J. DIFFERENTIAL GEOMETRY
44 (1996) 371-434

A BOUNDARY REGULARITY THEOREM FOR MEAN CURVATURE FLOW

ANDREW STONE

Abstract

We study singularity formation in the mean curvature flow of smooth, compact, embedded hypersurfaces of non-negative mean curvature in $\mathbf{R}^{n+1}$, with fixed smooth boundary, Γ. Then, subject to a so-called "Type I" hypothesis, and a certain geometrical constraint on Γ, we establish the following boundary regularity result:

Main Theorem (Boundary Regularity). *Suppose hypotheses A and B of Section 1 hold, and suppose that the hypersurfaces $\{M_t\}_{t\in[0,T)}$ are flowing by mean curvature as in* (1.1). *Then there is a fixed neighbourhood (in $\mathbf{R}^{n+1}$) of the boundary, Γ, in which all the surfaces M_t remain smooth, with uniform bounds on $|A|^2$ and all its derivatives, even as $t \nearrow T$.*

Note for instance that this result covers (modulo the Type I hypothesis) the model situation where Γ is some smooth (but possibly highly complicated) embedded submanifold of S^n which splits S^n into two "caps", and where M_0 is taken to be one of these caps (cf. Appendix C).

In the process of proving this theorem we also extend Huisken's classification of singularities (see [5]) to our setting with boundary, and refine his analysis along the lines of [10].

The principal ingredients used, to address these issues, are Allard's boundary regularity theory for varifolds, and also a certain "density function", whose definition is based on the analogue, for surfaces with boundary, of Huisken's important monotonicity formula for mean curvature flow.

Received May 2, 1994.

1. Introduction

Let M_0 be a smooth, compact, embedded hypersurface in $\mathbf{R}^{n+1}$ with smooth embedded boundary Γ^{n-1}, and suppose $\{\mathbf{F}_t : M_0 \to \mathbf{R}^{n+1}\}_{t\in[0,T)}$ is a family of embeddings which move M_0 by its mean curvature, holding the boundary fixed. That is, $\mathbf{F}_t = \mathbf{F}(\cdot, t)$ is a one-parameter family of embeddings satisfying $\mathbf{F} \in C^\infty(M_0 \times (0,T)) \cap C^{1,\alpha}(M_0 \times [0,T))$, and

$$(1.1)\qquad \begin{aligned} \frac{d}{dt}\mathbf{F}(p,t) &= \mathbf{H}(p,t) \quad \text{on} \quad M_0 \times [0,T) \\ \mathbf{F}(p,t) &\equiv p \quad \text{on} \quad \Gamma \times [0,T) \\ \mathbf{F}(p,0) &\equiv \mathbf{F}_0(p) = p \quad \text{on} \quad M_0\,. \end{aligned}$$

Here $\mathbf{H}(p,t) = -H(p,t)\nu(p,t)$ denotes the mean curvature vector of $M_t = \mathbf{F}_t(M_0)$ at $\mathbf{F}(p,t)$, and ν denotes a choice of unit normal for M_t.

Note that the Dirichlet boundary condition in (1.1) means, of course, that H must then satisfy, for all $t \in (0,T)$,

$$(1.2)\qquad H(\cdot, t) \equiv 0 \quad \text{on} \quad \Gamma\,.$$

Our main purpose in this paper is to study the formation of singularities in such a flow, as $t \to T$. To do so, however, we need to make three further assumptions. The latter two, stated in hypothesis B below, are geometric conditions, one about the boundary, Γ, and one about the mean curvature of the initial surface M_0. The first is a so-called "Type I hypothesis", about the rate of blow-up of the function $U(t) \equiv \max_{M_t} |A|^2$ as t approaches T, where A is the second fundamental form of $M_t = \mathbf{F}_t(M_0)$.

The first and third of these assumptions mirror those needed by Huisken in his analysis of singularity formation for surfaces without boundary, [5], which was in any case the inspiration for much of this work. Specifically, our hypotheses are;

Hypothesis A: (Type I hypothesis) There is a fixed constant C_0 such that the function $U(t) \equiv \max_{M_t} |A|^2(p,t)$ satisfies

$$(1.3)\qquad U(t) \le \frac{C_0}{2(T-t)}\,.$$

Hypothesis B: (Convexity hypotheses) The boundary, $\Gamma^{n-1} \subset \mathbf{R}^{n+1}$, lies on the boundary of some uniformly convex body $\mathcal{N}$, and moreover $M_0 \subset \bar{\mathcal{N}}$. (Actually this hypothesis may be weakened; see Remark 12.2(ii)).

Also the surface M_0 satisfies $H \ge 0$ everywhere.

1.1 Remarks and Notation. **(i)** Note that without loss of generality we may as well then assume, as we do henceforth, that M_0

in fact satisfies

$$(1.4) \qquad H \equiv 0 \quad \text{on} \quad \Gamma, \qquad\qquad H \geq 0 \quad \text{on} \quad \mathrm{int}(M_0)\,.$$

This is permissible because any smooth M_0 satisfying $H \geq 0$ and $\partial M_0 \subset \partial\mathcal{N}$ may be approximated in a suitable $C^{1,\alpha}$-sense (made precise in Appendix C) by a sequence of smooth surfaces $\{\Sigma_i\}$, of uniformly bounded non-negative mean curvature, which satisfy $\partial\Sigma_i \equiv \Gamma$ and $H|_{\partial\Sigma_i} \equiv 0$.

Then by use of appropriate $C^{1,\alpha}$-estimates it may be shown that the flow (1.1) with initial surface M_0 has a solution for at least a short time, which instantaneously becomes smooth and also satisfies $H(\cdot,t) \geq 0$, $H(\cdot,t)|_\Gamma \equiv 0$ for all $t > 0$ (see Appendix C). We may then simply replace M_0 in the discussion by one of the surfaces M_t, $t > 0$ small, and this "new M_0" will satisfy (1.4).

(ii) Recall that the first part of Hypothesis B means that we can find a uniform radius R such that, for any $\mathbf{x} \in \partial\mathcal{N}$, $\mathcal{N}$ may be enclosed in a sphere of radius R that touches $\partial\mathcal{N}$ precisely at $\mathbf{x}$ and nowhere else. Using that Γ is of class C^2 it thence follows that, at any point $\mathbf{x}_0 \in \Gamma$, we can find two distinct barrier hyperplanes, intersecting along $T_{\mathbf{x}_0}\Gamma$, that will then (by the maximum principle) enclose all the evolving hypersurfaces, M_t, in the wedge that they form. For each such $\mathbf{x}_0$ we henceforth let this wedge be denoted $V_{\mathbf{x}_0}$. Then we let $\hat{V}_{\mathbf{x}_0}$ denote the wedge in $\mathbf{R}^{n+1}$ obtained by rotating and translating $V_{\mathbf{x}_0}$ into "standard position", so that $T_{\mathbf{x}_0}\Gamma$ goes into $\mathbf{R}^{n-1} \times \{0\}$, the plane $\mathbf{R}^n \times \{0\}$ bisects $\hat{V}_{\mathbf{x}_0}$, and the surfaces M_t all lie in the $x_n \geq 0$ halfspace. Also we let $\hat{\mathcal{R}}_{\mathbf{x}_0}$ denote the (unique) rotation in $\mathbf{R}^{n+1}$ sending $V_{\mathbf{x}_0}$ (after translation) to $\hat{V}_{\mathbf{x}_0}$.

Finally, observe furthermore that, by the uniformity of the convexity of $\mathcal{N}$, all of the wedges $\{\hat{V}_{\mathbf{x}_0} : \mathbf{x}_0 \in \Gamma\}$ may be taken to have "wedge angles", $\varphi_{\mathbf{x}_0}$, satisfying $\varphi_{\mathbf{x}_0} \leq \pi - \delta_0$ for some fixed $\delta_0 > 0$. We let $\hat{V}^{(\delta_0)}$ denote the wedge in "standard position" in $\mathbf{R}^{n+1}$ with exactly this wedge angle, $\pi - \delta_0$ (so that each wedge $\hat{V}_{\mathbf{x}_0}$ then lies symmetrically inside $\hat{V}^{(\delta_0)}$).

1.2 Further Remarks. (i) It is well-known that, in the boundaryless case, if the surfaces M_t are developing a singularity at time T, in the sense that $\liminf_{t\to T} U(t) = +\infty$, then the function $U(t)$ must in fact satisfy (see, for instance, [5, Lemma 1.2]),

$$(1.5) \qquad U(t) \geq \frac{1}{2(T-t)}\,.$$

This was the motivation, in [5], for distinguishing the notions of

Type I and Type II singularities, the Type I assumption being "just right" to permit rescaling arguments based on the monotonicity identity discovered in that paper ([5, Theorem 3.1]). Such a rescaling procedure, based on an analogue of Huisken's monotonicity result, is also the key tool in our subsequent analysis.

(ii) The proof of the estimate (1.5) does not carry over directly to the case of surfaces with boundary, unfortunately. However, after additional work, an analogous "minimum blow-up rate" estimate can still be established in this setting (see [11]).

(iii) Hypothesis B (more precisely (1.4)), together with the maximum principle, yields:

1.3 Lemma. *For all* $t \in [0,T)$ *the surfaces* M_t *must satisfy either* $H \equiv 0$, *or* $H = 0$ *on* Γ, $H \geq 0$ *on* $\operatorname{int}(M_t)$.

Of course the first option here is the trivial case, for if this situation ever arises then the flow has successfully reached a minimal surface and stopped (with certainly no singularity formation at the later time T). Thus we always tacitly assume, henceforth, that this has not happened.

Now that we have set up the situation under consideration, and specified the three additional assumptions we need about the behaviour of the flow, we are in a position to outline the chief aims of this paper. They are essentially two-fold.

The principal one is to prove a boundary regularity theorem, that there is some neighbourhood of Γ in which the surfaces M_t remain regular as $t \to T$, regardless of whether or not $|A|^2$ may be blowing up elsewhere in the interior. This is established in two main stages.

The first is to show that, in a boundary strip, the Type I estimate for the maximum rate of blow-up of $|A|^2$, (1.3), may be improved to an estimate of the form $|A|^2(p,t) \leq \epsilon(t)/(T-t)$, where $\epsilon(t) \to 0$ as $t \to T$. This is the content of part (b) of Theorem 6.1, which is proven in Sections 7 and 8 (after the necessary preliminaries have been developed in Sections 2 to 5).

The second stage, carried out in Sections 9 to 12, is then to bootstrap up this improved estimate to a full boundedness result for $|A|^2$, in some boundary strip, as $t \to T$. To obtain this we employ the boundary regularity theory for varifolds of Allard, [1], as well as the techniques of [10, Section 4].

The second main aim of this paper is to extend Huisken's analysis of singularity formation (see [5]), for surfaces without boundary, to the present setting; and also to sharpen it along the lines of [10]. More precisely, recall the notions of "special" (those studied by Huisken) and "general" singular points, introduced in [10];

1.4 Definition. We say that $p \in M_0$ is a *special singular point* of the flow ("singular point in the sense of Huisken"), as $t \to T$, if there exists a fixed $\delta > 0$ such that, for some sequence of times $t_k \nearrow T$,

$$|A|^2(p, t_k) \geq \frac{\delta}{(T - t_k)} . \tag{1.6}$$

1.5 Definition. We say $p \in M_0$ is a *general singular point* of the flow, as $t \to T$, if there exists a fixed $\delta > 0$ such that, for some sequence of times $t_k \nearrow T$, and some sequence of points $p_k \in M_0$ with $p_k \to p$,

$$|A|^2(p_k, t_k) \geq \frac{\delta}{(T - t_k)} . \tag{1.7}$$

Also, we call such a sequence $\{(p_k, t_k)\}$ a *δ-essential blow-up sequence*, following the terminology of Altschuler ([2]).

Then our main result in this regard is that, in the above setting, just as in the boundaryless case (see [10, Theorem 3.1]), any "general" Type I singular point of the flow must actually be a "special" Type I singular point also. See Theorem 6.1(a).

The results described in this paper formed part of the author's doctoral dissertation at Stanford University.

2. The monotonicity identities

The essential tool in our analysis of singularity formation in the above setting will be the following analogue, for surfaces with boundary, of Huisken's important monotonicity formula for mean curvature flow ([5]). As there, we consider an "n-dimensional" backwards heat kernel centred at an arbitrary point $(\mathbf{x}_0, T)$ in spacetime $\mathbf{R}^{n+1} \times (0, \infty)$, and then study how the "amount of heat on our evolving manifold" changes as $t \nearrow T$. (We will be interested in the case where T is the time of first singularity formation).

2.1 Lemma. *Take any $(\mathbf{x}_0, T) \in \mathbf{R}^{n+1} \times (0, \infty)$. Suppose the surfaces M_t, all with fixed boundary Γ, are evolving by mean curvature as in* (1.1), *and set*

$$\rho_{\mathbf{x}_0} \equiv \frac{1}{(4\pi\tau)^{n/2}} \exp\Big(\frac{-|\mathbf{x} - \mathbf{x}_0|^2}{4\tau}\Big)$$

where here, and henceforth, $\tau \equiv (T - t)$. Then

(2.1)

$$\frac{d}{dt}\int_{M_t} \rho_{\mathbf{x}_0} dH^n = -\int_{M_t} \left|\mathbf{H} + \frac{(\mathbf{x} - \mathbf{x}_0)^{\perp}}{2\tau}\right|^2 \rho_{\mathbf{x}_0} dH^n - \int_{\Gamma} \Big\langle \frac{\mathbf{x} - \mathbf{x}_0}{2\tau}, \eta \Big\rangle \rho_{\mathbf{x}_0} dH^{n-1}$$

where η denotes the inward unit conormal to Γ with respect to M_t.

Note that (2.1) may be written in integrated form as

(2.2)

$$\int_{t_1}^{t_2}\int_{M_t}\left|\mathbf{H}+\frac{(\mathbf{x}-\mathbf{x}_0)^{\perp}}{2\tau}\right|^2\rho_{\mathbf{x}_0}dH^n dt = \left\{\int_{M_{t_1}}\rho_{\mathbf{x}_0}dH^n - \int_{M_{t_2}}\rho_{\mathbf{x}_0}dH^n\right\} - \int_{t_1}^{t_2}\int_{\Gamma}\left\langle\frac{\mathbf{x}-\mathbf{x}_0}{2\tau},\eta\right\rangle\rho_{\mathbf{x}_0}dH^{n-1}dt$$

for any $0 \le t_1 \le t_2 \le T$. In the event that $t_2 = T$, of course, we replace $\int_{M_T}\rho_{\mathbf{x}_0}dH^n$ in (2.2) by $\lim_{t\to T}\int_{M_t}\rho_{\mathbf{x}_0}dH^n$ (see also Section 5).

2.2 Remark. The presence of the boundary term in (2.1) means that actually the quantity $\int_{M_t}\rho_{\mathbf{x}_0}dH^n$ no longer need necessarily be (indeed won't be, as we shall see) monotone decreasing. However we shall still refer to (2.1) and (2.2) as the (unscaled) "Monotonicity Identities" because of their analogy with Huisken's monotonicity identities for the boundaryless case.

Proof. This proceeds almost exactly as in the boundaryless case - see [5, Theorem 3.1] for the details. The only difference is that, when the First Variation Formula is applied, we now get a boundary contribution (see [9, p. 46]), which gives the boundary integral in (2.1). Q.E.D.

The importance of this lemma is that it will allow us to carry out a rescaling analysis, just as in [5]. We now describe this rescaling procedure.

For any fixed point $\mathbf{x}_0 \in \mathbf{R}^{n+1}$ (and fixed $T \in (0,\infty)$) define the rescaled immersions $\tilde{\mathbf{F}}_{\mathbf{x}_0} : M_0 \times [-\frac{1}{2}\ln(T),\infty) \to \mathbf{R}^{n+1}$ by

$$\tilde{\mathbf{F}}_{\mathbf{x}_0}(p,s) \equiv \frac{1}{\sqrt{2\tau}}(\mathbf{F}(p,t)-\mathbf{x}_0) \quad , \quad s(t) \equiv -\frac{1}{2}\ln(T-t)\,. \tag{2.3}$$

Then, under this scaling, the surfaces $M_t = \mathbf{F}_t(M_0)$ go to

$$\tilde{M}_{\mathbf{x}_0,s} \equiv \frac{1}{\sqrt{2\tau}}(M_t - \mathbf{x}_0) \tag{2.4}$$

and it is readily checked that, just as in the boundaryless case, these new surfaces are evolving according to the equation

$$\frac{d}{ds}\tilde{\mathbf{F}}_{\mathbf{x}_0}(p,s) = \tilde{\mathbf{H}}(p,s) + \tilde{\mathbf{F}}_{\mathbf{x}_0}(p,s) \tag{2.5}$$

where $\tilde{\mathbf{H}}$ denotes the mean curvature vector of $\tilde{M}_{\mathbf{x}_0,s}$.

We can now state a rescaled version of Lemma 2.1.

2.3 Lemma. *Take any* $(\mathbf{x}_0, T) \in \mathbf{R}^{n+1} \times (0, \infty)$, *and set*

$$\tilde{\rho}(\mathbf{x}) = \exp\left(\frac{-|\mathbf{x}|^2}{2}\right). \tag{2.6}$$

Then, for the $\tilde{M}_{\mathbf{x}_0,s}$ *as above, we have (writing just* $\boldsymbol{H}$ *for* $\tilde{\boldsymbol{H}}$ *now)*

$$\frac{d}{ds}\int_{\tilde{M}_{\mathbf{x}_0,s}} \tilde{\rho}\, dH^n = -\int_{\tilde{M}_{\mathbf{x}_0,s}} |\mathbf{H} + \mathbf{x}^{\perp}|^2 \tilde{\rho}\, dH^n - \int_{\tilde{\Gamma}_{\mathbf{x}_0,s}} \langle \mathbf{x}, \eta \rangle \tilde{\rho}\, dH^{n-1} \tag{2.7}$$

or equivalently, in integrated form,

$$\begin{aligned} \int_{s_1}^{s_2}\int_{\tilde{M}_{\mathbf{x}_0,s}} |\mathbf{H} + \mathbf{x}^{\perp}|^2 \tilde{\rho}\, dH^n ds = &\left\{ \int_{\tilde{M}_{\mathbf{x}_0,s_1}} \tilde{\rho}\, dH^n - \int_{\tilde{M}_{\mathbf{x}_0,s_2}} \tilde{\rho}\, dH^n \right\} \\ &- \int_{s_1}^{s_2}\int_{\tilde{\Gamma}_{\mathbf{x}_0,s}} \langle \mathbf{x}, \eta \rangle \tilde{\rho}\, dH^{n-1} ds\,. \end{aligned} \tag{2.8}$$

Here $\tilde{\Gamma}_{\mathbf{x}_0,s}$ *denotes the rescaled boundary, namely* $\tilde{\Gamma}_{\mathbf{x}_0,s} \equiv \frac{1}{\sqrt{2\tau}}(\Gamma - \mathbf{x}_0)$.

Proof. We use 'tildes' to distinguish rescaled quantities. Then since, under the rescaling, we have the correspondence $\tilde{\mathbf{x}} = \frac{1}{\sqrt{2\tau}}(\mathbf{x} - \mathbf{x}_0)$, so we see that the measures, $d\tilde{\mu}_{s(t)}$ and $d\mu_t$, induced on M_0 by $\tilde{\mathbf{F}}_{\mathbf{x}_0,s(t)} \equiv \tilde{\mathbf{F}}_{\mathbf{x}_0}(\cdot, s(t))$ and $\mathbf{F}_t$ respectively, are related by $d\tilde{\mu}_s = \left(\sqrt{2\tau}\right)^{-n} d\mu_t$. Similarly the induced boundary measures, $d\tilde{\sigma}_{s(t)}$ and $d\sigma_t$, on Γ, are related by $d\tilde{\sigma}_s = \left(\sqrt{2\tau}\right)^{-(n-1)} d\sigma_t$.

But also, by direct computation, we have that

$$\rho_{\mathbf{x}_0} = \frac{1}{(4\pi\tau)^{n/2}} \exp\left(\frac{-|\tilde{\mathbf{x}}|^2}{2}\right) \quad , \quad \left|\mathbf{H} + \frac{(\mathbf{x} - \mathbf{x}_0)^{\perp}}{2\tau}\right|^2 = \frac{1}{2\tau}\left|\tilde{\mathbf{H}} + \tilde{\mathbf{x}}^{\perp}\right|^2, \tag{2.9}$$

and (since η does not change under scaling)

$$\left\langle \frac{(\mathbf{x} - \mathbf{x}_0)}{2\tau}, \eta \right\rangle = \frac{1}{\sqrt{2\tau}}\left\langle \tilde{\mathbf{x}}, \eta \right\rangle. \tag{2.10}$$

Putting all this in (2.1) then immediately yields, on multiplying by $(2\pi)^{n/2}$ (and writing now just $\mathbf{x}$ for $\tilde{\mathbf{x}}$, etc.), that

$$\frac{d}{dt}\int_{\tilde{M}_{\mathbf{x}_0,s}} \tilde{\rho}\, dH^n = -\frac{1}{2\tau}\int_{\tilde{M}_{\mathbf{x}_0,s}} |\mathbf{H} + \mathbf{x}^{\perp}|^2 \tilde{\rho}\, dH^n - \frac{1}{2\tau}\int_{\tilde{\Gamma}_{\mathbf{x}_0,s}} \langle \mathbf{x}, \eta \rangle \tilde{\rho}\, dH^{n-1}, \tag{2.11}$$

and then the proof is completed by noting that $\frac{ds}{dt} = 1/2\tau$, by (2.3). Q.E.D.

2.4 Remark. It turns out that in the subsequent analysis we need to keep careful track of the explicit numerical factor, $(2\pi)^{n/2}$, that arose in the above rescaling analysis. Therefore we note now, for future reference, that we just showed that

$$\int_{M_t} \rho_{\mathbf{x}_0} dH^n = \frac{1}{(2\pi)^{n/2}} \int_{\tilde{M}_{\mathbf{x}_0,s(t)}} \tilde{\rho} dH^n , \tag{2.12}$$

and also that

(2.13)
$$\int_{t_1}^{t_2} \int_{\Gamma} \left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2\tau}, \eta \right\rangle \rho_{\mathbf{x}_0} dH^{n-1} dt = \frac{1}{(2\pi)^{n/2}} \int_{s(t_1)}^{s(t_2)} \int_{\tilde{\Gamma}_{\mathbf{x}_0,s}} \langle \mathbf{x}, \eta \rangle \tilde{\rho} dH^{n-1} ds .$$

We would like now to use Lemma 2.3 to discuss the behaviour of sequences of rescalings $\{\tilde{M}_{\mathbf{x}_0,s_k}\}$, just as in the boundaryless case. However, before we can do this, we first need to prove two further results.

(i) We need to establish uniform control of all derivatives of $|\tilde{A}|^2$ for the rescaled flow (cf. [5, Proposition 2.3]), so as to allow us to extract limit surfaces; and

(ii) We need to obtain sufficiently good control over the boundary integral contribution to the right hand side of (2.2) (or equivalently (2.8)).

As regards the first of these, we have:

2.5 Lemma. *Suppose the Type I hypothesis holds. Then for each $m > 0$ there is a constant $\tilde{C}_m = \tilde{C}_m(n, m, C_0, \delta_0)$ such that the second fundamental form, $\tilde{A}$, of the rescaled surfaces satisfies*

$$\max_{\tilde{M}_{\mathbf{x}_0,s}} \left|\tilde{\nabla}^m \tilde{A}\right|^2 \leq \tilde{C}_m \tag{2.14}$$

uniformly in s. Of course these estimates are independent of the point, $\mathbf{x}_0$, chosen to rescale about.

Proof. Precisely such estimates were proven in [5, Proposition 2.3], in the boundaryless setting, using the maximum principle. The same argument then also works here, provided we can first establish estimates of the form of (2.14) on the boundary, Γ.

To obtain such boundary estimates, consider any point $\mathbf{x}_0 \in \Gamma$. Then, recalling the notation of 1.1, we have that all the surfaces $\{\hat{\mathcal{R}}_{\mathbf{x}_0}(M_t - \mathbf{x}_0)\}$ lie in the wedge $\hat{V}^{(\delta_0)}$, which has wedge angle $\pi - \delta_0$ (δ_0 independent of $\mathbf{x}_0$). But then, by the Type I hypothesis, (1.3), we can find uniform constants $r_1 = r_1(\delta_0, C_0)$ and $K_1 = K_1(\delta_0, C_0)$ such that, for each $t < T$, the hypersurface $\hat{\mathcal{R}}_{\mathbf{x}_0}(M_t - \mathbf{x}_0) \cap (B^{(n)}_{r_1(T-t)^{1/2}}(0) \times \mathbf{R})$ is a graph over (its projection into) the fixed hyperplane $\mathbf{R}^n \times \{0\}$, with gradient uniformly bounded above by K_1. (Here $B^{(n)}_\rho(0)$ denotes the ρ-ball about the origin in $\mathbf{R}^n \times \{0\}$).

Now, however, the desired boundary estimates at $\mathbf{x}_0$ follow, with constants independent of the choice of $\mathbf{x}_0 \in \Gamma$, by the parabolic theory of [8]. Since an analogous argument is given later in proving Lemma 10.8 we omit the details here. Q.E.D.

Turning now to the second of the main issues noted earlier, namely control of the boundary integral contribution in (2.2), it is convenient to distinguish two cases, $\mathbf{x}_0 \in \Gamma$ and $\mathbf{x}_0 \notin \Gamma$. In each instance we take this opportunity to prove results (see Lemmas 2.6(b) and 2.7(b) below) which are somewhat stronger than are required simply for the business of extracting limit surfaces from sequences of rescalings about a fixed point $\mathbf{x}_0$. For that, Lemmas 2.6(a) and 2.7(a) would suffice. However we shall need the stronger claims later, in Section 8, where we'll want to carry out a more involved rescaling analysis in which the centre, $\mathbf{x}_0$, of the rescalings, is also allowed to vary.

