - \lambda^2, t_0 + \lambda^2]$, then using the bound for $|b'|$ and (1.4) for $|a'|$,

$$|a(t) - a(t_0)| \leq c_0, \qquad |b(t) - b(t_0)| \leq C\lambda^2,$$

hence, if λ_0, c_0 are smaller than ϵ_1 then

$$|u - a_n(t_0)(x_n - b(t_0))^+| \leq C\epsilon_1\lambda \quad \text{in} \quad Q_\lambda \times [t_0 - \lambda^2, t_0 + \lambda^2], \tag{3.3}$$

with $|b(t_0)| \leq \lambda/2$. By applying interior gradient estimates in parabolic cylinders included in $\{u > 0\}$ we find from (3.3) that if

$$(x_0, t_0) \quad \text{with} \quad x_0 \in Q_\lambda, \quad t_0 > -c\lambda \quad \text{is in the region} \quad C\epsilon_1\lambda \leq u(x_0, t_0) \leq c\lambda$$

then

$$|\nabla u(x_0, t_0) - a_n(t_0)e_n| \leq (2K)^{-1}.$$

Finally, the main hypotheses of Theorem 1.2 can be written in terms of $\bar{u}$ as

$$|\bar{u} - (\bar{a}_n(t)y_n + \bar{b}(t))| \leq C\epsilon_1\bar{\lambda} \quad \text{in} \quad Q^+_{\bar{\lambda}} \times [-\bar{\lambda}, 0],$$

$$\bar{b}'(t) = g(\bar{a}_n(t)e_n), \qquad K^{-1} \leq \bar{a}_n \leq K,$$

$$\bar{\lambda} \leq \bar{\lambda}_1, \qquad |\bar{a}_n'| \leq \bar{c}_1\bar{\lambda}^{-2}.$$

Our purpose in this paper is to prove an improvement of flatness result for solutions of the nonlinear equation (3.1) as above, provided that ϵ_1, $\bar{\lambda}_1$, $\bar{c}_1$ are chosen small depending on n and K (see Proposition 4.1 in the next section). Then Theorem 1.2 can be obtained by iterating such statement.

3.2. **Assumptions on the nonlinear problem.** We consider solutions to the following problem (for simplicity of notation we drop the bars in our formulation, and we use x rather than y),

$$\begin{cases} u_t = F(\nabla u, D^2 u) & \text{in } \mathcal{C}_\lambda, \\ u_t = g(\nabla u) & \text{on } \mathcal{F}_\lambda. \end{cases} \tag{3.4}$$

We assume that F is linear in D^2u, that is $F(\nabla u, D^2u) = tr(A(\nabla u)D^2u)$ and $g_n > 0$.

We start by stating precisely the notion of viscosity solution. First we write it for continuous functions and then adapt it to include possibly multi-valued functions u whose graphs are compact sets of $\mathbb{R}^{n+2}$, which is relevant to our setting.

Definition 3.2. We say that a continuous function $u : \overline{\mathcal{C}}_\lambda \to \mathbb{R}$ is a *viscosity subsolution* to (3.4) if u cannot be touched by above at points in $\mathcal{C}_\lambda \cup \mathcal{F}_\lambda$ (locally, in parabolic cylinders) by a strict C^2 supersolutions φ of (3.4).

Precisely, we require that there do not exist points $(x_0, t_0) \in \mathcal{C}_\lambda \cup \mathcal{F}_\lambda$ and test functions $\varphi \in C^2(\mathcal{P}_r(x_0, t_0))$ that satisfy

$$\begin{cases} \varphi_t > F(\nabla \varphi, D^2 \varphi) & \text{in } \mathcal{P}_r(x_0,t_0) \\ \varphi_t > g(\nabla \varphi) & \text{on } \mathcal{F}_\lambda \cap \mathcal{P}_r(x_0,t_0). \end{cases} \tag{3.5}$$

such that

$$u(x_0,t_0) = \varphi(x_0,t_0), \quad u \le \varphi \quad \text{in} \quad \mathcal{P}_r(x_0,t_0). \tag{3.6}$$

Similarly we can define *viscosity supersolutions* and *viscosity solutions* to (3.4).

We extend this definition to multi-valued functions u, and require they still satisfy the comparison with respect to (single-valued) test functions φ.

Definition 3.3. Assume that $u : \overline{\mathcal{C}}_\lambda \to \mathbb{R}$ is a multi-valued function with compact graph in $\mathbb{R}^{n+2}$. We say that u is a *viscosity subsolution* to (3.4) if the definition above holds and (3.6) is understood as $\varphi(x_0,t_0) \in u(x_0,t_0)$ while the inequality $u \le \varphi \quad \text{in} \quad \mathcal{P}_r(x_0,t_0)$ means that $u(x,t) \le \varphi(x,t)$ for all possible values of u at (x,t), and for all $(x,t) \in \mathcal{P}_r(x_0,t_0)$.

We remark that this notion of viscosity solution for multi-valued functions is very weak. For example if we consider two single-valued functions $u_1 \le u_2$ with u_2 a subsolution and u_1 a supersolution, then the union of the 2 graphs, is a multi-valued solution according to Definition 3.3. In fact we can add to it any arbitrary closed set between the two graphs. However, in our analysis we only consider solutions which could be multi-valued near $\mathcal{F}_\lambda$ and single-valued farther away, which is a consequence of the flatness regime.

We define now a class of linear in x functions that we use throughout this paper to express the flatness condition.

Definition 3.4. We denote by $l_{a,b}(x,t)$ functions which for each fixed t are linear in the x variable, and whose coefficients in the x' variable are independent of t, and also so that $l_{a,b}$ satisfies the boundary condition in (3.4) on $\{x_n = 0\}$. More precisely,

$$l_{a,b}(x,t) := a(t) \cdot x + b(t),$$

with

$$a(t) := (a_1, \ldots, a_{n-1}, a_n(t)), \quad a_i \in \mathbb{R}, \;\; i = 1, \ldots, n-1,$$

and

$$b'(t) = g(a(t)).$$

Our main result is to show that if u is a viscosity solution of (3.4) which is possibly multi-valued near $\{x_n = 0\}$ and is well approximated by $l_{a,b}$ in a cylinder $\mathcal{C}_\lambda$, i.e.

$$|u - l_{a,b}| \le \epsilon\lambda \quad \text{in} \quad \mathcal{C}_\lambda,$$

then in a smaller cylinder $\mathcal{C}_{\tau\lambda}$ it can be approximated by another function $l_{\tilde{a},\tilde{b}}$ with an error $\epsilon_\tau = \epsilon\tau^\alpha$ that improved by a $C^{1,\alpha}$ scaling.

Before formulating this result rigorously in the next section, we state here the precise hypotheses on F and g. We assume that $F(p, M)$ is uniformly elliptic in M for each fixed slope $p \in \mathbb{R}^n$ with $p_n > 0$ and the ellipticity constants could degenerate as $p_n \to 0^+$ or $|p| \to \infty$. Precisely, for any given constant K large there exists Λ large depending on K such that

$$\Lambda I \ge D_M F(p, M) \ge \Lambda^{-1} I, \qquad \text{if} \quad p \in \mathcal{R}_K, \tag{3.7}$$

with

$$\mathcal{R}_K := B_K \cap \{p_n \ge K^{-1}\} \quad \subset \quad \mathbb{R}^n. \tag{3.8}$$

We choose K sufficiently large such that when p is restricted to the set above we also have

$$|D_p F| \le \Lambda |M|, \quad \|g\|_{C^1} \le \Lambda, \quad g_n \ge \Lambda^{-1}. \tag{3.9}$$

From now on we assume that the constants K and Λ have been fixed such that (3.7)-(3.9) hold. In fact, for notational simplicity, by possibly choosing K larger, we can assume that (3.7)-(3.9) hold with $\Lambda = K$. We consider the situation when u is well approximated in $\mathcal{C}_\lambda$ by a function $l_{a,b}$ as above with slopes $a(t)$ belonging to the region $\mathcal{R}_K$.

We suppose in addition that u satisfies the Harnack inequality from scale λ to scale $\sigma\lambda$ where σ is a small parameter. We denote this property for u as property $H(\sigma)$ which is defined in the following way.

Definition 3.5. Given a positive constant σ small, we say that

$$u \text{ has property } H(\sigma) \text{ in } \mathcal{C}_\lambda$$

if u (possibly multi-valued) satisfies the following version of interior Harnack inequality in parabolic cylinders of size $r \in [\sigma\lambda, \lambda]$.

Let l denote a linear function

$$l(x) := a \cdot x + b, \qquad \text{with} \quad a \in \mathbb{R}^n, \;\; b \in \mathbb{R}, \quad |a| \le K.$$

If

$$u \ge l \quad \text{in} \quad Q_r(x_0) \times [t_0 - r^2, t_0 + r^2] \qquad \subset \quad \mathcal{C}_\lambda,$$

with $r \ge \sigma\lambda$, and

$$(u - l)(x_0, t_0) \ge \mu, \quad \text{for some} \quad \mu \ge 0,$$

then

$$u - l \ge \kappa\mu \quad \text{in} \quad Q_{r/2}(x_0) \times \left[t_0 + \frac{1}{2}r^2, t_0 + r^2\right],$$

for some constant κ depending on n and K (but independent of σ).

Similarly, if $u \le l$ we require these inequalities to hold for $l - u$ instead of $u - l$.

Property $H(\sigma)$ for all $\sigma > 0$ is a consequence of the parabolic Harnack inequality in the case when u is a sigle-valued viscosity solution of (3.4), and in addition we know that $\nabla u \in \mathcal{R}_K$. Property $H(\sigma)$ for a multi-valued solution of (3.4) roughly states that u behaves as a single-valued function from scale λ up to scale $\sigma\lambda$. In fact we will show in Remark 4.2 below that property $H(\sigma)$ (for some appropriate σ small) is satisfied for multi-valued solutions u which are graphical with respect to the e_n direction and are well approximated by the functions $l_{a,b}$.

4. The iterative statement

In this section, we state our main improvement of flatness result Proposition 4.1, and we show how Theorem 1.2 can be deduced from it. We also describe the strategy of the proof of Proposition 4.1, and its connection to the corresponding linearized problem (4.7).

The improvement of flatness statement reads as follows (we use the notation from Subsection 3.2). The rest of the paper will be devoted to its proof.

Proposition 4.1 (Improvement of flatness)**.** *Fix $K > 0$ large, and assume F,g satisfy* (3.7)-(3.9)*. Assume that u is a viscosity solution to* (3.4) *possibly multi-valued, which satisfies property $H(\epsilon^{1/2})$ and*

$$|u - l_{a,b}| \le \epsilon\lambda \quad \text{in } \overline{\mathcal{C}}_\lambda, \text{ with} \quad b'(t) = g(a(t)), \tag{4.1}$$

$$a(t) \in \mathcal{R}_K, \quad |a_n'(t)| \le \delta\epsilon\lambda^{-2},$$

and

$$\epsilon \le \epsilon_0, \quad \lambda \le \lambda_0, \quad \lambda \le \delta\epsilon.$$

Then there exists $l_{\tilde a,\tilde b}$ such that

$$|u - l_{\tilde a,\tilde b}| \le \frac{\epsilon}{2}\tau\lambda \quad \text{in } \overline{\mathcal{C}}_{\tau\lambda}, \qquad \tilde b'(t) = g(\tilde a(t)),$$

with

$$|a(t) - \tilde a(t)| \le C\epsilon, \qquad |\tilde a_n'(t)| \le \frac{\delta\epsilon}{2}(\tau\lambda)^{-2}.$$

Here the constants $\epsilon_0, \lambda_0, \delta, \tau > 0$ small and C large depend only on n, and K.

For the remainder of the section constants depending only on n and K are called universal, and denoted by c_i, C_i.

