MEAN CURVATURE FLOW

FRIEDERIKE DITTBERNER

Contents

1. Mean curvature flow

Let $M_0 \subset \mathbb{R}^{n+1}$ be a smooth n-dimensional hypersurface without boundary, given by an immersion $X_0 : M^n \to \mathbb{R}^{n+1}$, where M^n is an abstract smooth manifold. We consider the family of embeddings $X : M^n \times [0,T) \to \mathbb{R}^{n+1}$ with

$$X(p,0) = X_0(p)$$

for all $p \in M^n$ and

$$\partial_t X(p,t) = \mathbf{H}(p,t) = -H(p,t)\boldsymbol{\nu}(p,t) = \Delta_{M_t} X(p,t) \tag{MCF}$$

for all $(p,t) \in M^n \times [0,T)$. We abbreviate $M_t := X(M^n,t)$. In the following, we will write $\Delta := \Delta_{M_t}$ and $\nabla := \nabla^{M_t}$. The parabolic ball with radius $r > 0$ and center $(x,t) \in \mathbb{R}^{n+1} \times \mathbb{R}$ is the product

$$P(x,t,r) := B_r(x) \times (t - r^2, t] \subset \mathbb{R}^{n+1} \times \mathbb{R}\,.$$

Given a family of subsets $\{M_t\}_{t \in I}$ the spacetime track is the set

$$\mathcal{M} := \bigcup_{t \in I} M_t \times \{t\} \subset \mathbb{R}^{n+1} \times \mathbb{R}\,.$$

Likewise, given a subset $\mathcal{M} \subset \mathbb{R}^{n+1} \times \mathbb{R}$, the time t slice of $\mathcal{M}$ is

$$M_t = \{x \in \mathbb{R}^{n+1} \mid (x,t) \in \mathcal{M}\}.$$

Example 1.1 (Shrinking spheres and cylinders)**.** (i) Let $M_t = \mathbb{S}^n_{r(t)}$, then (MCF) reduces to an ODE for the radius, namely

$$r' = -\frac{n}{r}.$$

The solution with $r(0) = r_0$ is

$$r(t) = \sqrt{r_0^2 - 2nt},$$

for $t \in (-\infty, r_0^2/2n)$.

(ii) The shrinking cylinders $M_t = \mathbb{S}^m_{r(t)} \times \mathbb{R}^{n-m}$ with $r(t) = \sqrt{r_0^2 - 2mt}$ exist for $t \in (-\infty, r_0^2/2m)$.

(iii) For $n = 1$ the so-called grim reaper is given by $M_t = \operatorname{graph}(u_t)$, where $u(x,t) = t - \log\cos x$ with $x \in (-\pi, \pi)$.

Remark 1.2 (Normal motion and tangential diffeomorphisms)**.** See [Eck04, Remark 2.2(3)]. We will often consider smoothly embedded hypersurfaces M_t satisfying

$$(\partial_t x)^{\perp} = \langle \partial_t x, \boldsymbol{\nu}(x)\rangle \boldsymbol{\nu}(x) = \mathbf{H}(x)$$

for $x \in M_t$, where $\perp$ denotes the projection onto the normal space of M_t. This equation is equivalent to (MCF) up to diffeomorphisms tangent to M_t. Indeed, let $\tilde{X}(\,\cdot\,,t) : M^n \to \mathbb{R}^{n+1}$ with $M_t = \tilde{X}(M^n,t)$ be a family of embeddings satisfying the equation

$$\Big(\partial_t \tilde{X}(q,t)\Big)^{\perp} = \tilde{\mathbf{H}}(q,t) := \mathbf{H}\Big(\tilde{X}(q,t)\Big)$$

for $q \in M^n$, where $\perp$ denotes the projection onto the normal space of $\tilde{X}(M^n,t)$. Let $\phi_t = \psi(\,\cdot\,,t)$ be a family of diffeomorphisms of M^n satisfying

$$\nabla \tilde{X}(\phi(p,t),t)\partial_t\phi(p,t) = -\Big(\partial_t \tilde{X}(\phi(p,t),t)\Big)^{\top},$$

where $\top$ denotes projection onto the tangent space of $\tilde{X}(M^n,t)$. The local existence of such a family is guaranteed by the assumptions on $\tilde{X}$. If we set

$$X(p,t) = \tilde{X}(\phi(p,t),t)$$

then $M_t = X(M^n,t) = \tilde{X}(M^n,t)$, and

$$\partial_t X(p,t) = \partial_t \tilde{X}(p,t) + \nabla \tilde{X}(\phi(p,t),t)\partial_t\phi(p,t) = \Big(\partial_t \tilde{X}(q,t)\Big)^{\perp} = \mathbf{H}(X(p,t)).$$

The previous remark results in the following theorem, see [Sch17a, Theorem 10.6].

Theorem 1.3. *Let $X : M^n \times [0,T) \to \mathbb{R}^{n+1}$ be a solution to* (MCF)*, that is $\langle \partial_t X, \boldsymbol{\nu}\rangle = -H$. Let $R \in O(n+1)$ be an orthonormal map and $\phi : M^n \times [0,T) \to M^n$ smooth. so that $\phi(\,\cdot\,,t)$ is a diffeomorphism. Then $\tilde{X}(p,t) := RX(\phi(p,t),t)$ evolves by*

$$\Big\langle \partial_t \tilde{X}(p,t), \tilde{\boldsymbol{\nu}}(p,t)\Big\rangle = -\tilde{H}(p,t),$$

where $\tilde{H}(p,t) = H(\phi(p,t),t)$ for all $p \in M^n$ and $t \in [0,T)$.

Lemma 1.4 (Evolution equations)**.** *Let* $(M_t)_{t\in[0,T)}$ *evolve by* (MCF)*. Then,*

$$\begin{aligned}
\partial_t g_{ij} &= -2Hh_{ij}\,,\\
\partial_t g^{ij} &= 2Hh^{ij}\,,\\
\partial_t d\mu_t^n &= -H^2\, d\mu_t^n\,,\\
\partial_t \boldsymbol{\nu} &= \nabla H\,,\\
\partial_t h_{ij} &= \nabla_i\nabla_j H - Hh_i^k h_{jk}\\
&= \Delta h_{ij} - 2Hh_i^k h_{jk} + |A|^2 h_{ij}\,,\\
\partial_t h_j^i &= \Delta h_j^i + |A|^2 h_j^i\,,\\
\partial_t H &= \Delta H + H|A|^2\,,\\
\partial_t |A|^2 &= \Delta |A|^2 - |\nabla A|^2 + 2|A|^4\,,\\
\partial_t |\nabla^m A|^2 &\leq \Delta |\nabla^m A|^2 - 2|\nabla^{m+1}A|^2\\
&\quad + C(m,n)\sum_{i+j+k=m} |\nabla^m A|\cdot|\nabla^i A|\cdot|\nabla^j A|\cdot|\nabla^k A|
\end{aligned}$$

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

Proof. See e. g. [Sch18, Section 3]. □

Corollary 1.5. *We have that*

$$\partial_t \mu_t^n(M_t) = -\int_{M_t} H^2 d\mu_t^n\,.$$

Moreover, (MCF) *is the negative* L^2 *gradient flow for the surface area functional.*

Proof. For arbitrary normal speeds $\partial_t X = -F\boldsymbol{\nu}$, we have that $\partial_t g_{ij} = -2Fh_{ij}$ and

$$\frac{d}{dt}\int_{M_t} d\mu_t^n = -\int_{M_t} FH d\mu_t^n \geq -\left(\int_{M_t} F^2 d\mu_t^n\right)^{1/2}\left(\int_{M_t} H^2 d\mu_t^n\right)^{1/2}$$

with equality if and only if $F=H$. □

Theorem 1.6 (Short time existence)**.** *Let* $M_0\subset\mathbb{R}^{n+1}$ *be a smooth, compact hypersurface given by an immersion* $X_0: M^n\to\mathbb{R}^{n+1}$*, there exists a unique, smooth solution of* (MCF) *in some positive time interval.*

Proof. See e.g. [Man11, Section 1.5]. □

Remark 1.7. See [Man11, Remark 1.5.4]. To proof existence and uniqueness for noncompact initial surfaces one needs estimates on the initial hypersurface (like similarly, on the initial datum in order to deal with the heat equation in all $\mathbb{R}^n$) to have existence in some positive interval of time. One possibility is to assume a uniform control on the norm of the second fundamental form of the initial hypersurface. Ecker and Huisken [EH89] showed that a uniform local Lipschitz condition on a hypersurface is sufficient to guarantee short time existence.

Theorem 1.8 (Comparison principle)**.** *Let* $X: M^n\times[0,T)\to\mathbb{R}^{n+1}$ *and* $Y: N^n\times[0,T)\to\mathbb{R}^{n+1}$ *be two hypersurfaces moving by MCF, where* M^n *is compact. Then the distance between them is nondecreasing in time.*

Proof. We follow the lines of [Man11, Theorem 2.2.1]. The distance between the two hypersurfaces $M_t = X(M^n,t)$ and $N_t = Y(N^n,t)$ at time t, is given by

$$d(t) := \inf_{p\in M^n, q\in N^n} |X(p,t)-Y(q,t)|\,.$$

This function is locally Lipschitz in time, as the curvature is locally bounded and the two hypersurfaces move by mean curvature. Hence it is differentiable almost everywhere. Assume that t is a differentiability point. Since M^n is compact, d is actually a minimum. Suppose that $d(t) > 0$ and let $(p_t, q_t) \in M^n \times N^n$ be points, where $d(t)$ is attained. Differentiating $|X(p,t) - Y(q,t)|$ with respect to $v = v_1 \oplus v_2 \in T_{X(p,t)}M_t \bigoplus T_{Y(q,t)}N_t$ yields that

$$0 = \left\langle \frac{X(p_t,t) - Y(q_t,t)}{d(t)}, \nabla_{v_1} X(p_t,t) - \nabla_{v_2} Y(q_t,t) \right\rangle,$$

so that $T_{X(p_t,t)}M_t$ and $T_{Y(q_t,t)}N_t$ have to be parallel. Hence, we can write M_t and N_t locally around $X(p_t,t)$ and $Y(q_t,t)$ as graphs of two functions $f, h : U \times (t-\varepsilon, t+\varepsilon) \to \mathbb{R}$, where $U \subset \mathbb{R}^n$. After rotation, we can assume that $\operatorname{span}(e_1, \ldots, e_n) \subset \mathbb{R}^{n+1}$ is such a tangent space with

$$X(p_t,t) = (0, f(0,t)), \quad Y(q_t,t) = (0, h(0,t)) \quad \text{and} \quad f(0,t) > h(0,t).$$

We calculate

$$\partial_t f = -H_M \langle \boldsymbol{\nu}_M, e_{n+1} \rangle = \Delta f - \frac{D_{ij} f D^i f D^j f}{1 + |Df|^2}$$

and

$$\partial_t h = -H_N \langle \boldsymbol{\nu}_N, e_{n+1} \rangle = \Delta h - \frac{D_{ij} h D^i h D^j h}{1 + |Dh|^2}.$$

The function $f - h$ has a spatial minimum at $x = 0$ at time t. Hence,

$$\Delta f(0,t) - \Delta h(0,t) \geq 0 \quad \text{and} \quad Df(0,t) = Dh(0,t) = 0$$

and so

$$-\langle H_M(p_t,t)\boldsymbol{\nu}_M(p_t,t) - H_N(q_t,t)\boldsymbol{\nu}_N(q_t,t), e_{n+1} \rangle = \Delta f(0,t) - \Delta h(0,t) \geq 0.$$

Since

$$\frac{X(p_t,t) - Y(q_t,t)}{d(t)} = e_{n+1}$$

we obtain at (p_t, q_t) that

$$\begin{aligned}
&\partial_t |X(p,t) - Y(q,t)| \\
&\quad = -\left\langle \frac{X(p_t,t) - Y(q_t,t)}{d(t)}, H_M(p_t,t)\boldsymbol{\nu}_M(p_t,t) - H_N(q_t,t)\boldsymbol{\nu}_N(q_t,t) \right\rangle \\
&\quad = -\langle e_{n+1}, H_M(p_t,t)\boldsymbol{\nu}_M(p_t,t) - H_N(q_t,t)\boldsymbol{\nu}_N(q_t,t) \rangle \geq 0.
\end{aligned}$$

This holds for every minimum so that $\partial_t d \geq 0$. □

Proposition 1.9 (Preservation of embeddedness)**.** *If M_0 is compact and embedded, then M_t is embedded for all $t \in (0,T)$.*

In particular, let

$$m(t) := \max_{(p,s) \in M^n \times [0,t]} |A(p,s)|$$

and

$$l(p,q,t) := \int_p^q |\dot\gamma(s)|_{g(t)} ds \quad \textit{for a minimizing geodesic } \gamma$$

and

$$\Omega_\varepsilon(t) := \left\{ (p,q) \in M^n \times M^n \,\middle|\, m(t) l(p,q,t) \leq \varepsilon \right\}$$

for $\varepsilon > 0$. Then there exists $\varepsilon > 0$ so that M_t is embedded on $\Omega_\varepsilon(t)$ and

$$d(t) := \min_{(p,q) \in (M^n \times M^n) \setminus \Omega_\varepsilon(t)} d(p,q,t) \geq \min\left\{ d(0), \frac{\sin(\varepsilon)}{m(t)} \right\}.$$

Proof. We follow similar lines to [Man11, Proposition 2.2.7]. If the hypersurface M_0 is embedded, then M_t is embedded for a small positive time, otherwise there is a sequence $(p_i, q_i, t_i)_{i\in\mathbb{N}}$ with $X(p_i, t_i) = X(q_i, t_i)$ and $t_i \to 0$. We have for a subsequence, that $p_i \to p$ and $q_i \to q$. If $p \neq q$ then $X(p, 0) = X(q, 0)$, which is a contradiction. If $p = q$, by the smooth existence of the flow, there exists an open neighbourhood $U \subset M^n$ of p so that the map $X(\cdot, t)|_U$ is one-to-one for $t \in [0, \varepsilon)$, which is in contradiction. Define the monotone nondecreasing function

$$m(t) := \max_{(p,s)\in M^n\times[0,t]} |A(p, s)|$$

and we choose a smooth, monotone nondecreasing function $m^* : [0, T) \to \mathbb{R}_+$ such that

$$m(t) \leq m^*(t) \leq 2m(t)$$

for every $t \in [0, T)$. Furthermore, define the geodesic intrinsic distance in the Riemannian manifold $(M^n, g(t))$

$$l(p, q, t) := \int_p^q |\dot\gamma(s)|_{g(t)} ds \qquad \text{for a minimizing geodesic } \gamma$$

and the extrinsic distances

$$d(p, q, t) := |X(p, t) - X(q, t)| \,.$$

Consider the following inscribed and outscribed balls

$$B_{\mathrm{in}}(p, t) := B_{1/m^*(t)}\left(X(p, t) - \frac{\boldsymbol{\nu}(p, t)}{m^*(t)}\right)$$

and

$$B_{\mathrm{out}}(p, t) := B_{1/m^*(t)}\left(X(p, t) + \frac{\boldsymbol{\nu}(p, t)}{m^*(t)}\right)$$

and the geodesic neighbourhood

$$U_\varepsilon(p, t) := \{q \in M^n \mid m^*(t)l(p, q, t) \leq \varepsilon\} \,.$$

Then there exists $\varepsilon \in (0, \pi/2)$ so that

$$X(U_\varepsilon(p, t), t) \cap B_{\mathrm{in}}(p, t) = X(U_\varepsilon(p, t), t) \cap B_{\mathrm{out}}(p, t) = \emptyset$$

Consider the open set

$$\Omega_\varepsilon(t) := \left\{(p, q) \in M^n \times M^n \,\middle|\, m^*(t)l(p, q, t) \leq \varepsilon\right\}$$

and the closed set

$$S(t) := \left\{(p, q) \in M^n \times M^n \,\middle|\, p \neq q \text{ and } X(p, t) = X(q, t)\right\} \,.$$

For embedded M_t,

$$\Omega_\varepsilon(t) \cap S(t) = \emptyset$$

and

$$d_{\partial\Omega_\varepsilon}(t) := \min_{(p,q)\in\partial\Omega_\varepsilon(t)} d(p, q, t) \geq \frac{2\sin(\varepsilon)}{m^*(t)} \,.$$

Assume that $t_0 \in (0, T)$ is the first time where the flow is no more embedded. Since $\Omega \cap S = \emptyset$ and $\partial\Omega_\varepsilon(t_0)$ is compact,

$$\min_{t\in[0,t_0]} d_{\partial\Omega_\varepsilon}(t) = \frac{2\sin(\varepsilon)}{m^*(t_0)} \geq \frac{\sin(\varepsilon)}{m^*(t)} =: c > 0 \,.$$

Furthermore, set

$$d(t) := \min_{(p,q)\in(M^n\times M^n)\setminus\Omega_\varepsilon(t)} d(p, q, t) \,.$$

Assume that there exists a time $t_1 \in (0, t_0)$ so that $d(t_1) < \min\{d(0), c\}$ for the first time. Then $d(t_1)$ is attained at points $(p_1, q_1) \in (M^n \times M^n) \setminus \Omega$. A geometric argument analogous to the one of the comparison principle, Theorem 1.8, shows that $\partial_t d(t) \geq 0$. Hence

$$d(t) \geq \min\{d(0), c\} > 0$$

on $[0, t_0]$, which is a contradiction. □

Theorem 1.10 (Huisken, [Hui84, Corollary 3.6(ii)]). *Let $(M_t)_{t\in[0,T)}$ be a family of closed hypersurfaces moving by* (MCF). *Assume $M_0 = X_0(M)$ closed and mean convex, i.e. $H \geq 0$. Then $H > 0$ for all $t \in (0,T)$.*

Proof. See [Sch17d, Theorem 2.1.2]. That $H \geq 0$ for $t \geq 0$ follows from the evolution equation of H and the parabolic maximum principle, Theorem D.3. Assume that $H(p_0, t_0) = 0$ for some $t_0 > 0$. The strong maximum principle then implies that $H = 0$ for all (p,t) and $0 \leq t \leq t_0$. But this is impossible since any closed hypersurface in $\mathbb{R}^{n+1}$ has points where $\lambda_1 > 0$. □

2. Homothetically shrinking solutions

Definition 2.1 (Homothetically shrinking solutions, Brakke [Bra78, Appendix C]). Let $\lambda : [t_0, T] \to \mathbb{R}_+$ be smooth and decreasing, $\lambda(t_0) = 1$ and $\lambda(T) = 0$. Let $x_0 \in \mathbb{R}^{n+1}$. A *homothetically shrinking* solution $X : M^n \times [t_0, T) \to \mathbb{R}^{n+1}$ to (MCF) satisfies

$$M_t = \lambda(t)(M_0 - x_0) + x_0$$

for all $t \in [t_0, T)$. This describes solutions of (MCF) which move by scaling about x_0.