Taking now the first of our two cases, we have:

2.6 Lemma. **(a)** *Suppose $\mathbf{x}_0 \notin \Gamma$. Then there is a constant $C_1 = C_1(n, \Gamma, \mathbf{x}_0)$ such that*

$$\int_0^T \int_\Gamma \left| \left\langle \frac{\mathbf{x} - \mathbf{x}_0}{2\tau}, \eta \right\rangle \right| \rho_{\mathbf{x}_0} \, dH^{n-1} dt \le C_1 \,. \tag{2.15}$$

(b). *In fact, for $\mathbf{x}_0 \notin \Gamma$ again, we can say more precisely that there is a function $\zeta : (0, \infty) \times [0, T) \to \mathbf{R}_+$, with $\zeta(d, t) \to 0$ as $t \to T$, for each d, such that, if* $\operatorname{dist}(\mathbf{x}_0, \Gamma) \ge d_0$, *then, for all $t_1 \in [0, T)$,*

$$\int_{t_1}^T \int_\Gamma \left| \left\langle \frac{\mathbf{x} - \mathbf{x}_0}{2\tau}, \eta \right\rangle \right| \rho_{\mathbf{x}_0} \, dH^{n-1} dt \le \zeta(d_0, t_1) \,. \tag{2.16}$$

(One can even explicitly take $\zeta(d, t) = \frac{C}{d^n} \int_{d/\sqrt{2(T-t)}}^{\infty} w^{n-1} e^{-w^2/2} dw$ for a suitable constant $C = C(n, \Gamma)$, if desired).

Proof. Clearly (a) follows from (b) by taking $t_1 = 0$, $d_0 = \operatorname{dist}(\mathbf{x}_0, \Gamma)$. Thus we need only prove (b). For this, note that, by hypothesis, for all

$t \in [0, T)$,

$$\int_{\Gamma} \frac{1}{(4\pi\tau)^{n/2}} \exp\Big(\frac{-|\mathbf{x}-\mathbf{x}_0|^2}{4\tau}\Big) \left|\left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2\tau}, \eta \right\rangle\right| dH^{n-1} \le (4\pi\tau)^{-n/2} e^{-\frac{d_0^2}{4\tau}} \left(\frac{C_2}{2\tau}\right) |\Gamma|$$

where $C_2 \equiv \max\{|\mathbf{y}-\mathbf{x}_0| : \mathbf{y} \in \Gamma\}$. But then, using the substitution $w(t) = d_0/\sqrt{2\tau}$, we may estimate that

$$\int_{t_1}^{T} \int_{\Gamma} \left|\left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2\tau}, \eta \right\rangle\right| \rho_{\mathbf{x}_0} dH^{n-1} dt \le \frac{(2\pi)^{-n/2} C_2 |\Gamma|}{d_0^n} \int_{d_0/\sqrt{2(T-t_1)}}^{\infty} w^{n-1} e^{-w^2/2} dw \tag{2.17}$$

and so we are done. Q.E.D.

A similar result holds in the case $\mathbf{x}_0 \in \Gamma$.

2.7 Lemma. (a) *Take any $\mathbf{x}_0 \in \Gamma$. Then there is a constant $C_3 = C_3(n, \Gamma)$ such that*

$$\int_{0}^{T} \int_{\Gamma} \left|\left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2\tau}, \eta \right\rangle\right| \rho_{\mathbf{x}_0} dH^{n-1} dt \le C_3\,. \tag{2.18}$$

(b) *In fact, we have the following uniformity in $\mathbf{x}_0$, that, for any $\epsilon > 0$, there is a time $t_\epsilon = t_\epsilon(n, \Gamma) < T$ such that, for any $\mathbf{x}_0 \in \Gamma$,*

$$\int_{t_\epsilon}^{T} \int_{\Gamma} \left|\left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2\tau}, \eta \right\rangle\right| \rho_{\mathbf{x}_0} dH^{n-1} dt \le \epsilon\,. \tag{2.19}$$

Proof. Clearly (a) follows from (b), so we need only prove the latter claim. As regards this, define

$$\Gamma_1(t) \equiv \Gamma \cap B_{r(t)}(\mathbf{x}_0), \qquad \Gamma_2(t) \equiv \Gamma \backslash \Gamma_1$$

where $r(t)$, with $r(t) \searrow 0$ as $t \to T$, is to be chosen explicitly later. Then set

$$I_1(t) = \int_{t}^{T} \frac{1}{2\tau} \int_{\Gamma_1} \left(\frac{1}{4\pi\tau}\right)^{n/2} |\langle \mathbf{x}-\mathbf{x}_0, \eta \rangle| \, dH^{n-1} dt\,, \tag{2.20}$$

$$(2.21) \qquad I_2(t) = \int_t^T \frac{1}{2\tau} \int_{\Gamma_2} \rho_{\mathbf{x}_0} \left|\langle \mathbf{x} - \mathbf{x}_0, \eta\rangle\right| dH^{n-1} dt\,.$$

Clearly we then have, for any $t_* < T$,

$$(2.22) \qquad \int_{t_*}^T \int_\Gamma \left|\left\langle \frac{\mathbf{x} - \mathbf{x}_0}{2\tau}, \eta\right\rangle\right| \rho_{\mathbf{x}_0} dH^{n-1} dt \le I_1(t_*) + I_2(t_*)\,.$$

We now want to estimate $I_1(t_*)$ and $I_2(t_*)$. The latter of these is the easier to handle, so we treat it first. Writing r for $r(t)$, observe that, on Γ_2, $|\mathbf{x} - \mathbf{x}_0| \ge r$, so

$$\rho_{\mathbf{x}_0} \le (4\pi\tau)^{-n/2} e^{-r^2/4\tau}\,.$$

Thus, since $|\langle \mathbf{x} - \mathbf{x}_0, \eta\rangle| \le |\mathbf{x} - \mathbf{x}_0| \le \operatorname{diam}(\Gamma)$, we may estimate that

$$(2.23) \qquad I_2(t_*) \le C_4 \int_{t_*}^T \tau^{-n/2-1} e^{-r^2/4\tau} dt\,,$$

where $C_4 = \frac{1}{2}(4\pi)^{-n/2} |\Gamma| \operatorname{diam}(\Gamma) = C_4(n, \Gamma)$.

To handle $I_1(t_*)$, first note that, since Γ is of class C^∞, there is a uniform constant $C_2(\Gamma)$, and radius $r_0(\Gamma)$, each determined only by the sup norm of the second fundamental form of $\Gamma^{n-1} \subset \mathbf{R}^{n+1}$, such that

$$(2.24) \qquad |(\mathbf{x} - \mathbf{x}_0)^\perp| \le C_5 |\mathbf{x} - \mathbf{x}_0|^2 \quad \text{for all} \quad \mathbf{x} \in \Gamma \cap B_{r_0}(\mathbf{x}_0)\,,$$

and such that also, for all $r_* \le r_0$,

$$(2.25) \qquad H^{n-1}(\Gamma \cap B_{r_*}(\mathbf{x}_0)) \le 2 r_*^{n-1} \omega_{n-1}\,.$$

Here ω_{n-1} is the volume of the $(n-1)$-dimensional unit ball. So then, set

$$(2.26) \qquad t_0 \equiv \min\{t : r(t) \le r_0(\Gamma) \ \text{for all}\ t \in [t_0, T)\}\,.$$

Then once we've chosen $r(t)$ explicitly, as we'll do shortly, we'll have $t_0 = t_0(\Gamma)$; and moreover we'll have, from (2.24), (2.25) and the definition of $\Gamma_1(t)$, that for all $t_* \in [t_0, T)$,

$$(2.27) \qquad |\langle \mathbf{x} - \mathbf{x}_0, \eta\rangle| \le C_5 r(t_*)^2 \quad \text{for all} \quad \mathbf{x} \in \Gamma_1(t_*)$$

and

$$(2.28) \qquad H^{n-1}(\Gamma_1(t_*)) \le 2 r(t_*)^{n-1} \omega_{n-1}\,.$$

In (2.27) we have used that $|\langle \mathbf{x} - \mathbf{x}_0, \eta\rangle| \le |(\mathbf{x} - \mathbf{x}_0)^\perp|$.

But then it follows immediately that, for all $t_* \in [t_0, T)$,

$$I_1(t_*) \leq \frac{C_5 \omega_{n-1}}{(4\pi)^{n/2}} \int_{t_*}^{T} \tau^{-n/2-1} r(t)^{n+1} dt \,. \tag{2.29}$$

Therefore, from (2.23) and (2.29) we obtain that, for all $t_* \in [t_0, T)$,

$$\begin{aligned} I_1(t_*) + I_2(t_*) \leq \; & C_4 \int_{t_*}^{T} \tau^{-n/2-1} e^{-r(t)^2/4\tau} dt \\ & + C_6 \int_{t_*}^{T} \tau^{-n/2-1} r(t)^{n+1} dt \end{aligned} \tag{2.30}$$

where C_4, C_6 are constants determined only by n and Γ.

So now, finally, choose

$$r(t) \equiv \sqrt{-2(n+2)\tau \ln(\tau)} \ ,$$

which is well-defined for $\tau < 1$ at least. Then (increasing $t_0(\Gamma)$ if necessary to ensure $T - t_0 < 1$), we get that, for all $t_* \in [t_0(\Gamma), T)$,
(2.31)

$$I_1(t_*) + I_2(t_*) \leq C_4 \int_{t_*}^{T} dt + C_6 (\sqrt{2n+4})^{n+1} \int_{t_*}^{T} \tau^{-1/2} (-\ln(\tau))^{\frac{n+1}{2}} dt \,.$$

But thence, on setting

$$C_7(n, \Gamma) = C_6 (\sqrt{2n+4})^{n+1} \max_{[t_0(\Gamma), T)} (-\tau^{\delta} \ln(\tau)),$$

where we have fixed $\delta \in (0, \frac{1}{n+1})$ arbitrarily, we get, for all $t_* \in [t_0(\Gamma), T)$, that

$$I_1(t_*) + I_2(t_*) \leq C_4 (T - t_*) + C_7 (T - t_*)^{\frac{1-(n+1)\delta}{2}} \ , \tag{2.32}$$

and this, in view of (2.22), immediately yields (b). Q.E.D.

3. Rescaling the flow

We are now in a position to mirror the rescaling analysis of Huisken in [5]. First, though, to simplify the later discussion, we introduce the important "limit-point" function, $\hat{} : M_0 \to \mathbf{R}^{n+1}$, given by

$$p \mapsto \hat{p} \equiv \lim_{t \to T} \mathbf{F}(p, t) \,. \tag{3.1}$$

That this limit exists follows easily from (1.3), which implies, for all $t_* \in [0, T)$, that

$$(3.2)\quad \left| \int_{t_*}^{T} \frac{d\mathbf{F}}{dt} dt \right| \leq \int_{t_*}^{T} |\mathbf{H}| dt \leq \int_{t_*}^{T} \sqrt{nC_0/2(T-t)}\, dt \leq \sqrt{2nC_0}(T-t_*)^{1/2}.$$

3.1 Remark. Clearly, $\hat{p}$ would be well-defined even if the Type I hypothesis were weakened merely to require that, for some fixed $\delta \in (0,1)$,

$$\max_{M_t} |A|^2 \leq \frac{C_0}{(T-t)^{1+\delta}}.$$

Some useful facts about the function $p \to \hat{p}$ are given in the following lemma.

3.2 Lemma. **(a)** *The function* $\hat{}\ : M_0 \to \mathbf{R}^{n+1}$ *given by* (3.1) *is continuous.*

(b) *For any* $p \in M_0$ *the surfaces* $\tilde{M}_{\hat{p},s}$ *all intersect* $B_{\sqrt{2nC_0}}(0)$. *Indeed* $\tilde{\mathbf{F}}_{\hat{p}}(p,s)$ *is always in this ball* (*cf.* [5, Lemma 3.3]).

Proof. **Of (a).** This is left to the reader (use (3.2)). Alternatively, see [11].

Proof. **Of (b).** This follows from (3.2) and the definition of the rescalings $\tilde{\mathbf{F}}_{\hat{p}}(\cdot, s)$. Q.E.D.

3.3 Remark. Typically if we rescale about an arbitrary point $\mathbf{x} \in \mathbf{R}^{n+1}$ then the rescaled surfaces will drift off to infinity as $s \to \infty$, but Lemma 3.2(b) says that this does not happen if we rescale about a limit-point, for time T, of the surface.

We are now in a position to state the main rescaling result we'll need, which is directly analogous to Theorems 3.4 and 3.5 of [5] for the boundaryless case.

3.4 Theorem. *Suppose* (*as always*) *that the Type I hypothesis holds. Take any* $p \in M_0$. *Then for every sequence of rescaled times* $s_j \nearrow \infty$, *corresponding to times* $t_j \nearrow T$, *there is a subsequence* $\{s_{j_k}\}$ *such that the surfaces* $\tilde{M}_{\hat{p},s_{j_k}}$ *converge smoothly on compact subsets of* $\mathbf{R}^{n+1}$ *to a non-empty, embedded limit-surface,* $\tilde{M}_{\hat{p},\infty}$.

If $\hat{p} \notin \Gamma$ *then* $\tilde{M}_{\hat{p},\infty}$ *has no boundary; while if* $\hat{p} \in \Gamma$ *then* $\tilde{M}_{\hat{p},\infty}$ *has boundary,* $\tilde{\Gamma}_{\hat{p},\infty}$, *an* $(n-1)$*-plane through the origin in* $\mathbf{R}^{n+1}$.

Finally, any such limit hypersurface $\tilde{M}_{\hat{p},\infty}$ *must satisfy the equation*

$$(3.3)\qquad H = \langle \mathbf{x}, \nu \rangle$$

where $\mathbf{x}$ *is the position vector,* H *is the mean curvature, and* ν *is the unit normal such that the mean curvature vector,* $\mathbf{H}$, *is given by* $\mathbf{H} = -H\nu$.

Proof. In view of Lemmas 2.3 and 2.6(a) (or 2.7(a) if $\hat{p} \in \Gamma$) we have

$$\int_{\tilde{M}_{\hat{p},s}} e^{-|x|^2/2} dH^n \leq \hat{C} \quad \text{for all} \quad s \in [-\frac{1}{2}\ln(T), \infty)$$

where $\hat{C} = \hat{C}(n, \hat{p}, M_0, \Gamma)$ is some constant.

Therefore for each $R > 0$ there is a uniform bound on $H^n(\tilde{M}_{\hat{p},s} \cap B_R(0))$, and then the proof of the convergence claim goes through exactly as in [5, Proposition 3.4] (which was itself based on a method of [7]), noting Lemma 2.5. The non-emptiness of the limit follows from Lemma 3.2(b), just as in the boundaryless case.

As for the claims regarding boundaries, these are clear, since the rescaling procedure will either send the boundary, Γ, to infinity (if $\hat{p} \notin \Gamma$), or will "straighten it out" to an $(n-1)$-plane, while holding it anchored through the origin (if $\hat{p} \in \Gamma$).

Finally, (3.3) also follows directly from Lemmas 2.3 (equation (2.8)), 2.5 and 2.6(a) (or 2.7(a) if $\hat{p} \in \Gamma$), exactly as in [5, Theorem 3.5]. Q.E.D.

3.5 Remarks. **(i)** As noted in [5, Proposition 3.4], a subsequence of the embeddings, $\tilde{\mathbf{F}}_{\hat{p}}(\cdot, s_j)$, need not necessarily converge to a limiting embedding; it may be necessary to "reparametrise" them first (see [7] for the details).

(ii) Although uniqueness of the limit $\tilde{M}_{\hat{p},\infty}$ remains open, it is shown in [11] that we have at least a "degree of uniqueness", namely uniqueness of the *shape* of any such blow-up (cf. [10, Theorem 3.11]). For instance, it is demonstrated in [11] that if one such blow-up (corresponding to one subsequence of times) looks like, say, a cylinder, $S^1(1) \times \mathbf{R}^{n-1}$, then so does any other such limit, though we cannot yet show that they need be the same cylinder.

4. Classification of the possible limit surfaces

The next important step is to classify all the possible "limit surfaces", satisfying (3.3), which could arise in Theorem 3.4 from rescaling. Recall here that we are restricting our attention to the case when any limit rescaling will also satisfy, at all points, $H \geq 0$.

Fortunately, for the case $\hat{p} \notin \Gamma$, so that any limit $\tilde{M}_{\hat{p},\infty}$ has no boundary, this was already done by Huisken (see [5] for the compact setting, and [6] for the general case). His result yields;

4.1 Theorem (Huisken). *Up to rotation in* $\mathbf{R}^{n+1}$ *there are precisely* $(n+1)$ *smooth, embedded surfaces,* $\tilde{M}_\infty$, *that satisfy* $\partial\tilde{M}_\infty = \phi$, $H \geq 0$, *and* $H = \langle \mathbf{x}, \nu \rangle$. *They are the surfaces* $\tilde{M}_\infty = \mathbf{R}^n \times \{0\}$, *or* $\tilde{M}_\infty =$

$S^m(\sqrt{m}) \times \mathbf{R}^{n-m}$ for $m \in \{1, \dots, n\}$, where $S^m(\sqrt{m})$ denotes the m-sphere in $\mathbf{R}^{m+1}$ of radius $\sqrt{m}$.

It remains then only to handle the case where the limit surface, $\tilde{M}_\infty$, has boundary $\mathbf{R}^{n-1} \times \{(0,0)\} \subset \mathbf{R}^{n+1}$. For this case we have;

4.2 Theorem. *Suppose $\tilde{M}_\infty$ is a smooth, embedded hypersurface in $\mathbf{R}^{n+1}$ satisfying $\partial\tilde{M}_\infty = \mathbf{R}^{n-1} \times \{(0,0)\} \subset \mathbf{R}^{n+1}$, $H \geq 0$, and $H = \langle \mathbf{x}, \nu \rangle$. Then $\tilde{M}_\infty$ is a half n-plane.*

For the proof we will use a First Variation Formula argument, with a choice of variation vector field suggested by (though somewhat different from) that used by Allard in [1, Lemma 5.1]. First, though, we need a preliminary result regarding L^2-approximation of characteristic functions of intervals on the real line.

4.3 Lemma. *Put $\mathcal{E} \equiv \{\gamma \in C^1([0,\infty)) : \gamma(r) = \mathrm{O}(e^{-r^2/2})$ as $r \to \infty\}$, and then define a class of functions by*

$$\mathcal{F} = \left\{ \psi : [0,\infty) \to \mathbf{R} \text{ such that } \psi(r) = \frac{\gamma'(r)}{r} + \gamma(r) \text{ for some } \gamma \in \mathcal{E} \right\}.$$

Then for any $R_0 > 0$ the characteristic function $\chi_{[0,R_0]}$ is in the L^2-closure of $\mathcal{F}$.

Proof of Lemma 4.3. Take any $\alpha_0 \in (0,1)$ and then set

$$\gamma(r) \equiv \begin{cases} 1 + \beta_1 e^{-r^2/2} & , 0 \leq r \leq \alpha_0 R_0 \\ \beta_2 - \beta_3 r^2 & , \alpha_0 R_0 \leq r < R_0 \\ \beta_4 e^{-r^2/2} & , r \geq R_0 \end{cases} \tag{4.1}$$

where the constants β_1 to β_4 satisfy

$$\beta_4 e^{-R_0^2/2} = 2\beta_3 \tag{4.2}$$

$$\beta_4 e^{-R_0^2/2} = \beta_2 - \beta_3 R_0^2 \tag{4.3}$$

$$\beta_2 - \beta_3 \alpha_0^2 R_0^2 = 1 + \beta_1 e^{-\alpha_0^2 R_0^2/2} \tag{4.4}$$

and

$$2\beta_3 = \beta_1 e^{-\alpha_0^2 R_0^2/2}\,. \tag{4.5}$$

(It may be readily seen that these simultaneous equations do have a (unique) solution).

Then $\gamma \in \mathcal{E}$, and the function $\psi(r) \equiv r^{-1}\gamma'(r) + \gamma(r)$ satisfies

$$\psi(r) \equiv \begin{cases} 1 & , 0 \leq r \leq \alpha_0 R_0 \\ -\beta_3 r^2 + (\beta_2 - 2\beta_3) & , \alpha_0 R_0 \leq r < R_0\,. \\ 0 & , r \geq R_0 \end{cases} \tag{4.6}$$

Yet clearly this $\psi(r) \in \mathcal{F}$ may be made as close as we please to $\chi_{[0,R_0]}$, in the L^2 sense, just by taking α_0 close enough to 1. Q.E.D.

Proof of Theorem 4.2. We may as well assume, by rotating $\tilde{M}_\infty$ if necessary, that $T_0\tilde{M}_\infty = \mathbf{R}^{n-1} \times \mathbf{R}_+ \times \{0\}$. Then, letting $\mathbf{e}_1, \ldots, \mathbf{e}_{n+1}$ be the standard basis for $\mathbf{R}^{n+1}$, take ν to denote the upward unit normal to $\tilde{M}_\infty$ locally about 0, so $\nu(0) = \mathbf{e}_{n+1}$.

Now let $\mathcal{R}$ denote the set of all radii, R, such that

$$\nu \text{ is well-defined on all of } \tilde{M}_\infty \cap B_R(0), \tag{4.7}$$

$$\langle \nu(\mathbf{x}), \mathbf{e}_{n+1}\rangle \geq 7/8 \quad \text{for all} \quad \mathbf{x} \in \tilde{M}_\infty \cap B_R(0), \tag{4.8}$$

$$|x_{n+1}| \leq |x_n| \quad \text{for all} \quad \mathbf{x} \in \tilde{M}_\infty \cap B_R(0), \tag{4.9}$$

and

$$x_n > 0 \quad \text{for all} \quad \mathbf{x} \in (\tilde{M}_\infty \cap B_R(0))\backslash \partial\tilde{M}_\infty\,. \tag{4.10}$$

Note that $\mathcal{R}$ then contains at least some interval $[0, R_0]$, in view of the assumption that $T_0\tilde{M}_\infty = \mathbf{R}^{n-1} \times \mathbf{R}_+ \times \{0\}$, and the fact that $x_n \equiv x_{n+1} \equiv 0$ on $\partial\tilde{M}_\infty$.

Next, observe that, to prove the theorem, it will suffice to show that, for any $R_* \in \mathcal{R}$, we must have

$$\tilde{M}_\infty \cap B_{R_*}(0) = \text{the half } n\text{-disc } \{(\mathbf{y}, 0) : \mathbf{y} \in \mathbf{R}^n, y_n \geq 0, |\mathbf{y}| < R_*\}\,. \tag{4.11}$$

This is because we would clearly then have to have $\sup(\mathcal{R}) = \infty$, and then we'd be done, in view of (4.11).

It remains, then, to establish (4.11).

So suppose $R_* \in \mathcal{R}$, and introduce a vector-field, $\mathbf{Z}(\mathbf{x})$, on $\tilde{M}_\infty$, defined by

$$\mathbf{Z}(\mathbf{x}) \equiv (0, \ldots, 0, -x_{n+1}, x_n)\,. \tag{4.12}$$

Geometrically $\mathbf{Z}(\mathbf{x})$ represents the projection of $\mathbf{x}$ onto the 2-plane $\Pi \equiv \{0\} \times \mathbf{R}^2 \subset \mathbf{R}^{n+1}$, followed by anticlockwise rotation by 90° in Π. Alternatively it may be written as $\mathbf{Z}(\mathbf{x}) = *(\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{n-1} \wedge \mathbf{x})$ where $*$ denotes the Hodge-star operator.

The three key facts about this vector-field that we need are then that

$$\langle \mathbf{x}, \mathbf{Z}(\mathbf{x})\rangle \equiv 0 \quad \text{for all} \quad \mathbf{x} \in \tilde{M}_\infty\,, \tag{4.13}$$

$$\mathbf{Z}(\mathbf{x}) \equiv 0 \quad \text{for all} \quad \mathbf{x} \in \partial\tilde{M}_\infty \equiv \mathbf{R}^{n-1} \times \{(0,0)\}\,, \tag{4.14}$$

and finally that

$$\operatorname{div}_{\tilde{M}_\infty} \mathbf{Z}(\mathbf{x}) \equiv 0\,. \tag{4.15}$$

Now, for convenience, set $r \equiv |\mathbf{x}|$, and let η denote the inward unit conormal to $\tilde{M}_\infty$ along $\partial\tilde{M}_\infty$. Then define the variation vector field, $\mathbf{X}(\mathbf{x})$, on $\tilde{M}_\infty$, by

$$\mathbf{X}(\mathbf{x}) \equiv \gamma(r)\mathbf{Z}(\mathbf{x}) \tag{4.16}$$

where $\gamma \in C^1([0,\infty);\mathbf{R})$ is, for the present, arbitrary.

Then, by (4.13) and (4.15), $\mathbf{X}(\mathbf{x})$ satisfies

$$\operatorname{div}_{\tilde{M}_\infty} \mathbf{X}(\mathbf{x}) = \frac{\gamma'(r)}{r}\langle \mathbf{x}^T, \mathbf{Z}(\mathbf{x})\rangle = -\frac{\gamma'(r)}{r}\langle \mathbf{x}^\perp, \mathbf{Z}(\mathbf{x})\rangle\,,$$

whence, since $\tilde{M}_\infty$ satisfies $\mathbf{H} = -\mathbf{x}^\perp$, we have

$$\operatorname{div}_{\tilde{M}_\infty} \mathbf{X}(\mathbf{x}) + \langle \mathbf{X}(\mathbf{x}), \mathbf{H}\rangle = -\left(\frac{\gamma'(r)}{r} + \gamma(r)\right)\langle \mathbf{x}^\perp, \mathbf{Z}(\mathbf{x})\rangle\,. \tag{4.17}$$

Also, by (4.14), we have $\langle \mathbf{X}(\mathbf{x}), \eta\rangle \equiv 0$ for all $\mathbf{x} \in \partial\tilde{M}_\infty$.

But then, by the First Variation Formula (see [9, p 46]), we obtain the identity that

$$-\int_{\tilde{M}_\infty} \left(\frac{\gamma'(r)}{r} + \gamma(r)\right)\langle \mathbf{x}^\perp, \mathbf{Z}(\mathbf{x})\rangle dH^n = 0 \tag{4.18}$$

for all $\gamma \in C^1([0,\infty);\mathbf{R})$ (of rapid decay, at least).