Remark 4.2. We apply the proposition above to the hodograph transform of a solution to the original Stefan problem, hence in our case u is graphical with respect to the e_n direction. Then (4.1) already implies our hypothesis that

$$u \text{ satisfies property } H(\epsilon^{1/2}) \text{ in } \mathcal{C}_\lambda.$$

Indeed, if $t \in [t_0 - \lambda^2, t_0 + \lambda^2]$, then using the bounds for $|a'|$, $|b'|$,

$$|a(t) - a(t_0)| \le \delta\epsilon, \qquad |b(t) - b(t_0)| \le C\lambda^2 \le C\delta\epsilon\lambda,$$

hence

$$|a(t_0)\cdot x + b(t_0) - l_{a,b}| \le C\delta\epsilon\lambda \quad \text{in} \quad Q_\lambda^+ \times [t_0 - \lambda^2, t_0 + \lambda^2]. \tag{4.2}$$

This shows that u is well approximated in each parabolic cylinder of size λ by a linear function which is constant in t,

$$|u - (a(t_0)\cdot x + b(t_0))| \le 2\epsilon\lambda \quad \text{in} \quad Q_\lambda^+ \times [t_0 - \lambda^2, t_0 + \lambda^2], \tag{4.3}$$

with $C \geq a_n(t_0) > c$. Since the graph of u coincides with the graph (in the e_n direction) of a solution to the heat equation, we can use the standard Harnack inequality for the heat equation and find that u satisfies property $H(C\epsilon)$ in $\mathcal{C}_\lambda$ (as we used interior regularity in Remark 3.1). Thus u satisfies property $H(\epsilon^{1/2})$ by choosing ϵ_0 smaller if necessary.

This argument shows that if u is graphical with respect to the e_n direction, then it is single-valued away from a $O(\epsilon\lambda)$ neighborhood of $\{x_n = 0\}$.

We now show that Proposition 4.1 implies Theorem 1.2, and the remainder of the paper will be devoted to prove Proposition 4.1.

Proof of Theorem 1.2. As discussed in Subsection 3.1, Theorem 1.2 is equivalent to obtaining $C^{1,\alpha}$ estimates on $\{x_n = 0\}$ for the hodograph transform. After relabeling constants if necessary, the hodograph transform does satisfy the hypotheses of Proposition 4.1 with $\epsilon = \epsilon_0$, $\lambda \leq \min\{\delta\epsilon_0, \lambda_0\}$, $a_0(t) = (0, 0, \dots, 0, (a_0)_n(t)) \in \mathcal{R}_{K/2}$. Now Proposition 4.1 can be applied indefinitely in the cylinders $\mathcal{C}_{\lambda_k}$, $\lambda_k := \lambda\tau^k$, with $\epsilon = \epsilon_k := \epsilon_0 2^{-k} = C(\lambda)\lambda_k^\alpha$. The hypothesis that $a_k(t) \in \mathcal{R}_K$ is satisfied (by choosing ϵ_0 smaller if necessary) since

$$|a_k(t) - a_{k-1}(t)| \leq C\epsilon_k, \qquad a_0(t) \in \mathcal{R}_{K/2},$$

from which we also deduce that

$$|a_k(t) - \nabla u(0, t)| \leq C\epsilon_k. \tag{4.4}$$

Hence

$$|u - l_{a_k,b_k}| \leq \epsilon_k\lambda_k \leq C(\lambda)\, \lambda_k^{1+\alpha} \qquad \text{in} \quad \mathcal{C}_{\lambda_k},$$

for all $k \geq 0$, and from (4.3) (applied for λ_k) and (4.4) we deduce that

$$|\nabla u(0, t) - \nabla u(0, s)| \leq C(\lambda)|t - s|^{\alpha/2},$$

which gives

$$|a_k(t) - a_k(s)| \leq C(\lambda)\lambda_k^{\alpha/2} \quad \text{if} \quad t, s \in [-\lambda_k, 0].$$

Using that $b_k' = g(a_k)$ we finally obtain

$$|u - (a_k(0) \cdot x + b_k'(0)t + b_k(0))| \leq C(\lambda)\lambda_k^{1+\frac{\alpha}{2}} \qquad \text{in} \quad \mathcal{C}_{\lambda_k},$$

which is the desired conclusion. □

4.1. **Strategy of the proof of the improvement of flatness.** We briefly explain the strategy of the proof of Proposition 4.1. The main idea is to linearize the equation near $l_{a,b}$. Define $w(x, t)$ the rescaled error by

$$u(x, t) := l_{a,b}(x, t) + \epsilon\lambda w\left(\frac{x}{\lambda}, \frac{t}{\lambda}\right), \quad (x, t) \in \mathcal{C}_\lambda. \tag{4.5}$$

Then w is defined in $\mathcal{C}_1$, possibly multi-valued near $\{x_n = 0\}$, and satisfies by hypothesis

$$|w| \leq 1 \quad \text{in} \quad \mathcal{C}_1,$$

and

$$\begin{cases} \lambda a_n'(\lambda t)x_n + b'(\lambda t) + \epsilon w_t(x, t) = F\left(a(\lambda t) + \epsilon\nabla w, \frac{\epsilon}{\lambda}D^2 w\right) & \text{in } \mathcal{C}_1, \\ \\ b'(\lambda t) + \epsilon w_t = g(a(\lambda t) + \epsilon\nabla w) & \text{on } \mathcal{F}_1. \end{cases} \tag{4.6}$$

We show that w is well approximated by a solution to the linear equation obtained formally by multiplying the first equation by $\lambda\epsilon^{-1}$ and the second by ϵ^{-1} and then letting $\epsilon \to 0$, $\delta \to 0$. Using $|a'| \leq \delta\epsilon\lambda^{-2}$, and $\lambda\epsilon^{-1} \leq \delta \to 0$ we obtain

$$\begin{cases} \lambda v_t = tr(A_\lambda(t) D^2 v) & \text{in } \mathcal{C}_1, \\ v_t = \gamma_\lambda(t) \cdot \nabla v & \text{on } \mathcal{F}_1, \end{cases} \tag{4.7}$$

with

$$A_\lambda(t) := A(a(\lambda t)), \qquad \gamma_\lambda(t) := \nabla g(a(\lambda t)).$$

Using that $A, g \in C^2(\mathcal{R}_K)$, and that $|a'| \ll \lambda^{-2}$ we find

$$|A'_\lambda(t)| \leq \lambda^{-1}, \qquad |\gamma'_\lambda(t)| \leq \lambda^{-1}.$$

The next sections are devoted to the study of the linear problem (4.7), and to obtain estimates which are uniform with respect to λ. To this aim, we introduce a distance d between points $(x, t) \in \mathbb{R}^{n+1}$

$$d((x,t),(y,s)) := \\ = \min\{|x' - y'| + |x_n - y_n| + |t - s|^{1/2}, \quad |x' - y'| + |x_n| + |y_n| + |t - s|\},$$

which is consistent with the scaling of the equation, so that d is equivalent with the standard Euclidean distance on the hyperplane $x_n = 0$ and with the standard parabolic distance far away from this hyperplane. The various Hölder estimates in the next section are written with respect to this distance d, or after a dilation of factor λ^{-1} with respect to the rescaled distance d_λ. In particular, this allows us to show that solutions v to the linear problem enjoy an improvement of flatness property in cylinders $\mathcal{C}_{\tau^k}$, which can be transferred further to the solutions of the nonlinear problem (4.6).

The relation between solutions w to (4.6) and v to (4.7) is made precise in the next proposition. It states that w satisfies essentially a comparison principle with C^2 subsolutions/supersolutions v of (4.7) which have bounded derivatives and second derivatives in x.

Proposition 4.3 (Comparison principle). *Let $v \in C^2(\overline{\Omega})$ with $\Omega \subset \mathcal{C}_1$ satisfy*

$$|\nabla v|, |D^2 v| \leq M,$$

for some large constant M and

$$\begin{cases} \lambda v_t \leq tr(A_\lambda(t) D^2 v) - C\delta & \text{in } \Omega, \\ v_t \leq \gamma_\lambda(t) \cdot \nabla v - \delta & \text{on } \mathcal{F}_1 \cap \overline{\Omega}, \end{cases} \tag{4.8}$$

with $A_\lambda(t)$, $\gamma_\lambda(t)$ as above.

Then v is a subsolution to (4.6), as long as C is sufficiently large, universal, and $\epsilon \leq \epsilon_1(\delta, M)$. In particular, if

$$v \leq w \quad \text{on} \quad \overline{\partial\Omega \setminus (\{t = 0\} \cup \{x_n = 0\})}$$

then

$$v \leq w \quad \text{in} \quad \Omega.$$

Similarly, we have the same result for supersolutions by replacing $\leq$ by $\geq$ and the $-$ signs in (4.8) by $+$.

Proof. It is straightforward to show that (4.8) implies the corresponding inequalities for v (in place of w) in (4.6). We need to use the hypotheses of Proposition 4.1 and that

$$\lambda\|a'\|_{L^\infty} + \|b'\|_{L^\infty} \le C, \quad |A(a(\lambda t) + \epsilon\nabla v) - A(a(\lambda t))| \le C\epsilon M,$$
$$|g(a(\lambda t) + \epsilon\nabla v) - g(a(\lambda t)) - \epsilon\nabla g(a(\lambda t)) \cdot \nabla v| \le C\epsilon^2 M^2.$$

□

As a consequence, we obtain that if the rescaled error w is close to a C^2 solution v of (4.7) on the *Dirichlet boundary* of a domain $\Omega \subset \mathcal{C}_1$ then v and w remain close to each other in the whole domain Ω.

Corollary 4.4. *Let w be a solution to* (4.6) *and $v \in C^2$ be a solution of* (4.7) *in a domain $\Omega \subset \mathcal{C}_1$, with*

$$|\nabla v|, |D^2 v| \le M.$$

If $\epsilon \le \epsilon_1(\delta, M)$ and

$$|v - w| \le \sigma \quad \text{on} \quad \overline{\partial\Omega \setminus (\{t = 0\} \cup \{x_n = 0\})}$$

then

$$|v - w| \le \sigma + C\delta \quad \text{in} \quad \Omega.$$

Proof. This follows immediately by applying Proposition 4.3 to

$$v \pm (C\delta(x_n^2 - t - 2) - \sigma).$$

□

We apply Proposition 4.3 and Corollary 4.4 to functions v for which M is large, universal. In order to apply Corollary 4.4 we need to show that w can be well approximated near the boundary of $\mathcal{C}_{1/2}$ by a solution v to (4.7) with bounded second derivatives in x. We prove that w has essentially a Hölder modulus of continuity (as $\delta \to 0$) with respect to the distance d_λ induced by d, and then we let v be the solution to the Dirichlet problem (4.7) in $\mathcal{C}_{1/2}$ with boundary data which is sufficiently close to w.

We conclude this section by stating a version of interior Harnack inequality for w with respect to constants, which is an immediate consequence of property $H(\epsilon^{1/2})$ of u in $\mathcal{C}_\lambda$, see Definition 3.5.

As in (4.2), the error between $l_{a,b}$ and a linear function independent of t in a time-interval of size $(\lambda r)^2$ is $C\delta\epsilon\lambda\, r^2$. Then Definition 3.5 implies the following property for $u - l_{a,b}$.

If for some constant ω

$$u - (\omega + l_{a,b}) \ge 0 \quad \text{in} \quad Q_{\lambda r}(x_0) \times [t_0 - (\lambda r)^2, t_0 + (\lambda r)^2] \subset \mathcal{C}_\lambda,$$

with $r \in [\epsilon^{1/2}, 1]$, and

$$(u - (\omega + l_{a,b}))(x_0, t_0) \ge \mu\epsilon\lambda, \quad \text{for some} \quad \mu \ge C\delta r^2,$$

then

$$u - (\omega + l_{a,b}) \ge \frac{\kappa}{2}\mu\epsilon\lambda \quad \text{in} \quad Q_{r\lambda/2}(x_0) \times \left[t_0 + \frac{1}{2}(\lambda r)^2, t_0 + (\lambda r)^2\right],$$

with κ the universal constant from Definition 3.5. In terms of w this can be written as follows.

Interior Harnack inequality for w. If

$$w \geq \omega \quad \text{in} \quad Q_r(x_0) \times [t_0 - \lambda r^2, t_0 + \lambda r^2] \subset \mathcal{C}_1,$$

with ω a constant, $r \geq \epsilon^{1/2}$, and

$$w(x_0, t_0) \geq \omega + \mu, \quad \text{for some} \quad \mu \geq C\delta r^2,$$

then

$$w \geq \omega + \frac{\kappa}{2}\mu \quad \text{in} \quad Q_{r/2}(x_0) \times \left[t_0 + \frac{\lambda}{2}r^2, t_0 + \lambda r^2\right]. \tag{4.9}$$

5. The linearized problem

In this section, we state various estimates for the linear problem (4.7) which are uniform in the parameter $\lambda \leq 1$ and we use them to prove our main result Proposition 4.1. We start with introducing the distance d_λ with respect to which our estimates are obtained.