Remark 2.2. See [Eck04, Examples 2.3(4)]. We can make the separation of variables ansatz

$$\tilde{X}(q,t) = \lambda(t)\tilde{X}(q,t_0)$$

for a family of embeddings $\tilde{X} : M^n \times [t_0, T) \to \mathbb{R}^{n+1}$ with $M_t = \tilde{X}(M^n, t)$ satisfying the evolution equation

$$\left(\partial_t \tilde{X}(q,t)\right)^{\perp} = \left\langle \partial_t \tilde{X}(q,t), \boldsymbol{\nu}(q,t)\right\rangle = \tilde{\mathbf{H}}(q,t)$$

for $q \in M^n$. In Remark 1.2, we saw that there are tangential diffeomorphisms $\phi_t : M^n \to M^n$, $t \in [t_0, T)$, with

$$\tilde{X}(q,t) = X\left(\phi_t^{-1}(q), t\right)$$

for $q \in M^n$, where the embeddings $X(\,\cdot\,, t) : M^n \to \mathbb{R}^{n+1}$ satisfy (MCF). This says that, up to tangential diffeomorphisms, the radial or homothetic motion of the hypersurfaces M_t (described by $\tilde{X}$) is equivalent to their normal motion along the mean curvature vector (described by X). For the shrinking sphere solution these two agree, but for the shrinking cylinder they differ. Since the mean curvature of the embeddings scales with factor $1/\lambda(t)$ we deduce

$$\partial_t \lambda(t)\left(\tilde{X}(q,t_0)\right)^{\perp} = \left(\partial_t \tilde{X}(q,t)\right)^{\perp} = \tilde{\mathbf{H}}(q,t) = \frac{1}{\lambda(t)}\tilde{\mathbf{H}}(q,t_0)$$

for $q \in M^n$. From this we infer that

$$\alpha \equiv 2\lambda(t)\partial_t\lambda(t) = \partial_t \lambda^2(t)$$

is independent of t. We therefore obtain under the assumption $\lambda(t_0) = 1$ that

$$\lambda(t) = \sqrt{1 + \alpha(t - t_0)}$$

for all t satisfying $t > t_0 - 1/\alpha$. Hence

$$\mathbf{H}(p,t) = \alpha\frac{\langle X(p,t), \boldsymbol{\nu}(p,t)\rangle}{2\lambda^2(t)}$$

for $(p,t) \in M^n \times (-\infty, T)$, where $T = t_0 - 1/\alpha$. This describes expanding homothetic solutions about 0 for $\alpha > 0$ and contracting homothetic solutions about 0 for $\alpha < 0$. Let us concentrate on $\alpha < 0$. If we set $\lambda(T) = 0$ for $T > t_0$, which requires the hypersurface to disappear at time T, then $\alpha = -1/(T - t_0)$ and thus

$$\lambda(t) = \sqrt{\frac{T-t}{T-t_0}}$$

and

$$\mathbf{H}(p,t) = \frac{\langle X(p,t), \boldsymbol{\nu}(p,t)\rangle}{2(T-t)}$$

for $(p,t) \in M^n \times (-\infty, T)$.

Lemma 2.3. *Let $(M_t)_{t\in(-\infty,0)}$ be an ancient solution of MCF. Then*

$$H(x) = \frac{\langle x, \boldsymbol{\nu}(x)\rangle}{-2t}$$

for all $x \in M_t$ and $t < 0$ if and only if $M_t = \sqrt{-t}M_{-1}$ for all $t < 0$.

Proof. Let $M_t = \sqrt{-t}M_{-1}$ for all $t < 0$. Then $H(x) = \langle x, \boldsymbol{\nu}(x)\rangle/(-2t)$ for all $x \in M_t$ and $t < 0$ follows by Remark 2.2.

On the other hand, let $H(x) = \langle x, \boldsymbol{\nu}(x)\rangle/(-2t)$ for all $x \in M_t$ and $t < 0$. Then

$$\langle \Delta_{M_t} X(p,t), \boldsymbol{\nu}(p,t)\rangle = -H(p,t) = -\frac{\langle X(p,t), \boldsymbol{\nu}(p,t)\rangle}{-2t}$$

and thus up to tangential motion $X(p,t) = \sqrt{-2t}X(p,t_0\rangle$. □

2.1. **Hypersurfaces.**

Theorem 2.4 (Huisken, [Hui90, Theorem 4.1] and [Hui93]). *Let $M \subset \mathbb{R}^{n+1}$ be a smooth, complete, embedded, mean convex hypersurface such that $H(x) = \langle x, \boldsymbol{\nu}\rangle/2$ at every $x \in M$ and there exists a constant $C > 0$ such that $|A| + |\nabla A| \le C$ and $\mu^n(M \cap B_R) \le Ce^R$, for every ball of radius $R > 0$ in $\mathbb{R}^{n+1}$. Then, up to a rotation in $\mathbb{R}^{n+1}$, M is of the form $\mathbb{S}^m_{\sqrt{2m}} \times \mathbb{R}^{n-m}$ for $m = 0, 1, \ldots, n$.*

Proof. See [Man11, Proposition 3.4.1]. We scale M by the factor $1/2$ so that $H(x) = \langle x, \boldsymbol{\nu}(x)\rangle$ at every $x \in M$. By covariant differentiation of the equation $H = \langle x, \boldsymbol{\nu}\rangle$ in an orthonormal frame $\{\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_n\}$ on M we get by the Weingarten equations $\nabla_i \boldsymbol{\nu} = \partial_i \boldsymbol{\nu} = h_i^j \partial_j x$ that

$$\nabla_j H = \langle x, \nabla_j \boldsymbol{\nu}\rangle = \langle x, \partial_k x\rangle h_j^k$$

and by the Gauss equations $\nabla_i \nabla_j x = -h_{ij}\boldsymbol{\nu}$ and Codazzi equations $\nabla_k h_{ij} = \nabla_k h_{ji} = \nabla_j h_{ik}$ at one fixed point where the Christoffel symbols vanish, that

$$\begin{aligned}
\nabla_i \nabla_j H &= g_{ik} h_j^k + \langle x, \nabla_i \nabla_k x\rangle h_j^k + \langle x, \partial_k K\rangle \nabla_i h_j^k \\
&= h_{ij} + \langle x, \boldsymbol{\nu}\rangle h_{ik} h_j^k + \langle x, \partial_k x\rangle g^{kl} \nabla_i h_{jl} \\
&= h_{ij} - H h_{ik} h_j^k + \langle x, \partial_k x\rangle g^{kl} \nabla_l h_{ij} \\
&= h_{ij} - H h_{ik} h_j^k + \langle x, \nabla h_{ij}\rangle .
\end{aligned} \tag{2.1}$$

Contracting with g^{ij} we have

$$\Delta H = H\left(1 - |A|^2\right) + \langle x, \nabla H\rangle . \tag{2.2}$$

From equation (2.2) and the strong maximum principle for elliptic equations, Theorem D.1, we see that, since M satisfies $H \ge 0$ by assumption and

$$\Delta H \le H + \langle x, \nabla H\rangle$$

we must either have that $H = 0$ or $H > 0$ on all M. Contracting (2.1) with h^{ij}, we have

$$h^{ij}\nabla_i\nabla_j H = |A|^2 - H\operatorname{tr}\bigl(A^3\bigr) + \frac{\langle x, \nabla|A|^2\rangle}{2}\,,$$

which implies, by Simons' identity (A.1),

$$\Delta h_{ij} = \nabla_i\nabla_j H + Hh_{ik}h^k_j - |A|^2 h_{ij}$$

that

$$\begin{aligned}
\Delta|A|^2 &= \Delta(h^{ij}h_{ij}) = h^{ij}\Delta h_{ij} + 2g^{mn}\nabla_m h^{ij}\nabla_n h_{ij} + h_{ij}\Delta h^{ij}\\
&= h^{ij}\Delta h_{ij} + 2g^{mn}g^{ki}g^{lj}\nabla_m h_{kl}\nabla_n h_{ij} + h_{ij}g^{ki}g^{jl}\Delta h_{kl}\\
&= 2h^{ij}\left(\nabla_i\nabla_j H + Hh_{ik}h^k_j - |A|^2 h_{ij}\right) + 2g^{mn}\nabla_m h^i_l\nabla_n h^l_i\\
&= 2|A|^2 - 2H\operatorname{tr}\bigl(A^3\bigr) + \langle x, \nabla|A|^2\rangle + 2H\operatorname{tr}\bigl(A^3\bigr) - 2|A|^4 + 2|\nabla A|^2\\
&= 2|A|^2(1-|A|^2) + \langle x, \nabla|A|^2\rangle + 2|\nabla A|^2\,.
\end{aligned}$$

Assume that $H = 0$. As M is complete and x is a tangent vectorfield on M by the equation $\langle x, \boldsymbol{\nu}\rangle = 0$, for every point $x \in M$ there is a unique solution of the ODE

$$\gamma'(s) = x(\gamma(s)) = \gamma(s)$$

passing through x and contained in M for every $s \in \mathbb{R}$, but such solution is simply the line in $\mathbb{R}^{n+1}$ passing through x and the origin. Thus, M has to be a cone and being smooth the only possibility is a hyperplane through the origin of $\mathbb{R}^{n+1}$.

Assume that $H > 0$ everywhere (so dividing by H and $|A|$ is allowed). For $R > 0$, define

$$\eta_R = \boldsymbol{\nu}_{\partial(M\cap B_R(0))}$$

to be the outward unit conormal to $M \cap B_R(0)$ along $\partial(M \cap B_R(0))$, which is a smooth boundary for almost every $R > 0$ (by Sard's theorem, see homework or Corollary C.3). Then, supposing that R belongs to the set $\mathcal{R} \subset \mathbb{R}^+$ of the regular values of the function $|\cdot|$ restricted to $M \subset \mathbb{R}^{n+1}$, from equation (2.2) and the divergence theorem, Theorem A.2, we compute

$$\begin{aligned}
\varepsilon_R &= \int_{\partial(M\cap B_R(0))} |A|\langle\nabla H, \eta_R\rangle \exp\left(-\frac{R^2}{2}\right)\,d\mu^{n-1}\\
&= \int_{M\cap B_R(0)} |A|\Delta H \exp\left(-\frac{|x|^2}{2}\right) + \left\langle \nabla\left(|A|\exp\left(-\frac{|x|^2}{2}\right)\right), \nabla H\right\rangle\,d\mu^n\\
&= \int_{M\cap B_R(0)} \left(|A|H\left(1-|A|^2\right) + |A|\langle x, \nabla H\rangle\right)\exp\left(-\frac{|x|^2}{2}\right)\,d\mu^n\\
&\quad+ \int_{M\cap B_R(0)} \left(\frac{1}{2|A|}\langle\nabla|A|^2, \nabla H\rangle - |A|\langle x, \nabla H\rangle\right)\exp\left(-\frac{|x|^2}{2}\right)\,d\mu^n\\
&= \int_{M\cap B_R(0)} \left(|A|H\left(1-|A|^2\right) + \frac{1}{2|A|}\langle\nabla|A|^2, \nabla H\rangle\right)\exp\left(-\frac{|x|^2}{2}\right)\,d\mu^n
\end{aligned}$$

and similarly

$$
\begin{aligned}
\delta_R &= \int_{\partial(M\cap B_R(0))} \frac{H}{|A|}\langle \nabla |A|^2, \eta_R\rangle \exp\Big(-\frac{R^2}{2}\Big)\, d\mu^{n-1}\\
&= \int_{M\cap B_R(0)} \frac{H}{|A|}\Delta |A|^2 \exp\Big(-\frac{|x|^2}{2}\Big) + \Big\langle \nabla\Big(\frac{H}{|A|}\exp\Big(-\frac{|x|^2}{2}\Big)\Big), \nabla |A|^2\Big\rangle\, d\mu^n\\
&= \int_{M\cap B_R(0)} \Big(2|A|H\left(1-|A|^2\right) + \frac{2H|\nabla A|^2}{|A|}\\
&\qquad\qquad + \frac{H}{|A|}\langle x, \nabla |A|^2\rangle\Big) \exp\Big(-\frac{|x|^2}{2}\Big)\, d\mu^n\\
&\quad + \int_{M\cap B_R(0)} \Big(\frac{\langle \nabla H, \nabla |A|^2\rangle}{|A|} - \frac{H|\nabla |A|^2|^2}{2|A|^3} - \frac{H}{|A|}\langle x, \nabla |A|^2\rangle\Big) \exp\Big(-\frac{|x|^2}{2}\Big)\, d\mu^n\\
&= \int_{M\cap B_R(0)} \Big(2|A|H\left(1-|A|^2\right) + \frac{2H|\nabla A|^2}{|A|} + \frac{\langle \nabla H, \nabla |A|^2\rangle}{|A|}\\
&\qquad\qquad - \frac{H|\nabla |A|^2|^2}{2|A|^3}\Big) \exp\Big(-\frac{|x|^2}{2}\Big)\, d\mu^n\,.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\sigma_R &= 2\delta_R - 4\varepsilon_R\\
&= \int_{M\cap B_R(0)} \Big(\frac{4H|\nabla A|^2}{|A|} - \frac{H|\nabla |A|^2|^2}{|A|^3}\Big) \exp\Big(-\frac{|x|^2}{2}\Big)\, d\mu^n\\
&= \int_{M\cap B_R(0)} \left(4|A|^2|\nabla A|^2 - \big|\nabla |A|^2\big|^2\right) \frac{H}{|A|^3} \exp\Big(-\frac{|x|^2}{2}\Big)\, d\mu^n\,.
\end{aligned}
$$

As we have

$$
4|A|^2|\nabla A|^2 \geq |\nabla |A|^2|^2
$$

the quantity σ_R is nonnegative and nondecreasing in R. If now we show that

$$
\liminf_{R\to\infty} \sigma_R = 0
$$

we can conclude that, at every point of M,

$$
4|A|^2|\nabla A|^2 = |\nabla |A|^2|^2\,. \tag{2.3}
$$

Getting back to the definitions of ε_R and δ_R, we have

$$
\begin{aligned}
|\sigma_R| &= \Big| -2\int_{\partial(M\cap B_R(0))} \frac{H}{|A|}\langle \nabla |A|^2, \eta\rangle \exp\Big(-\frac{R^2}{2}\Big)\, d\mu^{n-1}\\
&\qquad + 4\int_{\partial(M\cap B_R(0))} |A|\langle \nabla H, \eta\rangle \exp\Big(-\frac{R^2}{2}\Big)\, d\mu^{n-1}\Big|\\
&\leq 4\exp\Big(-\frac{R^2}{2}\Big)\int_{\partial(M\cap B_R(0))} \Big(\frac{H}{|A|}|\nabla |A|^2| + |A||\nabla H|\Big)\, d\mu^{n-1}\\
&\leq 8\exp\Big(-\frac{R^2}{2}\Big)\int_{\partial(M\cap B_R(0))} \left(H|\nabla A| + |A||\nabla H|\right)\, d\mu^{n-1}\\
&\leq C\exp\Big(-\frac{R^2}{2}\Big)\, \mu^{n-1}(\partial(M\cap B_R(0)))\,,
\end{aligned}
$$

by the estimates on A and ∇A in the hypotheses. Assume that the lefthand side does not go to zero. That is, suppose that for every R belonging to the set $\mathcal{R}\subset\mathbb{R}^+$ (which is of full measure) and R larger than some $R_0>0$ we have

$$
\mu^{n-1}(\partial(M\cap B_R(0))) \geq \delta\exp\Big(\frac{R^2}{2}\Big) \geq \delta R\exp\Big(\frac{R^2}{4}\Big)
$$

for some constant $\delta > 0$. Recall the area formula and divergence theorem, Theorems A.1 and A.2. As the function

$$R \mapsto \mu^n(M \cap B_R(0))$$

is monotone and continuous from the left and actually continuous at every value $R \in \mathcal{R}$, we can differentiate it almost everywhere in $\mathbb{R}^+$ and we have, for $R_0 < r < R$,

$$\begin{aligned}
\mu^n(M \cap B_R(0)) - \mu^n(M \cap B_r(0)) &= \int_r^R \frac{d}{d\xi}\mu^n(M \cap B_\xi(0))\, d\xi \\
&= \int_r^R \int_{M\cap B_\xi(0)} \operatorname{div}_{M\cap B_\xi(0)} \eta_\xi \, d\mu^{n-1}\, d\xi \\
&= -\int_r^R \int_{M\cap B_\xi(0)} \langle \eta_\xi, \mathbf{H}_{M\cap B_\xi(0)}\rangle \, d\mu^{n-1}\, d\xi \\
&\quad + \int_r^R \int_{\partial(M\cap B_\xi(0))} \langle \eta_\xi, \eta_\xi\rangle \, d\mu^{n-1}\, d\xi \\
&= \int_r^R \int_{\partial(M\cap B_\xi(0))} d\mu^{n-1}\, d\xi \\
&\geq \delta \int_r^R \xi \exp\left(\frac{\xi^2}{4}\right) d\xi = 2\delta\left(\exp\left(\frac{R^2}{4}\right) - \exp\left(\frac{r^2}{4}\right)\right) .
\end{aligned}$$