Thence in particular, in view of Lemma 4.3, we must clearly have that

$$\int_{\tilde{M}_\infty \cap B_{R_*}(0)} \langle \mathbf{x}^\perp, \mathbf{Z}(\mathbf{x})\rangle dH^n = 0\,,$$

or equivalently, since $-H\nu = \mathbf{H} = -\mathbf{x}^\perp$ on $\tilde{M}_\infty$,

$$\int_{\tilde{M}_\infty \cap B_{R_*}(0)} H\langle \nu, \mathbf{Z}(\mathbf{x})\rangle dH^n = 0\,. \tag{4.19}$$

But then, by (4.12) and the definition of $\mathcal{R}$,

$$\langle \nu, \mathbf{Z}(\mathbf{x})\rangle \equiv x_n\nu_{n+1} - x_{n+1}\nu_n \ge x_n(\nu_{n+1} - |\nu_n|) \ge \frac{3}{8}x_n$$

whence, on $(\tilde{M}_\infty \cap B_{R_*}(0))\backslash\partial\tilde{M}_\infty$, we have $\langle \nu, \mathbf{Z}(\mathbf{x})\rangle > 0$.

Therefore, since also $H \ge 0$ on $\tilde{M}_\infty$, we see from (4.19) that we must actually have

$$H(\mathbf{x}) \equiv 0 \quad \text{for all} \quad \mathbf{x} \in \tilde{M}_\infty \cap B_{R_*}(0)\,. \tag{4.20}$$

Yet $H = \langle \mathbf{x}, \nu\rangle$, so this implies $|\mathbf{x}^\perp| \equiv 0$ on $\tilde{M}_\infty \cap B_{R_*}(0)$ (i.e. $\tilde{M}_\infty$ must be a minimal cone, when restricted to $B_{R_*}(0)$); and then this immediately yields the desired result, (4.11), on noting that $\tilde{M}_\infty$ is smooth at 0. Q.E.D.

5. The limiting heat density function

We are now almost ready to state our first main result, relating to "partial boundary regularity" for the flow in (1.1), subject to Hypotheses A and B of Section 1. First, though, we need to introduce a function, $\hat{\Theta}$, which we have termed the "limiting heat density", which will play a crucial role in the subsequent analysis. This function was also used in [10] to study singularity formation in the boundaryless case.

5.1 Definition. For each $p \in M_0$, and $t \in [0, T)$, we set

$$\Theta(p,t) = \int_{M_t} \rho_{\hat{p}}(x,t)dH^n \, . \tag{5.1}$$

Then we set

$$\hat{\Theta}(p) = \lim_{t \to T} \Theta(p,t) \, . \tag{5.2}$$

5.2 Remark. It is not *a priori* clear that the limit in (5.2) need even exist, for a given $p \in M_0$, since, unlike in the boundaryless case (where we have genuine monotonicity), we do not here have that $\Theta(p,t)$ is monotonically decreasing in t for each p. Moreover we can no longer expect that $\hat{\Theta}$ will necessarily be *upper semicontinuous*, as it always is in the boundaryless case (again by monotonicity). In Lemma 5.7, however, we will see that in fact $\hat{\Theta}(p)$ *is* well-defined (i.e. the limit exists) for every $p \in M_0$, and furthermore that a certain semicontinuity property for $\hat{\Theta}$ can still be deduced.

First, though, it is convenient to give a lemma which makes specific a further reason, besides the above mentioned "partial semicontinuity", for the usefulness of the function, $\hat{\Theta}$, in what follows. This is that, in view of the results of §4, it can only take on a finite, discrete set of values.

5.3 Lemma. *At each point $p \in M_0$ the function $\hat{\Theta}$ must have one of only $n+2$ possible values. Moreover these values are explicitly computable. They are 1/2, 1, or one of the numbers $\Theta^{(m)}$, $m \in \{1, \dots, n\}$, given in Appendix A (all of which are distinct and strictly greater than 1). Furthermore, we have that*

$$\hat{\Theta}(p) = 1/2 \iff \hat{p} \in \Gamma \, . \tag{5.3}$$

Proof. In view of the computations of Appendix A, all the claims of the lemma follow provided we can establish that, if $\{t_j \nearrow T\}$ is any sequence of times, with $\{s_j \nearrow \infty\}$ the corresponding rescaled times, and if $\tilde{M}_{\hat{p},\infty}$ is a limit surface obtained from a subsequence of the surfaces $\tilde{M}_{\hat{p},s_j}$, as in Theorem 3.4, then

$$\hat{\Theta}(p) = \frac{1}{(2\pi)^{n/2}} \int_{\tilde{M}_{\hat{p},\infty}} e^{-|x|^2/2} dH^n \,. \tag{5.4}$$

This, however, is precisely the analogue of Lemma 3.10(b) of [10], and the proof here goes through just as there. In particular Lemma 2.9 of [10], which was the key element of the proof of Lemma 3.10(b), carries over to the present setting (with boundary), after replacing part (a) of it by the (much stronger) area growth bound of Lemma 10.1(b), proven later. (Note also that the reason we must have $\hat{\Theta}(p) = 1/2$ for the case $\hat{p} \in \Gamma$, rather than possibly some multiple of 1/2, is that for $\hat{p} \in \Gamma$ then $\tilde{M}_{\hat{p},\infty}$ can only ever be a multiplicity one half-plane. The proof that you cannot get higher multiplicity half-planes, or unions of half-planes, mimics the reasoning used later in Sections 10 and 12 to obtain improved boundary strip estimates, so we omit further discussion of this point here). Q.E.D.

5.4 Remark. The last part of Lemma 5.3 gives us a characterisation of those points, $p \in M_0$, which "end up on the boundary" (that is, have $\hat{p} \in \Gamma$). Geometrically it seems likely that the only such points will be the actual boundary points, $p \in \Gamma$. However we have not yet managed to rule out the possibility that there might be some points, $p \in \operatorname{int}(M_0)$, which satisfy $\hat{p} \in \Gamma$, and thus really "behave like part of the boundary".

From the point of view of the analysis that follows, we need to treat any such points differently from the rest of the interior of M_0. This leads us to introduce the set

$$\Gamma^+ \equiv \{p \in M_0 : \hat{p} \in \Gamma\} \equiv \{p \in M_0 : \hat{\Theta}(p) = 1/2\} \,.$$

5.5 Definition. We call Γ^+ the *extended boundary* of M_0 (w.r.t. the time T).

From Lemma 5.3, and the definition of Γ^+, we then have immediately;

5.6 Corollary. *The limiting heat density function, $\hat{\Theta}$, satisfies*

$$\begin{cases} \hat{\Theta}(p) \geq 1 & , p \in M_0 \backslash \Gamma^+ \\ \hat{\Theta}(p) \equiv 1/2 & , p \in \Gamma^+ \end{cases} . \tag{5.5}$$

Also, we can now state the lemma, mentioned earlier, relating to existence, and "partial semicontinuity", for the function $\hat{\Theta}$. These two issues are addressed in parts (a) and (b) of the lemma, while parts (c) and (d) give some useful "uniformity" results, for the convergence of the functions $\Theta(p,t)$ to $\hat{\Theta}(p)$, which we shall need in the sequel.

5.7 Lemma. (a) *The limit in the definition of* $\hat{\Theta}$ *(see (5.2)) does exist everywhere on* M_0.

(b) *On any compact subset* $K \subset (M_0\backslash\Gamma^+)$ *the limiting heat density function* $\hat{\Theta}$ *will be upper semicontinuous (so essentially* $\hat{\Theta}$ *can only "jump down" at the boundary).*

(c) *If* $\hat{\Theta} \equiv 1$ *on some compact set* $K \subset (M_0\backslash\Gamma^+)$*, then, for any* $\epsilon > 0$*, there is a* $t_{\epsilon,K} < T$ *such that, for any* $t \in [t_{\epsilon,K}, T)$,

$$\|\Theta(p,t) - \hat{\Theta}(p)\|_{L^\infty(K)} < \epsilon\,.$$

In other words, on any compact subset of $M_0\backslash\Gamma^+$ *on which* $\hat{\Theta} \equiv 1$*, then* $\Theta(p,t) \to \hat{\Theta}(p)$ *uniformly, as* $t \to T$.

(d) *On the extended boundary,* Γ^+, $\Theta(p,t) \to \hat{\Theta}(p)$ *uniformly, as* $t \to T$.

Proof. **Of (a).** For ease set $b : M_0 \times [0,T) \to \mathbf{R}$ ("b" for "boundary term") by

$$b(p,t) \equiv \int_t^T \int_\Gamma \left\langle \frac{\mathbf{x}-\hat{p}}{2\tau}, \eta \right\rangle \rho_{\hat{p}}\, dH^{n-1}\, dt\,. \tag{5.6}$$

Then (2.1) says precisely that, for any $p \in M_0$,

$$\frac{d}{dt}(\Theta(p,t) - b(p,t)) = -\int_{M_t} \left|\mathbf{H} + \frac{(\mathbf{x}-\hat{p})^\perp}{2\tau}\right|^2 \rho_{\hat{p}}\, dH^n \le 0\,. \tag{5.7}$$

Hence $(\Theta - b)(\cdot, t)$ has a pointwise limit everywhere on M_0.

It remains then only to show that the same is true of $b(\cdot,t)$ alone. But this is clear from Lemmas 2.6(b) (for $\hat{p} \notin \Gamma$) and 2.7(b) (for $\hat{p} \in \Gamma$), which in fact show that

$$\lim_{t\to T} b(p,t) \equiv 0 \quad \text{for all} \quad p \in M_0\,. \tag{5.8}$$

Of (b). Fix K any compact set in $M_0\backslash\Gamma^+$, and put $\hat{K} = \{\hat{p} : p \in K\}$. Then, by Lemma 3.2(a), $\hat{K}$ is also compact (in $\mathbf{R}^{n+1}$), and does not meet Γ. Hence $d_0 \equiv \operatorname{dist}(\hat{K}, \Gamma) > 0$.

Next observe that, restricted to K, the functions $(\Theta - b)(p,t)$ are all continuous in p, for each fixed $t < T$, since this is clearly true for each of Θ and b separately. (Note that, while this continuity is actually true for $\Theta(p,t)$ on the whole of M_0, this is *not* true for $b(p,t)$, which, it turns out, is discontinuous onto the extended boundary, for all t near to T. This is why we have to restrict to compact subsets $K \subset M_0 \backslash \Gamma^+$).

Therefore, by (5.7), the function $\lim_{t \to T}(\Theta - b)(\cdot, t)$ must be upper semi-continuous, when restricted to K. But now, by (5.8), this function is just the same as $\Theta(\cdot)\big|_K$, so we are done.

Of (c). Again fix K as in part (b) above, and let $\hat{K}$, d_0 be as there. Then, by Lemma 2.6(b), the functions $b(\cdot, t)\big|_K$ are in fact converging *uniformly* to 0, as $t \to T$.

Thus, to prove the claim, it will suffice to show uniform convergence of the functions $(\Theta - b)(\cdot, t)\big|_K$ to the constant function "1". Yet this is clear. For we certainly have such convergence pointwise, and therefore the functions $(\Theta - b)(\cdot, t)\big|_K$ form, by (5.7), a monotone decreasing family of functions converging to a continuous (indeed constant) limit, whence the convergence must be uniform, as desired.

Of (d). By part (a) above, along with Corollary 5.6 and Equation (5.8), the functions $\Theta(\cdot, t)|_{\Gamma^+}$ and $b(\cdot, t)|_{\Gamma^+}$ each separately have constant pointwise limits (namely "1/2" and "0" respectively), as $t \to T$. Thus also $(\Theta - b)(\cdot, t)|_{\Gamma^+}$ has a constant pointwise limit (viz. "1/2"), and therefore, in view of the monotonicity from (5.7), the functions $(\Theta - b)(\cdot, t)|_{\Gamma^+}$ must actually converge uniformly, as $t \to T$, to the function "1/2".

However we also know, by Lemma 2.7(b), that, in fact,

$$\|b(\cdot, t)\|_{L^\infty(\Gamma^+)} \to 0 \quad \text{as} \quad t \to T.$$

Hence, actually, we must have that $\|\Theta(\cdot, t) - 1/2\|_{L^\infty(\Gamma^+)} \to 0$ as $t \to T$. Q.E.D.

6. Statement of partial boundary regularity result

We can now state the main result we've been building up to, the first part of which is exactly analogous to Theorem 3.1 of [10]. The proof is given in Sections 7 and 8.

6.1 Theorem. *Suppose, as always, that Hypotheses A and B of Section 1 hold.*

(a) *Then if $p \in M_0$ is a general Type I singular point of the flow, as $t \to T$, then it is actually a special Type I singular point. (Recall that the notions of special and general singular points were defined in Section 1).*

(b) *Moreover, any such general Type I singular point must occur well in the interior of $M_0 \backslash \Gamma^+$, in that there is a fixed neighbourhood, $\mathcal{U}$, of the extended boundary, Γ^+, such that, for any $\delta > 0$, there is a time*

$\hat{t}_\delta < T$ with the property that any δ-essential blow-up sequence must leave the closure of $\mathcal{U}$ for good by time $\hat{t}_\delta$.

6.2 Remarks. **(i)** An equivalent formulation of part (b) here is that there is a fixed neighbourhood, $\mathcal{U}$, of Γ^+, and a function $\epsilon(t)$ satisfying $\epsilon(t) \searrow 0$ as $t \nearrow T$, such that, for all $t \in [0, T)$, $\max_{\mathcal{U}} |A|^2(\cdot, t) \leq \epsilon(t)/(T-t)$.

(ii) The key ingredient in the proof will be the monotonicity identities for surfaces with boundary, viz. (2.1) and (2.2), and their rescaled versions, (2.7) and (2.8).

Also, as in the boundaryless case, the proof will centre around study of the limit heat density function, $\hat{\Theta}$, on M_0. However the analysis is more complicated here because, as noted earlier, the functions $\Theta(p, t)$ no longer go monotonically down to $\hat{\Theta}$ as $t \to T$, for each p, nor is $\hat{\Theta}$ then globally upper semicontinuous. (Indeed we saw already, in Corollary 5.6, that $\hat{\Theta}$ satisfies $\hat{\Theta}(p) \geq 1$ on $M_0 \backslash \Gamma^+$, but $\hat{\Theta}(p) = 1/2$ on Γ^+).

This non-semicontinuity of $\hat{\Theta}$ means, of course, that the functions $\Theta(p, t)$ cannot here be assumed to converge uniformly to $\hat{\Theta}$, as $t \to T$, locally about non blow-up points, which was the key fact in the contradiction argument in the boundaryless case. However we shall see that it is enough to establish a form of "*one-sided uniformity*" result ("uniformity from above"), namely to show that:

6.3 Lemma. **(a)** *For any $\epsilon > 0$ there will be a time $t_\epsilon < T$ and an open neighbourhood $\mathcal{U}_\epsilon \subseteq M_0$ of Γ^+ such that, for all $t \in [t_\epsilon, T)$ and all $p \in \bar{\mathcal{U}}_\epsilon$,*

$$\Theta(p, t) \leq 1 + \epsilon\,. \tag{6.1}$$

(b) *Also, if $\hat{\Theta} \equiv 1$ on any given compact $K \subset M_0 \backslash \Gamma^+$, then the same sort of "uniform bound from above" is true on K, namely that, for any $\epsilon > 0$, there will be a time $t_{\epsilon,K} < T$ (just taken to be as in Lemma* 5.7(c)*) such that, for all $t \in [t_{\epsilon,K}, T)$ and all $p \in K$,*

$$\Theta(p, t) \leq 1 + \epsilon\,. \tag{6.2}$$

Moreover, if $\hat{\Theta} \equiv 1$ on all of $M_0 \backslash \Gamma^+$, then the same is true on the whole of M_0 (in this case we denote the relevant time by t_{ϵ, M_0}).

7. Outline of proofs of Lemma 6.3 and Theorem 6.1

Before giving the proofs of Lemma 6.3 and Theorem 6.1 we first give a brief description of the method we will use to attack these results. The 10-step procedure set out is then carried out in Section 8.

A. Preliminary results and outline of proof of Lemma 6.3

Step 1 : Show that for any $\epsilon > 0$ there exists a $t_\alpha(n, \Gamma, \epsilon) < T$ such that, for any point $p \in M_0 \backslash \Gamma^+$,

$$\int_{t_\alpha(\epsilon)}^{T} \int_{\Gamma} \left| \left\langle \frac{\mathbf{x} - \hat{p}}{2\tau}, \eta \right\rangle \right| \rho_{\hat{p}} dH^{n-1} dt \leq \frac{1}{2} + \epsilon \,. \tag{7.1}$$

More generally show that, for any $\epsilon \in (0, 1/2)$, there exists a $t_\beta(n, \Gamma, \epsilon) < T$ such that, for all $t_1 \in [t_\beta(\epsilon), T)$,

$$\frac{\operatorname{dist}(\hat{p}, \Gamma)}{\sqrt{2(T - t_1)}} \geq d_1 > 0 \quad \Longrightarrow$$

$$\int_{t_1}^{T} \int_{\Gamma} \left| \left\langle \frac{\mathbf{x} - \hat{p}}{2\tau}, \eta \right\rangle \right| \rho_{\hat{p}} dH^{n-1} dt \leq C_8(n) \int_{d_1/2}^{\infty} e^{-w^2/2} dw \, + \, \epsilon \,. \tag{7.2}$$

Note that (7.1) is saying that, after a certain time $t_\alpha(\epsilon) < T$, the "boundary contribution" in (2.2) (with "centre" $\hat{p}$) can give a "*kick up*" of at most $(1/2 + \epsilon)$ to $\Theta(p, t)$, as time progresses towards T, for *any* $p \in M_0 \backslash \Gamma^+$. Estimate (7.2) should be viewed as a (considerable) strengthening of Lemma 2.6(b).

Note also that these two estimates turn out to be solely properties of the boundary Γ (in particular the norm of its second fundamental form as a subset of $\mathbf{R}^{n+1}$), and could be discussed entirely in the absence of any associated flow of surfaces M_t (since the only property of the conormal η that we use is that it is perpendicular to Γ at each point).

Step 2 : Using Step 1 and the uniform convergence result on Γ^+ of Lemma 5.7(d), prove part (a) of the "one-sided uniformity" result, Lemma 6.3.

Step 3 : Using the uniform convergence result on compact subsets of $\operatorname{int}(M_0)$ of Lemma 5.7(c), together with the result of Lemma 6.3(a) just proven, establish part (b) of Lemma 6.3.

B. Outline of proof of Theorem 6.1

Now suppose $\{p_k, t_k\}$ is a "δ-essential blow-up sequence", for some fixed $\delta > 0$, with $p_k \to p$ as $k \to \infty$; so p is a general Type I singular point of the flow. Note that, by Lemma 3.2(a), we have then also that

$$\hat{p}_k \to \hat{p} \quad \text{as} \quad k \to \infty \,. \tag{7.3}$$

We now want to prove, first of all, that p must also be a special Type I singular point.

To this end, for each dimension n, let ϵ_n be the difference between 1 and the next largest possible value that $\hat{\Theta}$ can take. (Such a gap exists, and indeed is explicitly computable, by Lemma 5.3).

Step 4 : Establish, using Lemma 6.3(a) (and with the notation as there), that we must have

$$\hat{\Theta}(q) \equiv 1 \quad , \quad q \in \bar{\mathcal{U}}_{\epsilon_n/2} \backslash \Gamma^+ . \tag{7.4}$$

Step 5 : Using Lemma 5.7(d), prove that no subsequence of the points p_k can lie in Γ^+; so also without loss of generality we may assume, as we do henceforth, that

$$p_k \in M_0 \backslash \Gamma^+ \quad \text{for all} \quad k . \tag{7.5}$$

Now let $\mathcal{W}$ denote *any* open set in $M_0\backslash\Gamma^+$ containing the closed (possibly empty) set $\mathcal{B} = \{q : \hat{\Theta}(q) > 1\} \equiv \{q : \hat{\Theta}(q) \geq 1 + \epsilon_n\}$. By Step 4 this latter set stays well away from Γ^+, so such sets $\mathcal{W}$ exist. Then we want to establish;

7.1 Lemma. *The p_k must eventually lie in $\mathcal{W}$, for large enough k (depending of course on $\mathcal{W}$).*

To do this we argue by contradiction. Suppose instead that a subsequence (still denoted p_k) remained in the *closed* set $M_0\backslash\mathcal{W}$. Then we have to consider three cases.

Case (a) : ("The $\hat{p}_k$ stay away from Γ") i.e. $\hat{p} \notin \Gamma$.

Case (b) : ("The $\hat{p}_k$ go to Γ very fast") i.e. for some subsequence (still denoted p_k) we have

$$\frac{\operatorname{dist}(\hat{p}_k, \Gamma)}{\sqrt{2(T - t_k)}} \leq C_* \tag{7.6}$$

for some constant, C_*, independent of k.

Case (c) : ("The $\hat{p}_k$ go to Γ only slowly") i.e. we have $\hat{p} \in \Gamma$, but

$$\liminf_{k\to\infty} \frac{\operatorname{dist}(\hat{p}_k, \Gamma)}{\sqrt{2(T - t_k)}} = \infty . \tag{7.7}$$

Step 6 : Rule out Case (a), using Lemma 6.3(b) with $K = M_0\backslash\mathcal{W}$, along with a rescaling argument based on the "Monotonicity Identities".

Step 7 : Rule out Case (b), as follows. For each k let $\bar{p}_k$ be a point on Γ that achieves $\operatorname{dist}(\hat{p}_k, \Gamma)$. Then note that, on a suitable sequence of rescalings about the points $\bar{p}_k$, we would detect points, within a fixed radius ball (determined by C_* and the "Type I constant", C_0) about the origin, with second fundamental form at these points satisfying $|\tilde{A}|^2 \geq \delta$.

Now obtain a contradiction, however, by showing, via a rescaling argument based on the "Monotonicity Identities", that some subsequence of these rescalings must be converging to a half-plane.

Step 8 : Rule out Case (c), by combining estimate (7.2) and the "uniform convergence from above in a boundary neighbourhood" result of Lemma 6.3(a), again along with a rescaling argument.

This then finishes the proof of Lemma 7.1. Finally we then have;

Step 9 : Establish part (b) of Theorem 6.1, using Lemma 7.1 and Step 4.

Step 10 : Establish part (a) of Theorem 6.1 via a contradiction argument, using Lemmas 6.3(b) and 7.1, together with the upper semicontinuity result of Lemma 5.7(b).

8. Proofs of Lemma 6.3 and Theorem 6.1

We now carry out in detail the 10-step procedure described in Section 7, so proving Theorem 6.1.

Step 1 : To establish the claims in (7.1) and (7.2) we first need some notation. For all $\mathbf{x}_0 \in \Gamma$ let $\mathcal{R}_{\mathbf{x}_0}$ denote a rotation in $\mathbf{R}^{n+1}$ that sends $T_{\mathbf{x}_0}\Gamma$ to $\mathbf{R}^{n-1} \times \{(0,0)\}$. Then, for all $\mathbf{x} \in \mathbf{R}^{n+1}$, set

$$\mathbf{x}'_{\mathbf{x}_0} \equiv \mathcal{R}_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0)\,, \tag{8.1}$$

and also put

$$\Gamma'_{\mathbf{x}_0} \equiv \mathcal{R}_{\mathbf{x}_0}(\Gamma - \mathbf{x}_0)\,. \tag{8.2}$$

For later reference, note that then, for any $\mathbf{x} \in \mathbf{R}^{n+1}$, if $\mathbf{x}_0$ is a "closest point of Γ" to $\mathbf{x}$ (that is, satisfies $|\mathbf{x} - \mathbf{x}_0| = \operatorname{dist}(\mathbf{x}, \Gamma)$), then

$$\mathbf{x}'_{\mathbf{x}_0} \in \text{ the 2-plane } \{0\} \times \mathbf{R}^2 \subset \mathbf{R}^{n+1}\,. \tag{8.3}$$

Next observe that, for any $\epsilon > 0$, we can find a *uniform* radius $r_0 = r_0(\epsilon, \Gamma)$, and constants $C_9(\Gamma)$ and $C_{10}(\Gamma)$, all determined only by ϵ and the sup norm of the second fundamental form of $\Gamma^{n-1} \subset \mathbf{R}^{n+1}$, such that, for any $\mathbf{x}_0 \in \Gamma$, the following are true.

First, the set $\Gamma'_{\mathbf{x}_0} \cap (B^{(n-1)}_{r_0}(0) \times [-r_0, r_0] \times [-r_0, r_0])$ may be written as the graph of a function $\gamma \equiv \gamma_{\mathbf{x}_0}$, where

$$\gamma = (\gamma_n, \gamma_{n+1}) : B^{(n-1)}_{r_0}(0) \longrightarrow \mathbf{R}^2 \tag{8.4}$$

is smooth; also this function γ satisfies, for all $y \in B^{(n-1)}_{r_0}(0)$,

$$|\gamma_n(y)|\,,\ |\gamma_{n+1}(y)| \leq |y|\,, \tag{8.5}$$

$$|\gamma_n(y)|\,,\ |\gamma_{n+1}(y)| \leq C_9|y|^2\,, \tag{8.6}$$

$$\sqrt{1+\left|D^{R^{n-1}}\gamma_n(y)\right|^2+\left|D^{R^{n-1}}\gamma_{n+1}(y)\right|^2} \le 1+\epsilon\,, \tag{8.7}$$

and

$$\left\|P_{(T_{(y,\gamma(y))}\Gamma)^\perp}-P_{\{0\}\times R^2}\right\| \le C_{10}|y|\,. \tag{8.8}$$

In (8.8) we have used P_V to denote orthogonal projection onto a subspace, V, of $\mathbf{R}^{n+1}$, and then $\|\cdot\|$ is just the usual operator norm on $(n+1)\times(n+1)$ matrices.