5.1. **Definition of the distances d, d_λ and the family of balls $\mathcal{B}_r$, $\mathcal{B}_{\lambda,r}$.** We define the following distance in $\mathbb{R}^{n+1}$

$$d((x,t),(y,s)) : =$$
$$= \min\{|x' - y'| + |x_n - y_n| + |t - s|^{1/2}, \quad |x' - y'| + |x_n| + |y_n| + |t - s|\},$$

which interpolates between the parabolic distance and the standard one depending on how far points are from $\{x_n = 0\}$. It is not too difficult to check that d satisfies the triangle inequality.

For $r \leq 1$ and points (y,s) with $y_n \in [0,1]$, we define the family of "balls" of center (y,s) and radius r, which are backwards in time and restricted to $\{x_n \geq 0\}$, and which are consistent with the distance induced by d:

$$\mathcal{B}_r(y,s) := Q_r(y) \times (s - r^2, s), \qquad \text{if } r < |y_n|,$$
$$\mathcal{B}_r(y,s) := Q_r^+(y) \times (s - r, s), \qquad \text{if } 1 \geq r \geq |y_n|,$$

where we recall that

$$Q_r(y) := \{x \in \mathbb{R}^n |\, |x_i - y_i| < r\}, \qquad Q_r^+(y) := Q_r(y) \cap \{x_n \geq 0\}.$$

Notice that

$$(x,t) \in \mathcal{B}_{2r}(y,s) \setminus \mathcal{B}_r(y,s) \quad \Longrightarrow \quad d((x,t),(y,s)) \sim r.$$

A function $v : \overline{U} \to \mathbb{R}$, with $U \subset \mathcal{C}_1$, is Hölder with respect to the distance d if

$$[v]_{C^\alpha_d} := \sup_{(x,t) \neq (y,s)} |v(x,t) - v(y,s)|\; d((x,t),(y,s))^{-\alpha} < \infty.$$

Equivalently, $v \in C^\alpha_d(\overline{U})$ if and only if there exists M such that $\forall (x,t) \in \overline{U}$

$$osc\ v \leq M r^\alpha \quad \text{in} \quad \mathcal{B}_r(x,t) \cap \overline{U}.$$

Rescaling. Assume $\lambda \leq 1$ and we perform a dilation of factor λ^{-1} which maps Q^+_λ into Q^+_1. We use hyperbolic scaling for the rescaled distance d_λ of d

$$d_\lambda((x,t),(y,s)) := \frac{1}{\lambda} d(\lambda(x,t), \lambda(y,s))$$
$$= \min\{|x' - y'| + |x_n - y_n| + \lambda^{-1/2}|t - s|^{1/2}, |x' - y'| + |x_n| + |y_n| + |t - s|\}.$$

The corresponding family of balls induced by d_λ denoted by $\mathcal{B}_{\lambda,r}$ is obtained by dilating of a factor λ^{-1} the sizes of the balls $\mathcal{B}_r$ above and then relabeling $\lambda^{-1}r$ by r. We find

$$\mathcal{B}_{\lambda,r}(y,s) := Q_r(y) \times (s - \lambda r^2, s), \qquad \text{if } r < |y_n|,$$
$$\mathcal{B}_{\lambda,r}(y,s) := Q_r^+(y) \times (s - r, s), \qquad \text{if } \lambda^{-1} \geq r \geq |y_n|,$$

and notice that $\mathcal{B}_{\lambda,r}(y,s) = \mathcal{B}_r(y,s)$ if $y_n = 0$.

As above a function v is Hölder with respect to the distance d_λ in $\overline{U}$ and write $v \in C^\alpha_{d_\lambda}(\overline{U})$ if there exists M such that

$$osc\ v \leq Mr^\alpha \quad \text{in} \quad \mathcal{B}_{\lambda,r}(x,t) \cap \overline{U}.$$

5.2. **Estimates.** Having introduced the distance d_λ, we are now ready to state the estimates for the linear problem

$$\begin{cases} \lambda v_t = tr(A(t)D^2 v) & \text{in } \mathcal{C}_1, \\ v_t = \gamma(t) \cdot \nabla v & \text{on } \mathcal{F}_1, \end{cases} \tag{5.1}$$

with

$$K^{-1} I \leq A(t) \leq KI, \qquad K^{-1} \leq \gamma_n \leq K, \qquad |\gamma| \leq K$$
$$\lambda \in (0,1], \qquad |A'(t)| \leq \lambda^{-1}, \qquad |\gamma'(t)| \leq \lambda^{-1},$$

for some large constant K. Here constants depending on n and K are called universal.

We start with an interior regularity result (see Definition 3.4 of $l_{a,b}$).

Proposition 5.1 (Interior estimates)**.** *Let v be a viscosity solution to* (5.1) *such that $\|v\|_{L^\infty} \leq 1$. Then*

$$|\nabla v|,\ |D^2 v|\ \leq C \qquad \text{in} \quad \mathcal{C}_{1/2},$$

and for each $\rho \leq 1/2$, there exists $l_{\bar{a},\bar{b}}$ such that

$$|v - l_{\bar{a},\bar{b}}| \leq C\rho^{1+\alpha} \qquad \text{in } \mathcal{C}_\rho,$$

with

$$\bar{b}'(t) = \gamma(t) \cdot \bar{a}, \qquad |\bar{a}_n'| \leq C\rho^{\alpha-1}\lambda^{-1}, \qquad |\bar{a}| \leq C,$$

with α, C universal.

In terms of the Dirichlet problem for (5.1), we define the *Dirichlet boundary* of $\mathcal{C}_1$ as

$$\partial_D \mathcal{C}_1 := \partial \mathcal{C}_1 \cap \left(\{t = -1\} \cup \{x_n = 1\} \cup_{i=1}^{n-1} \{|x_i| = 1\}\right).$$

Notice that $\partial_D \mathcal{C}_1$ is different from the standard parabolic boundary since the points on $\mathcal{F}_1$ are also excluded.

Proposition 5.2 (The Dirichlet problem)**.** *Let ϕ be a continuous function on $\partial_D \mathcal{C}_1$. Then there exists a unique classical solution $v \in C^{2,1}(\mathcal{C}_1) \cap C^0(\bar{\mathcal{C}}_1)$ to the Dirichlet problem* (5.1) *with $v = \phi$ on $\partial_D \mathcal{C}_1$. Moreover,*

$$|\nabla v|, |D^2 v| \leq C(\sigma)\|v\|_{L^\infty} \quad \text{in} \quad C_1^\sigma := \{d_\lambda((x,t), \partial_D \mathcal{C}_1) \geq \sigma\},$$

and if ϕ is C^α with respect to the distance d_λ, then v is also C^α up to the boundary and

$$\|v\|_{C^\alpha_{d_\lambda}} \leq C\|\phi\|_{C^\alpha_{d_\lambda}},$$

with $C(\sigma)$, C universal constants (independent of λ).

Here

$$\|v\|_{C^\alpha_{d_\lambda}} := \|v\|_{L^\infty} + \sup_{(x,t)\neq(y,s)} |v(x,t)-v(y,s)|d_\lambda((x,t),(y,s))^{-\alpha}.$$

The proofs of Propositions 5.1 and 5.2 are based on a Harnack inequality for solutions to (5.1), which we provide in the next section. The Harnack inequality holds for more general equations of the same type with measurable coefficients. It applies also for solutions w to the nonlinear problem (4.6) up to scale $\epsilon^{1/2}$. To state it, we recall the definition of the maximal Pucci operators

$$\mathcal{M}^+_K(N) = \max_{K^{-1}I\leq A\leq KI} \quad tr\ AN, \qquad \mathcal{M}^-_K(N) = \min_{K^{-1}I\leq A\leq KI} \quad tr\ AN. \tag{5.2}$$

Theorem 5.3 (Hölder continuity)**.** *Let v be a viscosity solution to*

$$\begin{cases} \mathcal{M}^+_K(D^2v) \geq \lambda v_t \geq \mathcal{M}^-_K(D^2v) & in\ \mathcal{C}_1, \\ \\ K^{-1}v_n^- - Kv_n^+ - K|\nabla_{x'}v| \geq v_t \geq K^{-1}v_n^+ - Kv_n^- - K|\nabla_{x'}v| & on\ \mathcal{F}_1. \end{cases} \tag{5.3}$$

Then v is locally Hölder continuous in $\mathcal{C}_{1/2}$ with respect to the metric induced by d_λ, that is

$$\|v\|_{C^\alpha_{d_\lambda}(\mathcal{C}_{1/2})} \leq C\|v\|_{L^\infty(\mathcal{C}_1)}.$$

Moreover, if v is continuous up to the boundary and $v=\phi$ on $\partial_D\mathcal{C}_1$ with $\phi \in C^\alpha_{d_\lambda}$ then $v \in C^\alpha_{d_\lambda}$ up to the boundary and

$$\|v\|_{C^\alpha_{d_\lambda}} \leq C\|\phi\|_{C^\alpha_{d_\lambda}}.$$

The constants α and C depend only on n and K.

Proposition 5.4 (Harnack inequality for w)**.** *Assume that u satisfies the hypotheses of Proposition* 4.1 *and w is defined as in* (4.5)*. Then*

$$osc\ _{\mathcal{B}_{\lambda,r}(x_0,t_0)}\, w \leq Cr^\alpha, \qquad \forall (x_0,t_0)\in\mathcal{C}_{1/2}, \quad r\geq C(\delta)\epsilon^{1/2},$$

provided that $\delta \leq c'$ universal.

5.3. **Proof of Proposition** 4.1**.** Using the results above we can complete the proof of Proposition 4.1.

Proof of Proposition 4.1*.* We divide the proof in two steps.

Step 1. We prove that there exists a solution v to (4.7) which approximates w well in $\mathcal{C}_{1/2}$, that is

$$|v-w|\leq C\delta \quad \text{in } \mathcal{C}_{1/2},$$

provided that $\epsilon \leq \epsilon_1(\delta)$.

Indeed, by Proposition 5.4 we know that there exists a function ϕ defined in $\mathcal{C}_{1/2}$ such that

$$|w-\phi|\leq\delta, \qquad \|\phi\|_{C^\alpha_{d_\lambda}} \leq C. \tag{5.4}$$

Let v be the solution to (4.7) in $\mathcal{C}_{1/2}$ with $v=\phi$ on $\partial_D\mathcal{C}_{1/2}$, which exists in view of Proposition 5.2 and satisfies,

$$\|v\|_{C^\alpha_{d_\lambda}} \leq C. \tag{5.5}$$

Then, if $d_\lambda((x,t),\partial_D\mathcal{C}_{1/2}) \leq \delta^{1/\alpha}$, there exists (y,s) on $\partial_D\mathcal{C}_{1/2}$ so that (using (5.5) and (5.4)),

$$|v(x,t)-\phi(y,s)|\leq C\delta, \qquad |w(x,t)-\phi(y,s)|\leq C\delta,$$

thus,

$$|v-w| \le C\delta \quad \text{on} \quad \mathcal{C}_{1/2} \cap \{d_\lambda((x,t), \partial_D \mathcal{C}_{1/2}) \le \delta^{1/\alpha}\}. \tag{5.6}$$

In particular

$$|v-w| \le C\delta \quad \text{on} \quad \partial_D \Omega, \qquad \Omega := \mathcal{C}_{1/2} \cap \{d_\lambda((x,t), \partial_D \mathcal{C}_{1/2}) > \delta^{1/\alpha}\}.$$

On the other hand, by Proposition 5.2,

$$|\nabla v|, |D^2 v| \le C(\delta) \qquad \text{in} \quad \Omega.$$

Thus, using Corollary 4.4,

$$|v-w| \le C\delta \quad \text{in} \quad \Omega,$$

which gives the desired claim.