Then

$$\mu^n(M \cap B_R(0))e^{-R} \to \infty\,,$$

for $R \to \infty$, in contradiction with the hypotheses of the theorem. Hence, the

$$\liminf_{R\to\infty, R\in\mathcal{R}} \exp\left(-\frac{R^2}{2}\right) \mu^{n-1}(\partial(M \cap B_R(0))) = 0\,.$$

It follows that the same holds for $|\sigma_R|$ and equation (2.3) is proved. By Cauchy–Schwarz,

$$4|A|^2|\nabla A|^2 = |\nabla |A|^2|^2 = 4|A\nabla A|^2 \leq 4|A|^2|\nabla A|^2$$

or in coordinates

$$\begin{aligned}
4h^i_j h^j_i \nabla_k h^m_n \nabla^k h^n_m &= \nabla_k(h^i_j h^j_i)\nabla^k(h^m_n h^n_m) \\
&= 4h^i_j h^m_n \nabla_k h^j_i \nabla^k h^n_m \leq 4h^i_j h^j_i \nabla_k h^m_n \nabla^k h^n_m
\end{aligned}$$

with equality if and only if A and ∇A are linearly dependent, that is, at every point there exist constants c_k such that

$$\nabla_k h_{ij} = c_k h_{ij}$$

for every i, j. Contracting this equation with the metric g^{ij} and with h^{ij} we get

$$\nabla_k H = c_k H \quad \text{ and } \quad \nabla_k |A|^2 = 2c_k|A|^2\,,$$

hence

$$\nabla_k \log H = c_k \quad \text{ and } \quad \nabla_k \log |A|^2 = 2c_k\,.$$

This implies

$$\nabla_k \log\left(\frac{H}{|A|}\right) = 0 \quad \text{ so that } \quad |A| = \alpha H$$

for some constant $\alpha > 0$. By connectedness this relation has to hold globally on M. Suppose now that at a point $|\nabla H| \neq 0$, then

$$\nabla_k h_{ij} = c_k h_{ij} = \frac{\nabla_k H}{H} h_{ij} \tag{2.4}$$

which is a symmetric 3-tensor by the Codazzi equations, hence

$$h_{ij}\nabla_k H = h_{ik}\nabla_j H$$

at one point, where the Christoffel symbols vanish. Computing then in normal coordinates with an orthonormal basis $\{\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_n\}$ such that $\boldsymbol{\tau}_1 = \nabla H/|\nabla H|$, we have with $g^{ij} = \delta^{ij}$,

$$\begin{aligned}
0 &= |h_{ij}\nabla_k H - h_{ik}\nabla_j H|^2 \\
&= (h_{ij}\nabla_k H - h_{ik}\nabla_j H)g^{il}g^{jm}g^{kn}(h_{lm}\nabla_n H - h_{ln}\nabla_m H) \\
&= 2|\nabla H|^2|A|^2 - 2g^{il}g^{jm}g^{kn}h_{ij}h_{ln}\nabla_k H\nabla_m H \\
&= 2|\nabla H|^2|A|^2 - 2g^{il}h_i^m h_l^k \nabla_k H\nabla_m H \\
&= 2|\nabla H|^2|A|^2 - 2g^{il}h_i^1 h_l^1 \nabla_1 H\nabla_1 H \\
&= 2|\nabla H|^2 \left(|A|^2 - \sum_{i=1}^n \left(h_i^1\right)^2\right).
\end{aligned}$$

Hence, $|A|^2 = \sum_{i=1}^n \left(h_i^1\right)^2$ and

$$|A|^2 = \left(h_1^1\right)^2 + 2\sum_{i=2}^n \left(h_i^1\right)^2 + \sum_{i,j\neq 1}^n \left(h_i^j\right)^2$$

so $h_j^i = 0$ unless $i = j = 1$, which means that A has rank one. Thus, we have two possible (not mutually excluding) situations at every point of M, either A has rank one or $\nabla H = 0$.

If $\ker A \equiv \emptyset$ on M, A must have rank at least two as we assumed $n \geq 2$, then we have $\nabla H = 0$ which implies $\nabla A = 0$ and

$$h_{ij} = Hh_{ik}h_j^k = Hh_{ik}g^{kl}h_{lj}$$

by equation (2.1). This means that for an eigenvalue λ_m with eigenvector ξ_m,

$$h_{ij}\xi_m^j = Hh_{ik}g^{kl}h_{lj}\xi_m^j = Hh_{ik}g^{kl}\lambda_m g_{lj}\xi_m^j = \lambda_m H h_{ij}\xi_m^j$$

so that all the eigenvalues of A are 0 or $1/H$. As the kernel is empty

$$H = \sum_{i=1}^n \lambda_m = \frac{n}{H}$$

so that

$$H = \sqrt{n} \quad \text{and} \quad h_{ij} = \frac{g_{ij}}{\sqrt{n}}.$$

Then, the complete hypersurface M has to be the sphere $\mathbb{S}^n_{\sqrt{n}}$, indeed we compute

$$\begin{aligned}
\Delta|x|^2 = \Delta|x|^2 &= 2\nabla\langle x, \nabla x\rangle = 2n + 2\langle x, \Delta x\rangle \\
&= 2n - 2H\langle x, \boldsymbol{\nu}\rangle = 2n - 2H^2 = 0\,,
\end{aligned}$$

by means of the structural equation $H = \langle x, \boldsymbol{\nu}\rangle$. Hence, $|x|^2$ is a harmonic function on M. Looking at the point of M of minimum distance from the origin, by the strong maximum principle for elliptic equations, Theorem D.1, it must be constant on M and $M = \mathbb{S}^n_{\sqrt{n}}$.

Let now $\ker A(x) \neq \emptyset$ at some point $x \in M$, with $\dim\ker A(x) = (n-m)$ and $0 < m < n$ (as A is never zero), and let

$$v_1(x), \dots, v_{n-m}(x) \in T_xM \subset \mathbb{R}^{n+1}$$

be a family of unit orthonormal tangent vectors spanning $\ker A(x)$, that is,

$$h_{ij}(x)v_k^j(x) = 0$$

for $k = 1, \dots, n-m$. By (2.4), the geodesic $\gamma(s)$ from $x \in M$ (M is complete) with initial velocity $\partial_s\gamma(0) = v_k(x)$ satisfies

$$\nabla_{\partial_s\gamma}(h_{ij}\partial_s\gamma^j) = \frac{\langle\nabla H, \partial_s\gamma\rangle}{H}h_{ij}\partial_s\gamma^j$$

hence, by Gronwall's lemma there holds

$$h_{ij}(\gamma(s))\partial_s\gamma^j(s) = h_{ij}(\gamma(0))\partial_s\gamma^j(0)\exp\left(\int_0^s \frac{\langle \nabla H, \partial_\sigma\gamma\rangle}{H}\,d\sigma\right) = 0$$

for every $s \in \mathbb{R}$. Since γ is a geodesic in M, $\partial_s^2\gamma(s) \in (T_{\gamma(s)}M)^\perp$, that is, the normal to the curve in $\mathbb{R}^{n+1}$ is also the normal to M, then letting κ be the curvature of γ in $\mathbb{R}^{n+1}$, we have

$$\kappa = -\langle \boldsymbol{\nu}_M, \partial_s^2\gamma\rangle = h_{ij}\partial_s\gamma^i\partial_s\gamma^j = 0\,,$$

thus γ is a straight line in $\mathbb{R}^{n+1}$ and

$$x + \ker A(x) \subset M\,,$$

where $x + \ker A(x) \subset \mathbb{R}^{n+1}$ is an $(n-m)$-dimensional affine subspace. Let now $\sigma(s)$ be a geodesic from x to another point y parametrized by arclength and extend by parallel transport the vectors $v_k(x)$, $k = 1, \dots, n-m$, along σ, then

$$\nabla_{\partial_s\sigma}(h_{ij}v_k^j) = \frac{\langle \nabla H, \partial_s\sigma\rangle}{H}h_{ij}v_k^j$$

and again by Gronwall's lemma it follows that $h_{ij}(\gamma(s))v_k^j(\gamma(s)) = 0$ for every $s \in \mathbb{R}$ and $k = 1, \dots, n-m$, in particular $v_k(y) \in \ker A(y)$. Hence,

$$\dim\ker A \equiv n - m$$

on M with $0 < m < n$ (as A is never zero) and all the affine $(n-m)$-dimensional subspaces $x + \ker A(x) \subset \mathbb{R}^{n+1}$ are contained in M for every $x \in M$, that is,

$$M + \ker(M) \subset M\,.$$

Moreover, as $h_{ij}v_k^j = 0$ along the geodesic σ, we have

$$D^{\mathbb{R}^{n+1}}_{\partial_s\sigma} v_k = \nabla_{\partial_s\sigma}v_k + \langle \nabla_{\partial_s\sigma}v_k, \boldsymbol{\nu}_M\rangle\boldsymbol{\nu}_M = -h_{ij}v_k^j\partial_s\sigma^i\boldsymbol{\nu}_M = 0\,,$$

so the extended vectors v_k are constant in $\mathbb{R}^{n+1}$, which means that the parallel extension is independent of the geodesic σ, that the subspaces $\ker A(x)$ are all a common $(n-m)$-dimensional vector subspace of $\mathbb{R}^{n+1}$ and

$$M = M + \ker A\,.$$

Let $x \in M$. Then there exists $y \in M \cap (\ker A)^\perp$ and $v \in \ker A$ so that

$$x = y + v\,.$$

Define $f : M \to \ker A$ by

$$f(x) = v\,.$$

By Sard's theorem, Corollary C.3, there exists a vector $v \in \ker A$ such that

$$N(v) := f^{-1}(v) = M \cap \big(v + (\ker A)^\perp\big)$$

is a smooth, complete m-dimensional submanifold of $\mathbb{R}^{n+1}$. Since $M = M + \ker A$, $N(v) = N(w)$ for all $v, w \in \ker A$ and

$$M = N \times \ker A\,.$$

This implies that

$$L := N(0) = M \cap (\ker A)^\perp$$

is a smooth, complete m-dimensional submanifold of $(\ker A)^\perp = \mathbb{R}^{m+1}$ with

$$M = L \times \ker A\,.$$

Moreover, as $\ker A$ is in the tangent space to every point of L, the normal $\boldsymbol{\nu}_M$ to M at a point of L stays in $(\ker A)^\perp$ so it must coincide with the normal $\boldsymbol{\nu}_L$ to L in $(\ker A)^\perp$, then a simple computation shows that the mean curvature H_M of M at the points of L is equal to the mean curvature H_L of L as a hypersurface of $(\ker A)^\perp = \mathbb{R}^{m+1}$. This shows that L is a hypersurface in $\mathbb{R}^{m+1}$ satisfying $H_L(z) = \langle z, \boldsymbol{\nu}_L(z)\rangle$ for every $z \in L$. Finally, as by construction the second fundamental form of L has empty kernel, by the previous discussion we have $L = \mathbb{S}^m_{\sqrt{m}}$ and $M = \mathbb{S}^m_{\sqrt{m}} \times \mathbb{R}^{n-m}$ which proves the claim. $\square$

Theorem 2.5 (Colding–Minicozzi, [CM12, Theorem 10.1])**.** *If M^n, for $n \geq 2$, is an embedded hypersurface in $\mathbb{R}^{n+1}$, with non-negative mean curvature, satisfying $H = \langle x, \boldsymbol{\nu} \rangle/2$, then M^n is of the form $\mathbb{S}^m_{\sqrt{2m}} \times \mathbb{R}^{n-m}$ for $m = 0, 1, \dots, n$.*

2.2. **Curves.**

Theorem 2.6 (Abresch–Langer, [AL86])**.** *Let $\Sigma \subset \mathbb{R}^2$ be a smooth, complete, embedded curve satisfying $\kappa(x) = \langle x, \boldsymbol{\nu}(x) \rangle/2$ for every $x \in \Sigma$. Then Σ is either the line through the origin or the $\mathbb{S}^1_{\sqrt{2}}$.*

Proof. See [Man11, Proposition 3.4.1]. We scale the curve by the factor $1/2$ so that $\kappa = \langle x, \boldsymbol{\nu} \rangle$ for every $x \in \Sigma$. Fixing a reference point on a curve $\Sigma = X(I)$, $I \in \{\mathbb{S}^1, \mathbb{R}\}$, we have an arclength parameter s which gives a unit tangent vectorfield $\boldsymbol{\tau} = \partial_s X$ and a unit normal vectorfield $\boldsymbol{\nu} = (\boldsymbol{\tau}_2, -\boldsymbol{\tau}_1)$, which is the clockwise rotation of $\pi/2$ in $\mathbb{R}^2$ of the vector $\boldsymbol{\tau}$. Then the curvature is given by

$$\kappa = -\langle \partial_s \boldsymbol{\tau}, \boldsymbol{\nu} \rangle = \langle \boldsymbol{\tau}, \partial_s \boldsymbol{\nu} \rangle$$

so that

$$\partial_s \boldsymbol{\nu} = \kappa \boldsymbol{\tau} \quad \text{ and } \quad \partial_s \boldsymbol{\tau} = -\kappa \boldsymbol{\nu} \,.$$

The relation $\kappa = \langle x, \boldsymbol{\nu} \rangle$ implies the ODE for the curvature

$$\partial_s \kappa = \langle \boldsymbol{\tau}, \boldsymbol{\nu} \rangle + \langle x, \partial_s \boldsymbol{\nu} \rangle = \kappa \langle x, \boldsymbol{\tau} \rangle \,.$$

Suppose that at some point $\kappa = 0$, then also $\partial_s \kappa = 0$ at the same point. Hence, by the uniqueness theorem for ODE's we conclude that κ is identically zero so that Σ is a line. Since $\langle x, \boldsymbol{\nu} \rangle = 0$ for every $x \in \Sigma$, we conclude that $0 \in \Sigma$. So we suppose that κ is always nonzero and possibly reversing the orientation of the curve, we assume that $\kappa > 0$ at every point, that is, the curve is strictly convex. Computing the derivative of $|X|^2$,

$$\partial_s |X|^2 = 2\langle X, \boldsymbol{\tau} \rangle = 2\frac{\partial_s \kappa}{\kappa} = 2\partial_s \log \kappa$$

we get

$$\kappa = C \exp\left(\frac{|x|^2}{2} \right)$$

for some constant $C > 0$. Hence, κ is bounded from below by $C > 0$. Since Σ is convex, we can consider the coordinate $\vartheta = \arccos\langle e_1, \boldsymbol{\nu} \rangle$. (Note that ϑ is only locally continuous and jumps after a complete round). We have $\partial_s \vartheta = \kappa$ as well as

$$\partial_\vartheta \kappa = \frac{\partial_s \kappa}{\kappa} = \langle x, \boldsymbol{\tau} \rangle \quad \text{ and } \quad \partial^2_\vartheta \kappa = \frac{\partial_s \partial_\vartheta \kappa}{\kappa} = \frac{1 - \kappa \langle x, \boldsymbol{\nu} \rangle}{\kappa} = \frac{1}{\kappa} - \kappa \,. \tag{2.5}$$

Multiplying both sides of the last equation by $2\partial_\vartheta \kappa$ we get

$$0 = 2\partial_\vartheta \kappa \partial^2_\vartheta \kappa + 2\kappa \partial_\vartheta \kappa - \frac{2\partial_\vartheta \kappa}{\kappa} = \partial_\vartheta \left((\partial_\vartheta \kappa)^2 + \kappa^2 - \log \kappa^2 \right) \,,$$

so that,

$$(\partial_\vartheta \kappa)^2 + \kappa^2 - \log \kappa^2 \equiv E \geq 1$$

along all the curve. We have $E = 1$ if and only if $\kappa^2 \equiv 1$ along the curve, which is the unit circle centered at the origin of $\mathbb{R}^2$. When $E > 1$, it follows that κ is uniformly bounded from above, hence recalling that $\kappa = C \exp(|x|^2/2)$,

$$\Sigma \subset B_R(0)$$

for some $R > 0$ and by the embeddedness and completeness hypotheses, Σ must be closed, simple and strictly convex, as $\kappa > 0$ at every point.