So now fix $\epsilon\in(0,1)$ arbitrarily, and also, for each $p\in M_0\backslash\Gamma^+$, let p_* denote a point of Γ which *achieves* $\operatorname{dist}(\hat p,\Gamma)$; so that then, by (8.3), $\hat p'_{p_*}$ satisfies, for each such p,

$$\hat p'_{p_*}\in\{0\}\times\mathbf{R}^2\subset\mathbf{R}^{n+1}\,. \tag{8.9}$$

Then, first of all, we claim that there will be some *uniform* time $\bar t_\epsilon(n,\Gamma,\epsilon)<T$ such that, for any fixed $p\in M_0\backslash\Gamma^+$,

$$\int_{\bar t_\epsilon}^{T}\int_{\Gamma}\rho_{\hat p}\left|\left\langle\frac{\mathbf{x}-\hat p}{2\tau},\eta\right\rangle\right|dH^{n-1}dt \le \epsilon+ \tag{8.10}$$
$$(1+\epsilon)^2\int_{\bar t_\epsilon}^{T}\int_{R^{n-1}\times\{(0,0)\}}\frac{1}{(4\pi\tau)^{n/2}}\,e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat p'|^2}{4\tau}\right)}\left(\frac{|\hat p'|}{2\tau}\right)dy\,dt\,.$$

Here we have written $\hat p'$ for $\hat p'_{p_*}$. Henceforth, also, we will write $\mathbf{x}'$ for $\mathbf{x}'_{p_*}$, and will use the shorthand $\mathbf{R}^{n-1}$ for $\mathbf{R}^{n-1}\times\{(0,0)\}$ in integrals.

To see (8.10) (which is actually the crux of the proofs of both (7.1) and (7.2)), observe that, for any $p\in M_0\backslash\Gamma^+$, the left hand side of this inequality is clearly equivalent to the quantity $I(p,\bar t_\epsilon)$, where, for $t_*<T$, $I(p,t_*)$ is defined by

$$I(p,t_*)\equiv\int_{t_*}^{T}\int_{\Gamma'_{p_*}}\frac{1}{(4\pi\tau)^{n/2}}\,e^{-\left(\frac{|x'-\hat p'|^2}{4\tau}\right)}\left|\left\langle\frac{\mathbf{x}'-\hat p'}{2\tau},\eta'\right\rangle\right|dH^{n-1}dt\,. \tag{8.11}$$

Here η' denotes $\mathcal{R}_{p_*}(\eta)$. Thus, to establish (8.10), we need to show that we can find a uniform time $\bar t_\epsilon<T$ such that, for *any* $p\in M_0\backslash\Gamma^+$,

$$I(p,\bar t_\epsilon)\le\epsilon+(1+\epsilon)^2\int_{\bar t_\epsilon}^{T}\int_{R^{n-1}}\frac{1}{(4\pi\tau)^{n/2}}\,e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat p'|^2}{4\tau}\right)}\left(\frac{|\hat p'|}{2\tau}\right)dy\,dt\,. \tag{8.12}$$

For this we proceed in similar fashion to the proof of Lemma 2.7. First fix $p \in M_0 \backslash \Gamma^+$ arbitrarily. Then set

$$r(t) \equiv \sqrt{-2(n+2)\tau \ln(\tau)}\,. \tag{8.13}$$

Now, *a la* Lemma 2.7, split Γ'_{p_*} into pieces, this time as

$$\Gamma_1(t) = \Gamma'_{p_*} \cap \left(B^{(n-1)}_{r(t)}(0) \times [-r(t), r(t)]^2\right) \cap B_{r(t)}(\hat{p}')\,, \tag{8.14}$$

$$\Gamma_2(t) = \left(\Gamma'_{p_*} \cap \left(B^{(n-1)}_{r(t)}(0) \times [-r(t), r(t)]^2\right)\right) \Big\backslash B_{r(t)}(\hat{p}')\,, \tag{8.15}$$

$$\Gamma_3(t) = \Gamma'_{p_*} \backslash \left(\Gamma_1(t) \cup \Gamma_2(t)\right)\,. \tag{8.16}$$

Note that, for each $p \in M_0 \backslash \Gamma^+$, $\Gamma_1(t)$ will become empty at some time prior to T. However this time will, of course, depend on p. As we want an estimate that is uniform in p we therefore do not use this in the sequel.

Next, for $i = 1, 2, 3$, set

$$I_i(t_*) \equiv \int_{t_*}^{T} \int_{\Gamma_i(t)} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-\left(\frac{|x'-\hat{p}'|^2}{4\tau}\right)} \left| \left\langle \frac{\mathbf{x}' - \hat{p}'}{2\tau}, \eta' \right\rangle \right| dH^{n-1} dt\,, \tag{8.17}$$

so that then, writing $I(t_*)$ for $I(p, t_*)$, we have, for all $t_* \in [0, T)$,

$$I(t_*) = I_1(t_*) + I_2(t_*) + I_3(t_*)\,. \tag{8.18}$$

Now to handle $I_2(t_*)$ and $I_3(t_*)$ (the easy two of the three), observe that, since p_* was a "closest point of Γ" to $\hat{p}$, so 0 is a "closest point of Γ'_{p_*}" to $\hat{p}'$. Hence, for all $\mathbf{x}' \in \Gamma_3(t)$, we have that, for all $t \in [0, T)$,

(8.19)

$$|\mathbf{x}' - \hat{p}'| \geq \max\Big\{|\hat{p}'|, (|\mathbf{x}'| - |\hat{p}'|)\Big\} \geq \max\Big\{|\hat{p}'|, (r(t) - |\hat{p}'|)\Big\} \geq \frac{1}{2} r(t)\,.$$

Likewise, just from the definition of $\Gamma_2(t)$, we have that also, for all $\mathbf{x}' \in \Gamma_2(t)$ and $t \in [0, T)$,

$$|\mathbf{x}' - \hat{p}'| \geq r(t)\,. \tag{8.20}$$

Therefore, exactly as in Lemma 2.7 (cf. in particular (2.23) and (2.31)), we obtain that, for all $t_* \in [0, T)$, $I_2(t_*) + I_3(t_*) \leq C_{11}(n, \Gamma)(T - t_*)$; and hence, provided now that t_* is always assumed to be greater than some $t_a(n, \Gamma, \epsilon) < T$, we have that

$$I_2(t_*) + I_3(t_*) \leq \epsilon/2\,. \tag{8.21}$$

It remains to handle $I_1(t_*)$. For this, first let us restrict our attention only to those t_* such that

$$r(t) \le \max\{\epsilon, r_0(\epsilon,\Gamma)\} \quad \text{for all} \quad t \in [t_*, T)\,. \tag{8.22}$$

Clearly, from the definition of $r(t)$, this will hold for all $t_* \in [t_b, T)$, where $t_b < T$ is some constant depending only on n, Γ and ϵ.

Then note that, for all $t \in [t_*, T)$, we have, by (8.4), that $\Gamma_1(t) = graph(\gamma)$, where $\gamma = (\gamma_n, \gamma_{n+1}) : B^{(n-1)}_{r_0(\epsilon,\Gamma)}(0) \longrightarrow \mathbf{R}^2$ is smooth, and satisfies (8.5) to (8.8) above.

In particular, then, on $\Gamma_1(t)$ we have, for any $\mathbf{x}' = (y, \gamma(y))$, that

$$(\mathbf{x}' - \hat{p}') \equiv \Big((y,\gamma(y)) - \hat{p}'\Big) = (\mathbf{y} - \hat{p}') + (0, \gamma(y))\,, \tag{8.23}$$

where here, and henceforth, $\mathbf{y} \in \mathbf{R}^{n+1}$ denotes the vector $(y, 0, 0)$. Using this, together with (8.3), (8.6), (8.8) and the definition of $\Gamma_1(t)$, we may estimate that, for all $\mathbf{x}' = (y, \gamma(y)) \in \Gamma_1(t)$, $t \in [t_*, T)$,

$$\begin{aligned}
|\langle \mathbf{x}' - \hat{p}', \eta' \rangle| &\le \Big|P_{(T_{(y,\gamma(y))}\Gamma)^\perp}(\mathbf{x}' - \hat{p}')\Big| \\
&\le \Big|\Big(P_{(T_{(y,\gamma(y))}\Gamma)^\perp} - P_{\{0\}\times R^2}\Big)(\mathbf{x}' - \hat{p}')\Big| + \Big|P_{\{0\}\times R^2}(\mathbf{x}' - \hat{p}')\Big| \\
&\le C_{10}|y||\mathbf{x}' - \hat{p}'| + \Big|P_{\{0\}\times R^2}(\mathbf{y} - \hat{p}')\Big| + \Big|P_{\{0\}\times R^2}(0, \gamma(y))\Big| \\
&\le C_{10}|y| r(t) + |\hat{p}'| + 2C_9|y|^2 \le C_{12}(\Gamma) r(t)^2 + |\hat{p}'|\,.
\end{aligned} \tag{8.24}$$

Similarly, from (8.22) and (8.23), we may estimate that, for all $\mathbf{x}' = (y, \gamma(y)) \in \Gamma_1(t)$, $t \in [t_*, T)$,

$$\begin{aligned}
|\mathbf{x}' - \hat{p}'|^2 &\ge (1-\epsilon)|\mathbf{y} - \hat{p}'|^2 - (\epsilon^{-1} - 1)|\gamma(y)|^2 \\
&\ge (1-\epsilon)|\mathbf{y} - \hat{p}'|^2 - 2(\epsilon^{-1} - 1)C_9^2|y|^4 \\
&\ge (1-\epsilon)|\mathbf{y} - \hat{p}'|^2 - C_{13} r(t)^3
\end{aligned} \tag{8.25}$$

where again $C_{13} = C_{13}(\Gamma)$. Therefore, provided that t_* is, in addition, now always bigger than $t_c < T$, where $t_c(n, \Gamma)$ is such that

$$\exp\left(\frac{1}{4}(2(n+2))^{3/2} C_{13} \tau^{1/2} (-\ln(\tau))^{3/2}\right) \le 1 + \epsilon \tag{8.26}$$

for all $t \in [t_c, T)$, we obtain that also, for all $\mathbf{x}' = (y, \gamma(y)) \in \Gamma_1(t)$, $t \in [t_*, T)$,

$$e^{-\left(\frac{|x' - \hat{p}'|^2}{4\tau}\right)} \le (1+\epsilon) e^{-(1-\epsilon)\left(\frac{|(y,0,0) - \hat{p}'|^2}{4\tau}\right)}\,. \tag{8.27}$$

But then, bringing together (8.24) and (8.27), we may deduce that, for all $t \in [t_*, T)$,

$$\begin{aligned}
&\int\limits_{\Gamma_1(t)} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-\left(\frac{|x'-\hat{p}'|^2}{4\tau}\right)} \left|\left\langle \frac{\mathbf{x}'-\hat{p}'}{2\tau}, \eta' \right\rangle\right| dH^{n-1} \\
(8.28)\qquad &\leq (1+\epsilon)^2 \int\limits_{B^{(n-1)}_{r(t)}(0)} \frac{(C_{12}r(t)^2 + |\hat{p}'|)}{(2\tau)(4\pi\tau)^{n/2}}\, e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)}\, dy
\end{aligned}$$

where, to estimate the measure, we have also used (8.7). Therefore, for all times t_* sufficiently large (determined by n, Γ and ϵ, but *not* by p), we have (noting (8.13)) that

$$\begin{aligned}
I_1(t_*) &\leq (1+\epsilon)^2 \int\limits_{t_*}^{T} \int\limits_{R^{n-1}} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)} \frac{|\hat{p}'|}{2\tau}\, dy\, dt \\
&\qquad + (1+\epsilon)^2 C_{14}(n,\Gamma) \int\limits_{t_*}^{T} \tau^{-1/2}(-\ln(\tau))^{\frac{n+1}{2}}\, dt \\
(8.29)\qquad &\leq (1+\epsilon)^2 \int\limits_{t_*}^{T} \int\limits_{R^{n-1}} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)} \frac{|\hat{p}'|}{2\tau}\, dy\, dt \;+\; \epsilon/2\,.
\end{aligned}$$

Finally then, (8.29) and (8.21) in (8.18) immediately yield (8.12), and so (8.10) is proven.

But now, it turns out, we are nearly done! For, taking the integral on the right hand side of (8.10), we have, by (8.3), that

$$\begin{aligned}
&\int\limits_{\bar{t}_\epsilon}^{T} \int\limits_{R^{n-1}} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)} \left(\frac{|\hat{p}'|}{2\tau}\right) dy\, dt \\
&= \int\limits_{\bar{t}_\epsilon}^{T} (2\pi)^{-n/2} \left(\frac{|\hat{p}'|}{(2\tau)^{3/2}}\right) e^{-(1-\epsilon)\left(\frac{|\hat{p}'|^2}{4\tau}\right)} \left(\int\limits_{R^{n-1}} \frac{e^{-(1-\epsilon)\left(\frac{|y|^2}{4\tau}\right)}}{\left(\sqrt{2\tau}\right)^{n-1}}\, dy \right) dt \\
(8.30)\qquad &= \frac{1}{\sqrt{2\pi}}(1-\epsilon)^{-\left(\frac{n-1}{2}\right)} \int\limits_{\bar{t}_\epsilon}^{T} \left(\frac{|\hat{p}'|}{(2\tau)^{3/2}}\right) e^{-(1-\epsilon)\left(\frac{|\hat{p}'|^2}{4\tau}\right)}\, dt\,.
\end{aligned}$$

Hence, by the substitution $w \equiv \sqrt{1-\epsilon}\,|\hat{p}'|/\sqrt{2\tau}$, we obtain that

$$\int_{\bar{t}_\epsilon}^{T} \int_{R^{n-1}} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)} \left(\frac{|\hat{p}'|}{2\tau}\right) dy\, dt$$

$$= \frac{1}{\sqrt{2\pi}}(1-\epsilon)^{-n/2} \int_{\sqrt{1-\epsilon}\,|\hat{p}'|/\sqrt{2(T-\bar{t}_\epsilon)}}^{\infty} e^{-w^2/2}\, dw\,. \tag{8.31}$$

But then, in particular, we have that

$$\int_{\bar{t}_\epsilon}^{T} \int_{R^{n-1}} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)} \left(\frac{|\hat{p}'|}{2\tau}\right) dy\, dt$$

$$\leq \frac{1}{\sqrt{2\pi}}(1-\epsilon)^{-n/2} \int_{0}^{\infty} e^{-w^2/2}\, dw = \frac{1}{2}(1-\epsilon)^{-n/2}\,, \tag{8.32}$$

whence, by (8.32) in (8.10), it follows that, for any arbitrary $\epsilon \in (0,1)$, we have found a uniform time $\bar{t}_\epsilon(n,\Gamma,\epsilon) < T$ such that, for *any* $p \in M_0\backslash\Gamma^+$,

$$\int_{\bar{t}_\epsilon}^{T} \int_{\Gamma} \rho_{\hat{p}} \left|\left\langle \frac{\mathbf{x}-\hat{p}}{2\tau}, \eta \right\rangle\right| dH^{n-1} dt \leq \frac{(1+\epsilon)^2 (1-\epsilon)^{-n/2}}{2} + \epsilon\,. \tag{8.33}$$

This clearly proves the claim in (7.1). As for the claim in (7.2), we proceed similarly. Here we have $\epsilon \in (0, 1/2)$ arbitrary. Thence, by (8.31), for any $t_1 \in [\bar{t}_\epsilon, T)$ we have

$$\int_{t_1}^{T} \int_{R^{n-1}} \frac{1}{(4\pi\tau)^{n/2}} e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)} \left(\frac{|\hat{p}'|}{2\tau}\right) dydt \leq \frac{2^{n/2}}{\sqrt{2\pi}} \int_{\frac{\sqrt{1-\epsilon}\,|\hat{p}'|}{\sqrt{2(T-t_1)}}}^{\infty} e^{-w^2/2} dw. \tag{8.34}$$

But then, noting $|\hat{p}'| \equiv \operatorname{dist}(\hat{p},\Gamma)$, we obtain, for any $t_1 \in [\bar{t}_\epsilon, T)$, that

$$\frac{\operatorname{dist}(\hat{p},\Gamma)}{\sqrt{2(T-t_1)}} \geq d_1 > 0 \quad \Longrightarrow$$

$$\int_{t_1}^{T} \int_{R^{n-1}} \frac{1}{(4\pi\tau)^{n/2}}\, e^{-(1-\epsilon)\left(\frac{|(y,0,0)-\hat{p}'|^2}{4\tau}\right)} \left(\frac{|\hat{p}'|}{2\tau}\right) dy\, dt \tag{8.35}$$

$$\leq \frac{2^{n/2}}{\sqrt{2\pi}} \int_{\sqrt{1-\epsilon}\, d_1}^{\infty} e^{-w^2/2}\, dw\,.$$

Therefore, from (8.35) in (8.29), together with (8.21), (8.18) and (8.11), it follows that, for any $\epsilon \in (0,1/2)$, we have found a uniform $\bar{t}_\epsilon < T$ such that, for any $t_1 \in [\bar{t}_\epsilon, T)$,

$$\frac{\operatorname{dist}(\hat{p},\Gamma)}{\sqrt{2(T-t_1)}} \geq d_1 > 0 \quad \Longrightarrow$$

$$\int_{t_1}^{T}\int_{\Gamma}\left|\left\langle\frac{\mathbf{x}-\hat{p}}{2\tau},\eta\right\rangle\right|\rho_{\hat{p}}\,dH^{n-1}dt \leq C_8(n)\int_{d_1/2}^{\infty} e^{-w^2/2}dw + \epsilon\,. \tag{8.36}$$

Yet this is what we were trying to prove. Q.E.D. for Step 1.

Step 2 : To establish part (a) of Lemma 6.3, let $\epsilon > 0$ be arbitrary. Then, by Lemma 5.7(d), we can find a time $t_0(\epsilon) < T$ such that, for all $t \in [t_0(\epsilon), T)$,

$$\Theta(p,t) \leq \frac{1}{2} + \frac{\epsilon}{4} \quad \text{for all} \quad p \in \Gamma^+\,. \tag{8.37}$$

So now put $t_\epsilon \equiv \max\{t_0(\epsilon), t_\alpha(\epsilon/2)\}$, where $t_\alpha(\epsilon/2)$ is as in (7.1) with ϵ replaced by $\epsilon/2$. Also put

$$\mathcal{U}_\epsilon \equiv \left\{p \in M_0 : \Theta(p,t_\epsilon) < \frac{1}{2} + \frac{\epsilon}{2}\right\}. \tag{8.38}$$

Then, with these choices, we claim that Lemma 6.3(a) holds. For, by (8.37), along with the continuity of $\Theta(\cdot, t_\epsilon)$ as a function on M_0, $\mathcal{U}_\epsilon$ is then certainly an open neighbourhood of Γ^+ in M_0, and moreover

$$\Theta(p,t_\epsilon) \leq \frac{1}{2} + \frac{\epsilon}{2} \quad \text{for all} \quad p \in \bar{\mathcal{U}}_\epsilon\,. \tag{8.39}$$

But also, by (2.2), we have that for any $t_2 \in [t_\epsilon, T)$,

$$\Theta(p,t_2) - \Theta(p,t_\epsilon) \leq \left|\int_{t_\epsilon}^{t_2}\int_{\Gamma}\left\langle\frac{\mathbf{x}-\hat{p}}{2\tau},\eta\right\rangle\rho_{\hat{p}}\,dH^{n-1}dt\right| \tag{8.40}$$

whence, by Step 1, and the definition of t_ϵ,

$$\Theta(p,t_2) - \Theta(p,t_\epsilon) \leq \frac{1}{2} + \frac{\epsilon}{2} \quad \text{for all} \quad p \in M_0 \backslash \Gamma^+\,. \tag{8.41}$$

Thence, by (8.39) in (8.41) for the case $p \in \bar{\mathcal{U}}_\epsilon \backslash \Gamma^+$, and (8.37) for the case $p \in \Gamma^+$, we have that for all $t \in [t_\epsilon, T)$, and all $p \in \bar{\mathcal{U}}_\epsilon$, $\Theta(p,t) \leq 1 + \epsilon$. Q.E.D. for Step 2.

Step 3 : As noted, the first claim of Lemma 6.3(b) is immediate from Lemma 5.7(c), just by taking $t_{\epsilon,K}$ to be as there. (Indeed Lemma 5.7(c) actually proves a stronger uniform convergence result, not merely the uniform bound from above, (6.2)).

As for the second claim, let $\epsilon > 0$ be arbitrary, and then put $K_\epsilon \equiv M_0 \backslash \mathcal{U}_\epsilon$, $\mathcal{U}_\epsilon$ as in part (a) of Lemma 6.3. Then K_ϵ is a compact subset of $M_0 \backslash \Gamma^+$, so by above we can find a time $t_{\epsilon,K_\epsilon} < T$ such that $\Theta(p,t) \leq 1+\epsilon$, for all $p \in K_\epsilon$ and all $t \in [t_{\epsilon,K_\epsilon}, T)$.

But also, from part (a), we know that there is a time $t_\epsilon < T$ such that, for all $t \in [t_\epsilon, T)$, $\Theta(p,t) \leq 1+\epsilon$ for all $p \in \bar{\mathcal{U}}_\epsilon$.

So now just put $t_{\epsilon,M_0} \equiv \max\{t_{\epsilon,K_\epsilon}, t_\epsilon\}$, and the claim follows. Q.E.D. for Step 3.

Step 4 : By Lemma 6.3(a) we know that there is a time, $t_\alpha = t_\alpha(\epsilon_n/2)$, such that, for all $t \in [t_\alpha, T)$,

$$\Theta(q,t) \leq 1 + \epsilon_n/2 \quad \text{for all} \quad q \in \bar{\mathcal{U}}_{\epsilon_n/2} \,. \tag{8.42}$$

But then, by Lemma 5.3 and the definition of ϵ_n, we must have

$$\hat{\Theta}(q) = \frac{1}{2} \text{ or } 1 \quad \text{for all} \quad q \in \bar{\mathcal{U}}_{\epsilon_n/2} \,, \tag{8.43}$$

and then (7.4) follows, in view of Corollary 5.6. Q.E.D. for Step 4.

Step 5 : Now let $\{p_k, t_k\}$ be a "δ-essential blow-up sequence" satisfying (7.3), and suppose that some subsequence (still denoted $\{p_k\}$) of these points satisfied $p_k \in \Gamma^+$ for all k. We want to derive a contradiction.

To do so, consider the collection of rescalings $\tilde{M}_{\hat{p}_k,s}$. For each k these satisfy

$$0 \in \tilde{M}_{\hat{p}_k,s} \quad \text{for all } s \,, \tag{8.44}$$

and also, by hypothesis,

$$|\tilde{A}|^2(0,s_k) \geq \delta/2 \,, \tag{8.45}$$

where, as usual, s_k is the rescaled time corresponding to t_k.

But also, by Lemma 2.5, along with the evolution equation for $|\tilde{A}|^2$ (see, for instance, [3, Lemma 9.1]), we have the estimate, for all k, that

$$\max_{s\in[-\frac{1}{2}\ln(T),\infty)} \; \max_{\tilde{M}_{\hat{p}_k,s}} \left| \frac{d}{ds} |\tilde{A}|^2 \right| \leq C_{16} \tag{8.46}$$

where $C_{16} = C_{16}(n, C_0, \Gamma, \delta_0)$ is independent of k. Hence, in fact, (8.45) may be strengthened, to give that, for each k,

$$|\tilde{A}|^2(0,s) \geq \delta/4 \quad \text{for all} \quad s \in [s_k, s_k + \delta_1] \,, \tag{8.47}$$

where $\delta_1 \equiv \delta/4C_{16}$ is independent of k.

However, by the rescaled monotonicity identity, (2.8), we also have that, for all k,

$$\frac{1}{(2\pi)^{n/2}}\int_{s_k}^{s_k+\delta_1}\int_{\tilde{M}_{\hat{p}_k,s}}\left|\mathbf{H}+\mathbf{x}^{\perp}\right|^2\tilde{\rho}\,dH^n ds$$

(8.48)
$$= \{\Theta(p_k,t_k)-\Theta(p_k,t(s_k+\delta_1))\}+\{b(p_k,t(s_k+\delta_1))-b(p_k,t_k)\}$$

where the notation $b(\cdot,\cdot)$ is as in Lemma 5.7.

Then, by Lemma 5.7(d), the first bloc of (density) terms on the right hand side of (8.48) tend to 0, as $k\to\infty$ (independent of the fact that the points, p_k, are shifting around); and the same is true of the second bloc of (boundary integral) terms, by Lemma 2.7(b). Hence we have that also

$$\lim_{k\to\infty}\int_{s_k}^{s_k+\delta_1}\int_{\tilde{M}_{\hat{p}_k,s}}\left|\mathbf{H}+\mathbf{x}^{\perp}\right|^2\tilde{\rho}\,dH^n ds=0\,. \tag{8.49}$$

We are now in a position to obtain our contradiction. For, using also Lemma 2.5, we would then be able (just as in Theorem 3.4) to find a sequence of times $\{\tilde{s}_k\in[s_k,s_k+\delta_1]\}_{k=1}^{\infty}$, such that, for some subsequence of points (still denoted $\{p_k\}$), and times (still denoted $\{\tilde{s}_k\}$), the rescaled surfaces $\tilde{M}_{\hat{p}_k,\tilde{s}_k}$ converged smoothly on compact subsets to some limit surface, $\tilde{M}_\infty$, with boundary an $(n-1)$-plane through the origin. By (8.49) this limit surface would then also satisfy $\mathbf{H}=-\mathbf{x}^{\perp}$, and yet would have $|\tilde{A}|^2(0)\geq\delta/4$, in view of (8.47). This, however, is impossible, since, by Theorem 4.2, $\tilde{M}_\infty$ would have to be a half n-plane through the origin. Q.E.D. for Step 5.

Now, for the proofs of Steps 6,7 and 8, let the set-up be as described in Section 7; so, in particular, $\{p_k,t_k\}$ is again a δ-essential blow-up sequence, such that now (7.3) and (7.5) also hold, and $\mathcal{W}$ is as given there.

Step 6 : Then here we want to begin the proof of Lemma 7.1 by ruling out the possibility that we could have $p_k\in M_0\backslash\mathcal{W}$ for all k, yet have $\hat{p}\notin\Gamma$ (i.e. $p\notin\Gamma^+$).