Step 2. Applying Proposition 5.1, to the solution v above, we find that

$$|w - l_{\bar{a},\bar{b}}| \le C\rho^{1+\alpha} + C\delta \quad \text{in} \quad \mathcal{C}_\rho,$$

and

$$\bar{b}'(t) = \gamma_\lambda(t) \cdot \bar{a}, \quad |\bar{a}_n'| \le C\rho^{\alpha-1}\lambda^{-1}, \qquad |\bar{a}| \le C,$$

with $\gamma_\lambda(t) = \nabla g(a(\lambda t))$. We choose $\rho = \tau$ small, universal, and

$$\delta = \tau^{1+\frac{\alpha}{2}},$$

so that $\delta \le c'$ the constant from Proposition 5.4, and

$$|w - l_{\bar{a},\bar{b}}| \le \frac{1}{4}\tau \quad \text{in } \mathcal{C}_\tau, \qquad |\bar{a}_n'| \le \frac{1}{4}\,\delta\,\tau^{-2}\lambda^{-1}.$$

In terms of the original function u, this inequality implies

$$\left| u - \left(l_{a,b} + \epsilon\lambda l_{\bar{a},\bar{b}}\left(\frac{x}{\lambda}, \frac{t}{\lambda}\right)\right)\right| = \epsilon\lambda \left| w\left(\frac{x}{\lambda}, \frac{t}{\lambda}\right) - l_{\bar{a},\bar{b}}\left(\frac{x}{\lambda}, \frac{t}{\lambda}\right)\right| \le \frac{\epsilon}{4}\tau\lambda \quad \text{in} \quad \mathcal{C}_{\tau\lambda}.$$

Set

$$\tilde{a}(t) := a(t) + \epsilon\,\bar{a}\left(\frac{t}{\lambda}\right), \quad \hat{b}(t) := b(t) + \epsilon\lambda\,\bar{b}\left(\frac{t}{\lambda}\right),$$

then

$$|u - l_{\tilde{a},\hat{b}}| \le \frac{\epsilon}{4}\tau\lambda \quad \text{in} \quad \mathcal{C}_{\tau\lambda},$$

and

$$|\tilde{a}_n'| \le \frac{\epsilon\delta}{\lambda^2}\left(1 + \frac{1}{4\tau^2}\right) \le \frac{\epsilon\delta}{2(\tau\lambda)^2}.$$

Finally, we define $\tilde{b}$ by the ODE

$$\tilde{b}' = g(\tilde{a}), \qquad \tilde{b}(0) = \hat{b}(0),$$

and then we have

$$\hat{b}' = b' + \epsilon\bar{b}'\left(\frac{t}{\lambda}\right) = g(a(t)) + \epsilon\nabla g(a(t)) \cdot \bar{a}\left(\frac{t}{\lambda}\right) = g(\tilde{a}(t)) + O(\epsilon^2) = \tilde{b}' + O(\epsilon^2).$$

If $t \in [-\tau\lambda, 0]$ then

$$|(\tilde{b} - \hat{b})(t)| \le C\epsilon^2|t| \le \frac{\epsilon}{4}\tau\lambda,$$

which implies the desired conclusion

$$|u - l_{\tilde{a},\tilde{b}}| \le \frac{\epsilon}{2}(\lambda\tau) \quad \text{in} \quad \mathcal{C}_{\tau\lambda},$$

and $\tilde{a}, \tilde{b}$ satisfy the required bounds. □

6. Harnack inequality

In this section, we prove Theorem 5.3 and Proposition 5.4. The key ingredient is to establish a diminishing of oscillation property. As usual, universal constants depend on n, K.

Proposition 6.1. *Assume that v is a viscosity solution of (5.3) and $0 \leq v \leq 1$ in $\mathcal{C}_1$. Then*

$$osc_{\mathcal{C}_{1/2}} v \leq 1 - c,$$

with $c > 0$ universal.

In order to prove Proposition 6.1 we start with a lemma. Let Ω be a smooth domain in $\mathbb{R}^n$, $n \geq 2$, such that

$$\bar{Q}^+_{3/4} \subset \bar{\Omega} \subset \bar{Q}^+_{7/8},$$

and call

$$T := \{x_n = 0\} \cap Q_{3/4} \subset \partial\Omega.$$

Define $\eta(x')$ a standard bump function supported on $Q'_{5/8}$ and equal 1 on $Q'_{1/2}$ (here the prime denotes cubes in $\mathbb{R}^{n-1}$). Let ϕ satisfy (see (5.2) for the definition of the Pucci operator),

$$\mathcal{M}^-_K(D^2\phi) = 0 \quad \text{in } \Omega,$$

$$\phi = 0 \quad \text{on } \partial\Omega \setminus T, \quad \phi = \eta \quad \text{on } T,$$

and notice that $0 \leq \phi \leq 1$, $\phi \geq c$ on $Q^+_{1/2}$, and by Hopf lemma $\phi_n > 0$ on $\{x_n = 0\} \cap \{\phi = 0\}$. The following lemma holds.

Lemma 6.2. *Let $v \geq 0$ satisfy*

$$\begin{cases} \mathcal{M}^+_K(D^2v) \geq \lambda v_t \geq \mathcal{M}^-_K(D^2v) & \text{in } \mathcal{C}_1, \\ v_t \geq K^{-1}v^+_n - Kv^-_n - K|\nabla_{x'}v| & \text{on } \mathcal{F}_1, \end{cases} \tag{6.1}$$

in the viscosity sense. If for some $t_0 \in (-1, 0]$,

$$v(x, t_0) \geq s_0\, \phi(x) \quad \text{in } Q^+_1, \quad s_0 \geq 0,$$

then

$$v(x,t) \geq s(t)\, \phi(x) \quad \text{in } Q^+_1 \times [t_0, 0],$$

with

$$s'(t) = -C_0 s(t), \quad s(t_0) = s_0, \qquad C_0 \text{ large universal.}$$

Moreover, if $s_0 \leq c_0$ with c_0 small universal, and

$$v\left(\frac{1}{2}e_n, t_0 + \lambda/4\right) \geq \frac{1}{2}, \tag{6.2}$$

then

$$v(x, t_0 + \lambda) \geq (s_0 + c_0\lambda)\phi(x).$$

Proof. For the first part of the claim, since $v \geq 0$, it suffices to show that with our choice of s,

$$w(x,t) := s(t)\phi(x),$$

is a subsolution to (6.1) in $\Omega \times [t_0, 0]$, that is

$$\begin{cases} \lambda w_t \leq \mathcal{M}_K^-(D^2 w) & \text{in } \Omega \times (t_0, 0], \\ w_t \leq K^{-1} w_n^+ - K w_n^- - K|\nabla_{x'} w| & \text{on } \{x_n = 0\} \cap (\Omega \times (t_0, 0]). \end{cases}$$

The interior equation is immediately satisfied since $s' \leq 0$ and $s \geq 0$. On $\{x_n = 0\}$, we need to show that

$$C\phi + K^{-1}\phi_n^+ - K\phi_n^- - K|\nabla_{x'}\phi| \geq 0,$$

for some large C. By Hopf lemma $\phi_n > 0$ on $\{\phi = 0\} \cap \{x_n = 0\}$ and moreover $|\nabla_{x'}\phi| = 0$, thus

$$K^{-1}\phi_n^+ - K\phi_n^- - K|\nabla_{x'}\phi| = K^{-1}\phi_n > 0 \quad \text{on} \quad \{\phi = 0\} \cap \{x_n = 0\}.$$

The same holds in a neighborhood of this set by continuity, and then we can choose C sufficiently large so that the desired inequality holds.

For the second part, denote for simplicity

$$t_i := t_0 + i\frac{\lambda}{4}, \quad i = 1, \ldots, 4.$$

We define

$$D := \{x \in \Omega | \quad d(x, \partial\Omega) > c\} \subset \Omega,$$

with c small universal such that there exists a C^2 function $\psi \geq 0$ defined in $\Omega \setminus D$ satisfying

$$\mathcal{M}_K^-(D^2\psi) \geq 4 \quad \text{in} \quad \Omega \setminus D,$$

and

$$\psi = 0, \quad |\nabla\psi| \geq 1 \quad \text{on } \partial\Omega, \quad \psi \leq 1 \quad \text{on } \partial D.$$

An example of such a function is given by $\psi = d + Cd^2$ with C sufficiently large, where d is the distance function to $\partial\Omega$. In view of (6.2)

$$v\left(\frac{1}{2}e_n, t_1\right) \geq 1/2.$$

Thus, we can use Harnack inequality (after rescaling) to conclude that

$$v \geq 2c_1 \quad \text{on } D \times [t_2, t_4], \tag{6.3}$$

for some small c_1. We claim that at time $t = t_3$,

$$v(x, t_3) \geq s(t_3)\phi + c_1\psi \quad \text{in} \quad \Omega \setminus D. \tag{6.4}$$

For this we compare v in $(\Omega \setminus D) \times [t_2, t_3]$ with

$$q(x,t) := s(t_3)\phi + c_1\left(\psi + \frac{t - t_3}{t_3 - t_2}\right).$$

The inequality $q \leq v$ holds on the boundary of the domain. Indeed (recall that s is decreasing), on ∂D

$$q(x,t) \leq s(t_3)\phi + c_1 \leq s_0 + c_1 \leq 2c_1 \leq v,$$

where in the last inequality we used (6.3), and on $\partial\Omega$ or at $t = t_2$ we have $q \leq s(t_3)\phi \leq v$.

It remains to check that q is a subsolution for the interior equation. Indeed,

$$\lambda q_t = 4c_1 \leq c_1 \mathcal{M}_K^-(D^2\psi) \leq \mathcal{M}_K^-(D^2 q),$$

where we used that $\mathcal{M}_K^-(N_1) + \mathcal{M}_K^-(N_2) \leq \mathcal{M}_K^-(N_1 + N_2)$, and claim (6.4) is proved.

Next, in the domain $(\Omega \setminus D) \times [t_3, t_4]$ we compare v with the subsolution

$$z(x,t) := (s(t_3) + c_2(t - t_3))\phi(x) + c_1\psi(x),$$

with c_2 sufficiently small.

The inequality $v \geq z$ is satisfied at time $t = t_3$ by (6.4), and on ∂D we have

$$z \leq s_0 + c_2 + c_1 \leq 2c_1 \leq v,$$

while on $\partial\Omega \setminus \{x_n = 0\}$ we have $z = 0 \leq v$. We check that z is a subsolution of our problem. For the interior inequality we have

$$\lambda z_t = c_2 \lambda \phi \leq c_2 \leq c_1 \mathcal{M}_K^-(D^2\psi) \leq \mathcal{M}_K^-(D^2 z).$$

For the boundary condition, on $\{x_n = 0\}$ we get

$$z_t = c_2\phi \leq c_2 \leq \frac{c_1}{4} K^{-1}\psi_n, \tag{6.5}$$

where in the second inequality we have used that $\psi_n \geq 1$ on $\partial\Omega \cap \{x_n = 0\}$. Moreover, since $\phi_n \geq -C$ on $\partial\Omega \cap \{x_n = 0\}$, we get (for s_0, c_2 small enough),

$$z_n \geq -\left(s_0 + c_2\frac{\lambda}{4}\right) C + c_1\psi_n \geq \frac{c_1}{2}\psi_n,$$

and finally ($|\nabla_{x'}\psi| = 0$ on $\{x_n = 0\}$)

$$K|\nabla_{x'} z| \leq \left(s_0 + \frac{c_2}{4}\right) K|\nabla_{x'}\phi| \leq \frac{c_1}{4} K^{-1}\psi_n.$$

Together with (6.5), this gives

$$z_t = c_2\phi \leq c_2 \leq K^{-1} z_n - K|\nabla_{x'} z| \quad \text{on } \{x_n = 0\}.$$

In conclusion, at time $t = t_4$ we have $v \geq z$ in $\Omega \setminus D$ and $v \geq 2c_1$ in D which gives the desired claim by choosing c_0 sufficiently small. $\square$

Remark 6.3. In the proof above we only used the subsolution property for v

$$\mathcal{M}_K^+(D^2 v) \geq \lambda v_t, \tag{6.6}$$

in order to extend the inequality (6.2) from one point to (6.3) by applying the interior parabolic Harnack inequality. Alternately, it is sufficient to assume that the Harnack inequality holds for v only in a neighborhood of D and not necessarily up to $\{x_n = 0\}$.

The rest of the proof is based on comparing v with the explicit C^2 subsolutions w, q and z which all have bounded second derivatives in the x variable. Thus the hypothesis that v is a viscosity supersolution of (6.1) can be slightly relaxed, and require instead, that v only satisfies the comparison principle with respect to the explicit barriers above.

Remark 6.4. The hypothesis (6.6) can be removed completely if instead of (6.2) we assume a measure estimate

$$\left|\left\{v \geq \frac{1}{4}\right\} \cap \left(Q_1 \times \left[t_0, t_0 + \frac{\lambda}{4}\right]\right)\right| \geq \frac{1}{2}\left|Q_1 \times \left[t_0, t_0 + \frac{\lambda}{4}\right]\right|.$$

Then, the inequality (6.3) follows directly from the supersolution property for v and the weak Harnack inequality (see for example [W]).

We are now ready to prove Proposition 6.1.