Suppose that Σ is not a line. We follow the lines of [GH86, Lemma 5.7.9] and [Pih98, Lemma 7.23]. The system

$$\left\{1, \sqrt{2}\cos(n\vartheta), \sqrt{2}\sin(n\vartheta)\right\}_{n\in\mathbb{Z}} \tag{2.6}$$

forms an orthonormal basis of the periodic functions in the Hilbert space $C^2([0,2\pi])$ with respect to the L^2-inner product (see e.g. [HL99, p. 124]). We have $ds_t = d\vartheta/\kappa$ so that

$$\int_{\mathbb{S}^1} \frac{\sin(\vartheta)}{\kappa}\, d\vartheta = \int_{\mathbb{S}^1_{R_t}} \sin\left(\frac{s}{R_t}\right) ds_t = \cos(2\pi) - \cos(0) = 1 - 1 = 0$$

and

$$\int_{\mathbb{S}^1} \frac{\cos(\vartheta)}{\kappa}\, d\vartheta = \int_{\mathbb{S}^1_{R_t}} \cos\left(\frac{s}{R_t}\right) ds_t = \sin(2\pi) - \sin(0) = 0\,.$$

Furthermore, integration by parts yields

$$\begin{aligned}0 &= \int_{\mathbb{S}^1} \frac{\sin(\vartheta)}{\kappa}\, d\vartheta \int_{\mathbb{S}^1} \frac{1}{\kappa}\frac{\partial \cos}{\partial\vartheta}(\vartheta)\, d\vartheta \\ &= -\int_{\mathbb{S}^1} \cos(\vartheta)\frac{\partial}{\partial\vartheta}\left(\frac{1}{\kappa}\right) d\vartheta = \int_{\mathbb{S}^1} \cos(\vartheta)\frac{1}{\kappa^2}\frac{\partial\kappa}{\partial\vartheta}\, d\vartheta\end{aligned}$$

and

$$\begin{aligned}0 &= -\int_{\mathbb{S}^1} \frac{\cos(\vartheta)}{\kappa}\, d\vartheta = \int_{\mathbb{S}^1} \frac{1}{\kappa}\frac{\partial \sin}{\partial\vartheta}(\vartheta)\, d\vartheta \\ &= -\int_{\mathbb{S}^1} \sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\frac{1}{\kappa}\right) d\vartheta = \int_{\mathbb{S}^1} \sin(\vartheta)\frac{1}{\kappa^2}\frac{\partial\kappa}{\partial\vartheta}\, d\vartheta\,.\end{aligned}$$

Additionally, we have

$$0 = -\int_{\mathbb{S}^1} \frac{\partial}{\partial\vartheta}\left(\frac{1}{\kappa}\right) d\vartheta = \int_{\mathbb{S}^1} \frac{1}{\kappa^2}\frac{\partial\kappa}{\partial\vartheta}\, d\vartheta\,.$$

Hence, $1/\kappa^2 \frac{\partial}{\partial\vartheta}\kappa$ is orthogonal to the first five basis functions of the basis (2.6). Since all the other basis functions are zero at at least four points in $[0,2\pi]$ with distance $\leq \pi/2$, there exists a number $i_0 \geq 4$ and points $\vartheta_i \in \mathbb{S}^1$, $i \in \{0,\dots,i_0\}$, so that

$$\left(\frac{1}{\kappa^2}\frac{\partial\kappa}{\partial\vartheta}\right)(\vartheta_i,\tau) = 0$$

and

$$|\vartheta_i - \vartheta_{i+1}| \leq \frac{\pi}{2}$$

for $i \in \{0,\dots,i_0-1\}$ and

$$|\vartheta_{i_0} - (2\pi + \vartheta_0)| \leq \frac{\pi}{2}\,.$$

Since $1/\kappa^2 \frac{\partial}{\partial\vartheta}\kappa$ is periodic on $[0,2\pi]$, i_0 is odd. Define the intervals

$$I_i := [\vartheta_i, \vartheta_{i+1}]$$

for $i \in \{0,\dots,i_0-1\}$ and

$$I_{i_0} := [0,\vartheta_0] \cup [\vartheta_{i_0}, 2\pi)\,.$$

Then $|I_i| \leq \pi/2$ for all $i \in \{1,\dots,i_0\}$. Since $\partial_\vartheta^2\kappa = 1/\kappa - \kappa$, it holds that $\partial_\vartheta^2\kappa \neq 0$ when $\partial_\vartheta\kappa = 0$, otherwise this second-order ODE for κ would imply $\partial_\vartheta\kappa = 0$ everywhere, hence $\kappa = 1$ identically and we would be in the case of the unit circle. Suppose that Σ is neither a line nor a circle. By looking at the equation for the curvature (2.5) we can see easily that $\kappa < 1$ at a local minimum and $\kappa > 1$ at a local maximum. Suppose now that $\kappa(0)$ is a local maximum and $\kappa(\vartheta_0)$ is the first subsequent critical value for κ for $\vartheta_0 \leq \pi/2$ by the above. Then the curvature is strictly decreasing in the interval $[0, \vartheta_0]$. Also $\kappa(\vartheta_0) < 1$ must be a local minimum of the curvature, as every critical point is not degenerate. By a straightforward computation, starting by differentiating the equation $\partial^2_\vartheta \kappa = 1/\kappa - \kappa$, one gets

$$\begin{aligned}\partial^3_\vartheta \kappa^2 &= 2\partial^2_\vartheta(\kappa\partial_\vartheta\kappa) = 2\partial_\vartheta(\partial_\vartheta\kappa)^2 + 2\partial_\vartheta(\kappa\partial^2_\vartheta\kappa) = 6\partial_\vartheta\kappa\partial^2_\vartheta\kappa + 2\kappa\partial_\vartheta\partial^2_\vartheta\kappa\\ &= 6\frac{\partial_\vartheta\kappa}{\kappa} - 6\kappa\partial_\vartheta\kappa - 2\frac{\kappa}{\kappa^2}\partial_\vartheta\kappa - 2\kappa\partial_\vartheta\kappa = 4\frac{\partial_\vartheta\kappa}{\kappa} - 4\partial_\vartheta\kappa^2\end{aligned}$$

so that

$$\partial^3_\vartheta\kappa^2 + 4\partial_\vartheta\kappa^2 = 4\frac{\partial_\vartheta\kappa}{\kappa}\,.$$

We compute

$$\begin{aligned}4\int_0^{\vartheta_0} &\sin(2\vartheta)\frac{\partial_\vartheta\kappa}{\kappa}\,d\vartheta = \int_0^{\vartheta_0}\sin(2\vartheta)\left(\partial^3_\vartheta\kappa^2 + 4\partial_\vartheta\kappa^2\right)\,d\vartheta\\
&= \sin(2\vartheta)\partial^2_\vartheta\kappa^2\big|_0^{\vartheta_0} - 2\int_0^{\vartheta_0}\cos(2\vartheta)\partial^2_\vartheta\kappa^2\,d\vartheta + 4\int_0^{\vartheta_0}\sin(2\vartheta)\partial_\vartheta\kappa^2\,d\vartheta\\
&= 2\sin(2\vartheta_0)\left(\kappa(\vartheta_0)\partial^2_\vartheta\kappa(\vartheta_0) + (\partial_\vartheta\kappa)^2(\vartheta_0)\right) - 2\cos(2\vartheta)\partial_\vartheta\kappa^2\big|_0^{\vartheta_0}\\
&\quad - 4\int_0^{\vartheta_0}\sin(2\vartheta)\partial_\vartheta\kappa^2\,d\vartheta + 4\int_0^{\vartheta_0}\sin(2\vartheta)\partial_\vartheta\kappa^2\,d\vartheta\\
&= 2\sin(2\vartheta_0)\left(\kappa(\vartheta_0)\partial^2_\vartheta\kappa(\vartheta_0) + (\partial_\vartheta\kappa)^2(\vartheta_0)\right)\\
&\quad - 4\cos(2\vartheta_0)\kappa(\vartheta_0)\partial_\vartheta\kappa(\vartheta_0) + 4\kappa(0)\partial_\vartheta\kappa(0)\,.\end{aligned}$$

Now, since $\partial_\vartheta\kappa(0) = \partial_\vartheta\kappa(\vartheta_0) = 0$ using the equation for the curvature $\partial^2_\vartheta\kappa = 1/\kappa - \kappa$ we get

$$4\int_0^{\vartheta_0}\sin(2\vartheta)\frac{\partial_\vartheta\kappa}{\kappa}\,d\vartheta = 2\sin(2\vartheta_0)(1-\kappa^2(\vartheta_0))\,,$$

and this last term is nonnegative as $\kappa < 1$ at a local minimum and $0 < 2\vartheta_0 \leq \pi$. Looking at the left-hand integral we see instead that the factor $\sin(2\vartheta)$ is always nonnegative, since $2\vartheta_0 \leq \pi$ and $\partial_\vartheta\kappa$ is always nonpositive in the interval $[0, \vartheta_0]$, as we assumed that we were moving from a local maximum of κ at 0 to a local minimum of κ at ϑ_0 without crossing any other critical point of κ. This gives a contradiction so Σ must be the unit circle. □

3. Convex hypersurfaces with pinched second fundamental form

Definition 3.1 (Complete Riemannian manifold)**.** A *(geodesically) complete* manifold is a Riemannian manifold for which every maximal (inextendible) geodesic is defined on $\mathbb{R}$.

Definition 3.2 (Conformal map)**.** Two maps $X, Y : M^n \to \mathbb{R}^{n+1}$ are *conformal*, if there exists $\lambda : M^n \to \mathbb{R}$ with

$$g^X_{ij} = \lambda\, g^Y_{ij}\,.$$

We say X is *quasi-conformal* with respect to Y if

$$g^X_{ij} \geq \lambda\, g^Y_{ij}\,.$$

See [Ham94]. Suppose that $M = X(M^n) \subset \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ is written as a graph over a convex over a convex open set $U \subset \mathbb{R}^n$ of a strictly convex function

$$y = f(x_1, \ldots, x_n)$$

so that $y \to \infty$ as $x = (x_1, \ldots, x_n) \to \partial U$. By translating upwards if necessary, since y is bounded below, we can assume $y \geq e$ everywhere, so that $\log\log y \geq 0$. Let g_{ij} be the Riemannian metric induced on M so that

$$g_{ij} = \delta_{ij} + \frac{\partial y}{\partial x^i}\frac{\partial y}{\partial x^j}\,.$$

Theorem 3.3 (Hamilton, [Ham94, Theorem 2.1])**.** *The conformally equivalent metric*

$$\tilde{g}_{ij} = \frac{g_{ij}}{(y \log y)^2}$$

is complete with finite volume.

Proof. First, we show that $\tilde{g}_{ij}$ is complete. We have $\det(g_{ij}) \geq 1$. For any geodesic $\gamma : I \to M$ going to infinity, we have $\gamma^n \to \infty$. Therefore its length satisfies,

$$\begin{aligned}\tilde{L}(\gamma) = \int_I |\gamma'(t)|_{\tilde{g}}\, dt &\geq \int_{\gamma^n(a)}^{\infty} \sqrt{\det(\tilde{g}_{ij})}dy \\ &\geq \int_{\gamma^n(a)}^{\infty} \frac{dy}{y \log y} = \log\log y|_{\gamma^n(a)}^{\infty} = \infty\,.\end{aligned}$$

Since geodesics have constant speed, this is what we desired. To estimate the volume, we observe that, because y is a strictly convex function of x, outside a compact set we must have

$$\left|\frac{\partial y}{\partial x^i}\right| \geq \delta$$

for some $\delta > 0$ and at least one $i \in \{1, \dots, n\}$. Let dV denote the volume element on M in the induced metric g_{ij}, which in x coordinates is

$$dV = \sqrt{\det\left(\delta_{ij} + \frac{\partial y}{\partial x^i}\frac{\partial y}{\partial x^j}\right)} dx^1 \dots dx^n\,.$$

Let $k \in \mathbb{N}$ and

$$M^k := M \cap \{e + k - 1 \leq y \leq e + k\}$$

and let dV^k denote the volume element of the part of M^k. We can devide M^k into pieces $M_1^k, \dots, M_n^k$, where $\frac{\partial y}{\partial x^i}$ is largest on M_i^k, and estimate dV_i^k from above on each piece. For each $k \in \mathbb{N}$, on M_i^k, we take $x^1, \dots, x^{i-1}, y, x^{i+1}, \dots, x^n$ as coordinates. Since $\frac{\partial y}{\partial x^i}$ is larger than the other derivatives, and $\left|\frac{\partial y}{\partial x^i}\right| \geq \delta > 0$,

$$\sqrt{\det\left(\delta_{ij} + \frac{\partial y}{\partial x^i}\frac{\partial y}{\partial x^j}\right)} \leq C \left|\frac{\partial y}{\partial x^i}\right|$$

and thus

$$dV_i^k \leq C dx^1 \dots dx^{i-1} dy dx^{i+1} \dots dx^n$$

on M_i^k. By the gradient estimate shows that

$$|x| \leq Cy$$

for a suitable large constant. Let

$$U_i^k := \left\{x \in \mathbb{R}^n \;\middle|\; (x, f(x)) \in M_i^k\right\}\,.$$

We can integrate in every direction $x^1, \dots, x^{i-1}, x^{i+1}, \dots x^n$ and estmate

$$\int_{U_i^k} dV_i^k \leq C \int_{U_i^k} dx^1 \dots dx^{i-1} dy dx^{i+1} \dots dx^n \leq C \int_{U_i^k} y^{n-1} dy\,,$$

that is,

$$dV_y^i \leq C y^{n-1} dy\,.$$

Hence,

$$d\tilde{V}_y^i \leq \frac{C dy}{y \log^n y}$$

and

$$
\begin{aligned}
\tilde{V} &= \int_U d\tilde{V} = \sum_{k\in\mathbb{N}}\sum_{i=1}^{n}\int_{U_i^k} d\tilde{V}_i^k \le C\sum_{k\in\mathbb{N}}\sum_{i=1}^{n}\int_{U_i^k}\frac{dy}{y\log^n y}\\
&= C\int_e^\infty \frac{dy}{y\log^n y} = \frac{-C}{(n-1)\log^{n-1} y}\bigg|_e^\infty = \frac{C}{n-1} < \infty\,. \qquad \square
\end{aligned}
$$

Remark 3.4. (i) Let $p, q \in \mathbb{S}^n$. We rotate the sphere so that the north pole N lies on the geodesic between p and q with equal distance to both points. The stereographic projection $\varphi : \mathbb{S}^n \setminus \{N\} \to \mathbb{R}^n$, which is conformal, projects the sphere to the plane. We can choose the projection such that $\varphi(p), \varphi(q) \in \{x^n = 0\}$. By construction, $|\varphi(p)| = |\varphi(q)| = r$. Via the inverse stereographic projection $\psi : \mathbb{R}^n \to \mathbb{S}^n_r \setminus \{N\}$ we can conformally project the plane to the sphere of radius r. The points $\varphi(p)$ and $\varphi(q)$ are mapped antipodally to the equator. Hence, $1/r \circ \psi \circ \varphi : \mathbb{S}^n \setminus \{N\} \to \mathbb{S}^n \setminus \{N\}$ is a conformal map that, after rotation, maps p to the north pole and q to the south pole.

(ii) Let X be an embedding of the $\mathbb{S}^{n-1}$, Y be an embedding of the $\mathbb{S}^n$ and Z be an embedding of the cylinder $\mathbb{S}^{n-1} \times [-R, R]$, where

$$Y(x, \vartheta) = (X(x)\cos(\vartheta), \sin(\vartheta))$$

and

$$Z(x, \vartheta) = (X(x), z(\vartheta))$$

for $\vartheta \in [-\pi/2, \pi/2)$. Then

$$\left(g_{ij}^Y\right) = \begin{pmatrix} \cos^2(\vartheta)\, g_{ij}^X & 0\\ 0 & 1\end{pmatrix}$$

and

$$\left(g_{ij}^Z\right) = \begin{pmatrix} g_{ij}^X & 0\\ 0 & (z'(\vartheta))^2\end{pmatrix}.$$

For Y and Z to be conformal with $\left(g_{ij}^Y\right) = \lambda\left(g_{ij}^Z\right)$, we have to choose

$$\lambda(\vartheta) = \cos^2(\vartheta) \qquad \text{and} \qquad z'(\vartheta) = \frac{1}{\cos(\vartheta)}$$

for $\vartheta \in [-\pi/2 + \varepsilon, \pi/2 - \varepsilon]$, where $\varepsilon > 0$ and $R = R(\varepsilon)$, which is realized by

$$z(\vartheta) = \log\left(\tan\left(\frac{\theta}{2} + \frac{\pi}{4}\right)\right).$$

Theorem 3.5 (Hamilton, [Ham94]). *Let U be an open subset of the unit sphere $\mathbb{S}^n$ which is not empty and whose closure is not the whole sphere. Then there is no metric on U, conformal with respect to the round metric, which is complete with finite volume.*

Proof. By hypotheses we can find some point p_N which is contained in U, and some point p_S which avoids the closure of U. By Remark 3.4, we can assume that p_N is the north pole and p_S is the south pole. We can then find an $\varepsilon > 0$ so that the ε-ball around p_N lies in U,

$$B_\varepsilon(p_N) \subset U$$

while the ε-ball around p_S avoids U,

$$B_\varepsilon(p_S) \subset \mathbb{S}^n \setminus U\,.$$

By Remark 3.4, we can find a conformal map φ of the sphere $\mathbb{S}^n$ minus these two balls to the cylinder $\mathbb{S}^{n-1} \times [0, L]$,

$$\varphi : \mathbb{S}^n \to \mathbb{S}^{n-1} \times [0, L]$$

taking the boundary of the ε-ball around p_N to $\mathbb{S}^{n-1} \times \{0\}$

$$\varphi(\partial B_\varepsilon(p_N)) = \mathbb{S}^{n-1} \times \{0\}$$

and the boundary of the ε-ball around p_S to $\mathbb{S}^{n-1} \times \{L\}$,

$$\varphi(\partial B_\varepsilon(p_S)) = \mathbb{S}^{n-1} \times \{L\}\,.$$

The part of U outside the ε-ball around p_N will map to some relatively open subset

$$W := \varphi(U \setminus B_\varepsilon(p_N)) \subset \big(\mathbb{S}^{n-1} \times [0, L]\big) \setminus \big(\mathbb{S}^{n-1} \times \{L\}\big)$$

of the cylinder which contains $\mathbb{S}^{n-1} \times \{0\}$ and avoids $\mathbb{S}^{n-1} \times \{L\}$,

$$\mathbb{S}^{n-1} \times \{0\} \subset W\,.$$

The subset W will be a noncompact manifold with one compact boundary component $\mathbb{S}^{n-1}$. Any complete metric

$$g^U \quad \text{on} \quad U$$

with finite volume conformal to the round metric

$$g^{\mathbb{S}^n} \quad \text{on} \quad \mathbb{S}^n$$

would give a complete metric with finite volume on

$$g^W \quad \text{on} \quad W$$

conformal to the product metric

$$g^{\mathbb{S}^{n-1}\times[0,L]} \quad \text{on} \quad \mathbb{S}^{n-1} \times [0, L]\,.$$