So suppose we had this situation. Then, since $M_0\backslash\mathcal{W}$ is closed, we'd have $p\in M_0\backslash\mathcal{W}$ also. But then we'd clearly be able to find a compact set, $K\subset M_0\backslash\Gamma^+$, containing all the p_k, as well as p, and satisfying $K\subseteq M_0\backslash\mathcal{W}$, or in other words

$$\hat{\Theta}(q)\equiv 1\quad\text{for all}\quad q\in K\,. \tag{8.50}$$

However, by Lemma 5.7(c), we'd then have that

$$\|\Theta(\cdot,t)-1\|_{L^\infty(K)} \to 0 \quad \text{as} \quad t \to T\,. \tag{8.51}$$

Also, by Lemma 2.6(b), we'd have that

$$\|b(\cdot,t)\|_{L^\infty(K)} \to 0 \quad \text{as} \quad t \to T\,, \tag{8.52}$$

on noting that $d_0 \equiv \operatorname{dist}(\hat{K},\Gamma) > 0$, where $\hat{K}$ denotes $\{\hat{q} : q \in K\}$.

But then we'd have a contradiction just as in the proof of the previous step, via considering a sequence of rescalings about the (varying) centres $\{\hat{p}_k\}$. For, by proceeding exactly as there, we'd be able to find a subsequence of points, still denoted $\{p_k\}$, and a corresponding sequence of rescaled times $\tilde{s}_k \nearrow \infty$, satisfying that the rescaled surfaces $\tilde{M}_{\hat{p}_k,\tilde{s}_k}$ converge smoothly, on compact subsets of $\mathbf{R}^{n+1}$, to a limit surface $\tilde{M}_\infty$. This limit, $\tilde{M}_\infty$, would this time have no boundary, and would also satisfy $\mathbf{H} = -\mathbf{x}^\perp$ (since an exact analogue of (8.49) would also hold here, in view of (8.51) and (8.52)). Moreover, as in Step 5, it would also satisfy $|\tilde{A}|^2(z) \geq \delta/4$ for some $z \in B_{\sqrt{2nC_0}}(0)$ (cf. Lemma 3.2(b)).

But then, by Theorem 4.1, $\tilde{M}_\infty$ would have to be of the form $S^m \times \mathbf{R}^{n-m}$ for some $m \in \{1,\ldots,n\}$. Hence the $\tilde{M}_{\hat{p}_k,\tilde{s}_k}$ would be converging smoothly on compact subsets of $\mathbf{R}^{n+1}$ to some such cylinder, and therefore we'd have (for some $m \geq 1$)

(8.53)

$$\lim_{k\to\infty} \Theta(p_k, t(\tilde{s}_k)) \equiv \frac{1}{(2\pi)^{n/2}} \left\{ \lim_{k\to\infty} \int_{\tilde{M}_{\hat{p}_k,\tilde{s}_k}} \tilde{\rho}\, dH^n \right\} = \Theta^{(m)} \geq 1 + \epsilon_n\,,$$

where $\Theta^{(m)}$ is as in Appendix A (cf. (5.4)). Yet this would contradict (8.51), and this proves that "Case (a)" (as described in Section 7) cannot occur. Q.E.D. for Step 6.

Step 7 : Here, to show that "Case (b)" is also impossible, we argue just as outlined in the discussion of this step in Section 7. That is, for each k, let $\bar{p}_k$ be a point in Γ that achieves $\operatorname{dist}(\hat{p}_k,\Gamma)$. Then, if hypothesis (7.6) held, and passing to a subsequence if necessary, we'd have (noting $\hat{\bar{p}}_k \equiv \bar{p}_k$) that $\tilde{\mathbf{F}}_{\bar{p}_k}(p_k, s_k) \in B_{C_*+\sqrt{2nC_0}}(0)$ for each k. Also, for each k, the rescaled surfaces $\tilde{M}_{\bar{p}_k,\tilde{s}_k}$ would satisfy $\left|\tilde{A}\right|^2(\tilde{\mathbf{F}}_{\bar{p}_k}(p_k,s_k)) \geq \delta/2$.

But then this would lead to a contradiction, essentially exactly as in Step 5 above, save with the points $\{\bar{p}_k\}$ now playing the role played by the $\{p_k\}$ in that discussion. The only difference would be that, in the analogues of (8.45) and (8.47), we'd replace the origin by the point $\tilde{\mathbf{F}}_{\bar{p}_k}(p_k, s_k)$, for each k. Q.E.D. for Step 7.

Step 8 : Finally, to prove Lemma 7.1, we must establish that "Case (c)" is also impossible. The idea is again to study rescalings about the varying points $\{p_k\}$, as in Step 6, but here we must be more subtle, since the p_k now *do* approach Γ, whence we can no longer use Lemmas 5.7(c) or 2.6(b). In the argument that follows the first of these is replaced by Lemma 6.3(a), while the latter is replaced by the estimate (7.2), proven in Step 1.

To be more precise now, first set $d_k \equiv \operatorname{dist}(\hat{p}_k, \Gamma)/\sqrt{2(T-t_k)}$ for all k. Then we want to prove that we cannot simultaneously have that $p_k \in M_0 \backslash \mathcal{W}$ for all k, $p_k \to p \in \Gamma^+$ (so also $\hat{p}_k \to \hat{p} \in \Gamma$), and

$$\lim_{k\to\infty} d_k = +\infty\,. \tag{8.54}$$

So suppose instead that we did have such a situation. To derive a contradiction consider, as usual, the rescaled surfaces $\tilde{M}_{\hat{p}_k,s}$. Then, just as in Step 5, there would be some universal $\delta_1 = \delta_1(n, M_0, C_0, \delta_0, \Gamma, \delta)$ such that, for all k,

$$\max_{\tilde{M}_{\hat{p}_k,s}\cap B_{\sqrt{2nC_0}}(0)} |\tilde{A}|^2 \geq \delta/4 \quad \text{for all} \quad s \in [s_k, s_k+\delta_1]\,. \tag{8.55}$$

(Here the s_k, as always, are the rescaled times corresponding to the t_k, and we have used Lemma 3.2(b), as well as the estimates of Lemma 2.5).

But also, by (2.8) and (2.13), we'd have that, for each k,

$$\begin{aligned}
&\frac{1}{(2\pi)^{n/2}}\int_{s_k}^{s_k+\delta_1}\int_{\tilde{M}_{\hat{p}_k,s}} |\mathbf{H}+x^{\perp}|^2\tilde{\rho}H^n ds \\
&\leq \int_{s_k}^{\infty}\int_{\tilde{M}_{\hat{p}_k,s}} |\mathbf{H}+x^{\perp}|^2\tilde{\rho}H^n ds \\
&= \left\{\Theta(p_k,t_k)-\hat{\Theta}(p_k)\right\} + \int_{t_k}^{T}\int_{\Gamma}\left\langle\frac{\mathbf{x}-\hat{p}_k}{2\tau},\eta\right\rangle\rho_{\hat{p}_k} dH^{n-1}dt\,.
\end{aligned} \tag{8.56}$$

Yet, since the p_k approach Γ^+ as $k\to\infty$, by assumption, so for any given $\epsilon > 0$ we'd have also that, for all k sufficiently large (in terms of ϵ),

$$t_k > t_\epsilon \qquad \text{and} \qquad p_k \in \mathcal{U}_\epsilon\,, \tag{8.57}$$

where $t_\epsilon, \mathcal{U}_\epsilon$ are as in Lemma 6.3(a). Thence, by Lemma 6.3(a), and noting that $\hat{\Theta}(p_k) \equiv 1$ for all k (since by assumption $p_k \in M_0\backslash\mathcal{W}$ for all

k, and noting (7.5)), we'd have that, for any $\epsilon > 0$, there is a $k(\epsilon)$ such that, for all $k \geq k(\epsilon)$,

$$\Theta(p_k, t) - \hat{\Theta}(p_k) \leq \epsilon \quad \text{for all} \quad t \geq t_k \,. \tag{8.58}$$

Therefore, by (8.58) in (8.56), we'd have that, for all $k \geq k(\epsilon)$,

$$\frac{1}{(2\pi)^{n/2}} \int_{s_k}^{s_k+\delta_1} \int_{\tilde{M}_{\hat{p}_k,s}} \left|\mathbf{H} + \mathbf{x}^{\perp}\right|^2 \tilde{\rho} dH^n ds \leq \epsilon + \int_{t_k}^{T} \int_{\Gamma} \left\langle \frac{\mathbf{x} - \hat{p}_k}{2\tau}, \eta \right\rangle \rho_{\hat{p}_k} dH^{n-1} dt. \tag{8.59}$$

But also, by estimate (7.2), we'd have that, for all k sufficiently large that $t_k > t_\beta(\epsilon)$ ($\epsilon \in (0, 1/2)$ now),

$$\int_{t_k}^{T} \int_{\Gamma} \left| \left\langle \frac{\mathbf{x} - \hat{p}_k}{2\tau}, \eta \right\rangle \right| \rho_{\hat{p}_k} dH^{n-1} dt \leq \epsilon + C_8(n) \int_{d_k/2}^{\infty} e^{-w^2/2} dw \,.$$

Hence, on noting (8.54), we'd have, by increasing $k(\epsilon)$ suitably if necessary, that for any $\epsilon \in (0, 1/2)$ there is a $k(\epsilon)$ such that, for all $k \geq k(\epsilon)$,

$$\frac{1}{(2\pi)^{n/2}} \int_{s_k}^{s_k+\delta_1} \int_{\tilde{M}_{\hat{p}_k,s}} \left|\mathbf{H} + \mathbf{x}^{\perp}\right|^2 \tilde{\rho} dH^n ds \leq 3\epsilon \,. \tag{8.60}$$

But then, just as in Step 5, we'd be able to find a subsequence of points, still denoted $\{p_k\}$, and corresponding times $\{\tilde{s}_k \in [s_k, s_k + \delta_1]\}_{k=1}^{\infty}$, such that the rescaled surfaces $\tilde{M}_{\hat{p}_k, \tilde{s}_k}$ converged smoothly on compact subsets of $\mathbf{R}^{n+1}$ to a limit surface, $\tilde{M}_\infty$, which in this case would have no boundary (in view of (8.54)), and which once again, by (8.60), would satisfy $\mathbf{H} = -\mathbf{x}^{\perp}$. Moreover, by (8.55), it would also have $\max_{\tilde{M}_\infty \cap B_{\sqrt{2nC_0}}(0)} |\tilde{A}|^2 \geq \delta/4$.

Yet then, by Theorem 4.1, it would be of the form $S^m \times \mathbf{R}^{n-m}$ for some $1 \leq m \leq n$, and therefore, exactly as in Step 6 (cf. (8.53)), we'd have that

$$\lim_{k \to \infty} \Theta(p_k, t(\tilde{s}_k)) \geq 1 + \epsilon_n \,. \tag{8.61}$$

However this would contradict (8.58) (taking, say, $\epsilon = \frac{1}{2}\epsilon_n$ in (8.58), and observing that $t(\tilde{s}_k) \geq t(s_k) \equiv t_k$ for all k). This contradiction then proves that "Case (c)" also is impossible, and so completes the proof of Step 8, and of Lemma 7.1. Q.E.D. for Step 8.

Step 9 : Part (b) of Theorem 6.1 now follows readily. Simply take $\mathcal{U} \equiv \mathcal{U}_{\epsilon_n/2}$ ($\mathcal{U}_{\epsilon_n/2}$ as in Lemma 6.3(a)). Then, by Step 4 and Equation (5.3), $\hat{\Theta}(p) \leq 1$ on $\bar{\mathcal{U}}$, so $\mathcal{W} \equiv M_0 \backslash \bar{\mathcal{U}}$ is an open set in $M_0 \backslash \Gamma^+$ containing $\mathcal{B} \equiv \{p : \hat{\Theta}(p) > 1\}$. Hence, for any $\delta > 0$, there must be a time, $\hat{t}_\delta < T$, such that any δ-essential blow-up sequence must leave $\bar{\mathcal{U}}$ forever by time $\hat{t}_\delta < T$; otherwise we could find such a blow-up sequence which never ended up in $\mathcal{W}$, and this would violate Lemma 7.1. Q.E.D. for Step 9.

Step 10 : Finally, to prove part (a) of Theorem 6.1, let $\{p_k, t_k\}$ once again be a δ-essential blow-up sequence, with $p_k \to p$ as $k \to \infty$; so that p is a general Type I singular point of the flow. Observe that, by Step 9, p must then lie in the set $M_0 \backslash \bar{\mathcal{U}}$ ($\mathcal{U}$ as in part (b) of Theorem 6.1), on which $\hat{\Theta}(q) \geq 1$; so there must then be a neighbourhood, $\mathcal{V}_0$, of p, in $M_0 \backslash \bar{\mathcal{U}}$, such that

$$\hat{\Theta}(q) \geq 1 \quad \text{for all} \quad q \in \mathcal{V}_0 . \tag{8.62}$$

Now suppose that p were *not* a special Type I singular point. Then we'd have that

$$\limsup_{t \to T} (T - t)|A|^2(p, t) = 0 . \tag{8.63}$$

Hence, by Theorem 4.1, any limit blow-up about $\hat{p}$ would have to be an n-plane, and so we'd have

$$\hat{\Theta}(p) = 1 . \tag{8.64}$$

But then, by the "upper semicontinuity" result, Lemma 5.7(b) (applied with, say, $K = M_0 \backslash \mathcal{U}$), we'd have, in view of (8.62), that there would be a neighbourhood, $\mathcal{V}$, of p, in M_0, such that

$$\hat{\Theta}(q) \equiv 1 \quad \text{for all} \quad q \in \bar{\mathcal{V}} . \tag{8.65}$$

However then, by Lemma 7.1 (with $\mathcal{W} \equiv M_0 \backslash \bar{\mathcal{V}}$), we'd have that $p_k \notin \mathcal{V}$ for all k sufficiently large, and this is a contradiction, since by assumption the p_k converge to $p \in \mathcal{V}$. This contradiction proves that any such p must in fact also be a special Type I singular point of the flow, and we are done. Q.E.D. for Step 10.

This completes the proof of Theorem 6.1. Q.E.D.

9. An improved boundary strip estimate for $\mathbf{H}^2$

We would now like to extend the partial boundary regularity result of Theorem 6.1(b) to a full boundedness result, for $|A|^2$ and all its derivatives, in a boundary strip. Just such a form of strengthening was proven, for points not in the extended boundary, in [10, Section 4]. This was done via a localisation procedure, involving a careful choice of weighting function, together with the maximum principle. However the estimates on $|A|^2(p,t)$, thus obtained, deteriorate as $\operatorname{dist}(\hat{p},\Gamma) \to 0$. Moreover, given the use of the maximum principle, the method does not carry over directly to handle boundary points. Hence greater ingenuity is required here.

As regards this, recall that the approach used in [10, §4] involved essentially a two-stage procedure. The first was to show, subject to (say) a Type I hypothesis, that a local interior estimate of the form $|A|^2(p,t) \le \epsilon/(T-t)$, $\epsilon \in (0,1/2)$, could be improved to a local bound of the form $|A|^2(p,t) \le C/(T-t)^\beta$, $\beta = \beta(\epsilon) < 1$. The second was then to show that such an improved bound could, in turn, be bootstrapped up to a uniform local estimate, $|A|^2(p,t) \le C$.

Here, however, because we cannot avoid having to deal with boundary points, we must instead adopt a more roundabout, four-stage approach. The idea is to exploit that H is identically zero on Γ. This fact means that, in applying an analogue of the above method of estimation to H^2, rather than $|A|^2$, we do *not* encounter any additional problems in dealing with "boundary points", *vis-a-vis* those not in Γ^+. We can then, in turn, use improvements in our estimates for H^2 (in our boundary strip), together with Allard's boundary regularity theory for varifolds, to establish corresponding improvements in our bounds for $|A|^2$, and vice-versa. In this way the idea is alternately to upgrade our estimates for H^2 and $|A|^2$, one after the other, until eventually both have been uniformly bounded in a whole boundary strip.

To describe the above plan more precisely, recall first from [10, §4] the following (trivially modified) notion of "ϵ-boundedness" which was central to the discussion presented there;

9.1 Definition. For any given $\epsilon \in (0,1/2)$ we say that a point, $p \in M_0$, satisfies the *"ϵ-boundedness hypothesis"*, for time $t_0 < T$, if for some fixed $\alpha \in (0,1/4)$ we have that, for all $t \in [t_0,T)$,

$$\max_{M_t \cap B_{\tau^{\alpha/2}+\sqrt{2nC_0\tau}}(\hat{p})} |A|^2 \le \frac{\epsilon}{(T-t)} .$$

Note that the following is then an immediate consequence of Theorem 6.1(b);

9.2 Corollary. *There is a neighbourhood, $\mathcal{U}_0 \subset M_0$, of the extended boundary Γ^+, such that for any $\epsilon \in (0,1/2)$, and any $\alpha \in (0,1/4)$, the above "ϵ-boundedness hypothesis" holds for all $p \in \mathcal{U}_0$, with a uniform time $t_0(\alpha,\epsilon) < T$.*

With this established, our four step strategy for deriving full boundary regularity can now be sketched in more detail. (In this outline, "estimate" always means "estimate in a suitable boundary strip").

Step 1: Obtain, via the method of [10, §4], an improved estimate for H^2 of the form $H^2 \leq C_\beta/(T-t)^\beta$, $\beta \in (0,1)$ arbitrary.

Step 2: Noting the improved estimate for H^2 of Step 1, use Allard's boundary regularity theory for varifolds ([1]) to establish a corresponding improved estimate for $|A|^2$ of the form $|A|^2 \leq C_\beta/(T-t)^\beta$, $\beta \in (0,1)$ arbitrary. This is done via studying, for arbitrary boundary points $\mathbf{x}_0$, the rescaled surfaces $\hat{M}_t = \tau^{-\beta}(M_t - \mathbf{x}_0)$.

Step 3: Deduce a uniform bound for H^2 by bootstrapping up the argument of Step 1, making use of the improved estimate for $|A|^2$ of the previous step.

Step 4: Finally, bootstrap up also the improved bound of Step 2, for $|A|^2$, to a uniform estimate in a boundary strip (again via use of Allard's results).

Steps 2, 3 and 4 are carried out in Sections 10 to 12 respectively.

As regards implementing Step 1 of this plan, we begin with some notation, and a preliminary computational lemma. Let $\alpha \in (0,1/4)$ be (for the present) arbitrary, and let $\lambda \in C^\infty(\mathbf{R}^{n+1},\mathbf{R})$ denote the function $\lambda(x) \equiv \exp(-|\mathbf{x}|^2/2)$, so that

$$\lambda(\mathbf{x}) \in (0,1) \quad \text{for all} \quad x \quad , \qquad \lambda \geq 1/2 \quad \text{for all} \quad |x| \leq 1/2\,, \tag{9.1}$$

$$\lambda(x) = \lambda(R) \quad \text{where} \quad R \equiv |x| \quad , \qquad \frac{d\lambda}{dR} \leq 0 \quad \text{for all} \quad R\,, \tag{9.2}$$

$$|D\lambda(x)| = 2|\mathbf{x}|\lambda(x) \quad \text{for all} \quad x\,, \tag{9.3}$$

$$|D_{ij}\lambda(x)| \leq C_{17}(n)(1+|\mathbf{x}|^2)\lambda(x) \tag{9.4}$$

for all x, and for all $1 \leq i,j \leq n+1$, and for all x with $|x| \geq \tau^{\alpha/2}$, and for all $t \in [t_1,T)$, for some fixed $t_1(\alpha) < T$,

$$\frac{1}{\tau}\lambda\Big(\frac{x}{\tau^\alpha}\Big) \leq 1\,. \tag{9.5}$$

Next, for any fixed $p \in M_0$, set $\psi_p, \mu_p : M_0 \times [0,T) \to \mathbf{R}$ by

$$\psi_p(\tilde{p},t) = \lambda\Big(\frac{\mathbf{F}(\tilde{p},t)-\mathbf{F}(p,t)}{\tau^\alpha}\Big)\,, \quad \mu_p(\tilde{p},t) = \psi_p(\tilde{p},t)H^2(\tilde{p},t)\,. \tag{9.6}$$

9.3 Lemma. *The function μ_p satisfies the differential inequality that*

$$\begin{aligned}\tfrac{d}{dt}\mu_p \leq\ & \triangle(\mu_p) - 2\tfrac{\nabla\psi_p}{\psi_p}\cdot\nabla\mu_p \\ & + \left(2|A|^2 + C_{18}(n)(1+\tau^{-2\alpha})\max_{M_t}|\mathbf{H}|\right)\mu_p\,. \end{aligned}\tag{9.7}$$

Proof. We only sketch the derivation of (9.7) (cf. [10, Lemma 4.1] for an analogous computation, set out in greater detail). Note that the evolution equation for H^2 is $d_t(H^2) = \triangle(H^2) - 2|\nabla H|^2 + 2H^2|A|^2$. Thus, by property (9.2), we have directly that

$$\begin{aligned}\frac{d}{dt}\mu_p(\tilde{p},t) &= \psi_p(\triangle H^2 - 2|\nabla H|^2 + 2H^2|A|^2) \\ &\quad + H^2\left(\nabla\lambda\left(\frac{\mathbf{F}(\tilde{p},t)-\mathbf{F}(p,t)}{\tau^\alpha}\right)\right)\cdot \\ &\quad \left(\frac{\alpha(\mathbf{F}(\tilde{p},t)-\mathbf{F}(p,t))}{\tau^{1+\alpha}} + \frac{1}{\tau^\alpha}\frac{d\mathbf{F}}{dt}(\tilde{p},t) - \frac{1}{\tau^\alpha}\frac{d\mathbf{F}}{dt}(p,t)\right) \\ &\leq \psi_p(\triangle H^2 - 2|\nabla H|^2 + 2H^2|A|^2) \\ &\quad + \frac{H^2}{\tau^\alpha}\left(\nabla\lambda\left(\frac{\mathbf{F}(\tilde{p},t)-\mathbf{F}(p,t)}{\tau^\alpha}\right)\right)\cdot\Big(\mathbf{H}(\tilde{p},t)-\mathbf{H}(p,t)\Big) \\ &\leq \triangle(\psi_p H^2) - 2\nabla\psi_p\cdot\nabla H^2 - H^2\triangle\psi_p \\ &\quad + 2|A|^2\mu_p + \frac{2H^2}{\tau^\alpha}|\nabla\lambda|\left(\max_{M_t}|\mathbf{H}|\right).\end{aligned}\tag{9.8}$$

But now, at any point $\tilde{p}$ on M_t we have (cf. [10, (4.15)]) that

$$\triangle\psi_p = \triangle^{R^{n+1}}\psi_p - \nu(\nu(\psi_p)) - H\nu(\psi_p)\,,$$

whence, by properties (9.3) and (9.4), we may estimate that

$$|\triangle\psi_p(\tilde{p},t)| \leq C_{19}(n)(1+|\mathbf{H}(\tilde{p},t)|)\psi_p(\tilde{p},t)\,.\tag{9.9}$$

Therefore, by (9.9) in (9.8), together with property (9.3) again, we obtain that

$$\begin{aligned}\tfrac{d}{dt}\mu_p \leq\ & \triangle(\mu_p) - 2\tfrac{\nabla\psi_p}{\psi_p}\cdot\nabla\mu_p + \tfrac{2H^2}{\psi_p}|\nabla\psi_p|^2 \\ & + (2|A|^2 + C_{20}(n)(1+\tau^{-2\alpha})\max_{M_t}|\mathbf{H}|)\,\mu_p\,.\end{aligned}$$

Inequality (9.7) now follows by applying property (9.3) one last time to the third term on the right hand side of this estimate. Q.E.D.

We are now ready to carry out Step 1 of our plan, sketched above. This is accomplished by the following lemma.

9.5 Lemma. *Let $\mathcal{U}_0$ be the neighbourhood of Γ^+ in Corollary 9.2. Then for any $\beta \in (0,1)$ there is a constant $C_{21} = C_{21}(n, \beta, C_0, T)$ such that, on $\mathcal{U}_0$, we have the improved estimate for H^2 that, for all $t \in [0,T)$,*

$$H^2(p,t) \leq \frac{C_{21}}{(T-t)^\beta}.$$

Proof. Let $\alpha \in (0, 1/4), \beta \in (0,1)$ be arbitrary (α to be suitably fixed later), and set $\epsilon = \beta/3$. Then let t_0 be as in Corollary 9.2 (for our choices of α, ϵ), so that the ϵ-boundedness hypothesis holds for all $p \in \mathcal{U}_0$ with this uniform time $t_0(\alpha, \beta)$.

Now fix $p \in \mathcal{U}_0$ arbitrarily, and then set

$$\mu_{max}(t) = \max_{\tilde{p} \in M_0}(\mu_p(\tilde{p}, t)) \quad , \quad \mathcal{M}_t = \{\tilde{p} \in M_0 : \mu_p(\tilde{p}, t) = \mu_{max}(t)\}.$$

Then observe that, by property (9.5) and the Type I hypothesis, (1.3), we will have, for all $t \in [t_1, T)$, and all $\tilde{p} \in M_0$ with $|\mathbf{F}(\tilde{p},t) - \mathbf{F}(p,t)| \geq \tau^{\alpha/2}$, that $\mu_p(\tilde{p}, t) \leq nC_0$. Therefore if, at any such time t, we have $\mu_{max}(t) > nC_0$, then it must hold that $|\mathbf{F}(\tilde{p},t) - \mathbf{F}(p,t)| \leq \tau^{\alpha/2}$ for all $\tilde{p} \in \mathcal{M}_t$. Hence we must have (noting also (3.2)) that, for all $t \in [t_1, T)$ and for all such $\tilde{p}$,

$$\mu_{max}(t) > nC_0 \Longrightarrow \mathbf{F}(\tilde{p}, t) \in B_{\tau^{\alpha/2} + \sqrt{2nC_0\tau}}(\hat{p}).$$

But then, from Corollary 9.2, we may conclude that, for all $t \geq \max\{t_0, t_1\}$,

$$\mu_{max}(t) > nC_0 \Longrightarrow |A|^2(\tilde{p}, t) \leq \frac{\epsilon}{\tau} \tag{9.10}$$

for all $\tilde{p} \in \mathcal{M}_t$. Moreover, we must obviously have (since $H|_\Gamma \equiv 0$) that, for all t,

$$\mu_{max}(t) > 0 \Longrightarrow \mathcal{M}_t \cap \Gamma = \phi; \tag{9.11}$$

while also, by the Type I hypothesis, (1.3), we have the estimate that

$$\max_{M_t} |\mathbf{H}| \leq \sqrt{nC_0/2\tau}. \tag{9.12}$$

So now put $t_2 \equiv \max\{t_0, t_1, T-1\}$, and suppose $t \in [t_2, T)$ is such that $\mu_{max}(t) > nC_0$. Then from (9.10), (9.11) and (9.12) in (9.7) we see that, at any such instant, we'll have

$$\frac{d}{dt}\mu_{max}(t) \leq \frac{2\epsilon + C_{22}(n)\tau^{1/2-2\alpha}}{\tau}\mu_{max}(t). \tag{9.13}$$

So also now fix $\alpha = 1/8$, and set $t_3(n)$ to be the least time such that $C_{22}\tau^{1/4} \leq \epsilon$ for all $t \in [t_3, T)$. Then, putting $t_4 \equiv \max\{t_2, t_3\} = t_4(n, \beta, T)$, we see that we'll have, for $t \in [t_4, T)$, that

$$\frac{d}{dt}\mu_{max}(t) \leq \frac{3\epsilon}{\tau}\mu_{max}(t) \equiv \frac{\beta}{\tau}\mu_{max}(t) \tag{9.14}$$

whenever $\mu_{max}(t) > nC_0$.