Proof of Proposition 6.1. Assume that $0 \leq v \leq 1$, and for half of the values of

$$t_k := -1 + k\lambda, \qquad \text{so that} \quad t_k \in [-1, -1/2), \quad k = 0, 1, 2, \ldots,$$

we have

$$v\left(\frac{1}{2}e_n, t_k + \lambda/4\right) \geq \frac{1}{2}. \tag{6.7}$$

We apply Lemma 6.2 repeatedly to the sequence of times t_k and obtain

$$v(x, t_k) \geq s_k \phi, \quad s_k := s(t_k), \qquad s_0 = 0,$$

with ϕ given in Lemma 6.2, and

$$s_{k+1} \geq s_k + c_0 \lambda \qquad \text{if (6.7) holds and } s_k \leq c_0,$$

or

$$s_{k+1} \geq s_k(1 - C_0\lambda) \quad \text{otherwise.}$$

Now it follows that $s_k \geq c_1$ for the last value of k so that $t_k < -1/2$, for c_1 appropriately chosen depending on c_0, C_0. Then we apply the first part of Lemma 6.2 to obtain

$$v(x, t) \geq \bar{c}\phi \quad \text{for all } t \geq -1/2,$$

which gives the desired conclusion, since $\phi > c$ on $Q^+_{1/2}$. □

The same arguments show that a similar statement to that of Proposition 6.1 holds for a solution w of (4.6) defined in (4.5). Below is the key lemma which connects the linear and nonlinear problem and allows us to reduce our analysis mostly to the linear case.

Lemma 6.5. *Assume that u satisfies the hypotheses of Proposition* 4.1 *and let w be defined as in* (4.5)*, with $-1 \leq w \leq 1$. Then*

$$osc_{\,\mathcal{C}_{1/2}} w \leq 2(1 - c),$$

with c universal, provided that $\delta \leq c'$ and $\epsilon \leq \epsilon_1(\delta)$.

Proof. We may assume as above that $w(e_n/2, t_k + \lambda/4) \geq 0$ for more than half the values of k, and then show that w separates from the lower constraint -1. For this we apply the same argument as above to the function

$$\bar{w} := w + 1 + C\delta(2 + t - x_n^2) \geq 0,$$

for which the relaxed hypotheses of Remark 6.3 hold. Indeed, by (4.9), $\bar{w}$ satisfies the required Harnack inequality (6.2) $\Longrightarrow$ (6.3) and, by Proposition 4.3, it satisfies the comparison with the explicit barriers of Lemma 6.2.

We remark that we have only used that u has property $H(c'')$ in $\mathcal{C}_\lambda$ for some c'' small, universal. □

Before we proceed with the proofs of Theorem 5.3 and Proposition 5.4 we provide a boundary version of the diminishing of oscillation Proposition 6.1.

Lemma 6.6. *Assume that U is a space-time domain obtained by the intersection of $n+1$ half spaces in the $x_1, \dots, x_{n-1}, x_n$ and t variables,*

$$U := (-\infty, z_1) \times (-\infty, z_2) \times \cdots \times (-\infty, z_n) \times (-z_{n+1}, \infty) \quad \subset \mathbb{R}^{n+1},$$

with $z_i \in [0,1]$.

Assume that $v \geq 0$ satisfies

$$\begin{cases} \lambda v_t \geq \mathcal{M}_K^-(D^2 v) & \text{in } \mathcal{C}_1 \cap U, \\ v_t \geq K^{-1} v_n^+ - K v_n^- - K|\nabla_{x'} v| & \text{on } \mathcal{F}_1 \cap U, \\ v \geq \frac{1}{4} & \text{on } \partial U \cap \mathcal{C}_1. \end{cases} \tag{6.8}$$

If $\min z_i \leq \frac{7}{8}$, then

$$v \geq c \quad \text{in } \mathcal{C}_{1/2} \cap U, \qquad c \text{ universal.}$$

Proof. This follows easily from Lemma 6.2. Indeed, we work with the truncation $\tilde{v} := \min\{v, \frac{1}{4}\}$ extended by $\frac{1}{4}$ in $\mathcal{C}_1 \setminus U$. Then $\tilde{v}$ is a supersolution for our problem in $\mathcal{C}_1$.

If $z_{n+1} < 1$, then we can apply directly the first part of Lemma 6.2 for $\tilde{v}$ for some t_0 close to -1 and for s_0 universal, and obtain the desired conclusion.

On the other hand, if $z_{n+1} = 1$, then $z_i \leq \frac{7}{8}$ for some $i \leq n$ hence for each time $t \in [-1, 0]$ we find

$$\left| \left\{ \tilde{v} \geq \frac{1}{4} \right\} \cap Q_1 \right| \geq c|Q_1|.$$

Now the conclusion follows as before, see Remark 6.3. $\square$

We are now ready to prove Theorem 5.3.

Proof of Theorem 5.3. Notice that the rescaling of v

$$v_r(x,t) = v(r\,x, r\,t), \qquad r \leq 1,$$

satisfies again the hypotheses of Theorem 5.3 in $\mathcal{C}_1$ with the constant λ replaced by $\lambda_r = \lambda r$. Proposition 6.1 applied to v_r implies that

$$osc_{\,\mathcal{C}_{1/2}}\, v_r \leq (1-c) osc_{\,\mathcal{C}_1}\, v_r$$

which gives (recall that $\mathcal{B}_{\lambda,r}(y,s) = \mathcal{B}_r(y,s)$ if $y_n = 0$),

$$osc_{\,\mathcal{B}_{r/2}(0,0)}\, v \leq (1-c) osc_{\,\mathcal{B}_r(0,0)}\, v.$$

Similarly, if $(y,s) \in \overline{\mathcal{C}}_{1/2} \cap \{x_n = 0\}$, then by considering cylinders centered at (y,s) we obtain

$$osc_{\,\mathcal{B}_{r/2}(y,s)}\, v \leq (1-c) osc_{\,\mathcal{B}_r(y,s)}\, v, \qquad \forall r \leq 1/2, \tag{6.9}$$

which proves the desired oscillation decay on $\{x_n = 0\} \cap \overline{\mathcal{C}}_{1/2}$.

If $(y,s) \in \mathcal{C}_{1/2}$, then (6.9) applied at $((y',0),s)$ implies

$$osc_{\,\mathcal{B}_{\lambda,r/8}(y,s)}\, v \leq (1-c) osc_{\,\mathcal{B}_{\lambda,r}(y,s)}\, v, \qquad \text{if} \quad y_n \leq r \leq 1/4.$$

In the case when $r < y_n$, then the inequality above follows from the standard parabolic Harnack inequality applied to v in the interior cylinder $\mathcal{B}_{\lambda,r}(y,s)$.

The boundary version follows in the same way. Precisely, if $(y,s) \in \overline{\mathcal{C}}_1 \cap \{x_n = 0\}$ then we find

$$osc_{\,\mathcal{B}_{r/2}(y,s) \cap \overline{\mathcal{C}}_1}\, v \leq (1-c) osc_{\,\mathcal{B}_r(y,s) \cap \overline{\mathcal{C}}_1}\, v, \qquad \forall r \leq 1,$$

by applying either Proposition 6.1 or Lemma 6.6 depending whether or not $\mathcal{B}_{\lambda,r}(y,s)$ intersects the boundary $\partial_D \mathcal{C}_1$.

The inequality above can be deduced at all points $(y,s) \in \overline{\mathcal{C}}_1$ after replacing $r/2$ by $r/8$ on the left hand side. Indeed, if $r \geq y_n$ then it follows from the inequality above applied at the point $((y',0),s)$, and if $r < y_n$ then we can apply the standard parabolic Harnack inequality or its boundary version since $\mathcal{B}_{\lambda,r}(y,s)$ does not intersect $\{x_n = 0\}$. □

We conclude the section with the proof of Proposition 5.4, that is the Harnack inequality for w.

Proof of Proposition 5.4. By Lemma 6.5 we find that, in terms of u, we satisfy again the hypotheses of Proposition 4.1 in $\mathcal{C}_{\lambda/2}$ with λ replaced by $\lambda/2$, ϵ replaced by $2(1-c)\epsilon$, and with δ the same. The function a stays the same while b is modified by a small constant. Moreover, the property $H(\epsilon^{1/2})$ of u in $\mathcal{C}_\lambda$ implies that u satisfies property $H(2\epsilon^{1/2})$ in $\mathcal{C}_{\lambda/2}$. We can iterate this result k times as long as the scale parameter of the property $H(2^k\epsilon^{1/2})$ remains small, universal, and the hypotheses of Lemma 6.5 hold:

$$2^k\epsilon^{1/2} \leq c'', \qquad \delta \leq c', \qquad 2^k(1-c)^k\epsilon \leq \epsilon_1(\delta),$$

with c'' small, universal. This means that we can iterate k times if

$$2^k\epsilon^{1/2} \leq \epsilon_2(\delta), \qquad \delta \leq c'.$$

In terms of w, we obtain that its oscillation in $\mathcal{C}_{2^{-k}}$ is bounded by $2(1-c)^k$ as long as k satisfies the inequality above. On the other hand for the interior balls $\mathcal{B}_{\lambda,r}$, by (4.9), w satisfies a similar diminishing of oscillation up to scale $r \sim \epsilon^{1/2}$, and the conclusion follows. □

7. Proof of Proposition 5.1

In this section, we prove Proposition 5.1 by using Theorem 5.3 and the estimates for the one-dimensional problem which will be proved in Lemma 8.1 of the next section. The constants C in this proof depend on n and K.

Proof of Proposition 5.1. The proof is divided in four steps.

Step 1 - Interior Estimates. Let $(y,s) \in \mathcal{C}_{1/2}$. From Theorem 5.3 we know that

$$osc_{\ \mathcal{B}_r(y,s)} v \leq Cr^\alpha, \qquad r = y_n.$$

The rescaling

$$\tilde{v}(x,t) := v(y + rx, s + r^2\lambda t),$$

solves in $Q_1 \times (-1,0)$

$$\tilde{v}_t = tr(\tilde{A}(t)D^2\tilde{v}), \qquad \tilde{A}(t) := A(s + r^2\lambda t).$$

Since $|A'| \leq \lambda^{-1}$, we have $|\tilde{A}'(t)| \leq C$, and we find by interior estimates that $|\tilde{v}_n(0,0)| \leq C\ osc_{Q_1\times(-1,0)}\ \tilde{v}$, from which we deduce

$$|v_n(y,s)| \leq Cr^{\alpha-1} = Cy_n^{\alpha-1}.$$

On the other hand, we prove in appendix that the difference of two viscosity solutions is still a viscosity solution. Thus, the estimates for v can be extended to the derivatives of v in the x_i directions, $i = 1,\ldots,n-1$. Indeed, by applying the interior Hölder estimates to discrete differences in the x_i directions, and iterating this we find that

$$\|D^k_{x'}v\| \le C(k) \quad \text{in } \mathcal{C}_{1/2}, \quad \forall k \ge 1.$$

In particular, using also the estimate for v_n above, we obtain

$$\|D^2_{x'}v\| \le C, \qquad |v_{in}| \le C x_n^{\alpha-1} \quad \text{in } \mathcal{C}_{1/2}.$$

Step 2 - Reduction to 1D. Combining the interior estimates with our assumptions on γ, we obtain that when we restrict v to a two-dimensional space in which we freeze the x' variable, say for simplicity $x'=0$, then the function $v((0,x_n),t)$ solves in the x_n, t variables the equation

$$\begin{cases} v_t = \frac{1}{\lambda}\{a^{nn}(t)v_{nn} + h(x_n,t)\} & \text{in } \mathcal{C}_1, \\ v_t = \gamma_n(t)v_n + f(t) & \text{on } \mathcal{F}_1, \end{cases} \tag{7.1}$$

with

$$|h| \le C x_n^{\alpha-1}, \quad |f(t)| \le C,$$
$$h(x_n,t) := \sum_{(i,j)\neq(n,n)} a^{ij}(t)\ v_{ij}((0,x_n),t), \qquad f(t) := \sum_{i<n} \gamma_i(t)\ v_i(0,t).$$

The boundary condition on $\mathcal{F}_1$ is understood in the viscosity sense.