We show that such cannot exist. We introduce coordinates

$$\vartheta = (\vartheta_1, \ldots, \vartheta_{n-1}) \quad \text{on} \quad \mathbb{S}^{n-1}$$

and

$$t \quad \text{on} \quad [0, L]$$

Let $g^{\mathbb{S}^{n-1}}$ denote the metric on $\mathbb{S}^{n-1}$ and $d\mu$ the volume form. Then

$$g := g^{\mathbb{S}^{n-1}\times[0,L]} = \begin{pmatrix} g^{\mathbb{S}^{n-1}} & 0 \\ 0 & 1 \end{pmatrix}$$

is the product metric on $\mathbb{S}^{n-1} \times [0, L]$ and

$$dV = d\mu dt$$

is the product volume form. For every $\vartheta \in \mathbb{S}^{n-1}$, there will be a first point

$$t = h(\vartheta)$$

where the pair (ϑ, t) is no longer in W. Of course h may not be a continuous function and the pair may reenter W for larger values of t. This does not matter. Any quasi-conformally equivalent metric on W is given by

$$\tilde{g} = \lambda(\vartheta, t) g$$

for some funtion λ defined at least for $0 \le t \le h(\vartheta)$. The corresponding volume form is

$$d\tilde{V} = \lambda^n d\mu dt\,.$$

If the total volume $\tilde{V}$ of W in the conformally equivalent metric is finite, we have

$$\iint_W \lambda^n d\mu dt = \tilde{V} < \infty\,.$$

By Hölder's inequality

$$\iint_W \lambda\, d\mu dt \le \left(\iint_W \lambda^n d\mu dt\right)^{1/n} \left(\iint_W d\mu dt\right)^{(n-1)/n}$$

and surely

$$\iint_W d\mu dt \le L|\mathbb{S}^{n-1}| < \infty\,.$$

Therefore

$$\iint_{0\le t<h(\vartheta)} \lambda(\vartheta,t) d\mu dt < \infty\,.$$

On the other hand, if we integrate first in t, we see that

$$\int_{\mathbb{S}^{n-1}} \left(\int_0^{h(\vartheta)} \lambda(\vartheta,t)\,dt \right) d\mu \ge |\mathbb{S}^{n-1}| \inf_{\vartheta\in\mathbb{S}^{n-1}} \int_0^{h(\vartheta)} \lambda(\vartheta,t)\,dt$$

and therefore

$$\inf_{\vartheta\in\mathbb{S}^{n-1}} \int_0^{h(\vartheta)} \lambda(\vartheta,t)\,dt < \infty\,.$$

But along a path where ϑ is constant we have $\tilde g = \lambda$. Thus there is some ϑ where the path from $(\vartheta,0)$ to $(\vartheta,h(\vartheta))$ has finite length. This shows that the metric is not complete and proves the theorem. □

Theorem 3.6 (Hamilton, [Ham94, Theorem 1.1])**.** *Let M be a smooth strictly convex hypersurface bounding a region in $\mathbb{R}^{n+1}$, $n\ge 2$. Suppose that its second fundamental form is ε-pinched in the sense that*

$$h_{ij} \ge \varepsilon H g_{ij}$$

for some $\varepsilon>0$. Then M is compact.

Proof. Assume that M is noncompact. By Theorem 3.3, M has a conformally equivalent metric $\tilde g_{ij}$ which is complete with finite volume. Observe that the Gauss map $\boldsymbol{\nu}: M\to\mathbb{S}^n$ gives a diffeomorphism of the convex hypersurface M onto an open subset $U=\boldsymbol{\nu}(M)$ of the sphere $\mathbb{S}^n$ which lies in a hemisphere. Thus U is not empty and its closure is not all of $\mathbb{S}^n$. By Theorem 3.5, there is no metric $\hat g_{ij}$ on U, quasi-conformal with respect to the round metric, which is complete with finite volume. However, the pinching condition implies

$$\varepsilon H\delta_i^k \le h_i^k \le H\delta_i^k$$

so that

$$\varepsilon H\partial_i = \varepsilon H\delta_i^k\partial_k \le h_i^k\partial_k = \partial_i\boldsymbol{\nu} \le H\delta_i^k\partial_k = H\partial_i\,.$$

We define

$$\hat g_{ij} := \langle \partial_i\boldsymbol{\nu}, \partial_j\boldsymbol{\nu}\rangle$$

and observe that

$$(\varepsilon H)^2 g_{ij} = \left(\varepsilon\tilde H\right)^2 \tilde g_{ij}\,.$$

Hence,

$$\left(\varepsilon\tilde H\right)^2 \tilde g_{ij} \le \hat g_{ij} \le \tilde H^2 \tilde g_{ij}\,.$$

If $\tilde g_{ij}$ is complete, $\hat g_{ij}$ is, by the first inequality. If $\tilde g_{ij}$ has finite Volume, $\hat g_{ij}$ must have by the second inequality. This is a contradiction. □

4. Singularities

Definition 4.1 (Singularities, see [Eck04, Definitions 3.5 and 5.1])**.** We say that a solution $(M_t)_{t\in[0,T)}$ of (MCF) *reaches* a point $x_0\in\mathbb{R}^{n+1}$ at time $T\le\infty$ if there exists a sequence $(p_k,t_k)_{k\in\mathbb{N}}$ in $M^n\times[0,T)$ with $t_k\nearrow T$ so that $X(p_k,t_k)\to x_0$ for $k\to\infty$.

Let $\mathcal{S}$ be the set of points $x\in\mathbb{R}^{n+1}$ so that there exists a sequence $(p_k,t_k)_{k\in\mathbb{N}}$ with $t_k\nearrow T$ and $X(p_k,t_k)\to x$ for $k\to\infty$. We call $\mathcal{S}$ the set of *reachable points.*

A point $x_0\in\mathbb{R}^2$ is called a *singular* or *blow-up point* of the flow at time T if $(M_t)_{t\in[0,T)}$ reaches x_0 at time T and has no smooth extension beyond time T in any neighbourhood of x_0. The sequence $(p_k,t_k)_{k\in\mathbb{N}}$ is called *blow-up sequence.*

All other points (which includes those not reached by the solution) are called *regular points.*

We want to investigate singularities of the flow.

Proposition 4.2. *Let $T < \infty$. If $|A|^2 \leq C_0$ on $M^n \times [0,T)$, then $|\nabla^m A|^2 \leq C_m$ on $M^n \times [0,T)$, where $C_m = C_m(n, M_0, C_0)$.*

Proof. See [Sch17d, Proposition 2.1.5]. By Lemma 1.4,

$$\partial_t |\nabla^m A|^2 \leq \Delta |\nabla^m A|^2 - 2|\nabla^{m+1} A|^2 + C(n,m) \sum_{i+j+k=m} |\nabla^i A||\nabla^j A||\nabla^k A||\nabla^m A| \,.$$

We give a proof by induction. The case $m = 0$ is trivially true. So we assume that for $m > 0$ we have $|\nabla^l A|^2 \leq C_l$ for $0 \leq l \leq m-1$. Thus

$$\partial_t |\nabla^{m-1} A|^2 \leq \Delta |\nabla^{m-1} A|^2 - 2|\nabla^m A|^2 + B_{m-1}$$

and

$$\partial_t |\nabla^m A|^2 \leq \Delta |\nabla^m A|^2 - B_m \left(1 + |\nabla^m A|^2\right) \,.$$

We consider the function $f := |\nabla^m A|^2 + B_m |\nabla^{m-1} A|^2$, which satisfies

$$\begin{aligned} \partial_t f &\leq \Delta f - B_m \left(1 + |\nabla^m A|^2\right) - 2B_m |\nabla^m A|^2 + B_{m-1} B_m \\ &\leq \Delta f - B_m f + B_m^2 |\nabla^{m-1} A|^2 + B_{m-1} B_m \\ &\leq \Delta f - B_m f + B \,. \end{aligned}$$

Define $\tilde{f} := \exp(B_m t) f - \exp(B_m T) B t$. Then

$$\begin{aligned} \partial_t \tilde{f} &\leq \exp(B_m t)(B_m f + \partial_t f) - \exp(B_m T) B \\ &\leq \exp(B_m t)(\Delta f + B) - \exp(B_m T) B \leq \Delta \tilde{f} \end{aligned}$$

which implies $\tilde{f}(\cdot, t) \leq \max_M \tilde{f}(\cdot, 0)$ and thus

$$f(\cdot, t) \leq \exp(-B_m t) \left(\max_M \tilde{f}(\cdot, 0) + \exp(B_m T) B t \right) \leq C \,. \qquad \square$$

Theorem 4.3. *Let $T < \infty$ and $(M_t)_{t \in [0,T)}$ be a family of smooth, immersed hypersurfaces evolving by* (MCF) *with*

$$M_t \cap B_R(0) \neq \emptyset$$

for some $R > 0$ and all $t \in [0,T)$ and there exists $C_0 < \infty$ such that

$$\sup_{t \in [0,T)} \sup_{M_t} |A| \leq C_0 \,.$$

Then M_T is smooth.

Proof. By Proposition 4.2,

$$\sup_{t \in [0,T)} \sup_{M_t} |\nabla^m |A|| \leq C_m$$

for all $m \in \mathbb{N} \cup \{0\}$. By Lemma 1.4,

$$\partial_t \boldsymbol{\nu} = \nabla H$$

so that the rotation of the normal is uniformly bounded in small space-time neighbourhoods. That is, there exist $t_0 \in [0,T)$, $r > 0$ and $\varepsilon > 0$ so that for each $p \in M^n$ there exists an open neighbourhood

$$U_{r,t_0}(p) = X^{-1}(B_r(X(p,t_0)), t_0) \subset \mathbb{R}^n \,,$$

where B_r is the geodesic ball in M_{t_0}, so that, after rotation and translation,

$$\boldsymbol{\nu}(q,t) \in \mathbb{S}^n \cap \{x^n \geq \varepsilon\}$$

for all $q \in U_{r,t_0}(p)$ and $t \in [t_0, T)$. For $R_0 \geq R$, there exist finitely many points $\{p_i\}_{i=1}^{N_0}$ so that

$$M_t \cap B_{R_0}(0) \subset \bigcup_{i=1}^{N_0} X(U_{r,t_0}(p_i), t)$$

for all $t \in [t_0, T)$. For $p \in \{p_i\}_{i=1}^{N_0}$ we can write

$$M_t \cap X(U_{r,t_0}(p), t)$$

as a graph of a function $f : U_{r,t_0}(p) \times [t_0, T) \to \mathbb{R}$ with $|D^m f|$ uniformly bounded on $U_{r,t_0}(p) \times [t_0, T)$ for all $m \in \mathbb{N} \cap \{0\}$. Let $(t_k)_{k\in\mathbb{N}}$ with $t_k \nearrow T$. By Arzelá–Ascoli, for each $m \in \mathbb{N} \cap \{0\}$, the sequence

$$(f_k^m := D^m f(\cdot, t_k))_{k\in\mathbb{N}}$$

converges uniformly along a subsequence to a continuous limit

$$f_\infty^m = D^m f_\infty = D^m f(\cdot, T)\,.$$

Hence, $f(\cdot, T)$ is smooth. This can be done for each $i \in \{1, \ldots, N_0\}$. We define

$$X_k := X(\cdot, t_k)\,.$$

Locally, we can describe X_k via f_k. Thus $X(\cdot, T)$ is smooth on $\bigcup_{i=1}^{N_0} U_{r,t_0}(p_i)$ and so is $M_T \cap B_{R_0}(0)$. Let now be $(R_l)_{l\in\mathbb{N}}$ be a sequence of radii with $R \leq R_l \nearrow \infty$. For each $l \in \mathbb{N}$, there exist finitely many points $\{p_i\}_{i=1}^{N_l}$ so that

$$M_t \cap B_{R_l}(0) \subset \bigcup_{i=1}^{N_l} X(U_{r,t_0}(p_i), t)$$

for all $t \in [t_0, T)$. Define

$$X_k^l := X^l(\cdot, t_k)$$

locally via f_k^l. By the same argument as above, $X_\infty^l = X^l(\cdot, T) : \bigcup_{i=1}^{N_l} U_{r,t_0}(p_i) \to \mathbb{R}^{n+1}$ and $M_T \cap B_{R_l}(0)$ is smooth for every $l \in \mathbb{N}$. We now pick a diagonal sequence to obtain a smooth limit $X_\infty^\infty = X(\cdot, T) : M^n \to \mathbb{R}^{n+1}$ with image M_T which coincides with X_∞^l on every ball $B_{R_l}(0)$. Since $M_t \to M_T$ continuously for $t \to T$, the smooth convergence holds for $t \to T$. □

Corollary 4.4. *If $T < \infty$, then $\limsup_{t\to T} \max_{M_t} |A|^2 = \infty$.*

Proof. See [Sch17d, Corollary 2.1.6]. Let us assume to the contrary that $|A|^2 \leq C_0$ on $M^n \times [0, T)$. By Proposition 4.2 all higher derivatives of A are uniformly bounded on $M^n \times [0, T)$. By Theorem 4.3, $X(\,\cdot\,, T)$ is a smooth immersion. By short-time existence this implies that we can extend the solution further, which contradicts the assumption that T is maximal. □

Lemma 4.5 (Hamilton's trick [Ham86, Lemma 3.5])**.** *Let $f : [a, b] \times (0, T) \to \mathbb{R}$ be in C^1. Then $f_{\max}(t) := \max_{p\in[a,b]} f(p, t)$ is locally Lipschitz for $t \in (0, T)$ and at a differentiable time,*

$$\frac{d}{dt} f_{\max}(t) \leq \sup \left\{ \partial_t f(p, t) \;\middle|\; p \in [a, b] \text{ with } f(p, t) = f_{\max}(t) \right\}\,.$$

Proposition 4.6 (Huisken, [Hui90, Lemma 1.2])**.** *If $T < \infty$, then $\max |A|^2(t) \to \infty$ for $t \to T$ where*

$$\max |A|^2(t) \geq \frac{1}{\sqrt{2(T - t)}}\,.$$

Proof. By Corollary 4.4, $|A|_{\max}(t) \to \infty$ for $t \to T$. For $t \in (0,T)$, let $p \in M^n$ so that $|A|^2(p,t) = |A|^2_{\max}(t)$. Then

$$\operatorname{Hess}|A|^2(p,t) \preceq 0\,.$$

By Lemma 1.4

$$\partial_t |A|^2 = \Delta |A|^2 - |\nabla A|^2 + 2|A|^4 \le 2|A|^4$$

at (p,t). Since $|A|^2_{\max}$ is Lipschitz we obtain by Rademacher's theorem, Theorem A.3, that $\partial_t |A|^2_{\max}$ exists for almost every $t \in (0,T)$. By Hamilton's trick, Lemma 4.5,

$$\begin{aligned}\partial_t |A|^2_{\max}(t) &\le \max\left\{\partial_t |A|^2(p,t) \,\middle|\, p \in M^n \text{ with } |A|^2(p,t) = |A|^2_{\max}(t)\right\}\\ &\le \max\left\{2|A|^4(p,t) \,\middle|\, p \in M^n \text{ with } |A|^2(p,t) = |A|^2_{\max}(t)\right\} = 2|A|^4_{\max}(t)\end{aligned}$$

for almost every $t \in (0,T)$. Assume that there exists a time $t_0 \in [0,T)$ where $|A|^2_{\max} = 0$. Then M_{t_0} is a plane segment in $\mathbb{R}^{n+1}$ which contradicts that $T < \infty$. Hence, $|A|^2_{\max}(t) > 0$ for all $t \in [0,T)$ and $|A|^{-2}_{\max}$ is Lipschitz as well. Rademacher's theorem implies that $\partial_t |A|^{-2}_{\max}(t)$ exists for almost every $t \in (0,T)$. Thus,

$$\partial_t |A|^{-2}_{\max} = -|A|^{-4}_{\max} \partial_t |A|^2_{\max} \ge -2 \tag{4.1}$$

for almost every $t \in (0,T)$. Since $|A|^{-2}_{\max}$ is Lipschitz, we can integrate (4.1) over an interval $[t, t_k] \subset [0,T)$ to obtain

$$\frac{1}{|A|^2_{\max}(t_k)} - \frac{1}{|A|^2_{\max}(t)} \ge -2(t_k - t)\,. \tag{4.2}$$

Let $t \in [0,T)$ and $(t_k)_{k \in \mathbb{N}}$ be a sequence with $t_k \in (t,T)$ for all $k \in \mathbb{N}$, $t_k \nearrow T$ and $|A|^2_{\max}(t_k) \to \infty$ for $k \to \infty$. Taking the limit $k \to \infty$ in (4.2) yields

$$\frac{1}{|A|^2_{\max}(t)} \le 2(T-t)$$

for all $t \in [0,T)$. □

Example 4.7. (i) The curvature of the spheres $\mathbb{S}^n_{r(t)}$ blows up in the exact rate.
(ii) A dumbbell with a small neck develops a singularity at the neck before the surface disappears.

We distinguish between two types of singularities.

Definition 4.8 (Type-I and type-II singularities)**.** We say that a singularity is of *type I*, if there exists a constant $C_0 > 1$ so that

$$|A|_{\max}(t) \le \frac{C_0}{\sqrt{T-t}} \tag{4.3}$$

for all $t \in [0,T)$, and of *type II*, if such a constant does not exist, that is,

$$\limsup_{t \to T} |A|_{\max}(t)\sqrt{T-t} = \infty\,. \tag{4.4}$$

Remark 4.9 (Parabolic rescaling)**.** Let $\lambda > 0$ and $t_0 \in (0,T)$. Consider the rescaled flow $X_\lambda : M^n \times [-\lambda^2 t_0, t_0) \to \mathbb{R}^2$ with

$$X_\lambda(p,\tau) = \lambda\left(X\left(p, t_0 + \frac{\tau}{\lambda^2}\right) - x_0\right)\,.$$

and define

$$M^\lambda_\tau := \lambda\left(M_{t_0 - \tau/\lambda^2} - x_0\right)\,.$$

Then $\tau = \lambda^2(t - t_0)$, $\partial_\tau = \frac{1}{\lambda^2}\partial_t$, $g^\lambda_{ij} = \lambda^2 g_{ij}$ and $h^\lambda_{ij} = \lambda h_{ij}$ so that

$$|A_\lambda| = \frac{1}{\lambda}|A| \qquad \text{and} \qquad H_\lambda = \frac{1}{\lambda}H$$

so that

$$\partial_\tau X_\lambda = \frac{1}{\lambda}\partial_t X = -\frac{1}{\lambda}H\boldsymbol{\nu} = -H_\lambda \boldsymbol{\nu}$$

again flows by mean curvature flow.