So now define

$$\mathcal{G} = \{t \in [t_4, T) : \mu_{max}(t) > nC_0\} \quad , \quad \zeta(t) = \max(\mu_{max}(t), nC_0)\,.$$

Then, by (9.14), we'll have (at least a.e.) that

$$\begin{cases} \dfrac{d}{dt}\zeta(t) \leq \dfrac{\beta}{\tau}\zeta(t)\,, & t \in \mathcal{G} \\ \dfrac{d}{dt}\zeta(t) = 0\,, & t \in [t_4, T)\backslash\mathcal{G}\,. \end{cases}$$

Hence, for any $t \in [t_4, T)$, we'll have by integration from t_4 to t that $\zeta(t) \leq \frac{C_{23}}{\tau^\beta}$, where, by the Type I hypothesis, $C_{23} = C_{23}(n, \beta, C_0, T)$ satisfies that

$$\begin{aligned} C_{23} = \zeta(t_4)(T - t_4)^\beta \quad &\leq \Big(\max\big(\max_{M_{t_4}} |A|^2, nC_0\big)\Big)(T - t_4)^\beta \\ &\leq nC_0(T - t_4)^{\beta - 1}\,. \end{aligned}$$

To finish the argument it remains only to notice that then, since $\lambda(0) = 1$, so we obtain that, for all $t \in [t_4(n, \beta, T), T)$,

$$H^2(p,t) = \mu_p(p,t) \leq \mu_{max}(t) \leq \zeta(t) \leq \frac{C_{23}}{\tau^\beta}\,. \tag{9.15}$$

The result then follows, from (9.15) and the Type I hypothesis, since $p \in \mathcal{U}_0$ was arbitrary here. Q.E.D.

10. An improved boundary strip estimate for $|A|^2$

Turning now to Step 2 of our strategy, we accomplish this via a series of lemmas. The key element is first to obtain the improved estimate for $|A|^2$ on the boundary, Γ. It is then not difficult to extend it to a whole boundary strip, via a method very similar to that employed in Step 1 above.

We begin by establishing an "area ratio bound" that holds uniformly for all the surfaces M_t, $t \in [T/2, T)$. This will be needed shortly to allow us to extract varifold limits from sequences of suitable rescalings of the surfaces, $\{M_t\}$. The notation in this lemma (cf. Lemma 2.1) is that, for any $\mathbf{x}_0 \in \mathbf{R}^{n+1}$, and any $T_0 \in [0, \infty)$, we set

$$\rho_{\mathbf{x}_0,T_0}(\mathbf{x},t) \equiv \frac{1}{\left(4\pi(T_0-t)\right)^{n/2}} \exp\left(\frac{-|\mathbf{x}-\mathbf{x}_0|^2}{4(T_0-t)}\right),$$

$$B_{\mathbf{x}_0,T_0}(t) \equiv \int_0^t \int_\Gamma \rho_{\mathbf{x}_0,T_0}(\mathbf{x},t) \left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2(T_0-t)}, \eta \right\rangle dH^{n-1} dt\,.$$

10.1 Lemma. (a). *There is a uniform constant, $C_{24} = C_{24}(n, \Gamma, T)$, independent of the choice of $\mathbf{x}_0$ and T_0, such that, for all $t \in [0, \min\{T_0, T\})$, $|B_{\mathbf{x}_0,T_0}(t)| \le C_{24}$.*

(b). *There is a uniform constant, $C_{25} = C_{25}(n, \Gamma, |M_0|, T)$, such that, for all $\mathbf{x}_0 \in \mathbf{R}^{n+1}$, all $t_0 \in [T/2, T)$, and all $r > 0$, we have $r^{-n} |(M_{t_0} \cap B_r(\mathbf{x}_0))| \le C_{25}$.*

Proof. **Of (a).** Consider the uniform estimate, (7.1). This was proven (see Section 8) in the case where T is as in the Type I hypothesis, (1.3), but it is clear from the proof that it holds equally well in the event that "T" is replaced by any $T_0 \in (0, \infty)$. (After all, (7.1) is a statement purely about the boundary, and could be considered even in the absence of any related "flow of surfaces". In particular the Type I hypothesis is of course never needed). Moreover, although the proof is given in terms of points of the form $\hat{p}$, $p \in M_0 \backslash \Gamma^+$, it is also evident that the argument holds, unaltered, for the case of any point $\mathbf{x}_0 \notin \Gamma$ in place of $\hat{p}$.

Thus we get that, for any $\epsilon > 0$, and any $T_0 \in (0, \infty)$, there exists a time $t^*(n, \Gamma, \epsilon, T_0) < T_0$ such that, for any point $\mathbf{x}_0 \notin \Gamma$, and for all $t_2 \in [t^*, \min\{T_0, T\})$,

$$\int_{t^*}^{t_2} \int_\Gamma \rho_{\mathbf{x}_0,T_0}(\mathbf{x},t) \left| \left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2(T_0-t)}, \eta \right\rangle \right| dH^{n-1} dt \le \frac{1}{2} + \epsilon\,.$$

Furthermore, it is clear that, for any $\epsilon > 0$, the time $t^*(n, \Gamma, \epsilon, T_0)$ may be taken to be of the form $T_0 - \sigma^*(n, \Gamma, \epsilon)$, for some $\sigma^* > 0$ independent of T_0. Simply observe that $\rho_{\mathbf{x}_0,T_0}(\mathbf{x},t) \equiv \rho_{\mathbf{x}_0,T_1}(\mathbf{x}, t + (T_1 - T_0))$ for any $T_0, T_1 \in (0, \infty)$.

Hence, taking say $\epsilon = 1/2$, we get that, for any $T_0 \in (0, \infty)$, there exists a $\sigma^{**}(n, \Gamma) > 0$ such that, for any $\mathbf{x}_0 \notin \Gamma$, and for all $t_2 \in [T_0 - \sigma^{**}, \min\{T_0, T\})$,

$$\int_{T_0-\sigma^{**}}^{t_2} \int_\Gamma \rho_{\mathbf{x}_0,T_0}(\mathbf{x},t) \left| \left\langle \frac{\mathbf{x}-\mathbf{x}_0}{2(T_0-t)}, \eta \right\rangle \right| dH^{n-1} dt \le 1\,. \tag{10.1}$$

But also then it is evident that, for any $T_0 \in (0, \infty)$, and any $\mathbf{x}_0 \notin \Gamma$, we can find a constant $C_{26} = C_{26}(n, \Gamma, T)$ such that

$$(10.2) \qquad \int_0^{\min\{T, T_0 - \sigma^{**}\}} \int_\Gamma \rho_{\mathbf{x}_0, T_0}(\mathbf{x}, t) \left| \left\langle \frac{\mathbf{x} - \mathbf{x}_0}{2(T_0 - t)}, \eta \right\rangle \right| dH^{n-1} dt \le C_{26} \,.$$

Combining now (10.1) and (10.2) then yields the desired estimate, for the case of points $\mathbf{x}_0 \notin \Gamma$. The case of boundary points, $\mathbf{x}_0 \in \Gamma$, follows similarly, this time by inspection of Lemma 2.7 and its proof (rather than (7.1)).

Of (b). The proof of this part was shown to us (in the boundaryless setting) by Brian White. Fix $\mathbf{x}_0 \in \mathbf{R}^{n+1}, t_0 \in [T/2, T)$ and $r > 0$ arbitrarily. Then recall that the monotonicity identity, (2.2), holds for any $T_0 \in (0, \infty)$, not just the first singular time T. So now apply (2.2) with T replaced by $T_0 \equiv t_0 + r^2$. (Note that this time T_0 may conceivably be greater than the critical time, T, but this does not matter as we shall only employ (2.2) for times $t_1, t_2 < t_0 < T$). This yields (with the choices $t_1 = 0$, $t_2 = t_0$) that

$$(10.3) \qquad \int_{M_{t_0}} \rho_{\mathbf{x}_0, T_0} dH^n \le \int_{M_0} \rho_{\mathbf{x}_0, T_0} dH^n + |B_{\mathbf{x}_0, T_0}(t_0)| \,.$$

But now, writing this out explicitly, and noting the result of part (a), this says that

$$\frac{1}{(4\pi r^2)^{n/2}} \int_{M_{t_0}} \exp\left(\frac{-|\mathbf{x} - \mathbf{x}_0|^2}{4r^2} \right) dH^n \le \frac{1}{(4\pi(t_0 + r^2))^{n/2}} |M_0| + C_{24} \,.$$

Yet this, noting $t_0 > T/2$, implies precisely the desired estimate, that

$$(10.4) \qquad r^{-n} |M_{t_0} \cap B_r(\mathbf{x}_0)| \le e^{1/4} (2^{n/2} T^{-n/2} |M_0| + C_{24}) \,.$$

We next need a series of abstract lemmas about varifolds. For their statement, and the subsequent analysis in this section, recall the notation of 1.1.

10.2 Lemma (Allard). *Let S^1 denote the unit circle inside $\{0\} \times \mathbf{R}^2 \subset \mathbf{R}^{n+1}$, let P_W denote orthogonal projection onto any subspace $W \subset \mathbf{R}^{n+1}$, and let $\zeta : \mathbf{R}^{n+1} \backslash (\mathbf{R}^{n-1} \times \{0\}) \to \{0\} \times \mathbf{R}^2$ denote $P_{\{0\} \times R^2}$ followed by anticlockwise rotation in $\{0\} \times \mathbf{R}^2$ by $\pi/2$. Suppose C is an n-dimensional varifold cone in $\mathbf{R}^{n+1}$, stationary away from $\mathbf{R}^{n-1} \times \{0\}$, and satisfying $\Theta^n(\|C\|, \mathbf{x}) \ge 1$ for $\|C\|$-almost every $\mathbf{x} \in \mathbf{R}^{n+1} \backslash (\mathbf{R}^{n-1} \times$*

$\{0\})$. *Finally let $\mathcal{T}_C$ denote the linear functional on $C^\infty(S^1)$ given by*

$$\mathcal{T}_C(\varphi) \equiv \int_{(B_1(0)\setminus(R^{n-1}\times\{0\}))\times G(n+1,n)} \varphi\left(\frac{P_{\{0\}\times R^2}(\mathbf{x})}{|P_{\{0\}\times R^2}(\mathbf{x})|}\right) \frac{|P_{T_x C}(\zeta(\mathbf{x}))|^2}{|P_{\{0\}\times R^2}(\mathbf{x})|^2}\, d\|C\|(\mathbf{x})\,.$$

Here $G(n+1,n)$ is the Grassmannian of n-planes in $\mathbf{R}^{n+1}$. Then

(a) *$\mathcal{T}_C$ is a multiple of $H^1 \llcorner S^1$.*

(b) *If $\mathcal{T}_C = 0$ then $P_{\{0\}\times R^2}(spt\|C\|) \cap S^1$ is a finite set of points.*

(c) *If Σ is a rectifiable, integer multiplicity n-varifold in $B_1(0)$, the unit ball in $\mathbf{R}^{n+1}$, and if Σ is stationary away from $(\mathbf{R}^{n-1}\times\{0\})\cap B_1(0)$, and if $spt(\|\Sigma\|) \subset \hat{V}^{(\delta_0)}$ for some fixed $\delta_0 > 0$, then any tangent cone, C, to Σ at 0 must be a sum of integer multiplicity half n-spaces, all lying in the wedge $\hat{V}^{(\delta_0)}$.*

Proof. **Of (a) and (b).** These are special cases of Lemma 5.1 of Allard, [1].

Of (c). This follows by the same reasoning as in the first part of the proof of Lemma 5.2 of Allard, [1]. Simply construct the functional $\mathcal{T}_{\hat{C}}$ corresponding to the varifold cone $\hat{C}$ obtained via the reflection principle (see [1, Lemma 3.2]) from the cone C. Then by part (a), together with the wedge condition, we must have that $\mathcal{T}_{\hat{C}} = 0$. The result then follows from part (b). Q.E.D.

10.3 Proposition. *Suppose $\{\Sigma_j\}$ is a family of smooth, orientable hypersurfaces in $B_1(0)$, with smooth boundaries $\{\partial\Sigma_j\}$, satisfying that*

(i) *$\Sigma_j \subset \hat{V}^{(\delta_0)}$ for all j , $0 \in \partial\Sigma_j$ for all j;*
(ii) *the $\partial\Sigma_j$ converge smoothly to $B_1(0)\cap(\mathbf{R}^{n-1}\times\{0\})$ as $j\to\infty$; and*
(iii) *there is a constant C_{27} such that, for all j, $\sup_{\Sigma_j}|\mathbf{H}_j| \le C_{27}$, where $\mathbf{H}_j(\mathbf{x})$ denotes the mean curvature vector of Σ_j at $\mathbf{x}$.*

Suppose that the Σ_j converge as varifolds to an n-varifold, Σ, in $B_1(0)$. Then Σ is integer multiplicity and rectifiable (and if in fact we have $\sup_{\Sigma_j}|\mathbf{H}_j| \to 0$ as $j\to\infty$ then Σ is also stationary away from $B_1(0)\cap(\mathbf{R}^{n-1}\times\{0\})$). Moreover, any tangent cone to Σ at 0 is a single, multiplicity one, half n-plane.

Proof. That Σ is integer multiplicity and n-rectifiable follows from [9, Theorem 42.7 and Remark 42.8], as does the stationarity assertion. To see the tangent cone claim, let C denote any such cone; say C is the varifold limit as $k\to\infty$ of $\lambda_k^{-1}\Sigma$ for some sequence of real numbers $\lambda_k \searrow 0$. Then, since Σ is itself the limit of the Σ_j, observe that, corresponding to this sequence $\{\lambda_k\}$, we can find a related sequence $\{j(k) \nearrow \infty\}$ of integers such that, in terms of varifold convergence, we have

$$\lambda_k^{-1}\Sigma_{j(k)} \to C\,. \tag{10.5}$$

In other words, C may be viewed as the limit of a sequence of smooth hypersurfaces $\tilde{\Sigma}_k \equiv \lambda_k^{-1}\Sigma_{j(k)}$, satisfying now that

(i*) $\tilde{\Sigma}_k \subset \hat{V}^{(\delta_0)}$ for all k , $0 \in \partial\tilde{\Sigma}_k$ for all k ;

(ii*) the $\partial\tilde{\Sigma}_k$ converge smoothly to $\mathbf{R}^{n-1} \times \{0\}$ as $k \to \infty$; and

(iii*) $\sup_{\tilde{\Sigma}_k} |\tilde{\mathbf{H}}_k| \to 0$ as $k \to \infty$, where $\tilde{\mathbf{H}}_k(\mathbf{x})$ denotes the mean curvature vector of $\tilde{\Sigma}_k$ at $\mathbf{x}$.

Yet also then we know that C must be stationary away from $\mathbf{R}^{n-1} \times \{0\}$, and so by Lemma 10.2(c) must have the form of a sum of half n-planes with boundary $\mathbf{R}^{n-1} \times \{0\}$, $\mathcal{H}_i$, all lying in the wedge $\hat{V}^{(\delta_0)}$, and appearing with some integer multiplicities; say

$$C = \sum_{i=1}^{l} n_i \mathcal{H}_i \quad , \quad n_i \in \mathbf{Z}\,. \tag{10.6}$$

We want to show that we must have $l = 1, n_1 = 1$.

The first step towards establishing this is to show, more weakly, that we must at least have

$$\sum_{i=1}^{l} n_i = \text{odd}\ . \tag{10.7}$$

To prove this note that, instead of considering varifold convergence of the $\tilde{\Sigma}_k$, we can also view the $\tilde{\Sigma}_k$ as rectifiable, integer multiplicity *currents*. Then, as currents, the $\tilde{\Sigma}_k$ must also converge weakly to some limit current $\hat{C}$, which is an integer multiplicity rectifiable cone with support in $\hat{V}^{(\delta_0)}$.

But then clearly $\hat{C}$ must simply be the same as C, except that the $\mathcal{H}_i$ are now endowed with orientations, and the multiplicities may now be different, since we may get cancellation of oppositely oriented half n-planes in the limit. Evidently though the half n-planes can only cancel in pairs, so we must have that

$$\hat{C} = \sum_{i=1}^{l} m_i [\![\mathcal{H}_i]\!] \quad , \quad m_i \in \mathbf{Z} \tag{10.8}$$

where $[\![\mathcal{H}_i]\!]$ denotes $\mathcal{H}_i$ with a choice of orientation, and where, for each $1 \le i \le l$,

$$n_i - m_i = \text{even}\,. \tag{10.9}$$

But also now, as currents, the boundaries of the $\tilde{\Sigma}_k$ must in addition be converging to $\partial\hat{C} = \sum_{i=1}^{l} m_i \partial[\![\mathcal{H}_i]\!]$. Furthermore, letting $[\![Q]\!]$ denote the current $[\![\mathbf{R}^{n-1} \times \{0\}]\!]$ with multiplicity one and orientation given by $\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{n-1}$ at each point, we clearly have first that, by (ii*),

$$\lim_{k\to\infty} \partial\tilde{\Sigma}_k = [\![Q]\!] , \tag{10.10}$$

and also, for each $1 \le i \le l$, $\partial[\![\mathcal{H}_i]\!] = \pm[\![Q]\!]$, so that

$$\partial\hat{C} = \left\{\sum_{i=1}^{l}(-1)^{\kappa_i} m_i\right\} [\![Q]\!] \tag{10.11}$$

for some numbers $\kappa_i \in \{0, 1\}$, $i = 1, \ldots, l$.

Thus we must have that $\sum_{i=1}^{l}(-1)^{\kappa_i} m_i = 1$, and hence $\sum_{i=1}^{l} m_i =$ odd; and then this, together with (10.9), immediately implies (10.7).

Now, however, we are in good shape, since we can then argue as follows. For any vector field $\mathbf{X} \in C_c^\infty(B_1(0); \mathbf{R}^{n+1})$, we have, by the usual First Variation formula for smooth surfaces (see [9, p. 46]), that, for each k,

$$\int_{\tilde{\Sigma}_k} \mathrm{div}_{\tilde{\Sigma}_k} \mathbf{X}\, dH^n = -\int_{\partial\tilde{\Sigma}_k} \langle \mathbf{X}, \tilde{\eta}_k\rangle dH^{n-1} - \int_{\tilde{\Sigma}_k} \langle \mathbf{X}, \tilde{\mathbf{H}}_k\rangle dH^n . \tag{10.12}$$

Here $\tilde{\eta}_k$ denotes the inward unit conormal to $\tilde{\Sigma}_k$ along $\partial\tilde{\Sigma}_k$.

Then the second term on the right hand side of this equality tends to zero as $k \to \infty$, by (iii*). Also, by virtue of the smooth convergence of the $\partial\tilde{\Sigma}_k$ to $\mathbf{R}^{n-1} \times \{0\}$, of (ii*), we have that there is some vector field, $\tilde{\eta}(\mathbf{x})$, along $\mathbf{R}^{n-1} \times \{0\}$, such that

$$\|\tilde{\eta}(\mathbf{x})\| \le 1 \quad \text{for all } \mathbf{x} , \tag{10.13}$$

and

$$\lim_{k\to\infty} \int_{\partial\tilde{\Sigma}_k} \langle \mathbf{X}, \tilde{\eta}_k\rangle dH^{n-1} = \int_{\mathbf{R}^{n-1}\times\{0\}} \langle \mathbf{X}, \tilde{\eta}\rangle dH^{n-1} . \tag{10.14}$$

On the other hand, though, in view of (10.5), we have that, as $k \to \infty$, the left hand side of (10.12) converges to $\int_C \mathrm{div}_C \mathbf{X} dH^n$.

Putting all this, together with (10.6), into (10.12), we thence get that,

$$\sum_{i=1}^{l} n_i \int_{\mathcal{H}_i} \mathrm{div}_{\mathcal{H}_i} \mathbf{X}\, dH^n = -\int_{\mathbf{R}^{n-1}\times\{0\}} \langle \mathbf{X}, \tilde{\eta}\rangle dH^{n-1} . \tag{10.15}$$

Yet clearly, letting η_i denote the (constant) inward unit conormal to $\mathcal{H}_i$ along $\partial\mathcal{H}_i \equiv \mathbf{R}^{n-1} \times \{0\}$, we have, for each i, that $\eta_i \in \hat{V}^{\delta_0}$, and

$$(10.16) \qquad \int_{\mathcal{H}_i} \mathrm{div}_{\mathcal{H}_i}\mathbf{X}\, dH^n = - \int_{\mathbf{R}^{n-1}\times\{0\}} \langle \mathbf{X}, \eta_i\rangle dH^{n-1}\,.$$

Thence, by (10.16) in (10.15), we have that, for any $\mathbf{X} \in C_c^\infty(B_1(0); \mathbf{R}^{n+1})$,

$$(10.17) \qquad \int_{\mathbf{R}^{n-1}\times\{0\}} \left\langle \mathbf{X}, \left(\tilde{\eta} - \sum_{i=1}^{l} n_i\eta_i\right)\right\rangle dH^{n-1} = 0\,.$$

However, unless $l = 1$ and $n_1 = 1$, we have, from (10.7) in Lemma B.1 of Appendix B, that $\sum_{i=1}^{l} n_i\eta_i$ would be a vector in $\hat{V}^{\delta_0}$ of length strictly greater than one. But this observation, together with (10.13), would give a contradiction in (10.17); and this proves that we must indeed have $l = 1$ and $n_1 = 1$, as desired. Q.E.D.

As a final preliminary we recall the following special case of Allard's boundary regularity theorem for varifolds. This version is for *a priori* smooth hypersurfaces with boundary in $\mathbf{R}^{n+1}$, which, while much weaker than Allard's general result, is all we shall need in this paper (see [1, Corollary on p. 419]).

10.4 Theorem (Allard). *Suppose $\mathcal{B}$ is a smooth $(n-1)$-dimensional submanifold of $B_1(0)$, the open unit ball in $\mathbf{R}^{n+1}$, and suppose $\kappa > 0$ bounds the curvature of $\mathcal{B}$ in the sense that, for every point $\mathbf{b} \in \mathcal{B}$, and every unit normal, ξ, to $\mathcal{B}$ at $\mathbf{b}$, we have, for all $\mathbf{y} \in \mathcal{B}$,*

$$(10.18) \qquad |\langle \mathbf{y} - b, \xi\rangle| \le \kappa|\mathbf{y} - b|^2/2\,.$$

Suppose M is a smooth n-dimensional submanifold of $B_1(0)\backslash\mathcal{B}$, relatively closed in $B_1(0)\backslash\mathcal{B}$, with mean curvature vector $\mathbf{H}(x)$.

Then, given any q, ϵ with $n < q < \infty$ and $0 < \epsilon < 1$, there is a $\delta = \delta(n, q, \epsilon) > 0$, independent of M, such that, provided we have

(i) $\mathrm{dist}(0, M) \le \delta \quad , \quad \kappa < \delta$ *; and*

(ii) $H^k(M) \le (1 + \delta)\omega_n/2$ *; and*

(iii) $\left(\int_M |\mathbf{H}(\mathbf{x})|^q dH^n\right)^{1/q} \le \delta$ *,*

then the following hold:

(a) *$\mathcal{B} \cap B_1(0)$ is contained in the closure of M relative to $B_1(0)$; and*

(b) *after a suitable rotation in $\mathbf{R}^{n+1}$, sending M to M^* say, then $M^* \cap (B^{(n)}_{1-r(\epsilon)}(0) \times \mathbf{R})$ may be written as the graph of a $C^{1,1-n/q}$ function*

$f : B^{(n)}_{1-r(\epsilon)}(0) \cap P_{R^n\times\{0\}}(M^*) \to \mathbf{R}$ satisfying $f(0) = 0$, $|Df(0)| = 0$, and

$$|Df(\mathbf{z}_2) - Df(\mathbf{z}_1)| \le \epsilon|\mathbf{z}_2 - \mathbf{z}_1|^{1-n/q} \quad \textit{for all } \mathbf{z}_1, \mathbf{z}_2 \,. \tag{10.19}$$

Here $P_{R^n\times\{0\}}$ denotes orthogonal projection onto $\mathbf{R}^n \times \{0\}$, as in Lemma 10.2, $B^{(n)}_\rho(0)$ denotes the ρ-ball about 0 in $\mathbf{R}^n \times \{0\}$, and $r(\epsilon)$ is a function with $r(\epsilon) \searrow 0$ as $\epsilon \searrow 0$.

We are now finally ready to return to our task of obtaining an improved boundary estimate for $|A|^2(p,t)$ for our evolving surfaces, M_t. As a start we establish a number of results leading up to a certain "graphical representation theorem", locally about boundary points $\mathbf{x}_0 \in \Gamma$, for certain rescalings of the M_t (see Lemma 10.7).

Throughout the remainder of this section, let $\beta \in (0, 1/2)$ be fixed arbitrarily, and then let us adopt the notation, for any $\mathbf{x}_0 \in \Gamma$ and any $t_* \in [0, T)$, that $\Sigma_{\mathbf{x}_0,t_*}$ denotes the hypersurface $[(T-t_*)^{-\beta}(M_{t_*} - \mathbf{x}_0)] \cap B_1(0)$.