Indeed, if a C^1 function $\varphi(x_n,t)$ touches $v(0,x_n,t)$ by above/below, say at $(0,0)$, in $\mathcal{B}_r(0,0) \subset \mathbb{R}^2$, then

$$\varphi(x_n,t) + \sum_{i<n} v_i(0,0)x_i \pm C|(x,t)|^{1+\alpha}$$

touches v by above/below at the origin in $\mathcal{B}_r(0,0) \subset \mathbb{R}^{n+1}$. This follows from the C^α continuity of v_i, $i<n$, which implies

$$\left| v(x,t) - \left(v(0,x_n,t) + \sum_{i<n} v_i(0,0)x_i \right) \right| \le C|(x,t)|^{1+\alpha}. \tag{7.2}$$

Now, we can use Lemma 8.1 a) for $v(0,x_n,t)$, where we establish $C^{1,\alpha}$ estimates for the 1D problem (7.1). We obtain

$$|v((0,x_n),t) - v(0,t) - v_n(0,t)x_n| \le C x_n^{1+\alpha},$$

which together with (7.2) gives

$$\left| v - \left(v(0,t) + v_n(0,t)x_n + \sum_{i=1}^{n-1} v_i(0,0)x_i \right) \right| \le C\rho^{1+\alpha} \quad \text{in } \mathcal{C}_\rho.$$

This means that

$$|v - l_{a,b}| \le C\rho^{1+\alpha} \quad \text{in } \mathcal{C}_\rho,$$

with

$$a(t) := (v_1(0,0), \dots, v_{n-1}(0,0), v_n(0,t)), \quad b(t) := v(0,t),$$

and

$$b' = \gamma_n(t)a_n + f(t) = \gamma(t)\cdot a + \sum_{i<n} \gamma_i(t)(v_i(0,t) - v_i(0,0)).$$

Step 3 - Modifying the linear approximation. Next, we modify a and b slightly into $\bar a$, $\bar b$ so that

$$|v - l_{\bar a,\bar b}| \le C\rho^{1+\alpha} \quad \text{in } \mathcal{C}_\rho,$$

and we also satisfy

$$|\bar{a}'(t)| \le C\lambda^{-1}\rho^{\alpha-2}, \qquad \bar{b}' = \gamma(t)\cdot\bar{a}. \tag{7.3}$$

By Lemma 8.1 we know that

$$|a_n(t) - a_n(s)| \le C\lambda^{-\frac{\alpha}{2}}|t-s|^{\frac{\alpha}{2}}, \tag{7.4}$$

and by the Hölder continuity of the v_i's,

$$|b' - \gamma(t)\cdot a| \le \sum_{i<n} |\gamma_i||v_i(0,t) - v_i(0,0)| \le C|t|^\alpha. \tag{7.5}$$

Thus, a_n oscillates $C\rho^\alpha$ in an interval of length $\lambda\rho^2$. We define $\bar{a}$ by averaging a over intervals of this length. More precisely, let η be a standard mollifier in $\mathbb{R}$ with compact support in $[-1,1]$, and η_τ denote its rescaling with support of size τ. We extend $a_n(t)$ to be constant for $t \ge 0$ and define

$$\bar{a}_n := a_n * \eta_{\lambda\rho^2}, \qquad \bar{a}_i := a_i, \quad i = 1,\dots,n-1.$$

Then (7.4) implies the inequality (7.3) for $\bar{a}'$ and also

$$|a - \bar{a}| \le C\rho^\alpha. \tag{7.6}$$

We define $\bar{b}(t)$ for $t \le 0$ as

$$\bar{b}' = \gamma(t)\cdot\bar{a}, \qquad \bar{b}(0) = b(0).$$

Then, (7.5), (7.6) imply

$$|(\bar{b} - b)'| \le C\rho^\alpha \quad \Longrightarrow \quad |\bar{b} - b| \le C\rho^{1+\alpha} \quad \text{in } [-\rho, 0],$$

and the desired conclusion follows.

Step 4 - Conclusion. The tangential derivatives v_i, with $i < n$, satisfy the same estimates as v. We find from Step 2 applied to v_i that the mixed derivatives v_{in} must be bounded by a universal bound. This improves the initial estimate in Step 1, which in turn improves the regularity of f and h in Step 2. More precisely, by Lemma 8.1 we find that v_{in} satisfies the estimate (8.2). This holds also for the tangential derivatives of order up to 2. Then the functions $h(x,t)$ and $f(t)$ in (7.1) satisfy the hypotheses of part b) of Lemma 8.1. This gives that the remaining second derivative v_{nn} is bounded as well, and (7.4) holds for $\alpha + 1$ instead of α. Thus we can replace α by $\alpha + 1$ in the bound (7.3) above, and the proposition is proved. □

8. Estimates for the 1D case

In this section, we provide the necessary estimates for solutions to the 1D linear problem. The difference with the higher dimensional case is that now, in the 1D case, the Hölder estimates and the subsequent $C^{1,\alpha}$ and $C^{2,\alpha}$ estimates can be iterated in parabolic cylinders

$$\mathcal{P}_\rho := (0,\rho) \times (-\rho^2, 0],$$

and we can use the standard Hölder parabolic norms with respect to the standard parabolic distance: $d((x,t),(y,s)) := |x-y| + |t-s|^{1/2}$. Following Krylov [K], we denote the corresponding Hölder spaces with respect to this distance with $C^{k,\alpha}_{x,t}$.

Precisely, we prove the following.

Lemma 8.1 (1D-Estimates)**.** *Assume that $\lambda \leq 1$ and $w(x,t)$ is a viscosity solution in $\mathcal{C}_1 \subset \mathbb{R}^2$ of the equation*

$$\begin{cases} w_t = \frac{1}{\lambda}\{A(t)w_{xx} + h(x,t)\} & \text{in } \mathcal{C}_1, \\ w_t = \gamma(t)\, w_x + f(t) & \text{on } \mathcal{F}_1, \end{cases} \tag{8.1}$$

with

$$\|w\|_{L^\infty} \leq 1, \quad K^{-1} \leq A(t),\, \gamma(t) \leq K, \qquad |A'(t)| \leq K\lambda^{-1}.$$

a) If

$$|h| \leq Kx^{\alpha-1}, \quad |f(t)| \leq K,$$

then $w \in C^{1,\alpha}$ in the x variable, $w \in C^1$ on $\{x = 0\}$, and the free boundary condition is satisfied in the classical sense. More precisely, in $\mathcal{C}_{1/2}$ we have

$$|w(x,t) - (w(0,t) + xw_x(0,t))| \leq Cx^{1+\alpha}, \qquad |w_x| \leq C,$$

and

$$|w(y,t) - w(z,s)| \leq C\left(|y-z|^\alpha + \lambda^{-\frac{\alpha}{2}}|t-s|^{\frac{\alpha}{2}}\right),$$

$$|w_x(y,t) - w_x(z,s)| \leq C\left(|y-z|^\alpha + \lambda^{-\frac{\alpha}{2}}|t-s|^{\frac{\alpha}{2}}\right), \tag{8.2}$$

with C depending only on K and α.

b) If in addition in $\mathcal{C}_{3/4}$

$$|h(y,t) - h(z,s)| \leq K\left(|y-z|^\alpha + \lambda^{-\frac{\alpha}{2}}|t-s|^{\frac{\alpha}{2}}\right),$$

$$|\gamma(t) - \gamma(s)| \leq K\lambda^{-\frac{\alpha}{2}}|t-s|^{\frac{\alpha}{2}}, \qquad |f(t) - f(s)| \leq K\lambda^{-\frac{\alpha}{2}}|t-s|^{\frac{\alpha}{2}},$$

then in $\mathcal{C}_{1/2}$

$$|w_x(0,t) - w_x(0,s)| \leq C\lambda^{-\frac{1+\alpha}{2}}|t-s|^{\frac{1+\alpha}{2}}, \qquad |w_{xx}| \leq C. \tag{8.3}$$

After subtracting $F(t) := \int_0^t f(s)ds$ from w and replacing h by $h - \lambda f(t)$ we may assume that $f \equiv 0$. We work with $v(x,t) = w(x, \lambda t)$, and after relabeling λt by t in the arguments of A and h, we obtain

$$\begin{cases} v_t = A(t)v_{xx} + h(x,t) & \text{in } (0,1) \times (-\lambda^{-1}, 0], \\ v_t = \lambda\gamma(t)\, v_x & \text{on } \{x = 0\}, \end{cases} \tag{8.4}$$

with

$$K^{-1} \leq A(t),\, \gamma(t) \leq K, \qquad |A'(t)| \leq K, \quad |h| \leq Kx^{\alpha-1}. \tag{8.5}$$

Lemma 8.1 is equivalent to the Lemma 8.2 below, where we establish the corresponding estimates for v using parabolic scaling.

Lemma 8.2. *Assume that v is a viscosity solution of* (8.4) *in $\mathcal{P}_1$ with $\lambda \leq 1$, and coefficients that satisfy* (8.5)*. Then*

$$\|v\|_{C^{1,\alpha}_{x,t}(\mathcal{P}_{1/2})} \leq C(\|v\|_{L^\infty(\mathcal{P}_1)} + 1), \tag{8.6}$$

and the free boundary condition is satisfied in the classical sense. If in addition

$$\|h\|_{C^{0,\alpha}_{x,t}}, \quad \|\gamma\|_{C^{\frac{\alpha}{2}}_t} \leq K,$$

then

$$\|v\|_{C^{2,\alpha}_{x,t}(\mathcal{P}_{1/2})} \leq C(\|v\|_{L^\infty(\mathcal{P}_1)} + 1),$$

with C depending only on n, K and α.

Proof. If v solves (8.4) in $\mathcal{P}_\rho$ then the rescaling

$$\tilde{v}(x,t) := \rho^{-\beta} v(\rho x, \rho^2 t)$$

solves (8.4) in $\mathcal{P}_1$ with coefficients

$$\tilde{A}(t) = A(\rho^2 t), \quad \tilde{h}(x,t) = \rho^{2-\beta} h(\rho x, \rho^2 t), \quad \tilde{\lambda} = \rho\lambda, \quad \tilde{\gamma}(t) = \gamma(\rho^2 t). \tag{8.7}$$

Notice that the hypotheses on the coefficients are preserved as long as $\beta \le 1+\alpha$, and moreover $\tilde{\lambda} \to 0$ as $\rho \to 0$.

We divide the proof in four steps.

Step 1: Hölder estimates. We show that

$$\|v\|_{C^{0,\beta}_{x,t}(\mathcal{P}_{1/2})} \le C\left(\|v\|_{L^\infty(\mathcal{P}_1)} + 1\right),$$

for some $\beta > 0$ small.

Notice that after an initial dilation, we may assume that $\lambda \le \lambda_0$ is small. It suffices to prove the following claim.

If v is a viscosity solution of (8.4) then

$$osc_{\,\mathcal{P}_1}\ v \le 2 \quad \Longrightarrow \quad osc_{\,\mathcal{P}_\rho}\ v \le \frac{3}{2}, \qquad \text{with } \rho = c_0 \text{ small, universal.} \tag{8.8}$$

The Hölder estimate is obtained by iterating this claim in parabolic cylinders centered on the t axis, while for the interior parabolic cylinders (included in $\{x > 0\}$) we can apply directly the diminishing of oscillation for parabolic equations.

In order to prove (8.8), we let $g(x,t)$ be the solution to the 1D heat equation on the real-line

$$g_t = K^{-1} g_{xx}, \quad g(x,0) = \chi_{(0,\infty)} - \chi_{(-\infty,0)}. \tag{8.9}$$

Notice that for all $t > 0$, in $x = 0$ we have

$$g(0,t) = 0, \qquad g_x(0,t) \le C t^{-1/2},$$

and

$$g_t \le 0, \quad \text{for } x > 0.$$

We want to show that if $|v| \le 1$ in $\mathcal{P}_1$, then we can improve the upper bound or lower bound by a fixed amount in the interior, depending on the value of v at $(0,-1)$, i.e.

$$|v| \le 1 \text{ in } \mathcal{P}_1 \text{ and } v(0,-1) \le 0, \text{ then } v \le 1/2 \text{ in } \mathcal{P}_\rho, \text{ with } \rho = c_0.$$

In $\mathcal{P}_1$ we compare v with

$$G(x,t) := C_1 g(x, t+1) + \frac{1}{4}(t+1)^{1/2} - C_2 x^{1+\alpha}.$$

We choose C_2 and then C_1 sufficiently large such that G is a classical supersolution to (8.4) and $G \ge 1$ on the boundary $(0,1] \times \{-1\}$ and $\{1\} \times [-1,0]$, while $G(0,0) = 1/4$. Then we find $v \le G$ in $\mathcal{P}_1$, which gives the claim (8.8) by choosing c_0 sufficiently small.

Step 2: $C^{1,\alpha}$ estimates. We show that (8.6) holds by first establishing a pointwise $C^{1,\alpha}$ estimate at the origin.

After an initial dilation and after dividing by a large constant, we may assume that $\lambda \le \delta$, $|h| \le \delta x^{\alpha-1}$ for some small δ, and $\|v\|_{L^\infty(\mathcal{P}_1)}$ is sufficiently small.