Theorem 4.10. *Let $T < \infty$ and $k \in \mathbb{N}$. Let $\emptyset \neq J_k \subset J_{k+1}$ be a sequence of intervals and $(M^k_\tau)_{\tau \in J_k}$ be families of smooth, immersed hypersurfaces evolving by* (MCF) *for each $k \in \mathbb{N}$ with*

$$M^k_\tau \cap B_R(0) \neq \emptyset$$

for some $R > 0$ and for all $k \in \mathbb{N}$ and all $\tau \in J_k$, and there exists $C_0 < \infty$ such that

$$\sup_{k\in\mathbb{N}} \sup_{\tau \in J_k} \sup_{M^k_\tau} |A_k| \leq C_0 \,.$$

Then there exists a subsequence $\big((M^k_\tau)_{\tau\in J_k}\big)_{k\in\mathbb{N}}$ that converges on compact subsets of J_∞ and in $\mathbb{R}^{n+1}$ to a smooth, immersed limit flow $(M^\infty_\tau)_{\tau \in J_\infty}$ evolving by (MCF).

Proof. By Proposition 4.2,

$$\sup_{k\in\mathbb{N}} \sup_{\tau \in J_k} \sup_{M^k_\tau} |\nabla^m |A_k|| \leq C_m$$

for all $m \in \mathbb{N} \cup \{0\}$. Let $R_0 \geq R$, $k_0 \in \mathbb{N}$ and $\tau_0 \in J_k$ for $k \geq k_0$. Since $M^k_{\tau_0}$ is smooth and

$$\tilde{M}^k_{\tau_0} := M^k_{\tau_0} \cap B^{n+1}_{R_0}(0) \neq \emptyset$$

for every $k \in \mathbb{N}$, there exists a subsequence $(\tilde{M}^k_{\tau_0})_{k\in\mathbb{N}}$ with continuous limit

$$\tilde{M}^\infty_{\tau_0} \subset B^{n+1}_{R_0}(0) \,.$$

Moreover, there exists $r > 0$ so that for every $x \in \tilde{M}^\infty_{\tau_0}$,

$$\tilde{M}^\infty_{\tau_0, r}(x) := \tilde{M}^\infty_{t_0} \cap B^{n+1}_r(x)$$

can be written as a graph of some function $g : B^n_r(x) \subset P(x) \to \mathbb{R}$ over some affine tangent plane $P(x)$ at x. By the convergence, there exists a subsequence $(\tilde{M}^k_{t_0})_{k\in\mathbb{N}}$ so that, for k big enough,

$$\tilde{M}^k_{\tau_0} \cap B^{n+1}_r(x)$$

can be written as graphs of some function $g_k : B^n_{r/2}(x) \to \mathbb{R}$ over the same affine plane $P(x)$. By the uniform bounds on $|A_k|$, $|D^m g_k|$ is uniformly bounded for all $m \in \mathbb{N}$ and g_k is smooth for every $k \geq k_0$. Furthermore, there exists $\delta, \varepsilon > 0$ so that, after rotation and translation,

$$\boldsymbol{\nu}_k(y) \in \mathbb{S}^n \cap \{x^n \geq \varepsilon\}$$

for all $y \in \tilde{M}^k_\tau \cap B^{n+1}_r(x)$ and $\tau \in (\tau_0 - \delta, \tau_0 + \delta)$, so that $\tilde{M}^k_\tau \cap B^{n+1}_r(x)$ can be written as graphs of the functions $f_k : B^n_{r/2}(x) \times (t_0 - \delta, t_0 + \delta) \to \mathbb{R}$. Since all time derivatives can be expressed in terms of spatial derivatives, f is smooth in time. By Arzelá–Acsoli, $(f_k)_{k\in\mathbb{N}}$ converges along a subsequence to a smooth limit f_∞. Like in the proof of Theorem 4.3, we can repeat this process this for a sequence $(R_l)_{l\in\mathbb{N}}$ with $R \geq R_l \to \infty$, and after picking a diagonal sequence we obtain a smooth limit $M^\infty_\tau \subset \mathbb{R}^{n+1}$. Note that a subsequence of the $X_k(\cdot, \tau)$ does not necessarily converge to a limiting immersion; it will be necessary to "reparametrize" $X_k(\cdot, \tau)$ (see [Lan85,] for details). □

5. Typ-I singularities

We want to rescale the surface M_t near a type-I singularity as $t \to T < \infty$. The following rescaling technique was introduced in [HS99b, Remark 4.6].

Definition 5.1 (Type-I rescaling)**.** Let $(p_k, t_k)_{k\in\mathbb{N}}$ be a blow-up sequence in $M^n \times [0,T)$ with $t_k \nearrow T$ for $k \to \infty$ and

$$|A|^2(p_k,t_k) = \max_{p\in M^n} |A|^2(p,t_k) = \max_{M^n\times[0,t_k]} |A|^2(p,t)$$

for each $k \in \mathbb{N}$. We set

$$\lambda_k^2 := |A|^2(p_k,t_k) \qquad \text{and} \qquad \alpha_k := -\lambda_k^2 T$$

and define the rescaled embeddings $X_k : M^n \times [\alpha_k, 0) \to \mathbb{R}^2$ by

$$X_k(p,\tau) := \lambda_k \left(X\left(p, T + \frac{\tau}{\lambda_k^2}\right) - x_0 \right) . \tag{5.1}$$

Lemma 5.2 (Properties of the type-I rescaling)**.** *Let $X : M^n \times (0,T) \to \mathbb{R}^2$ be a solution of* (MCF) *with $T < \infty$. For the type-I rescaling 5.1 in case of a type-I singularity,*

$$\lambda_k \to \infty \qquad \textit{and} \qquad \alpha_k \to -\infty$$

for $k \to \infty$. Furthermore,

$$X_k(0,\tau_k) \in B_{3C_0^2}(0) \qquad \textit{and} \qquad |A_k|^2(0,\tau_k) = 1\,,$$

where

$$\tau_k := -\lambda_k^2(T - t_k) \in \left[-\frac{C_0^2}{2}, -\frac{1}{2} \right]$$

and, for $\delta > 0$,

$$\max_{M^n\times[\alpha_k,-\delta^2]} |A_k| \le \frac{C_0}{\delta}$$

for all $k \in \mathbb{N}$.

Proof. We follow [MB14, Corollary 4.8, Lemma 7.1.8 and Proposition 7.1.10]. Let $x_0 \in \mathbb{R}^{n+1}$ be a singular point with corresponding blow-up sequence $(p_k, t_k)_{k\in\mathbb{N}}$ in $M^n \times [0,T)$. By the definition (4.3) of a type-I singularity, we calculate for $p \in M^n$ and $t_k, t_l \in [0,T)$,

$$\begin{aligned} |X(p,t_l) - X(p,t_k)| &\le \int_{t_k}^{t_l} \left| \frac{\partial X}{\partial t}(p,t) \right| dt \le \int_{t_k}^{t_l} |H(p,t)|\, dt \\ &\le 2\int_{t_k}^{t_l} |H|_{\max}(t)\, dt \le 2\int_{t_k}^{t_l} \frac{C_0}{\sqrt{2(T-t)}}\, dt \\ &= C_0 \left(-\sqrt{2(T-t_l)} + \sqrt{2(T-t_k)} \right) \le C_0\sqrt{2(T-t_k)}\,. \end{aligned} \tag{5.2}$$

Since the sequence $(p_k)_{k\in\mathbb{N}}$ is bounded, there exist a point $p_0 \in M^n$ and a subsequence with

$$p_k \to p_0 \tag{5.3}$$

for $k \to \infty$. We employ (5.2) for $p = p_l$, and obtain

$$|X(p_l,t_l) - X(p_l,t_k)| \le C_0\sqrt{2(T-t_k)} \tag{5.4}$$

for all $k, l \in \mathbb{N}$. By Definition 5.1, we can choose $l_0 = l_0(k)$ large enough so that, for fixed $k \in \mathbb{N}$,

$$|X(p_l,t_l) - x_0| \le C_0\sqrt{2(T-t_k)} \tag{5.5}$$

for all $l \ge l_0$. Estimates (5.4) and (5.5) imply

$$\begin{aligned} |X(p_l,t_k) - x_0| &\le |X(p_l,t_k) - X(p_l,t_l)| + |X(p_l,t_l) - x_0| \\ &\le 3C_0\sqrt{2(T-t_k)} \end{aligned} \tag{5.6}$$

for fixed $k \in \mathbb{N}$ and for all $l \geq l_0(k)$. For given $\varepsilon > 0$, choose $k_0 = k_0(\varepsilon)$ large enough, so that

$$3C_0\sqrt{2(T-t_k)} < \frac{\varepsilon}{2}\,.$$

for all $k \geq k_0$. Then (5.6) yields

$$|X(p_l, t_k) - x_0| < \frac{\varepsilon}{2}$$

for all $k \geq k_0(\varepsilon)$ and $l \geq l_0(k)$. By the convergence (5.3) and the continuity of the immersion X in its spatial argument, we can further choose l_0 large enough, so that also

$$|X(p_0, t_k) - X(p_l, t_k)| < \frac{\varepsilon}{2}$$

for $l \geq l_0$. Hence,

$$|X(p_0, t_k) - x_0| \leq |X(p_0, t_k) - X(p_{l_0}, t_k)| + |X(p_{l_0}, t_k) - x_0| < \varepsilon$$

for all $k \geq k_0(\varepsilon)$. Since $\varepsilon > 0$ was chosen arbitrarily, we obtain

$$X(p_0, t_k) \to x_0 \tag{5.7}$$

for $k \to \infty$. Definition 5.1 and the type-I condition (4.3) yield

$$\lambda_k = |A(p_k, t_k)| \leq \frac{C_0}{\sqrt{2(T-t_k)}}$$

and the estimate (5.2) implies

$$|X(p_0, t_l) - X(p_0, t_k)| \leq 2C_0\sqrt{2(T-t_k)} \leq \frac{2C_0^2}{\lambda_k}\,.$$

We send $l \to \infty$ in the above inequality and obtain with (5.7),

$$\lambda_k |x_0 - X(p_0, t_k)| \leq 2C_0^2$$

for all $k \in \mathbb{N}$. The definition (5.1) of the rescaled embedding provides, for $\tau_k := \lambda_k^2(t_k - T)$,

$$|X_k(p_0, \tau_k)| = \lambda_k \left| X\left(p_0, T + \frac{\tau_k}{\lambda_k^2}\right) - x_0 \right| \leq 2C_0^2$$

for all $k \in \mathbb{N}$. By the convergence (5.3), for given $\delta > 0$, there exists $k_1 \in \mathbb{N}$ so that $|p_k - p_0| < \delta$ for all $k \geq k_0$. By the continuity of the rescaled embedding, for given $\varepsilon > 0$, there exists $\delta > 0$ so that, for $|p_k - p_0| < \delta$, we have

$$|X_k(p_k, \tau_k) - X_k(p_0, \tau_k)| < \varepsilon\,.$$

Hence, for given $\varepsilon > 0$, there exists $k_1 \in \mathbb{N}$ so that

$$|X_k(0, \tau_k)| = |X_k(p_k, \tau_k)| \leq |X_k(p_k, \tau_k) - X_k(p_0, \tau_k)| + |X_k(p_0, \tau_k)| < \varepsilon + 2C_0^2$$

for all $k \geq k_1$. Choosing $\varepsilon = C_0^2$ yields $X_k(0, \tau_k) \in B_{3C_0^2}(0)$ for all $k \geq k_1$. To bound the sequence

$$\left(\tau_k = -\lambda_k^2(T - t_k)\right)_{k \in \mathbb{N}}\,,$$

we estimate

$$\alpha_k = -\lambda_k^2 T < -\lambda_k^2 T + \lambda_k^2 t_k = \tau_k < 0$$

for all $k \in \mathbb{N}$. The rescaling behaviour from Remark 4.9 of the curvature yields

$$|A_k|^2(0, \tau_k) = |A_k|^2(p_k, \tau_k) = \frac{1}{\lambda_k^2}|A|^2\left(p_k, T + \frac{\tau_k}{\lambda_k^2}\right) = \frac{1}{\lambda_k^2}|A|^2(p_k, t_k) = 1\,.$$

Using Definition 5.1 and the lower blow-up rate from Proposition 4.6, we estimate

$$\tau_k = -\lambda_k^2(T - t_k) = -|A|^2(p_k, t_k)(T - t_k) \leq -\frac{(T - t_k)}{2(T - t_k)} = -\frac{1}{2}$$

and, by the type-I assumption (4.3),

$$\tau_k = -\lambda_k^2(T-t_k) = -|A|^2(p_k,t_k)(T-t_k) \geq -\frac{C_0^2(T-t_k)}{2(T-t_k)} = -\frac{C_0^2}{2}$$

for all $k \in \mathbb{N}$. For the curvature estimate, let $\delta > 0$, $k \in \mathbb{N}$, $\tau \in [\alpha_k, -\delta^2]$ and $p \in M^n$. Then, the type-I condition (4.3) rescales to

$$|A_k(p,\tau)| = \frac{1}{\lambda_k}\left|A\left(p, T + \frac{\tau}{\lambda_k^2}\right)\right| \leq \frac{1}{\lambda_k}\frac{C_0}{\sqrt{-2\tau/\lambda_k^2}} \leq \frac{C_0}{\sqrt{-\tau}}\,.$$

Hence,

$$\max_{M^n\times[\alpha_k,-\delta^2]} |A_k| \leq \frac{C_0}{\delta}$$

for each $k \in \mathbb{N}$. □

Theorem 5.3 (Convergence of rescalings)**.** *Let $(M_t)_{t\in[0,T)}$ be a smooth, immersed solution of* (MCF) *with $T < \infty$. For the type-I rescaling 5.1 in case of a type-I singularity, there exists a sequence of rescaled immersions*

$$\Big((M_\tau^k)_{\tau\in[\alpha_k,0)}\Big)_{k\in\mathbb{N}}$$

that converges for $k \to \infty$ along a subsequence, uniformly and smoothly on compact subsets of $(-\infty, 0)$ and $\mathbb{R}^{n+1}$ to a maximal, smooth limit solution $(M_\tau^\infty)_{\tau\in(-\infty,0)}$ which satisfies

$$M_{\tau_\infty}^\infty \cap B_{3C_0^2}(0) \neq \emptyset \qquad \text{and} \qquad |A_\infty|^2(x) = 1 \;\text{ for some }\; x \in M_{\tau_\infty}^\infty\,,$$

where $\tau_\infty \in [-C_0^2/2, -1/2]$ and, for $\delta > 0$,

$$\sup_{\tau\in(-\infty,-\delta)}\sup_{M_\tau^\infty}|A_\infty| \leq \frac{C_0}{\delta^2}\,.$$

Moreover, if $(M_t)_{t\in[0,T)}$ is embedded, then $(M_\tau^\infty)_{\tau\in(-\infty,0)}$ is embedded.

Proof. The convergence follows from Theorem 4.10 and Lemma 5.2 yields the properties. By Proposition 1.9, M_τ^k is embedded for all $k \in \mathbb{N}$ and all $\tau \in [\alpha_k, 0)$. Furthermore,

$$d_k(\tau) \geq \min\left\{d_k(\alpha_k), \frac{\sin(\varepsilon)}{m_k(\tau)}\right\} \geq \min\left\{\lambda_k d(0), \frac{\sin(\varepsilon)\delta^2}{C_0}\right\}$$

is uniformly bounded in k for $\tau \leq \delta < 0$. □

5.1. **Huisken's monotonicity formula.** For $x_0 \in \mathbb{R}^{n+1}$ and $t_0 \in \mathbb{R}$, define the *backward heat kernel* $\Phi_{(x_0,t_0)} : \mathbb{R}^{n+1} \times (-\infty, t_0) \to \mathbb{R}$ by

$$\Phi_{(x_0,t_0)}(x,t) := \frac{1}{(4\pi(t_0-t))^{n/2}}\exp\left(-\frac{|x-x_0|^2}{4(t_0-t)}\right).$$

Let $x, x_0, y_0 \in \mathbb{R}^{n+1}$, $t_0, \tau_0 \in \mathbb{R}$, $t \in (-\infty, t_0)$, $\lambda > 0$ and $\tau_0 > \lambda^2(t - t_0)$. Then

$$\Phi_{(y_0,\tau_0)}\big(\lambda(x-x_0), \lambda^2(t-t_0)\big) = \frac{1}{\lambda^n}\Phi_{(x_0+y_0/\lambda,\,t_0+\tau_0/\lambda^2)}(x,t)\,.$$

For the rescaled flow $(M_\tau^\lambda)_{\tau\in[-\lambda^2T,0)}$,

$$d\mu_\lambda^n = \sqrt{\det(g_{ij}^\lambda)}\,dp = \sqrt{\lambda^{2n}\det(g_{ij})}\,dp = \lambda^n d\mu\,.$$

Hence, the integral

$$\int_{M_\tau^\lambda}\Phi_{(y_0,\tau_0)}\,d\mu_\lambda^n = \int_{M_{T-\tau/\lambda^2}}\Phi_{(x_0+y_0/\lambda,\,T+\tau_0/\lambda^2)}\,d\mu^n$$

is scaling invariant, which makes it a useful quantity. In the following, we set $H(x,t) = H(p,t)$ and $\boldsymbol{\nu}(x,t) = \boldsymbol{\nu}(p,t)$ for $x = X(p,t)$.