10.5 Result. *For any $\epsilon_0 > 0$, any sequence of times $\{t_j \nearrow T\}$, and any sequence of boundary points $\{\mathbf{x}_j \in \Gamma\}$, there exists a subsequence of the corresponding hypersurfaces Σ_{x_j,t_j}, say $\{\Sigma_{x_{j_k},t_{j_k}}\}_{k=1}^\infty$, and an integer k_0, and a radius $\rho_0 > 0$, such that, for all $k \ge k_0$, and for all $\rho \in (0, \rho_0]$,*

$$|\Sigma_{x_{j_k},t_{j_k}} \cap B_\rho(0)| \le (1+\epsilon_0)\frac{\omega_n \rho^n}{2} \,. \tag{10.20}$$

Proof. Consider the sequence of surfaces $\{\Sigma_{x_j,t_j}\}$ viewed as n-varifolds. Note that, by Lemma 9.5, these surfaces satisfy that

$$\sup_{\Sigma_{x_j,t_j}} |\mathbf{H}(\mathbf{x})| \to 0 \quad \text{as } j \to \infty \,. \tag{10.21}$$

Now, by Lemma 10.1(b), we have a uniform area estimate on all these surfaces (at least once $t_j > T/2$). Thus, by compactness (see [9, Theorem 42.7]), there must be a subsequence of them, say $\{\Sigma_{x_{j_k},t_{j_k}}\}_{k=1}^\infty$, that converge as varifolds to some limit, Σ, as $k \to \infty$. Henceforth in this proof, for ease, we write Σ_k for $\Sigma_{x_{j_k},t_{j_k}}$.

Now, by Proposition 10.3, Σ must be an integer multiplicity rectifiable varifold, stationary away from $(\mathbf{R}^n \times \{0\}) \cap B_1(0)$, and more importantly must have density $1/2$ at 0. But then, for any $\epsilon_1 > 0$, we must be able to find a radius $\rho_1 = \rho_1(\epsilon_1) > 0$ such that

$$|\Sigma \cap B_{\rho_1}(0)| \le (1+\epsilon_1/2)\frac{\omega_n \rho_1^n}{2} \,;$$

and then in turn, since the Σ_k are converging as varifolds to Σ, we must be able to find a $k_1 \in \mathbf{N}$ such that, for all $k \geq k_1$

$$|\Sigma_k \cap B_{\rho_1}(0)| \leq (1+\epsilon_1)\frac{\omega_n \rho_1^n}{2}\,. \tag{10.22}$$

Note that this almost completes the proof (with $\epsilon_0 = \epsilon_1$), but not quite, since we have only obtained (10.22) for the particular radius ρ_1, and not yet for all $\rho \in (0, \rho_1]$.

To overcome this, however, we can just invoke Theorem 10.4 as follows. In that theorem take $\epsilon = 1/2$ and (say) $q = 2n$, and then let δ_1 denote the delta corresponding to these choices. Then set $\epsilon_1 = \delta_1/2$, and put ρ_1 to be the radius in (10.22) corresponding to this choice of ϵ_1. Finally take the surface M in the theorem to be successively the surfaces $\rho_1^{-1}\Sigma_k \cap B_1(0)$, $k \geq k_1$.

Then it is clear, from (10.21) and (10.22), that, for all k sufficiently large, these hypersurfaces satisfy the hypotheses (i), (ii) and (iii) of Theorem 10.4. Hence there is some $k_0 \in \mathbf{N}$, and some fixed radius $\rho_2 > 0$, such that, for all $k \geq k_0$, all the surfaces $\rho_1^{-1}\Sigma_k \cap B_{\rho_2}(0)$ have graphical representations as graphs of $C^{1,1/2}$ functions, f_k, over appropriate pieces of hyperplanes through the origin, with all the f_k satisfying a uniform $C^{1,1/2}$-norm estimate.

But then this clearly implies that we can find a uniform radius $\rho_0 > 0$ such that, for all $k \geq k_0$, the surfaces Σ_k satisfy that

$$|\Sigma_k \cap B_\rho(0)| \leq (1+\epsilon_0)\frac{\omega_n \rho^n}{2} \quad \text{for all } \rho \in (0, \rho_0]\,,$$

and this completes the proof. Q.E.D.

10.6 Corollary. *Given any $\epsilon_2 > 0$ there is a uniform radius $\sigma_0(\epsilon_2) > 0$, and a uniform time $t_5(\epsilon_2) < T$, such that, for every $\mathbf{x}_0 \in \Gamma$ and every $t_* \in [t_5, T)$, the corresponding surface $\Sigma_{\mathbf{x}_0,t_*} = [(T-t_*)^{-\beta}(M_{t_*} - \mathbf{x}_0)] \cap B_1(0)$ satisfies that*

$$|\Sigma_{\mathbf{x}_0,t_*} \cap B_\rho(0)| \leq (1+\epsilon_2)\frac{\omega_n \rho^n}{2} \quad \text{for all } \rho \in (0, \sigma_0]\,. \tag{10.23}$$

Proof. Suppose the result were false. Then for some $\epsilon_2 > 0$ we could find a sequence of radii $\{\rho_j \searrow 0\}$, and corresponding sequences of boundary points $\{\mathbf{x}_j \in \Gamma\}$ and times $\{t_j \nearrow T\}$, such that, for each j,

$$|\Sigma_{x_j,t_j} \cap B_{\sigma_j}(0)| > (1+\epsilon_2)\frac{\omega_n \sigma_j^n}{2} \quad \text{for some } \sigma_j \in (0, \rho_j]\,.$$

But then, no matter how we fixed ρ_0 in Result 10.5, we'd have, for the sequence of surfaces $\{\Sigma_{x_j,t_j}\}$, that regardless of what subsequence of them, say $\{\Sigma_k \equiv \Sigma_{x_{j_k}, t_{j_k}}\}_{k=1}^{\infty}$, we considered, then for all k sufficiently large (so that $\rho_{j_k} < \rho_0$) there would be a radius $\sigma_{j_k} < \rho_0$ such that

$$|\Sigma_k \cap B_{\sigma_{j_k}}(0)| > (1+\epsilon_2)\frac{\omega_n \sigma_{j_k}^n}{2}.$$

Yet this would contradict Result 10.5. Q.E.D.

Using results 10.6 and 10.4, we can now give a form of "uniform graphical representation at boundary points" lemma. To state it we adopt the following notation.

For any $\mathbf{x}_0 \in \Gamma$ recall that $\hat{\mathcal{R}}_{\mathbf{x}_0}$ denotes the rotation in $\mathbf{R}^{n+1}$ sending the wedge $V_{\mathbf{x}_0}$ to $\hat{V}_{\mathbf{x}_0}$ (see 1.1). Then, for any $\mathbf{x}_0 \in \Gamma$, any $t_* < T$, and any radius $\rho > 0$, we set

(10.24)

$$\bar{\Sigma}_{\mathbf{x}_0, t_*, \rho} \equiv \hat{\mathcal{R}}_{\mathbf{x}_0}\left(\rho^{-1}(T-t_*)^{-\beta}(M_{t_*} - \mathbf{x}_0)\right) \cap (B_1^{(n)}(0) \times \mathbf{R}) \cap B_2(0).$$

Finally we let $\delta_0 > 0$ also be as in 1.1.

10.7 Lemma. *Let $\beta \in (0, 1/2)$ and $\epsilon_3 \in (0,1)$ be arbitrary. Then there is a fixed time $t_6 = t_6(n, C_0, T, \Gamma, \beta, \epsilon_3, \delta_0) < T$, and a fixed radius $\rho_* = \rho_*(\beta, \epsilon_3, \delta_0) \in (0,1)$, such that, for any $\mathbf{x}_0 \in \Gamma$ and any $t_* \in [t_6, T)$,*

(a) *the surface $\bar{\Sigma}_{\mathbf{x}_0, t_*, \rho_*}$ may be written as the graph of a $C^{1,1/2}$ function, $f_{\mathbf{x}_0, t_*}$, over (its projection into) the fixed hyperplane $\mathbf{R}^n \times \{0\}$; and*

(b) *each such function $f_{\mathbf{x}_0, t_*}$ moreover satisfies that, for all $\mathbf{z}_1, \mathbf{z}_2$ in its domain,*

(10.25)

$$|Df_{\mathbf{x}_0, t_*}(\mathbf{z}_1)| \le C_{28} \quad , \quad |Df_{\mathbf{x}_0, t_*}(\mathbf{z}_2) - Df_{\mathbf{x}_0, t_*}(\mathbf{z}_1)| \le \epsilon_3 |\mathbf{z}_2 - \mathbf{z}_1|^{1/2}$$

where $C_{28} = C_{28}(\delta_0, \beta, \epsilon_3)$ is independent of the choice of $\mathbf{x}_0$ and t_*.

10.8 Remark. It will be clear from the proof that, if we preferred, we could, for any $\kappa < 1$, prove the same result with "$C^{1,1/2}$" replaced by "$C^{1,\kappa}$" (and with (10.25) also modified appropriately), provided we then let t_6 and ρ_* depend also on κ.

Proof. Take any $\mathbf{x}_0 \in \Gamma$ and $t_* < T$. Let $\epsilon_4 > 0$ be a small constant, to be chosen later in terms of ϵ_3 and δ_0, and let $q = 2n$ (say). Then let $\delta_1 = \delta_1(n, \epsilon_3, \delta_0)$ denote the delta from Theorem 10.4 corresponding to this ϵ_4 and q. Finally put $\sigma_0 = \sigma_0(n, \epsilon_3, \delta_0)$ and $t_5 = t_5(n, \epsilon_3, \delta_0)$ to be, respectively, the radius and time from Corollary 10.6 corresponding to the choice $\epsilon_2 = \delta_1$.

Then observe that, by Corollary 10.6, the surface $\bar{\Sigma}_{\mathbf{x}_0, t_*, \sigma_0}$ will satisfy hypothesis (ii) of Theorem 10.4 (with $\delta = \delta_1$), provided $t_* \ge t_5$. (Indeed this will be true of all the surfaces $\bar{\Sigma}_{\mathbf{x}_0, t_*, \rho}$, $\rho \in (0, \sigma_0]$). But also then, noting Lemma 9.5, we can clearly find a uniform time $t_7 = t_7(n, C_0, T, \Gamma, \beta, \delta_1) < T$ such that, provided $t_* \ge t_7$, the surface $\bar{\Sigma}_{\mathbf{x}_0, t_*, \sigma_0}$

will additionally satisfy hypotheses (i) and (iii) of Theorem 10.4 (with $\delta = \delta_1$).

So now put $t_6 = \max\{t_5, t_7\}$, which, as required, is independent of our particular choice of $\mathbf{x}_0$. Then, by Theorem 10.4, we obtain, in view of the way we defined δ_1, that, provided $t_* \geq t_6$, the surface $\bar{\Sigma}_{\mathbf{x}_0,t_*,\sigma_0}$ may be written as the graph of a $C^{1,1/2}$ function, $g_{\mathbf{x}_0,t_*,\sigma_0}$, over (its projection into) the n-dimensional ball of radius $1 - r(\epsilon_4)$ in a suitable hyperplane, $\Pi_{\mathbf{x}_0,t_*,\sigma_0}$, through the origin in $\mathbf{R}^{n+1}$; and moreover this function $g_{\mathbf{x}_0,t_*,\sigma_0}$ satisfies that $g_{\mathbf{x}_0,t_*,\sigma_0}(0) = 0$, $|Dg_{\mathbf{x}_0,t_*,\sigma_0}(0)| = 0$, and

$$\left|Dg_{\mathbf{x}_0,t_*,\sigma_0}(\mathbf{z}_2) - Dg_{\mathbf{x}_0,t_*,\sigma_0}(\mathbf{z}_1)\right| \leq \epsilon_4 |\mathbf{z}_2 - \mathbf{z}_1|^{1/2} \tag{10.26}$$

for all $\mathbf{z}_1, \mathbf{z}_2$ in its domain.

Note in particular that (10.26) implies (taking $\mathbf{z}_2 = 0$) that

$$\left|Dg_{\mathbf{x}_0,t_*,\sigma_0}(\mathbf{z}_1)\right| \leq \epsilon_4 \tag{10.27}$$

for all $\mathbf{z}_1$ in the domain of $g_{\mathbf{x}_0,t_*,\sigma_0}$.

But also now it is clear that the hyperplane $\Pi_{\mathbf{x}_0,t_*,\sigma_0}$ must contain $\mathbf{R}^{n-1} \times \{0\}$ and be lying in the wedge $\hat{V}^{(\delta_0)}$. (Indeed since $|Dg_{\mathbf{x}_0,t_*,\sigma_0}(0)|$ is zero we see that $\Pi_{\mathbf{x}_0,t_*,\sigma_0}$ is just the tangent plane to $\bar{\Sigma}_{\mathbf{x}_0,t_*,\sigma_0}$ at the origin).

So now simply select $\epsilon_4 = \epsilon_4(\epsilon_3, \delta_0)$ to satisfy first of all that $\epsilon_4 \leq \epsilon_3$; secondly that $1 - r(\epsilon_4) \geq 1/2$; and thirdly (noting (10.27)) that it is small enough in terms of δ_0 that, within the cylinder $B^{(n)}_{R_0}(0) \times \mathbf{R}$ of radius $R_0(\delta_0) \equiv \frac{1}{2}\sin(\delta_0/2)$, the surface $\bar{\Sigma}_{\mathbf{x}_0,t_*,\sigma_0}$ is still a graph over its projection into $\mathbf{R}^n \times \{0\}$. See Figure 1.

Then with this choice of ϵ_4 it is clear, on noting the arbitrariness of $\mathbf{x}_0 \in \Gamma$ in the above discussion, that we are done, on simply setting $\rho_* \equiv R_0\sigma_0$. Q.E.D.

Now at last, armed with this "graphical representation" result, we can derive the improved boundary estimate for $|A|^2(p,t)$ that we have been seeking.

10.9 Theorem. *For any $\beta \in (0, 1/2)$ there is a uniform constant $C_{29} = C_{29}(n, C_0, T, \Gamma, \beta, \delta_0)$ such that*

$$\max_{p\in\Gamma} |A|^2(p,t) \leq \frac{C_{29}}{(T-t)^{2\beta}}\,. \tag{10.28}$$

Proof. Fix $\mathbf{x}_0 \in \Gamma$ and $\beta \in (0, 1/2)$ arbitrarily. Then let t_6, ρ_* be as in Lemma 10.7 with $\epsilon_3 = 1/2$, and fix any $\tilde{t} \in [t_6, T)$. Then, for this choice of $\tilde{t}$, put

$$t_* \equiv \frac{1}{2}(T + \tilde{t}) \quad , \quad r_* \equiv \rho_*(T - t_*)^{\beta} \tag{10.29}$$

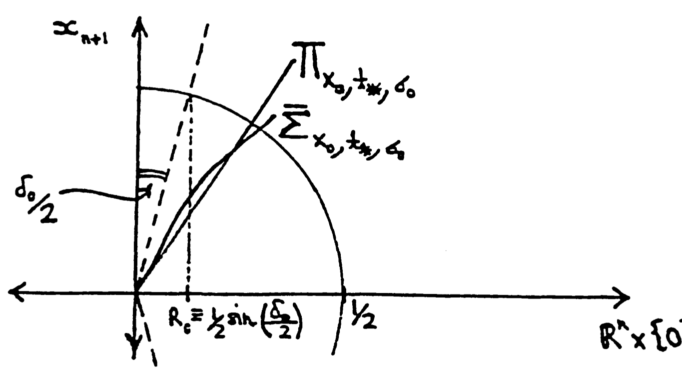

FIGURE 1. $\bar{\Sigma}_{\mathbf{x}_0,t_*,\sigma_0}$ may be written as a graph over (its projection into) the 1/2-ball about 0 in the hyperplane $\Pi_{\mathbf{x}_0,t_*,\sigma_0} \subset \hat{V}^{(\delta_0)}$, whose slope with respect to $\Pi_{\mathbf{x}_0,t_*,\sigma_0}$ is very small. Hence, within some cylinder $B^{(n)}_{R_0}(0) \times \mathbf{R}$, with R_0 determined by δ_0, $\bar{\Sigma}_{\mathbf{x}_0,t_*,\sigma_0}$ may also be written as a graph over (its projection into) the fixed hyperplane $\mathbf{R}^n \times \{0\}$.

and, for each $t \in [t_6, t_*]$, write $\Sigma_t \equiv \hat{\mathcal{R}}_{\mathbf{x}_0}(M_t - \mathbf{x}_0) \cap (B^{(n)}_{r_*}(0) \times \mathbf{R}) \cap B_{2r_*}(0)$. Then, by Lemma 10.7, we have that, for each $t \in [t_6, t_*]$,

$$\Sigma_t = \operatorname{graph}(u(\mathbf{x}, t)) \,, \tag{10.30}$$

where $u^{(t)}(\cdot) \equiv u(\cdot, t) : P_{R^n \times \{0\}}(\Sigma_t) \cap B^{(n)}_{r_*}(0) \to \mathbf{R}$ is a $C^{1,1/2}$ function, satisfying that $u^{(t)}(0) = 0$ and, for all $\mathbf{z}_1, \mathbf{z}_2$ in its domain,

(10.31)
$$|Du^{(t)}(\mathbf{z}_1)| \le C_{28}(\delta_0, \beta) \quad , \quad |Du^{(t)}(\mathbf{z}_2) - Du^{(t)}(\mathbf{z}_1)| \le \frac{1}{2}|\mathbf{z}_2 - \mathbf{z}_1|^{1/2} \,.$$

Note furthermore that, since $\Gamma_* \equiv \hat{\mathcal{R}}_{\mathbf{x}_0}(\Gamma - \mathbf{x}_0) \cap (\bar{B}^{(n)}_{r_*}(0) \times \mathbf{R}) \cap B_{2r_*}(0)$ is remaining unchanging, the domains $P_{R^n \times \{0\}}(\Sigma_t) \cap B^{(n)}_{r_*}(0)$ are actually fixed, independent of $t \in [t_6, t_*]$, so we write $\Omega'' \equiv P_{R^n \times \{0\}}(\Sigma_t) \cap B^{(n)}_{r_*}(0)$.

But also now it is clear that the surfaces Σ_t are still flowing by mean curvature, unaffected by the rotation $\hat{\mathcal{R}}_{\mathbf{x}_0}$, and so the function $u : \Omega'' \times [t_6, t_*] \to \mathbf{R}$ satisfies the *linear* differential equation (see, for instance, [4, equation (2)]) that

$$\frac{d}{dt}u - a_{ij}(\mathbf{x}, t) D_i D_j u(\mathbf{x}, t) = 0 \,, \tag{10.32}$$

where

$$a_{ij}(\mathbf{x}, t) \equiv \delta_{ij} - \frac{D_i u(\mathbf{x}, t) D_j u(\mathbf{x}, t)}{1 + |Du(\mathbf{x}, t)|^2} \,, \tag{10.33}$$

and we are using the Einstein summation convention in (10.32).

Note that, in treating (10.32) as a linear equation, we are using the usual "doublethink" of ignoring that the functions $D_i u$ in the definition of $a_{ij}(\mathbf{x}, t)$ are related to our solution, u, of (10.32), and are treating them just as given functions on $\Omega'' \times [t_6, t_*]$.

Now we are done, however, simply by invoking the parabolic machinery of [8]. For observe that, by (10.31), our differential equation (10.32) is uniformly parabolic on $\Omega'' \times [t_6, t_*]$ (indeed the least eigenvalue of $a_{ij}(\mathbf{x}, t)$ is bounded below, on $\Omega'' \times [t_6, t_*]$, by $(1 + C_{28})^{-1}$). Moreover the coefficient functions $a_{ij}(\mathbf{x}, t)$ are of class $C^{0,1/2}$ in $\Omega'' \times [t_6, t_*]$, with a uniform estimate on their $C^{0,1/2}$-norms determined only by C_{28}; and the part of $\partial\Omega''$ given by $\Gamma''_* \equiv P_{R^n \times \{0\}}(\Gamma_*)$ is of class $C^{2,1/2}$ (with a uniform estimate also, determined only by the geometry of Γ, and independent of the choice of $\mathbf{x}_0$).

Therefore we can apply to (10.32) the spatially scale-invariant version of estimate (10.5) of [8, Theorem 10.1, p. 351-352]. Doing so we obtain the estimate that

$$
\begin{aligned}
\|u\|_{C^{2,1/2}(Q')} \leq C_{30}(\Gamma, C_{28}) &\left(\|u(\cdot, t_6)\|_{C^{2,1/2}(\Omega'')} + \|\Phi\|_{C^{2,1/2}(\Gamma''_* \times (t_6, t_*))}\right) \\
&+ C_{31}(\Gamma, C_{28}) \left(\max_{t \in [t_6, t_*]} \|u(\cdot, t_6)\|_{L^2(\Omega'')} + \|Du\|_{L^2(Q'')}\right),
\end{aligned} \tag{10.34}
$$

where $\Phi(\mathbf{x})$ denotes the (time-independent) function $u\big|_{\Gamma''_*}$, where

$$
\Omega' \equiv \Omega'' \cap B^{(n)}_{r_*/2}(0) \quad , \quad Q' \equiv \Omega' \times (t_6, \tilde{t}\,) \quad , \quad Q'' \equiv \Omega'' \times (t_6, t_*) \, ,
$$

and where the spatially scale-invariant norms used in (10.34) are explicitly given by

$$
\begin{aligned}
\|u\|_{C^{2,1/2}(Q')} \equiv &\left(\frac{r_*}{2}\right)^{5/2} \sup_{t \in (t_6, \tilde{t})} \left[D^2 u^{(t)}\right]_{1/2;\Omega'} + \left(\frac{r_*}{2}\right)^{1/2} \sup_{t \in (t_6, \tilde{t})} \left[d_t u^{(t)}\right]_{1/2;\Omega'} \\
&+ \left(\frac{r_*}{2}\right)^{2} \sup_{x \in \Omega'} \left[D^2 u(\mathbf{x}, \cdot)\right]_{1/4;(t_6, \tilde{t})} + \sup_{x \in \Omega'} \left[d_t u(\mathbf{x}, \cdot)\right]_{1/4;(t_6, \tilde{t})} \\
&+ \left(\frac{r_*}{2}\right)^{2} \left|D^2 u\right|_{0;Q'} + \left|d_t u\right|_{0;Q'} + \left(\frac{r_*}{2}\right) \left|Du\right|_{0;Q'} + \left|u\right|_{0;Q'}
\end{aligned} \tag{10.35}
$$

and

$$
\begin{aligned}
\|u(\cdot, t_6)\|_{C^{2,1/2}(\Omega'')} \equiv\ & r_*^{5/2} \left[D^2 u(\cdot, t_6)\right]_{1/2;\Omega''} + r_*^2 \left|D^2 u(\cdot, t_6)\right|_{0;\Omega''} \\
&+ r_* \left|Du(\cdot, t_6)\right|_{0;\Omega''} + \left|u(\cdot, t_6)\right|_{0;\Omega''}
\end{aligned} \tag{10.36}
$$

and

$$
\begin{aligned}
(10.37)\qquad \|\Phi\|_{C^{2,1/2}(\Gamma''_* \times (t_6,t_*))} &\equiv \|\Phi\|_{C^{2,1/2}(\Gamma''_*)} \\
&\equiv r_*^{5/2}\big[\tilde{D}^2\Phi\big]_{1/2;\Gamma''_*} + r_*^2\big|\tilde{D}^2\Phi\big|_{0;\Gamma''_*} \\
&\quad + r_*\big|\tilde{D}\Phi\big|_{0;\Gamma''_*} + \big|\Phi\big|_{0;\Gamma''_*} \,.
\end{aligned}
$$

In (10.37) we have used that Φ is independent of t (since it is the function whose graph is the fixed boundary Γ_*); and also the derivatives $\tilde{D}\Phi$ and $\tilde{D}^2\Phi$ on the right hand side of (10.37) represent derivatives only along Γ''_*.

But now (10.34) immediately implies that we have $\|u\|_{C^{2,1/2}(\Omega'\times(t_6,\tilde{t}\,))} \le C_{32}$ for some constant $C_{32} = C_{32}(\Gamma, C_{28}, t_6)$, whence in particular we obtain that

$$
(10.38)\qquad \big|D^2u\big|_{0;\Omega'\times(t_6,\tilde{t}\,)} \le \frac{4}{r_*^2}C_{32}\,.
$$

This, however, obviously yields (noting the definition of r_* in (10.29)) that $|D^2u(0,\tilde{t})| \le 2^{2+2\beta}\rho_*^{-2}C_{32}(T-\tilde{t}\,)^{-2\beta}$, and this in turn implies that

$$
(10.39)\qquad |A|^2(\mathbf{x}_0,\tilde{t}) \le C_{29}(T-\tilde{t}\,)^{-2\beta}\,,
$$

for some constant $C_{29} = C_{29}(n,\beta,\rho_*,C_{32})$. The result now follows by checking the dependencies of the constants ρ_* and C_{32}, and noting that $\mathbf{x}_0 \in \Gamma$ and $\tilde{t} \in [t_6, T)$ were arbitrary here (along with using (1.3) to handle the case of times less than t_6). Q.E.D.

10.10 Remark. The higher derivative analogues of estimate (10.34) also then hold, of course, by the usual bootstrapping method.

Now that we have this improved boundary estimate for $|A|^2$ it is straightforward to extend it to the whole boundary strip, $\mathcal{U}_0$, along the lines of the proof of Lemma 9.5.

10.11 Lemma. *Let $\mathcal{U}_0$ be the neighbourhood of Γ^+ in Corollary 9.2. Then, for any $\beta \in (0,1/2)$ there is a constant $C_{33} = C_{33}(n,\beta,C_0,T,\Gamma,\delta_0)$ such that, on $\mathcal{U}_0$, we have the improved estimate for $|A|^2$ that, for all $t \in [0,T)$,*

$$
|A|^2(p,t) \le \frac{C_{33}}{(T-t)^{2\beta}}\,.
$$

Proof. As in Lemma 9.5, let $\alpha \in (0,1/4)$, $\beta \in (0,1/2)$ be arbitrary (α to be chosen later), and set $\epsilon \equiv 2\beta/3$. Then let $t_0(\alpha,\beta)$ be as in Corollary 9.2, so that the ϵ-boundedness hypothesis (see 9.1) holds for all $p \in \mathcal{U}_0$ with this uniform time t_0.