Claim. If a function l_0 (linear in x) of the form

$$l_0 = a_0 x + b_0(t), \qquad b_0' = \lambda\gamma(t) a_0, \qquad |a_0| \le 1, \tag{8.10}$$

approximates v in $\mathcal{P}_\rho$ to order $1+\alpha$, i.e.

$$|v - l_0| \le \rho^{1+\alpha} \qquad \text{in } \mathcal{P}_\rho, \qquad \rho \le \delta,$$

then we can approximate v to order $1+\alpha$ in $\mathcal{P}_{c_1\rho}$ by a function l_1 as above, with $|a_1-a_0| \le C\rho^\alpha$, and c_1 small universal. Then the claim can be iterated indefinitely by starting with $l_0 \equiv 0$ in $\mathcal{P}_\delta$.

We prove the claim by compactness. Notice that $v-l_0$ solves (8.4) with a slightly modified h that satisfies $|h| \le \delta x^{\alpha-1} + C\delta$. This means that the rescaled error

$$\tilde{v}(x,t) := \rho^{-(1+\alpha)}(v-l_0)(\rho x, \rho^2 t),$$

satisfies (8.4) with coefficients as in (8.7). Since $\|\tilde{v}\|_{L^\infty} \le 1$, by Step 1 we know that

$$\|\tilde{v}\|_{C^{0,\beta}_{x,t}(\mathcal{P}_{1/2})} \le C.$$

This means that if we consider a sequence of $\delta_n \to 0$ and corresponding solutions v_n in $\mathcal{P}_{\rho_n}$, then we can extract a uniformly convergence subsequence of the rescalings $\tilde{v}_n$ in $\mathcal{P}_{1/2}$ such that

$$\tilde{v}_n \to \bar{v}.$$

Then the Hölder continuous limit function $\bar{v}$ is a viscosity solution of

$$\begin{cases} \bar{v}_t = \bar{A}\,\bar{v}_{xx} & \text{in } \mathcal{P}_{1/2}, \\ \bar{v}_t = 0 & \text{on } \{x=0\}, \end{cases}$$

with $\bar{A}$ constant. Since $\bar{v}$ is constant on the boundary $\{x=0\}$, the C^2 estimate for the standard heat equation implies

$$|\bar{v} - (\bar{a}x + \bar{b})| \le C\tau^2 \le \frac{1}{2}\tau^{1+\alpha} \qquad \text{in} \quad \mathcal{P}_\tau, \quad \tau \le c_1.$$

This shows that if δ is chosen sufficiently small, then the rescaling $\tilde{v}$ satisfies the inequality above instead of $\bar{v}$ which implies

$$|v - (a_1 x + b(t))| \le \frac{3}{4}(\tau\rho)^{1+\alpha} \qquad \text{in} \quad \mathcal{P}_{\tau\rho}, \quad \tau = c_1,$$

with

$$a_1 = a_0 + \rho^\alpha \bar{a}, \quad b(t) = b_0(t) + \rho^{1+\alpha}\bar{b}.$$

We define $b_1(t)$ so that l_1 has the form as in (8.10), that is

$$b_1'(t) = \lambda\gamma(t)a_1, \quad b_1(0) = b(0).$$

Then

$$|(b_1-b)'| \le C|\bar{a}\rho^\alpha| \le C\rho^\alpha \quad \Longrightarrow \quad |b_1 - b| \le C\rho^\alpha(\tau\rho)^2 \le \frac{1}{4}(\tau\rho)^{1+\alpha} \quad \text{in} \quad \mathcal{P}_{\tau\rho},$$

where we used $\rho \le \delta$ sufficiently small. In conclusion,

$$|v - l_1| \le (\tau\rho)^{1+\alpha} \quad \text{in} \quad \mathcal{P}_{\tau\rho}, \qquad l_1 = a_1 x + b_1(t),$$

and the claim is proved.

We remark that the oscillation of $b_0(t)$ which appears in the approximation function l_0 in (8.10) is less than $C\rho^2$ in $\mathcal{P}_\rho$. Thus we can modify b_0 to be constant in (8.10) and take l_0 to be linear, and then adjust the error $\rho^{1+\alpha}$ by $C\rho^{1+\alpha}$. This pointwise $C^{1,\alpha}$ estimate can be applied at other points on $\{x=0\}$, which combined with interior $C^{1,\alpha}$ estimates for parabolic equations implies the desired conclusion (8.6).

Step 3. Boundary regularity. We check that v is C^1 on $\{x=0\}$ and the boundary condition is satisfied in the classical sense.

For this assume by contradiction that there exists a sequence $t_k \to 0^-$ such that

$$\frac{1}{t_k}(v(0,t_k)-v(0,0)) < \mu := \lambda\gamma(0)(v_x(0,0)-\eta), \quad \text{for some } \eta>0. \tag{8.11}$$

For each k, we look at the contact point where the graph of v is touched by below by a translation of the graph of the classical strict subsolution to (8.4)

$$g(x,t) := v(0,0) + \mu t + x\left(v_x(0,0) - \frac{1}{2}\eta\right) + Cx^{1+\alpha},$$

in the domain $D_k := [0, c(\eta)] \times [t_k, 0]$.

We choose $c(\eta)$ small such that $g_x(x,t) < v_x(x,t)$ in the domain D_k for all large k. This implies that the contact point must occur on $D_k \cap \{x=0\}$. On the other hand, (8.11) gives

$$v(0,t_k) - v(0,0) > g(0,t_k) - g(0,0)$$

which shows that the contact point is different than $(0,t_k)$ and we reach a contradiction.

Step 4. $C^{2,\alpha}$ estimates. On $\{x=0\}$ we know that $v_x, \gamma \in C^{\alpha/2}$, and the boundary condition implies that $v(0,t) \in C^{1,\alpha/2}$. Now we can apply the standard $C^{2,\alpha}$ Schauder estimates up to the boundary for the heat equation. □

9. Viscosity solutions for the linear problem

In this section, we collect some general facts about viscosity solutions for the linear problem (5.1) and establish the existence and uniqueness claim in Proposition 5.2 by Perron's method. Similar results for different types of boundary conditions were established by G. Lieberman (see for example [L]). However, we are not aware of an existence result that applies directly to the linear problem (5.1). Therefore, for completeness we provide the details in this case.

Recall that $v \in C(\mathcal{C}_1)$ satisfies

$$\begin{cases} \lambda v_t \le tr(A(t)D^2 v) & \text{in } \mathcal{C}_1, \\ v_t \le \gamma(t)\cdot\nabla v & \text{on } \mathcal{F}_1, \end{cases} \tag{9.1}$$

in the viscosity sense if v cannot be touched by above at any point $(x_0,t_0) \in \mathcal{C}_1 \cup \mathcal{F}_1$ in a small neighborhood $\mathcal{B}_r(x_0,t_0)$ by a classical strict supersolution $w \in C^2(\overline{\mathcal{B}_r(x_0,t_0)})$. As usually, this definition is equivalent to the one where we restrict w to belong to the class of quadratic polynomials rather than to the class of C^2 functions.

Another equivalent way is to say that v is a viscosity subsolution of the parabolic equation in $\mathcal{C}_1$, and a viscosity subsolution of the boundary condition on $\mathcal{F}_1$. This last condition means that we cannot touch v locally by above at any point $(x_0,t_0) \in \mathcal{F}_1$ by a function $w \in C^1(\overline{\mathcal{B}_r(x_0,t_0)})$ (or say w is a linear function) that satisfies

$$w_t(x_0,t_0) > \gamma(t_0)\cdot\nabla w(x_0,t_0).$$

The two definitions are the same since, if $w \in C^1$ is as above, and say $(x_0,t_0) = (0,0)$, then a vertical translation of the quadratic polynomial

$$w(0) + (w_t(0)-\epsilon)t + (\nabla w(0) + \epsilon e_n)\cdot x + M(|x'|^2 - nK^2 x_n^2),$$

must touch v by above at some interior point $(x,t) \in \mathcal{B}_r$. Here r is chosen sufficiently small and M large, appropriately, and then the polynomial is a strict supersolution in $\mathcal{B}_r$.

We state the comparison principle for viscosity solutions.

Lemma 9.1. *Assume v_1 is a viscosity subsolution, and v_2 a viscosity supersolution to* (5.1) *in $\overline{\mathcal{C}_1}$. If $v_1 \le v_2$ on $\partial_D \mathcal{C}_1$ then $v_1 \le v_2$ in $\mathcal{C}_1$.*

Corollary 9.2. *The difference of two viscosity solutions of* (5.1) *is also a viscosity solution of* (5.1).

We work with the rescaling $w(x,t) = v(x, \lambda t)$.

First we prove a preliminary result on the evolution in time of a Lipschitz "trace" $w((x',0),t)$ under specific growth assumptions.

Lemma 9.3. *Assume that $w \le 1$ satisfies*

$$\begin{cases} w_t \le \mathcal{M}^+_K(D^2 w) + 1 & \text{in } (Q_1 \cap \{x_n > 0\}) \times (0,T], \\ \frac{1}{\lambda} w_t \le K w_n^+ - K^{-1} w_n^- + K|\nabla_{x'} w| & \text{on } \{x_n = 0\}, \end{cases} \tag{9.2}$$

and

$$w((x',0),0) \le |x'|^2.$$

Then

$$w(0,t) \le C\lambda(t^{1/2} + t) \quad \text{for } t \ge 0,$$

with C depending on n and K.

Proof. We compare w with

$$G(x,t) := g(x_n,t) + C\lambda(t^{1/2}+t) + |x'|^2 + C(2x_n - x_n^2),$$

where $g(x_n,t)$ is the solution to the 1D heat equation on the real-line (see (8.9))

$$g_t = K^{-1} g_{nn}, \quad g(x_n, 0) = \chi_{(0,\infty)} - \chi_{(-\infty,0)}.$$

It is easy to check that G is a classical supersolution which is above w on the boundary of our domain, and that gives the desired result. □

Lemma 9.4. *Assume that $w \le 1$ satisfies* (9.2) *in $\mathcal{C}_1$ and the trace of w on $\{x_n = 0\}$ is Lipschitz, i.e.*

$$|\nabla_{x'} w| \le 1 \quad \text{on } \{x_n = 0\}.$$

Then

$$w((x',0),t) \ge w((x',0),0) - C\lambda^{\frac{2}{3}}\,|t|^{\frac{1}{2}} \qquad \text{if } x' \in Q'_{1/2}.$$

Proof. We prove the inequality for $x'=0$. Since w is Lipschitz the parabola

$$w(0,t) + Cr^2 + r^{-2}|x'|^2$$

is greater than $w((x',0),t)$, with r to be specified later. Now we can apply the previous lemma to the rescaling

$$\tilde{w}(y,s) := w(ry, t + r^2 s) - w(0,t) - Cr^2,$$

which solves (9.2) with $\tilde{\lambda} = \lambda r$, and obtain that

$$\tilde{w}(0,s) \le C\tilde{\lambda}(s^{1/2} + s).$$

This gives

$$C(r^2 + \lambda|t|^{\frac{1}{2}} + \lambda r^{-1}|t|) \ge w(0,0) - w(0,t),$$

and we choose $r = (\lambda|t|)^{1/3}$ to get

$$w(0,t) \geq w(0,0) - C(\lambda|t|^{\frac{1}{2}} + (\lambda|t|)^{2/3}) \geq w(0,0) - C\lambda^{\frac{2}{3}}\,|t|^{\frac{1}{2}}.$$

□

Remark 9.5. The proof of Lemma 9.4 shows that we can construct a supersolution $\bar{G}(x,t)$ in $\mathcal{C}_1$ such that $\bar{G}((x',0),-1) = |x'|$, $\bar{G} \geq 1$ on the remaining part of $\partial_D \mathcal{C}_1$, and so that $\bar{G}(0,t) \leq C\lambda^{\frac{2}{3}}\,|t|^{\frac{1}{2}}$. Similarly, given $\alpha > 0$, we can construct a supersolution with $\bar{G}((x',0),-1) = |x'|^\alpha$, $\bar{G} \geq 1$ on the remaining of $\partial_D \mathcal{C}_1$ and such that $\bar{G}(0,t) \leq C(\lambda|t|)^\beta$, for some β depending on α.

We are now ready to prove our main lemma.