Theorem 5.4 (Monotonicity formula, Huisken [Hui90, Theorem 3.1]). *Let $X : M^n \times (0,T) \to \mathbb{R}^{n+1}$ be a solution of* (MCF). *Then*

$$\frac{d}{dt}\left(\int_{M_t} \Phi_{(x_0,t_0)}\, d\mu_t^n\right) = -\int_{M_t} \left|H - \frac{\langle x - x_0, \boldsymbol{\nu}\rangle}{2(t_0 - t)}\right|^2 \Phi_{(x_0,t_0)}\, d\mu_t^n$$

for $t_0 \in (0,T]$ and $t \in (0,t_0)$.

Proof. We follow the lines of [Hui90, Theorem 3.1]. We set $x_0 = 0$ and $t_0 = 0$. Since $x = x(t)$ with $\partial_t x(t) = \mathbf{H}$, we derive

$$\begin{aligned}\frac{d}{dt}\Phi_{(0,0)} &= \left(\frac{(n/2)4\pi}{-4\pi t} - \frac{2\langle x, \mathbf{H}\rangle}{-4t} - \frac{|x|^2}{4t^2}\right)\Phi_{(0,0)} \\ &= \left(\frac{n}{-2t} + H\frac{\langle x, \boldsymbol{\nu}\rangle}{-2t} - \frac{|x|^2}{4t^2}\right)\Phi_{(0,0)}\end{aligned}$$

so that

$$\frac{d}{dt}\left(\int_{M_t} \Phi_{(0,0)}\, d\mu_t^n\right) = \int_{M_t}\left(\frac{n}{-2t} + H\frac{\langle x, \boldsymbol{\nu}\rangle}{-2t} - \frac{|x|^2}{4t^2} - H^2\right)\Phi_{(0,0)}\, d\mu_t^n$$

Observe that

$$-H^2 + H\frac{\langle x, \boldsymbol{\nu}\rangle}{-t} - \frac{\langle x, \boldsymbol{\nu}\rangle^2}{4t^2} = -\left|H - \frac{\langle x, \boldsymbol{\nu}\rangle}{-2t}\right|^2$$

and

$$|x|^2 = \langle x, \boldsymbol{\nu}\rangle^2 + g^{ij}\langle x, \partial_i X\rangle\langle x, \partial_j X\rangle\,.$$

Hence,

$$\begin{aligned}&\frac{n}{-2t} + H\frac{\langle x, \boldsymbol{\nu}\rangle}{-2t} - \frac{|x|^2}{4t^2} - H^2 \\ &\quad = \frac{1}{-2t}\left(n - H\langle x, \boldsymbol{\nu}\rangle - \frac{1}{-2t}g^{ij}\langle x, \partial_i X\rangle\langle x, \partial_j X\rangle\right) - \left|H - \frac{\langle x, \boldsymbol{\nu}\rangle}{-2t}\right|^2\,. \qquad (5.8)\end{aligned}$$

For $x \in M_t$,

$$\operatorname{div}_{M_t} x = \operatorname{div}_{M^n} X(p,t) = n$$

and by the divergence theorem,

$$-\int_{M_t} H\langle x, \boldsymbol{\nu}\rangle\Phi_{(0,0)}\, d\mu_t^n = \int_{M_t}\langle x, \mathbf{H}\rangle\Phi_{(0,0)}\, d\mu_t^n = -\int_{M_t}\operatorname{div}_{M_t}\big(x\,\Phi_{(0,0)}\big)\, d\mu_t^n\,,$$

where

$$\operatorname{div}_{M_t}\big(x\,\Phi_{(0,0)}\big) = \Phi_{(0,0)}\operatorname{div}_{M_t} x + \left\langle x, \nabla^{M_t}\Phi_{(0,0)}\right\rangle\,.$$

We calculate on M_t,

$$\nabla^{M_t}\Phi_{(0,0)} = -\Phi_{(0,0)}g^{ij}\frac{2\langle x, \partial_i x\rangle}{-4t}\partial_j X = -\Phi_{(0,0)}g^{ij}\frac{\langle x, \partial_i X\rangle}{-2t}\partial_j X\,.$$

so that

$$\operatorname{div}_{M_t}\big(x\,\Phi_{(0,0)}\big) = n - \frac{1}{-2t}g^{ij}\langle x, \partial_i X\rangle\langle x, \partial_j X\rangle\Phi_{(0,0)}$$

which proves the claim. □

Theorem 5.5 (Weighted monotonicity formula, [Eck04, Theorem 4.13]). *Let $X : M^n \times (0,T) \to \mathbb{R}^{n+1}$ be a solution of* (MCF) *and $\varphi : \mathbb{R}^{n+1} \times (0,T) \to \mathbb{R}$ in $C^{2;1}$. Then*

$$\begin{aligned}\frac{d}{dt}\int_{M_t}\varphi\,\Phi_{(x_0,t_0)}\, d\mu_t^n = &-\int_{M_t}\left|\mathbf{H} + \frac{(x - x_0)^\perp}{2(t_0 - t)}\right|^2 \varphi\,\Phi_{(x_0,t_0)}\, d\mu_t^n \\ &+ \int_{M_t}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi\,\Phi_{(x_0,t_0)}\, d\mu_t^n\end{aligned}$$

for $t_0 \in (0,T]$ and $t \in (0,t_0)$.

Proof. The proof is like the one for Theorem 5.4 with one additional step. When applying the divergence theorem, Theorem A.2, we now use the vector $v = x\varphi\,\Phi_{(0,0)}$ instead and deduce

$$\int_{M_t} \langle x, H\boldsymbol{\nu}\rangle\, \varphi\, \Phi_{(0,0)}\, d\mu_t^n = \int_{M_t} \operatorname{div}_{M_t}\big((x)\varphi\, \Phi_{(0,0)}\big)\, d\mu_t^n\,,$$

where

$$\operatorname{div}_{M_t}\big(x\varphi\,\Phi_{(0,0)}\big) = n\varphi\,\Phi_{(0,0)} + \varphi\left\langle x, \nabla^{M_t}\Phi_{(0,0)}\right\rangle + \left\langle x, \nabla^{M_t}\varphi\right\rangle \Phi_{(0,0)}\,.$$

Since $\nabla^{M_t}\varphi = \boldsymbol{\tau}_i(\varphi)\boldsymbol{\tau}_i$ we can utilise the gradient of $\Phi_{(0,0)}$ again to find

$$\frac{\left\langle x, \nabla^{M_t}\varphi\right\rangle}{-2t}\Phi_{(0,0)} = -\left\langle \nabla^{M_t}\Phi_{(0,0)}, \nabla^{M_t}\varphi\right\rangle$$

so that integration by parts yields the extra term

$$\int_{M_t} \frac{\left\langle x, \nabla^{M_t}\varphi\right\rangle}{-2t}\Phi_{(0,0)}\, d\mu_t^n = \int_{M_t} \Delta_{M_t}\varphi\, \Phi_{(0,0)}\, d\mu_t^n\,.$$

The minus sign comes from the operation in (5.8). □

Remark 5.6 (see [Eck04, Remark 4.8])**.** If M_t is only defined locally, say in $B_{\sqrt{4n}\rho}(x_0) \times (t_0 - \rho^2, t_0)$, then we can use the cut-off function

$$\varphi^{\rho}_{(x_0,t_0)}(x,t) = \left(1 - \frac{|x-x_0|^2 + 2n(t_0 - t)}{\rho^2}\right)^3_+$$

where $(\partial_t - \Delta_{M_t})\varphi \leq 0$. Thus we still get the monotonicity inequality

$$\frac{d}{dt}\int_{M_t} \varphi^{\rho}_{(x_0,t_0)}\, \Phi_{(x_0,t_0)}\, d\mu_t^n \leq -\int_{M_t} \left|\mathbf{H} + \frac{(x-x_0)^{\perp}}{2(t_0-t)}\right|^2 \varphi^{\rho}_{(x_0,t_0)}\, \Phi_{(x_0,t_0)}\, d\mu_t^n$$

for $t \in (0, t_0)$.

Theorem 5.7. *Let M_0 be compact, convex and embedded. Then, every limit flow obtained by the type-I rescaling 5.1 around a type-I singularity, up to a rotation in $\mathbb{R}^{n+1}$, must be either the skrinking spheres $\big(\mathbb{S}^n_{\sqrt{-2n\tau}}\big)_{\tau\in(-\infty,0)}$ or one of the shrinking cylinders $\big(\mathbb{S}^m_{\sqrt{-2m\tau}} \times \mathbb{R}^{n-m}\big)_{\tau\in(-\infty,0)}$ for $0 < m < n$.*

Proof. Let $x_0 \in \mathbb{R}^{n+1}$ be arbitrary. For $t \in [0, T)$, define the monotonicity quantity

$$\Theta_{(x_0,T)}(t) := \int_{M_t} \Phi_{(x_0,T)}(x,t)\, d\mu_t^n\,.$$

The monotonicity formula, Theorem 5.4, yields

$$\partial_t \Theta_{(x_0,T)}(t) \leq 0 \tag{5.9}$$

for $t \in (0, T)$. Hence, the monotonicity quantity is monotonically decreasing and strictly positive, so that the limit

$$\lim_{t\to T} \Theta_{(x_0,T)}(t)$$

exists and for any sequence $(t_k)_{k\in\mathbb{N}}$ with $t_k \nearrow T$ for $k \to \infty$,

$$\lim_{t\to T} \Theta_{(x_0,T)}(t) = \lim_{k\to\infty} \Theta_{(x_0,T)}(t_k)\,. \tag{5.10}$$

For $k \in \mathbb{N}$, $y = \lambda_k(x - x_0) \in \mathbb{R}^{n+1}$ and $\tau = \lambda_k^2(t - T) \in [\alpha_k, 0)$, the backward heat kernel rescales according to

$$\Phi_{(0,0)}(y,\tau) = \frac{1}{\lambda_k^n}\Phi_{(x_0+0/\lambda_k, T+0/\lambda_k^2)}(x,t) = \frac{1}{\lambda_k^n}\Phi_{(x_0,T)}(x,t)\,.$$

Let $\tau \in (-\infty, 0)$ and $k_0 \in \mathbb{N}$ so that $\tau \in [\alpha_k, 0)$ for $k \geq k_0$. Let $(\lambda_k)_{k\in\mathbb{N}}$ be a sequence of positive real numbers with $\lambda_k \to \infty$ for $k \to \infty$. We rescale the flow according to the type-I rescaling 5.1 with respect to the sequence $(\lambda_k)_{k\in\mathbb{N}}$ and consider the rescaled flow $(M^k_\tau)_{\tau\in[\alpha_k,0)}$. We receive a factor of λ_k^n from the scaling behaviour of the area element, and a factor of $1/\lambda_k^n$ from the scaling behaviour of the backward heat kernel. Hence, the monotonicity quantity translates, for $t_k := T + \tau/\lambda_k^2$, by

$$\begin{aligned}\Theta_{(x_0,T)}(t_k) &= \int_{M_{t_k}} \Phi_{(x_0,T)}(x,t_k)\, d\mu^n_{t_k} \\ &= \int_{M^k_\tau} \Phi_{(0,0)}(y,\tau)\, d\mu^n_{k,\tau} =: \Theta^k_{(0,0)}(\tau)\,.\end{aligned}$$

Corollary 4.4 implies that there exist $p_0 \in M^n$, $x_0 \in \mathbb{R}^{n+1}$ and $(p_k,t_k)_{k\in\mathbb{N}}$ with

$$X(p_k,t_k) \to x_0 \quad \text{ and } \quad |A(p_k,t_k)| = \max_{M^n} |A(\cdot,t_k)| \to \infty$$

for $k \to \infty$. We rescale according to Definition 5.1 with respect to x_0 and $(p_k,t_k)_{k\in\mathbb{N}}$ and consider the rescaled embeddings $X_k : M^n \times [\alpha_k, 0) \to \mathbb{R}^{n+1}$. We apply the monotonicity formula 5.4 and estimate similar to [Bak10, Proposition 6.6] or [Coo11, Proposition 5.8],

$$\begin{aligned}0 \le \int_{\tau_1}^{\tau_2}\int_{M^k_\tau} \left|\mathbf{H}_k + \frac{y^\perp}{-2\tau}\right|^2 \Phi_{(0,0)}\, d\mu^n_\tau d\tau &\le \Theta^k_{(0,0)}(\tau_1) - \Theta^k_{(0,0)}(\tau_2) \\ &= \Theta_{(x_0,T)}\left(T + \frac{\tau_1}{\lambda_k^2}\right) - \Theta_{(x_0,T)}\left(T + \frac{\tau_2}{\lambda_k^2}\right)\end{aligned} \tag{5.11}$$

for all $k \ge k_0$. Since

$$T + \frac{\tau_i}{\lambda_k^2} \to T$$

for $k \to \infty$ and $i = 1, 2$, and by the existence of the limit (5.10), the right-hand side of (5.11) converges to 0 for $k \to \infty$. By Theorem 5.3, the sequence $\big((M^k_\tau)_{\tau\in[\tau_1,\tau_2]}\big)_{k\in\mathbb{N}}$ converges smoothly along a subsequence and on compact subsets of $\mathbb{R}^{n+1}$ to a smooth flow $(M^\infty_\tau)_{\tau\in[\tau_1,\tau_2]}$. Let $R > 0$. By the smooth convergence, there exists a $k_0 \in \mathbb{N}$ so that for all $k \ge k_0$, $M^k_\tau \cap B_R(0)$ can be parametrized over $M^\infty_\tau \cap B_R(0)$. That is, there exist embeddings $Y_k : M^\infty_\tau \cap B_R(0) \to \mathbb{R}^{n+1}$ with

$$M^k_\tau \cap B_R(0) = Y_k(M^\infty_\tau \cap B_R(0))$$

and $Y_k \to \mathrm{id}$ for $k \to \infty$. For $\tau \in [\tau_1, \tau_2]$, Fatou's lemma, Lemma A.4, implies

$$\begin{aligned}0 &= \liminf_{k\to\infty} \int_{M^k_\tau \cap B_R(0)} \left|\mathbf{H}_k + \frac{y^\perp}{-2\tau}\right|^2 \Phi_{(0,0)}\, d\mu^n_{k,\tau} \\ &= \liminf_{k\to\infty} \int_{M^\infty_\tau \cap B_R(0)} \left|\mathbf{H}_k + \frac{Y_k^\perp}{-2\tau}\right|^2 \Phi_{(0,0)} \sqrt{\det(DY_k)}\, dx \\ &\ge \int_{M^\infty_\tau \cap B_R(0)} \liminf_{k\to\infty} \left(\left|\mathbf{H}_k + \frac{Y_k^\perp}{-2\tau}\right|^2 \Phi_{(0,0)} \sqrt{\det(DY_k)}\right) dx \\ &= \int_{M^\infty_\tau \cap B_R(0)} \left|\mathbf{H}_\infty + \frac{Y_\infty^\perp}{-2\tau}\right|^2 \Phi_{(0,0)} \sqrt{\det(DY_\infty)}\, dx \\ &= \int_{M^\infty_\tau \cap B_R(0)} \left|\mathbf{H}_\infty + \frac{y^\perp}{-2\tau}\right|^2 \Phi_{(0,0)}\, d\mu^n_{\infty,\tau} \ge 0\,.\end{aligned}$$

Thus also

$$\int_{\tau_1}^{\tau_2}\int_{M^\infty_\tau \cap B_R(0)} \left|\mathbf{H}_\infty + \frac{y^\perp}{-2\tau}\right|^2 \Phi_{(0,0)}\, d\mu^n_t d\tau = 0\,.$$

Since $R > 0$ was chosen arbitrarily, we deduce

$$\int_{\tau_1}^{\tau_2}\int_{M^\infty_\tau} \left|\mathbf{H}_\infty + \frac{y^\perp}{-2\tau}\right|^2 \Phi_{(0,0)}\, d\mu^n_t d\tau = 0\,.$$

Since the convergence is smooth, and sending $\tau_1 \to -\infty$ and $\tau_2 \to 0$ yields

$$\left| \mathbf{H}_\infty + \frac{y^\perp}{-2\tau} \right|^2 = 0$$

for every $\tau \in (-\infty, 0)$ and every $y \in M^\infty_\tau$.