Next fix $p \in \mathcal{U}_0$ arbitrarily. Then put

$$
(10.40)\qquad \chi_p(\tilde{p},t) \equiv (T-t)^{2\beta}\psi_p(\tilde{p},t)|A|^2(\tilde{p},t)\,,
$$

where ψ_p is as in (9.6), and furthermore set

(10.41)
$$\chi_{max}(t) \equiv \max_{\tilde{p}\in M_0} (\chi_p(\tilde{p},t)) \,, \ \mathcal{M}_t^* \equiv \{\tilde{p} \in M_0 : \chi_p(\tilde{p},t) = \chi_{max}(t)\} \,.$$

Now, by Theorem 10.9, we have that, for all $t \in [0,T)$,

$$\max_{\Gamma} \chi_p(\cdot,t) \leq C_{29} \,. \tag{10.42}$$

But thence, combining property (9.5) of the function λ, the Type I hypothesis, (1.3), estimate (3.2), and Corollary 9.2, together with (10.42), we may readily deduce (proceeding exactly as in the proof of Lemma 9.5) that we can find a uniform time $t_2 = t_2(\alpha,\beta) \geq T-1$ such that, for all $t \in [t_2,T)$,

$$\chi_{max}(t) \geq \max\{C_0, C_{29}\} \Longrightarrow \mathcal{M}_t^* \cap \Gamma = \phi \tag{10.43}$$

and

(10.44)
$$\chi_{max}(t) \geq \max\{C_0, C_{29}\} \Longrightarrow |A|^2(\tilde{p},t) \leq \frac{\epsilon}{T-t} \quad \text{for all } \ \tilde{p} \in \mathcal{M}_t^* \,.$$

Yet also we can derive, for our function χ_p, an analogue of the differential inequality (9.7) for μ_p. Computing as in the proof of Lemma 9.3, we get that

$$\begin{aligned}\frac{d}{dt}\chi_p \leq \ & \triangle\chi_p - 2\frac{\nabla\psi_p}{\psi_p}\cdot\nabla\chi_p \\ & + \left(2|A|^2 - \frac{2\beta}{\tau} + C_{34}(n)(1+\tau^{-2\alpha})\max_{M_t}|\mathbf{H}|\right)\chi_p \,.\end{aligned}$$

Therefore, from (10.43) and (10.44) in this inequality, together with the estimate (9.12) for $\max_{M_t}|\mathbf{H}|$, we obtain that, for all $t \in [t_2,T)$,

$$\begin{aligned}\chi_{max}(t) \ & \geq \max\{C_0,C_{29}\} \Longrightarrow \frac{d}{dt}\chi_{max}(t) \\ & \leq \frac{2\epsilon - 2\beta + C_{35}(n)\tau^{\frac{1}{2}-2\alpha}}{\tau}\chi_{max}(t) \,.\end{aligned} \tag{10.45}$$

So now, as in Lemma 9.5, fix $\alpha = 1/8$, and set $t_3(n)$ to be the least time such that $C_{35}\tau^{1/4} \leq \epsilon$ for all $t \in [t_3,T)$. Then, putting $t_4 = \max\{t_2,t_3\} = t_4(n,\beta,T)$, we get, for $t \in [t_4,T)$, that, whenever $\chi_{max}(t) \geq \max\{C_0,C_{29}\}$, we'll have

$$\frac{d}{dt}\chi_{max}(t) \leq \frac{3\epsilon - 2\beta}{\tau}\chi_{max}(t) \leq 0 \,. \tag{10.46}$$

The result of the lemma now follows easily, noting the definition of χ_p, and the arbitrariness of our choice of $p \in \mathcal{U}_0$ (cf. the proof of Lemma 9.5). Q.E.D.

To conclude this section it is convenient to re-express the result of Lemma 10.11 as follows. Let $\mathcal{U}_1 \subset \mathcal{U}_0 \subset M_0$ be the sub-neighbourhood of the extended boundary, Γ^+, defined by

$$\mathcal{U}_1 \equiv \{p \in M_0 : \text{dist}(\hat{p}, \Gamma) < d_*/2\}\,, \tag{10.47}$$

where $d_* > 0$ is set to be $\min_{p \in M_0 \setminus \mathcal{U}_0}\{\text{dist}(\hat{p}, \Gamma)\}$. Then Lemma 10.11 trivially implies;

10.12 Corollary. *Let $\mathcal{U}_1$ be as above. Then for any $\alpha \in (0, 1/4)$ there is a uniform time $t_8 = t_8(\alpha, C_0, d_*) < T$ such that, for any $\beta \in (0, 1/2)$, and any $t \in [t_8, T)$, the estimate*

$$\max_{M_t \cap B_{\tau^{\alpha/2} + \sqrt{2nC_0\tau}}(\hat{p})} |A|^2 \leq \frac{C_{33}}{(T-t)^{2\beta}} \tag{10.48}$$

holds for all $p \in \mathcal{U}_1$, where $C_{33}(n, \beta, C_0, T, \Gamma, \delta_0)$ is as in Lemma 10.11.

11. Boundedness of $\mathbf{H}^2$ in a boundary strip

Having established the improved estimate for $|A|^2(p,t)$ of the previous section, it is now a relatively simple matter to bootstrap up the boundary strip estimate for H^2, of Lemma 9.5, to a full boundedness result.

11.1 Lemma. *Let $\mathcal{U}_1$ be as in (10.47). Then there is a constant $C_{36}(n, C_0, T, \Gamma, \delta_0)$ such that, on $\mathcal{U}_1$, we have the improved estimate for H^2 that, for all $t \in [0, T)$,*

$$H^2(p,t) \leq C_{36}\,.$$

Proof. Fix $p \in \mathcal{U}_1$ arbitrarily. Then proceed exactly as in the proof of Lemma 9.5, up until immediately prior to equation (9.13) (so in particular the quantities $\alpha, \lambda, \psi_p, \mu_p, \mu_{max}$ and $\mathcal{M}_t$, as well as t_0 and t_2, are as there), with the exception that we alter (9.10) as follows. In that equation, in the proof of Lemma 9.5, we used Corollary 9.2 to gain control of $|A|^2(\tilde{p}, t)$ for points $\tilde{p} \in \mathcal{M}_t$, $t \geq \max\{t_0, t_1\}$. Now, however, by virtue of our improved estimate for $|A|^2$ on the boundary strip $\mathcal{U}_1$, of Corollary 10.12, we may instead replace (9.10) by the stronger estimate that, for all $t \geq \max\{t_1, t_8\}$,

$$\mu_{max}(t) > nC_0 \Longrightarrow |A|^2(\tilde{p}, t) \leq C_{37}(n, C_0, T, \Gamma, \delta_0)\,\tau^{-1/2} \tag{11.1}$$

for all $\tilde{p} \in \mathcal{M}_t$. Here we have taken $\beta = 1/4$ in Corollary 10.12 (though actually any $\beta \in (0, 1/2)$ would serve equally well for our purposes), and t_8 is as there.

Then, by using (11.1) rather than (9.10) in writing down the analogue of (9.13), we now get instead that

$$\frac{d}{dt}\mu_{max}(t) \leq \frac{2C_{37}\tau^{1/2} + C_{22}(n)\tau^{1/2-2\alpha}}{\tau}\mu_{max}(t)\,, \tag{11.2}$$

at any instant $t \in [t_9, T)$ at which $\mu_{max}(t) > nC_0$. Here $t_9 \equiv \max\{t_1, t_8, T-1\}$.

But thus, taking $\alpha = 1/8$ again, we obtain now that, for every $t \in [t_9, T)$ for which $\mu_{max}(t) > nC_0$,

$$\frac{d}{dt}\mu_{max}(t) \leq C_{38}(n, C_0, T, \Gamma, \delta_0)\,\tau^{-3/4}\mu_{max}(t)\,. \tag{11.3}$$

In other words, letting $\mathcal{G}$ and $\zeta(t)$ also be as in the proof of Lemma 9.5 (with "t_4" replaced by "t_9"), we get now that (at least a.e., as usual),

$$\begin{cases} \dfrac{d}{dt}\zeta(t) \leq C_{38}\,\tau^{-3/4}\zeta(t)\,, & t \in \mathcal{G} \\ \dfrac{d}{dt}\zeta(t) = 0\,, & t \in [t_9, T)\backslash\mathcal{G}\,, \end{cases}$$

and then the boundedness claim easily follows by direct integration (and noting that $p \in \mathcal{U}_1$ was arbitrary here). Q.E.D.

12. Full boundary regularity

To complete the four-step program for establishing full boundary regularity, outlined in Section 9, it remains to extend the boundedness result for H^2, in the boundary neighbourhood $\mathcal{U}_1$, of Lemma 11.1, to a corresponding boundedness result in $\mathcal{U}_1$ for $|A|^2$. This, however, may be established by precisely the same procedure as used in Section 10, so we spare the reader a repetition of the argument here.

Noting that the higher derivative estimates on A in our boundary strip can also now readily be established, for instance by graphical representation methods (cf. Remark 10.10), we thence obtain;

12.1 Theorem. *Suppose hypotheses A and B of Section 1 hold, and suppose that the hypersurfaces $\{M_t\}_{t\in[0,T)}$ are flowing by mean curvature as in (1.1). Then there is a fixed neighbourhood, $\mathcal{U}_1 \subset M_0$, of the extended boundary Γ^+, and a constant $C_{39} = C_{39}(n, C_0, T, \Gamma, \delta_0)$, such that, for all $p \in \mathcal{U}_1$, and all $t \in [0, T)$,*

$$|A|^2(p,t) \leq C_{39}\,.$$

Furthermore, corresponding boundary strip estimates hold also on all derivatives of the second fundamental form in time and space.

Finally this clearly implies the full boundary regularity result stated at the outset.

12.2 Remarks. **(i)** Note that, while the Type I hypothesis was used strongly at several points in proving Theorem 6.1, for the discussion in Sections 9 to 12 it may be replaced by the much weaker assumption that, for some $\delta \in (0,1)$, we have $\max_{M_t} |A|^2 \leq C_\delta/(T-t)^{1+\delta}$. Simply modify the arguments along the lines discussed in [10, Section 4] for the boundaryless case.

(ii) Observe that, throughout the proof of the Boundary Regularity Theorem, the first condition in Hypothesis B that Γ lie on the boundary of some uniformly convex body $\mathcal{N}$, with moreover $M_0 \subset \bar{\mathcal{N}}$, was needed just to give us a "wedge condition" at each point $\mathbf{x}_0 \in \Gamma$, namely that there exists some wedge of two half-hyperplanes, V_{x_0}, with opening angle less than $\pi - \delta_0$, $\delta_0 > 0$, which contains each M_t, $t < T$. This wedge condition in turn was required to study sequences of blow-ups of the surfaces M_t about points $\mathbf{x}_0 \in \Gamma$ (in particular to apply Allard's lemma, Lemma 10.2, as well as Proposition 10.3 and Lemma B.1).

In view of these observations we see that the first part of Hypothesis B may in fact be weakened to the requirement merely that Γ lie on some hypersurface, $\mathcal{H}$, which is uniformly mean-convex; i.e. satisfies $H \geq \epsilon$ everywhere, for some $\epsilon > 0$. The reason this will suffice is that in this setting one can still recover the required "wedge condition" at each boundary point.

This may be done by taking the pieces into which $\mathcal{H}$ is split by Γ and flowing each by mean curvature flow for a short time, while holding them fixed along Γ. By the maximum principle the collection of hypersurfaces so produced will enclose each M_t, $t < T$, and moreover, by the mean-convexity condition, will satisfy that for each $\mathbf{x}_0 \in \Gamma$ the two hypersurfaces which meet at $\mathbf{x}_0$ will have tangent planes at $\mathbf{x}_0$ which will form a wedge of angle less than $\pi - \delta_0$ for some $\delta_0 > 0$. Furthermore, while this wedge will not necessarily enclose each of the surfaces M_t, it (or rather a translate of it to the origin) will enclose any limit rescaling of them about $\mathbf{x}_0$. Since this is all that we required in the original arguments we see that this weaker form of of the first part of Hypothesis B will still suffice in the Boundary Regularity Theorem. I am grateful to Prof. Gerhard Huisken, and to the referees, for pointing this out as a possible improvement on the original version of the theorem.

Appendix A - The possible "limiting heat density" values

Recall that, for smooth n-dimensional surfaces without boundary in

$\mathbf{R}^{n+1}$ satisfying $H \geq 0$, there are (by Theorem 4.1) precisely $n+1$ possible embedded "limit surface" solutions to $\mathbf{H} = -\mathbf{x}^{\perp}$, up to rotations in $\mathbf{R}^{n+1}$. These are

$$\tilde{M}_{\infty}^{(n,0)} \equiv \mathbf{R}^n$$

and

$$\tilde{M}_{\infty}^{(n,m)} \equiv S^m(\sqrt{m}) \times \mathbf{R}^{n-m} \quad , \quad m = 1, 2, \ldots, n\,.$$

Similarly, by Theorem 4.2, there is (again up to rotation) only one possible smooth, embedded hypersurface in $\mathbf{R}^{n+1}$ satisfying both $H \geq 0$ and $\mathbf{H} = -\mathbf{x}^{\perp}$, and with boundary an $(n-1)$-plane through the origin. This is (for want of better notation)

$$\tilde{M}_{\infty}^{(n,-1)} \equiv \{(x_1, \ldots, x_{n+1}) : x_{n+1} = 0\,,\ x_n \geq 0\}\,.$$

Now, for each m, n, $m \leq n$, set (cf. Lemma 5.3, and in particular (5.4))

$$\Theta^{(n,m)} \equiv \frac{1}{(2\pi)^{n/2}} \int_{\tilde{M}_{\infty}^{(n,m)}} e^{-|x|^2/2} dH^n(x)\,. \tag{A.1}$$

Then an explicit formula for these values, $\Theta^{(n,m)}$, was found in [10, Appendix A]. Indeed the computations there yield the following lemma, which encapsulates all the main facts we need to know about these numbers:

A.1 Lemma. (a) *For each m, n, $m \leq n$, the numbers $\Theta^{(n,m)}$ are given, in terms of the Gamma function, by*

$$\Theta^{(n,m)} = \begin{cases} 1/2, & m = -1 \\ 1, & m = 0 \\ 2\sqrt{\pi}\,(m/2e)^{m/2}\Big(\frac{1}{\Gamma(\frac{m+1}{2})}\Big), & m = 1, \ldots, n\,. \end{cases} \tag{A.2}$$

Note that each $\Theta^{(n,m)}$ depends only on m, and not on n. Thus we henceforth write simply $\Theta^{(m)}$ in place of $\Theta^{(n,m)}$.

(b). *For each fixed $n \in \mathbf{N}$ the values of $\Theta^{(m)}$, $m = -1, 0, 1, \ldots, n$, are all distinct, and bigger than 1 for $m > 0$. Indeed the numbers $\{\Theta^{(m)} : m = 1, 2, \ldots\}$ form a strictly decreasing sequence in m, with $\lim_{m\to\infty} \Theta^{(m)} = \sqrt{2}$.*

A.2 Remark. Some sample values may easily be computed from (A.2). For instance we have that, for all n,

$$\Theta^{(1)} = \sqrt{2\pi/e} \simeq 1.520 \quad , \quad \Theta^{(2)} = 4/e \simeq 1.472$$

$$\Theta^{(3)} = 2\sqrt{\pi}\,(3/2e)^{3/2} \simeq 1.453 \quad , \quad \Theta^{(4)} = 32/3e^2 \simeq 1.444\,.$$

Appendix B - A geometrical lemma

B.1 Lemma. *Let V denote any wedge in $\mathbf{R}^2$ with wedge angle strictly less than π, say*

$$V \equiv \{(x,y) : y \geq 0 \ \ and \ \ |y/x| \geq \delta_0\}$$

for some $\delta_0 > 0$. Suppose $\{\eta_1, \ldots, \eta_{2k+1}\}$ is any collection of an odd number of unit vectors all lying in the wedge V. Set $\eta = \sum_{i=1}^{2k+1} \eta_i$. Then $\eta \in V$ and $\|\eta\| \geq 1$, with a strict inequality unless $k = 0$.

Proof. The following argument was pointed out to us by Prof. Paul Cohen. We begin by proving a preliminary claim, namely that if ξ_1, ξ_2, ξ_3 are three vectors in V, with ξ_2 lying between (here understood to include the possibility of being parallel to) ξ_1 and ξ_3, and if

$$\|\xi_1\| = \|\xi_3\| = 1 \quad , \quad \|\xi_2\| \geq 1\,,$$

then $\xi \equiv \xi_1 + \xi_2 + \xi_3$ is a vector in V lying between ξ_1 and ξ_3, and with $\|\xi\| > \|\xi_2\| \geq 1$.

To see this claim just observe that $\xi_1 + \xi_3$ will be a vector in V of some positive length. Also, it will lie along the bisector of ξ_1 and ξ_3, and so will have positive inner product with ξ_2. But then clearly $\xi \in V$, and, as desired,

$$\|\xi\|^2 = \|\xi_2\|^2 + 2\langle \xi_2, \xi_1 + \xi_3 \rangle + \|\xi_1 + \xi_3\|^2 > \|\xi_2\|^2\,.$$

Having proven the claim, the lemma is now easily established by induction, as follows. The case $k = 0$ is trivial, and the case $k = 1$ is covered by the claim. Now suppose it has been proven true for families of $(2l - 1)$ unit vectors in V, $l \geq 2$, and suppose $\{\eta_1, \ldots, \eta_{2l+1}\}$ is any family of $(2l + 1)$ unit vectors in V. Without loss of generality we may take it that $\eta_2, \ldots, \eta_{2l}$ lie between η_1 and η_{2l+1}.

Then, by the inductive hypothesis, $\tilde{\eta} \equiv \sum_{i=2}^{2l} \eta_i$ is a vector clearly lying between η_1 and η_{2l+1}, and satisfying $\|\tilde{\eta}\| > 1$. Yet then, by the claim proven at the outset, we'll have that $\sum_{i=1}^{2l+1} \eta_i \in V$, and $\|\sum_{i=1}^{2l+1} \eta_i\| \equiv \|\eta_1 + \tilde{\eta} + \eta_{2l+1}\| > 1$, as desired. Q.E.D.

Appendix C - A remark on short-time existence

We do not wish here to discuss in detail short-time existence for the flow (1.1) (which is checked by first solving geometrically for the flow as a graph over M_0 for some short time, then solving an associated system of ODEs to find a continuous family of reparametrisations of M_0 which turn this graphical solution into a solution of the parametrised evolution, (1.1)).

However, in this regard, two serious issues arise concerning the applicability of our boundary regularity theorem to situations such as the case, described in the abstract, where the initial surface is a "cap" of a sphere. Both relate to the fact that, in the original flow of the "cap", H is discontinuous on Γ at time zero, jumping instantaneously down to be zero on the boundary (i.e. (1.2) is not satisfied right back to $t = 0$). Precisely, these issues are as follows.

(i) Do we really have short-time existence for the flow in such a setting, say in the class $C^{1,\alpha}(M_0 \times [0,T)) \cap C^{2,\alpha}(M_0 \times (0,T))$ for some $T > 0$. Note that this question is not easily addressed directly from the standard theory of, say, [8], since the discussion there of short-time existence for Dirichlet problems always assumes a so-called "first order compatibility condition", which here translates precisely to the condition that $H\big|_\Gamma \equiv 0$ on M_0.

(ii) Even if we do have a solution for a short time might the flow not also instantaneously develop regions (near Γ) where H is negative, so meaning that the mean convexity ($H \geq 0$) hypothesis for the boundary regularity theorem is not satisfied by any of the surfaces at times shortly after the flow has begun.

Fortunately both these potential problems can be handled, and we have:

C.1 Theorem. *Suppose $\mathcal{N} \subset \mathbf{R}^{n+1}$ is a smooth uniformly convex domain, and suppose M_0 is a smooth n-dimensional hypersurface contained in $\mathcal{N}$, with smooth embedded boundary $\partial M_0 \equiv \Gamma \subset \partial\mathcal{N}$ of finite $(n-1)$-dimensional Hausdorff measure. Then for this initial surface M_0 the flow* (1.1) *has a solution in the class $C^{1,\alpha}(M_0 \times [0,T)) \cap C^{2,\alpha}(M_0 \times (0,T))$, $\alpha \in (0,1)$ arbitrary, at least for some short time T depending only on $\mathcal{N}$, M_0 and Γ. Moreover the surfaces M_t, for each $t \in (0,T)$, all satisfy $M_t \subset \mathcal{N}$, $H\big|_\Gamma \equiv 0$, and $H \geq 0$ everywhere.*

C.2 Remark. The proof of this theorem is based on an approximation argument, together with suitable $C^{1,\alpha}$ estimates (as opposed to the usual $C^{2,\alpha}$ bounds) for the flow. (The key point is that in these estimates we want dependence now only on $C^{1,\alpha}$-norms of the initial data). We content ourselves with outlining the proof below, and refer the reader to [11] for the details.

Sketch of Proof. The two main parts to the proof are as follows.

Step 1: Find a sequence of approximating initial surfaces $\{M_0^{(j)}\}$ which satisfy that:

(i) Each $M_0^{(j)}$ has boundary $\Gamma \equiv \partial M_0$, and is expressable as the graph over M_0 of a smooth function $f^{(j)}$; i.e.

$$M_0^{(j)} = \left\{\mathbf{x} + f^{(j)}(\mathbf{x})\nu(\mathbf{x}) : \mathbf{x} \in M_0 \subset \mathbf{R}^{n+1}, f^{(j)}(\mathbf{x}) \in \mathbf{R}\right\}$$

where $\nu(\mathbf{x})$ denotes a (smoothly varying) choice of unit normal for M_0 at $\mathbf{x}$, and where $f^{(j)}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \Gamma$.

(ii) The surfaces $M_0^{(j)}$ converge to M_0 in the $C^{1,\alpha}$-sense as $j \to \infty$; that is

$\lim_{j\to\infty} \|f^{(j)}\|_{C^{1,\alpha}(M_0)} = 0$.

(iii) The mean curvature, $H^{(j)}$, of $M_0^{(j)}$ is uniformly bounded independent of j, and satisfies that $H^{(j)} \geq 0$ everywhere and $H^{(j)}\big|_{\Gamma} \equiv 0$.

The key point about these approximating surfaces is that, while tending to M_0 in a $C^{1,\alpha}$-sense, they will also all retain the property of having non-negative mean curvature everywhere under the flow, since they do each satisfy the conditions that $H \geq 0$ and $H\big|_{\partial M_0^{(j)}} \equiv 0$.

Step 2: Next establish suitable *a priori* estimates, for smooth graphical solutions of (1.1), depending only on the $C^{1,\alpha}$-norm of the initial and boundary data.

Using these derive first that there is a uniform minimum time of existence for the flows of each of the initial surfaces $M_0^{(j)}$ constructed in Step 1. Also thence deduce (by viewing M_0 as the $C^{1,\alpha}$-limit of the cauchy sequence of initial surfaces $\{M_0^{(j)}\}$) that a solution exists for a short time for the flow (1.1) with initial data M_0, at least in the class $C^{1,\alpha}(M_0 \times [0,T)) \cap C^{2,\alpha}(M_0 \times (0,T))$.

Finally, noting that on $M_0 \backslash \Gamma$ the solutions corresponding to the $\{M_0^{(j)}\}$ will actually be converging smoothly (on any fixed small time interval, as $j \to \infty$) to the solution corresponding to M_0, establish that the surfaces M_t obtained from flowing M_0 must also satisfy $H \geq 0$ everywhere, as desired, since this non-negativity under the flow does hold for $H^{(j)}$, for each j (as observed in Step 1 above). Q.E.D.

Acknowledgements

I would like to thank Prof. Gerhard Huisken for introducing me to the problem which motivated these investigations, Prof. Brian White for showing me the proof of Lemma 10.1(b), and the referees for their useful comments. I would also like to express my gratitude to Adelaide University, and to the ETH in Zürich, for their hospitality during the period when the bulk of this work was carried out.

Most of all, however, I would like to thank my advisor, Prof. Leon Simon, for his invaluable guidance over the course of numerous discussions. On a specific note he suggested the use of Allard's boundary regularity theory in Section 10 to extend the improved boundary strip estimates for H^2 to corresponding improved estimates for $|A|^2$. More

generally his patience and insight, as well as his encouragement, contributed greatly to the whole of this paper.

References

[1] W. Allard, *On the first variation of a varifold: boundary behaviour*, Ann. of Math. **101** (1975) 418-446.

[2] S. Altschuler, *Singularities of the curve shrinking flow for space curves*, J. Differential Geom. **34** (1991) 491-514.

[3] G. Huisken, *Flow by mean curvature of convex surfaces into spheres*, J. Differential Geom. **20** (1984) 237-266.

[4] _____, *Non-parametric mean curvature evolution with boundary conditions*, J. Differential Equations **77** (1989) 369-378.

[5] _____, *Asymptotic behaviour for singularities of the mean curvature flow*, J. Differential Geom. **31** (1990) 285-299.

[6] _____, *Local and global behaviour of hypersurfaces moving by mean curvature*, Proc. Symp. Pure Math. **54** Part 1, (1993) 175-191.

[7] J. Langer, *A compactness theorem for surfaces with L^p-bounded second fundamental form*, Math. Ann. **270** (1985) 223-234.

[8] O. Ladyženskaja, V. Solonnikov and N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, Transl. of Math. Monographs **23**, Amer. Math. Soc., Providence, Rhode Island, 1968.

[9] L. Simon, *Lectures on geometric measure theory*, Proc. Centre Math. Anal., Canberra, 1983.

[10] A. Stone, *A density function and the structure of singularities of the mean curvature flow*, Calc. of Variations **2** (1994) 443-480.

[11] _____, PhD Thesis, Stanford University, 1994.

AUSTRALIAN COMMONWEALTH TREASURY, CANBERRA, AUSTRALIA