Proof of Lemma 9.1. Let $w_i(x,t) = v_i(x,\lambda t)$, $i = 1,2$, so that w_1 is a subsolution and w_2 a supersolution of

$$\begin{cases} w_t = tr(A(t)D^2 w) & \text{in } \{x_n > 0\}, \\ \frac{1}{\lambda} w_t = \gamma(t) \cdot \nabla w & \text{on } \{x_n = 0\}, \end{cases}$$

and we want to show that w_1 cannot touch w_2 strictly by below at an interior point. Assume by contradiction that this is the case.

The standard viscosity theory of parabolic equations implies that the contact point cannot occur in $\{x_n > 0\}$. Below we denote by C, c various constants that may depend on w_i and λ.

After a translation and a dilation we may assume that in $\mathcal{C}_1$

$$w_1 \leq w_2 + \mu t, \qquad w_1(0,0) = w_2(0,0) = 0,$$

for some $\mu > 0$ small. Without loss of generality we may also assume that w_1/w_2 has a semiconvex/semiconcave trace in the x' variable, that is

$$D^2_{x'} w_1 \geq -I, \qquad D^2_{x'} w_2 \leq I, \tag{9.3}$$

and also

$$\|w_i\|_{L^\infty} \leq 1 \tag{9.4}$$

and each w_i solves the parabolic equation in the interior. This is achieved in the following way. First we replace a subsolution w with the standard regularization using the sup-convolutions in the x' variable

$$w_\epsilon(x,t) = \max_y \left\{ w(y,t) - \frac{1}{2\epsilon}|y' - x'|^2 \right\},$$

then we divide w_ϵ by a large constant, and in the end we solve the parabolic equation in the interior of $\mathcal{C}_1$ by keeping the same boundary values on the parabolic boundary. All these operations maintain the subsolution property of w, and justify the extra assumptions (9.3)-(9.4).

Moreover, after subtracting from each w_i a function of the type $a' \cdot x' + b(t)$ with $\frac{d}{dt} b(t) = \lambda a' \cdot \gamma(t)$ we may assume in addition that

$$w_i(0,0) = 0, \quad \nabla_{x'} w_i(0,0) = 0, \tag{9.5}$$

and the interior parabolic equations have the form

$$\partial_t w_i = tr(A(t)D^2 w_i) + h(t), \qquad |h| \leq C.$$

We show that $w_i(0,t)$ are differentiable at the origin in the t variable, and that the derivative of w_1 is less than the derivative of w_2, which would contradict our hypothesis that $w_1 \leq w_2 + \mu t$.

To achieve this we apply Lemma 9.4 several times. By (9.3)-(9.4)-(9.5) and Lemma 9.4 we find that

$$w_1 \geq -Cr \quad \text{and} \quad w_2 \leq Cr \qquad \text{on} \quad \mathcal{P}_r \cap \{x_n = 0\}. \tag{9.6}$$

Since $w_1 \leq w_2$, we can use the pointwise C^α parabolic estimates at the origin and find that, given any $\alpha < 1$, we have

$$osc_{\mathcal{P}_r} w_i \leq Cr^\alpha \quad \text{for all } r > 0. \tag{9.7}$$

We can iterate this argument, by working with the rescaling

$$\tilde{w}_1(x,t) = r^{-\alpha} w_1(rx, r^2 t),$$

which satisfies a similar equation with $\tilde{\lambda} = \lambda r$, and is such that (9.3)-(9.4)-(9.5) hold for $\tilde{w}_1$. Again by Lemma 9.4 we find

$$\tilde{w}_1((x',0),t) \geq -Cr^{2/3} \qquad \text{if } x' \in Q'_{1/2},$$

hence we improve the estimate (9.6) as

$$w_1 \geq -Cr^{\alpha + \frac{2}{3}} \quad \text{on} \quad \mathcal{P}_r \cap \{x_n = 0\}. \tag{9.8}$$

The same holds for w_2 with $\leq$ instead of $\geq$ and $Cr^{\alpha+\frac{2}{3}}$ instead of $-Cr^{\alpha+\frac{2}{3}}$. This in turn shows that w_i are pointwise $C^{\alpha+\frac{2}{3}}$ at the origin.

We modify again each w_i by subtracting the corresponding function $\partial_n w_i(0) x_n + b_i(t)$, with $\frac{d}{dt} b_i = \lambda \gamma_n \partial_n w_i(0)$. Using that $\partial_n(w_1 - w_2)(0) \leq 0$, we find that the inequality $w_1 \leq w_2 + \mu t$ is still valid on $\{x_n = 0\}$, while (9.7) holds with $r^{\alpha+2/3}$ instead of r^α. The same argument as above implies that (9.8) holds again with $r^{\alpha+4/3}$ instead of $r^{\alpha+2/3}$. Since $\alpha + 4/3 > 2$, this means that $w_1(0,t) \geq -C|t|^{1+\beta}$ and $w_2(0,t) \leq C|t|^{1+\beta}$ for all small $t < 0$, which contradicts $w_1(0,t) \leq w_2(0,t) + \mu t$. □

We can finally conclude the proof of Proposition 5.2.

Proof of Proposition 5.2. The interior C^2 estimates in the x variable and the Hölder estimates up to the boundary were already proved in Proposition 5.1 and Theorem 5.3. It remains to prove existence by Perron's method.

We assume for simplicity that the boundary data ϕ is Lipschitz, and the general case follows by approximation. As usual, we define

$$v(x,t) := \sup_{w \in \mathcal{A}} w(x,t),$$

where $\mathcal{A}$ is the class of continuous subsolutions on $\overline{\mathcal{C}}_1$ which have boundary data below ϕ on $\partial_D \mathcal{C}_1$. The conclusion that v solves our problem is easily checked once its continuity has been established.

Claim. For each $(x_0, t_0) \in \partial_D \mathcal{C}_1$ there exists a subsolution $w_{(x_0,t_0)}$ which vanishes at (x_0,t_0), is below the cone $-|(x,t) - (x_0,t_0)|$ on $\partial_D \mathcal{C}_1$ and has a Hölder modulus of continuity at (x_0,t_0).

This can be deduced from the proof of Theorem 5.3, where the Hölder continuity at the boundary was achieved using explicit barriers. More precisely, as in Lemma

6.2 and Lemma 6.6, for all $r \le 1/2$ we can construct a subsolution ϕ_r defined in $\mathcal{B}^{\pm}_{\lambda,r}(x_0,t_0) \cap \mathcal{C}_1$, where

$$\mathcal{B}^{\pm}_{\lambda,r}(x_0,t_0) := \{(x,t)| \quad d_\lambda((x,t),(x_0,t_0)) < r\},$$

so that

$$\phi_r = 0 \quad \text{on} \quad \partial\mathcal{B}^{\pm}_{\lambda,r}(x_0,t_0) \setminus (\partial_D \mathcal{C}_1 \cup \mathcal{F}_1), \qquad \phi_r \le 1 \quad \text{on} \quad \partial\mathcal{B}^{\pm}_{\lambda,r}(x_0,t_0) \cap \partial_D \mathcal{C}_1$$

and

$$\phi_r \ge c_0 \quad \text{on} \quad \partial\mathcal{B}^{\pm}_{\lambda,r/2}(x_0,t_0).$$

Then $w_{(x_0,t_0)}$ is obtained by superposing appropriate multiples of ϕ_r for a dyadic sequence of $r = 2^{-m}$. We omit the details.

Using the claim we can construct a subsolution $\underline{\phi}$ and supersolution $\overline{\phi}$ which are Hölder continuous on $\partial_D \mathcal{C}_1$ and agree with the boundary data ϕ. Thus we can restrict the class $\mathcal{A}$ of subsolutions to satisfy

$$\underline{\phi} \le w \le \overline{\phi}. \tag{9.9}$$

This shows that the limit v achieves the boundary data ϕ continuously. Moreover, using (9.9) we can replace each $w \in \mathcal{A}$ by its maximum among appropriate x' translations

$$\max_{y'} \{w(x - (y',0),t) - C|y'|^\alpha\},$$

and remain in the same class. Therefore we may assume that $\mathcal{A}$ contains only subsolutions which are uniformly Hölder continuous in the x' variable. Using this together with Remark 9.5, we find that the trace of v on $\{x_n = 0\}$ is locally Hölder continuous in the x', t variables. This means that the solution $\bar{v}$ to the interior parabolic equation in $\mathcal{C}_1$ with boundary data v is continuous up to the boundary. By the maximum principle $\bar{v} \ge w$ for any $w \in \mathcal{A}$, and it is straightforward to check that $\bar{v} \in \mathcal{A}$, hence $v = \bar{v}$ is continuous in $\overline{\mathcal{C}}_1$. □

Acknowledgement.

The authors would like to thank Sandro Salsa and Fausto Ferrari for fruitful conversations on the topic of this paper. N.F. wishes to thank the Department of Mathematics of Columbia University for its warm hospitality.

References

[A] I. Athanasopoulos, *Free boundary regularity in Stefan type problems,* Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni, Serie 9, Vol. 15 (2004), n.3-4, p. 345–355.

[ACS1] I. Athanasopoulos, L. Caffarelli, S. Salsa, *Regularity of the free boundary in parabolic phase-transition problems*, Acta Math. Volume 176, Number 2 (1996), 245–282.

[ACS2] I. Athanasopoulos, L. Caffarelli, S. Salsa, *Phase transition problems of parabolic type: Flat free boundaries are smooth*, Comm. on Pure and Applied Math, Volume 51, Issue 1 (1998),77–112.

[C1] L. A. Caffarelli, *A Harnack inequality approach to the regularity of free boundaries. Part I: Lipschitz free boundaries are* $C^{1,\alpha}$, Rev. Mat. Iberoamericana **3** (1987) no. 2, 139–162.

[C2] L. A. Caffarelli, *A Harnack inequality approach to the regularity of free boundaries. Part II: Flat free boundaries are Lipschitz*, Comm. Pure Appl. Math. **42** (1989), no.1, 55–78.

[CS] L.A. Caffarelli, S. Salsa, *A Geometric Approach to Free Boundary Problems,* Graduate Studies in Mathematics, vol. 68.

[CG] H. A. Chang-Lara, N. Guillen, *From the free boundary condition for Hele-Shaw to a fractional parabolic equation*, arXiv:1605.07591.

[CE] L.A. Caffarelli, C. Evans, *Continuity of the temperature in two-phase Stefan problems,* Arch. Rational Mech., 81, 1983, 199-220.

[CK] S. Choi, I. C. Kim, *Regularity of one-phase Stefan problem near Lipschitz initial domain*, American Journal of Mathematics Vol. 132, No. 6 (December 2010), pp. 1693-1727.

[CV] L. A. Caffarelli, J. L. Vazquez, *A Free-Boundary Problem for the Heat Equation Arising in Flame Propagation*, Transactions of the American Mathematical Society, Vol. 347, No. 2 (Feb., 1995), pp. 411-441

[D] D. De Silva, *Free boundary regularity for a problem with right hand side*, Interfaces and free boundaries **13** (2011), 223–238.

[DFS] D. De Silva, F. Ferrari, S. Salsa, *Two-phase problems with distributed sources: regularity of the free boundary*, Anal. PDE Volume 7, Number 2 (2014), 267-310.

[DR] D. De Silva, J-M. Roquejoffre, *Regularity in a one-phase free boundary problem for the fractional Laplacian.* Preprint 2010. To appear in Annales de l'Institut Henri Poincare.

[DSV] G. De Philippis, L. Spolaor, B. Velichkov, *Regularity of the free boundary for the two-phase Bernoulli problem,* arXiv:1911.02165.

[F] A. Friedman, *The Stefan problem in several space variables,* Trans. Amer. Math. Soc., 133 (1968), 51-87.

[K] N. V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces,* vol. 12 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 1996.

[LS] P. Laurence, S. Salsa, *Regularity of the free boundary of an American option on several assets*, Communications on Pure and Applied Mathematics 62(7), 969 - 994 (2009).

[L] G. Lieberman, *Mixed boundary value problems for elliptic and parabolic differential equations of second order*, Journal of Mathematical Analysis and Applications Volume 113, Issue 2 (1986), Pages 422-440.

[R] L. Rubinstein, *The Stefan Problem,* American Mathematical Society, Providence, R.I., 1971.

[W] L. Wang, *On the regularity theory of fully nonlinear parabolic equations. I,* Comm. Pure Appl. Math., 45 (1992), pp. 27–76.

Department of Mathematics, Barnard College, Columbia University, New York, NY 10027

E-mail address: desilva@math.columbia.edu

Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato, 5, 49126, Bologna, Italy

E-mail address: nicolo.forcillo2@unibo.it

Department of Mathematics, Columbia University, New York, NY 10027

E-mail address: savin@math.columbia.edu