For the area estimate, let again be $R > 0$ and $\tau \in (-\infty, 0)$. Then there exists again $k_0 \in \mathbb{N}$ so that $\tau \in [\alpha_k, 0)$ and

$$T - \frac{\tau}{\lambda_k^2} \geq \frac{T}{2}$$

for all $k \geq k_0$. Like in Corollary 1.5,

$$\partial_t \mu^n_t (M_t \cap B_R) = - \int_{M_t \cap B_R} H^2 d\mu^n_t \,,$$

the area is decreasing locally also locally. By (5.9), the monotonicity quantity is decreasing in time and we can estimate with the definition of the backward heat kernel and the behaviour of the area of the hypersurfaces,

$$\begin{aligned}
&\int_{M^k_\tau \cap B_R(0)} \Phi_{(0,0)}(y,\tau)\, d\mu^n_{k,\tau} \\
&\leq \int_{M_{T-\tau/\lambda_k^2} \cap B_R(x_0)} \Phi_{(x_0,T)} \left(x, T - \frac{\tau}{\lambda_k^2} \right) d\mu^n_{T-\tau/\lambda_k^2} \\
&\leq \int_{M_{T/2} \cap B_R(x_0)} \Phi_{(x_0,T)} \left(x, \frac{T}{2} \right) d\mu^n_{T/2} \\
&= \frac{1}{(4\pi(T - T/2))^{n/2}} \int_{M_{T/2} \cap B_R(x_0)} \exp\left(- \frac{|x - x_0|^2}{4(T - T/2)} \right) d\mu^n_{T/2} \\
&\leq C(n,T) \mu^n_{T/2}(M_{T/2} \cap B_R(x_0)) \leq C(n,T) \mu^n_0 (M_0 \cap B_R(x_0))
\end{aligned}$$

Like before, Fatou's lemma implies

$$\begin{aligned}
&\liminf_{k\to\infty} \int_{M^k_\tau \cap B_R(0)} \Phi_{(0,0)}\, d\mu^n_{k,\tau} \\
&\quad = \liminf_{k\to\infty} \int_{M^\infty_\tau \cap B_R(0)} \Phi_{(0,0)} \sqrt{\det(DY_k)}\, dx \\
&\quad \geq \int_{M^\infty_\tau \cap B_R(0)} \liminf_{k\to\infty} \left(\Phi_{(0,0)} \sqrt{\det(DY_k)} \right) dx \\
&\quad = \int_{M^\infty_\tau \cap B_R(0)} \Phi_{(0,0)}\, d\mu^n_{\infty,\tau} \,.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\int_{M^\infty_\tau \cap B_R(0)} \Phi_{(0,0)}(y,\tau)\, d\mu^n_{\infty,\tau} \\
&\quad = \frac{1}{(-4\pi\tau)^{n/2}} \int_{M^\infty_\tau \cap B_R(0)} \exp\left(- \frac{|y|^2}{-4\tau} \right) d\mu^n_{\infty,\tau} \\
&\quad \geq \frac{1}{(-4\pi\tau)^{n/2}} \int_{M^\infty_\tau \cap B_R(0)} \exp\left(- \frac{R^2}{-4\tau} \right) d\mu^n_{\infty,\tau} \\
&\quad = \frac{1}{(-4\pi\tau)^{n/2}} \exp\left(- \frac{R^2}{-4\tau} \right) \mu^n_{\infty,\tau}(M^\infty_\tau \cap B_R(0)) \,.
\end{aligned}$$

so that

$$\mu^n(M^\infty_\tau \cap B_R(0)) \leq C(n,T,\tau) \mu^n_0 (M_0 \cap B_R(x_0)) \exp\left(\frac{R^2}{-4\tau} \right)$$

holds for all $\tau \in (-\infty, 0)$.

For every fixed $\tau \in (-\infty, 0)$, by Theorem 5.3, $|A|$ is not identically zero and $|\nabla^m A| \leq C_m$, for every $m \in \mathbb{N}$. Theorem 2.4 yields that

$$M_\tau^\infty = \mathbb{S}^m_{\sqrt{-2m\tau}} \times \mathbb{R}^{n-m}$$

where $0 < m \leq n$. Since the flow is smooth, the claim follows. □

5.2. **Gaussian density.**

Definition 5.8 (Gaussian density, [Sch17d, p. 26])**.** We define the *Gaussian density ratios* of the flow $\mathcal{M} = (M_t)_{t\in[0,T)}$ with respect to (x,t) as

$$\Theta(\mathcal{M},(x,t),r) := \int_{M_{t-r^2}} \Phi_{(x,t)}\, d\mu^n_{t-r^2}\,.$$

Note that the monotonicity formula implies that $\Theta(\mathcal{M},(x,t),r)$ is increasing in r. In case the flow is only defined locally as in Remark 5.6 we set

$$\Theta^\rho(\mathcal{M},(x,t),r) := \int_{M_{t-r^2}} \varphi^\rho_{(x,t)} \Phi_{(x,t)}\, d\mu^n_{t-r^2}\,.$$

Hence as $r \searrow 0$, the limit exists, so we can set

$$\Theta(\mathcal{M},(x,t)) := \lim_{r\searrow 0} \Theta(\mathcal{M},(x,t),r)\,,$$

called the *Gaussian density* of $\mathcal{M}$ at (x,t).

Remark 5.9. Let $\mathcal{M} = (M_t)_{t\in[0,T)}$ be a smooth mean curvature flow. We say that (x,t) is a smooth point of the flow, if in a space-time neighbourhood of (x,t) the flow $\mathcal{M}$ is smooth. One can show that at a smooth point (x,t) in the support of $\mathcal{M}$ one has $\Theta(\mathcal{M},(x,t)) = 1$, and thus at each singular point $\Theta \geq 1$. Similarly, any point reached by the flow has $\Theta \geq 1$. Furthoermore, if $\mathcal{M}$ is a smooth mean curvature flow such that (x,t) is a smooth point of the flow, then that $\Theta(\mathcal{M},(x,t),r) = 1$ for all $r > 0$ if and only if $\mathcal{M}$ is a multiplicity one plane containing (x,t).

Theorem 5.10 (Local regularity, White [Whi05, Theorem 1.1] see also [Eck04, Theorem 5.6])**.** *There exist universal constants $\varepsilon > 0$ and $C < \infty$ with the following property: If $\mathcal{M}$ is a smooth mean curvature flow of hypersurfaces in a parabolic ball $P(x_0,t_0,2(n+1)\rho)$ with*

$$\sup_{(x,t)\in P(x_0,t_0,r)} \Theta^\rho(\mathcal{M},(x,t),r) < 1+\varepsilon$$

for some $r \in (0,\rho)$, then

$$\sup_{P(x_0,t_0,r/2)} |A| \leq \frac{C}{r}\,.$$

Proof. See [HK17, Theorem C.1]. Suppose the assertion fails. Then there exists a sequence of smooth flows $\mathcal{M}^j$ in $P(0,0,2(n+1)\rho_j)$ for some $\rho_j > 1$ such that

$$\sup_{(x,t)\in P(0,0,1)} \Theta^{\rho_j}(\mathcal{M}^j,(x,t),1) < 1 + \frac{1}{j}$$

but such that there are points $(x_j,t_j) \in P(0,0,1/2)$ with $|A|(x_j,t_j) > j$. We can find $(\bar{x}_j,\bar{t}_j) \in P(0,0,3/4)$ with $\lambda_j = |A|(\bar{x}_j,\bar{t}_j) > j$ such that

$$\sup_{(x,t)\in P(\bar{x}_j,\bar{t}_j,j/10\lambda_j)} |A|(x,t) \leq 2\lambda_j \tag{5.12}$$

by the following technique, called point selection. Fix j. If $(x_j^0,t_j^0) = (x_j,t_j)$ already satisfies (5.12) with $\lambda_j^0 = |A|(x_j^0,t_j^0)$, we are done. Otherwise, there is a point $(x_j^1,t_j^1) \in P(x_j^0,t_j^0,j/10\lambda_j^0)$ with $\lambda_j^1 = |A|(x_j^1,t_j^1) > 2\lambda_j^0$. If (x_j^1,t_j^1) satisfies (5.12), we are done. Otherwise, there is a point $(x_j^2, t_j^2) \in P(x_j^1, t_j^1, j/10\lambda_j^1)$ with $\lambda_j^2 = |A|(x_j^2, t_j^2) > 2\lambda_j^1$, etc.. Note that

$$\frac{1}{2} + \frac{1}{10\lambda_j^0}\left(1 + \frac{1}{2} + \frac{1}{4} + \dots\right) < \frac{3}{4}\,.$$

By smoothness, the iteration terminates after a finite number of steps, and the last point of the iteration lies in $P(0,0,3/4)$ and satisfies (5.12). Now let $\hat{\mathcal{M}}^j$ be the flows obtained by shifting (x_j, t_j) to the origin and parabolically rescaling by $\lambda_j = |A|(x_j, t_j) \to \infty$. Since $|A|(0,0) = 1$ and $\sup_{P(0,0,j/10)} |A| \leq 2$, we can pass smoothly to a nonflat global limit, with

$$1 \leq \Theta^{\hat{\rho}_j}(\hat{\mathcal{M}}^j, (0,0), \lambda_j) < 1 + \frac{1}{j} \to 1$$

where $\hat{\rho}_j = \lambda_j \rho_j \to \infty$. On the other hand, like in the proof of Theorem 5.7, the limit is a flat plane. This is a contradiction. □

6. Typ-II singularities

The rescaling technique for type-II singularities was introduced in [Ham95a, Proof of Theorem 16.4] for Ricci flow, and applied to type-II singularities of MCF in [HS99b, p. 11].

Definition 6.1 (Type-II rescaling)**.** Let $(p_k, t_k)_{k\in\mathbb{N}}$ be a sequence in $M^n \times [0, T - 1/k]$ with

$$H^2(p_k, t_k)\left(T - \frac{1}{k} - t_k\right) = \max_{(p,t)\in M^n\times[0,T-1/k]}\left(H^2(p,t)\left(T - \frac{1}{k} - t\right)\right)$$

for each $k \in \mathbb{N}$. We set

$$\lambda_k^2 := |A|^2(p_k, t_k)\,, \qquad \alpha_k := -\lambda_k^2 t_k \qquad \text{and} \qquad T_k := \lambda_k^2\left(T - \frac{1}{k} - t_k\right)\,.$$

and define the rescaled embeddings $X_k : M^n \times [\alpha_k, T_k] \to \mathbb{R}^2$ by

$$X_k(p,\tau) := \lambda_k\left(X\left(p, t_k + \frac{\tau}{\lambda_k^2}\right) - X(p_k, t_k)\right)\,.$$

Lemma 6.2 (Properties of the type-II rescaling, [HS99b, Lemma 4.3])**.** *Let $X : M^n \times (0,T) \to \mathbb{R}^2$ be a solution of* (MCF) *with $T < \infty$. For the type-II rescaling 6.1 in case of a type-II singularity,*

$$\lambda_k \to \infty\,, \qquad \alpha_k \to -\infty \qquad \text{and} \qquad T_k \to \infty$$

for $k \to \infty$. Moreover,

$$X_k(0,0) = 0 \qquad \text{and} \qquad |A_k|^2(0,0) = 1$$

for every $k \in \mathbb{N}$ and for any $\varepsilon > 0$ and any $\bar{T} > 0$, there exists a $k_0 \in \mathbb{N}$ such that

$$\max_{M^n\times[\alpha_k,\bar{T}]} |A_k|^2 < 1 + \varepsilon$$

for all $k \geq k_0$.

Proof. We follow the lines of [HS99b, Lemma 4.3]. By definition, $X_k(0,0) = X_k(p_k, 0) = 0$ and

$$|A_k|^2(0,0) = \frac{1}{\lambda_k^2}|A|^2(p_k, t_k) = 1$$

for each $k \in \mathbb{N}$. Let $m > 0$ be arbitrary. By the definition (4.4) of a type-II singularity, there exist $\bar{t} \in [0,T)$ and $\bar{p} \in M^n$ so that

$$|A|^2(\bar{p}, \bar{t})(T - \bar{t}) > 2m\,.$$

We fix $\bar{t}$ and choose $k_0 \in \mathbb{N}$, so that $\bar{t} < T - 1/k$ and $|A|^2(\bar{p}, \bar{t})/k < m$ for all $k \geq k_0$. Then

$$|A|^2(\bar{p}, \bar{t})\left(T - \frac{1}{k} - \bar{t}\right) = |A|^2(\bar{p}, \bar{t})(T - \bar{t}) - \frac{1}{k}|A|^2(\bar{p}, \bar{t}) > m$$

and Definition 6.1 yields

$$T_k = |A|^2(p_k, t_k)\left(T - \frac{1}{k} - t_k\right) \geq |A|^2(\bar{p}, \bar{t})\left(T - \frac{1}{k} - \bar{t}\right) > m\,.$$

Since m was chosen arbitrarily, it follows that $T_k \to \infty$ and thus also $\lambda_k = |A|^2(p_k, t_k) \to \infty$ for $k \to \infty$. Since $t_k \nearrow T$, we conclude that $\alpha_k = -\lambda_k^2 t_k \to -\infty$ for $k \to \infty$. For the curvature estimate, it again follows from Definition 6.1 that

$$|A|^2(p, t)\left(T - \frac{1}{k} - t\right) \leq |A|^2(p_k, t_k)\left(T - \frac{1}{k} - t_k\right) = T_k \tag{6.1}$$

for all $p \in M^n$, $t \in [0, T - 1/k]$ and $k \in \mathbb{N}$. Let $\varepsilon > 0$ and $\bar{T} > 0$ be given. Since $T_k \to \infty$, there exists again $k_1 \in \mathbb{N}$ so that, for all $k \geq k_1$, $\bar{T} < T_k$ and

$$0 < \frac{\bar{T}}{T_k - \bar{T}} < \varepsilon\,.$$

For $\tau \in [\alpha_k, \bar{T}]$, it is $t := t_k + \tau/\lambda_k^2 \in [0, T - 1/k)$, and we can use the scaling behaviour of the curvature and (6.1) to estimate

$$\begin{aligned}|A_k|^2(p, \tau) &= \frac{1}{\lambda_k^2}|A|^2\left(p, t_k + \frac{\tau}{\lambda_k^2}\right) \leq \frac{T - 1/k - t_k}{T - 1/k - (t_k + \tau/\lambda_k^2)} \\ &= \frac{T_k}{T_k - \tau} \leq \frac{T_k}{T_k - \bar{T}} = 1 + \frac{\bar{T}}{T_k - \bar{T}} < 1 + \varepsilon\end{aligned}$$

for all $p \in M^n$ and $k \geq k_1$. Hence,

$$\max_{M^n \times [\alpha_k, \bar{T}]} |A_k|^2 < 1 + \varepsilon$$

for all $k \geq \max\{k_0, k_1\}$. □

Theorem 6.3. *Let $(M_t)_{t\in[0,T)}$ be a smooth, immersed solution of* (MCF) *with $T < \infty$. For the type-II rescaling 6.1 in case of a type-II singularity, there exists a sequence of rescaled immersions*

$$\Big(\big(M_\tau^k\big)_{\tau\in[\alpha_k, T_k]}\Big)_{k\in\mathbb{N}}$$

that converges for $k \to \infty$ along a subsequence, uniformly and smoothly on compact subsets of $\mathbb{R}$ and $\mathbb{R}^{n+1}$ to a maximal, smooth limit solution $(M_\tau^\infty)_{\tau\in\mathbb{R}}$ which satisfies again (MCF) *and*

$$0 \in M_0^\infty \qquad \textit{and} \qquad \sup_{\mathbb{R}\times\mathbb{R}} |A_\infty| = |A_\infty(0)| = 1\,.$$

Moreover, if $(M_t)_{t\in[0,T)}$ is embedded, then $(M_\tau^\infty)_{\tau\in(-\infty,0)}$ is embedded.

Proof. The convergence follows from Theorem 4.10. Lemma 6.2 implies $0 \in M_0^\infty$ and $|A_\infty(0)| = 1$ and that for any $\varepsilon > 0$ and any $\bar{T} > 0$,

$$\sup_{\mathbb{R}\times(-\infty,\bar{T}]} |A_\infty|^2 \leq 1 + \varepsilon\,.$$

Sending $\bar{T} \to \infty$ and $\varepsilon \to 0$ yields

$$\sup_{\mathbb{R}\times\mathbb{R}} |A_\infty| \leq 1 = |A_\infty(0)|\,.$$

By Proposition 1.9, M_τ^k is embedded for all $k \in \mathbb{N}$ and all $\tau \in [\alpha_k, T_k]$. Furthermore,

$$d_k(\tau) \geq \min\left\{d_k(\alpha_k), \frac{\sin(\varepsilon)}{m_k(\tau)}\right\} \geq \min\{\lambda_k d(0), \sin(\varepsilon)\}$$

is uniformly bounded in k for $\tau \in \mathbb{R}$. □

Remark 6.4. In the following chapters, we will show that the eternal solution obtained in Theorem 6.3 is convex and translating.

7. Convex hypersurfaces

Theorem 7.1 (Huisken, [Hui84, Corollary 4.2])**.** *Assume $M_0 = X_0(M)$ closed and convex, i.e. $h_{ij} \succeq 0$. Then $h_{ij} \succ 0$ for all $t \in (0, T)$.*

Proof. By Lemma 1.4 and Simons' identity (A.1),

$$\partial_t h_{ij} = \Delta h_{ij} - 2H g^{km} h_{ik} h_{jm} + |A|^2 h_{ij} \,.$$

Use Theorem D.5 for $m_{ij} = h_{ij}$, $u^k \equiv 0$ and $b_{ij} = -2H h_{il} g^{lm} h_{mj} + |A|^2 h_{ij}$. □

Corollary 7.2. *There is some $\varepsilon > 0$ such that $h_{ij} \succeq \varepsilon H g_{ij}$ holds on $M \times (0, T)$.*

Theorem 7.3 (Huisken, [Hui84, Theorem 4.3])**.** *If $\varepsilon H g_{ij} \preceq h_{ij} \preceq \beta H g_{ij}$, and $H > 0$ at $t = 0$ for some constants $0 < \varepsilon \leq 1/n < \beta < 1$, then this remains so on $(0, T)$.*

Proof. To prove the first inequality, we want to apply Theorem D.5 with

$$m_{ij} = \frac{h_{ij}}{H} - \varepsilon g_{ij} \,, \quad u^k = \frac{2}{H} g^{kl} \nabla_l H \,, \quad b_{ij} = 2\varepsilon H h_{ij} - 2h_{im} g^{ml} h_{lj} \,.$$

With this choice the evolution equation in Theorem D.5 is satisfied since

$$\partial_t \left(\frac{h_{ij}}{H} \right) = \frac{1}{H^2} \left(H \Delta h_{ij} - h_{ij} \Delta h_{ij} \right) - 2h_{im} g^{ml} h_{mj}$$

and

$$\Delta \left(\frac{h_{ij}}{H} \right) = \frac{1}{H^2} \left(H \Delta h_{ij} - h_{ij} \Delta h_{ij} \right) - \frac{2}{H} g^{kl} \nabla_k H \nabla_l \left(\frac{h_{ij}}{H} \right) .$$

It remains to check that b_{ij} is nonnegative on the null-eigenvectors of m_{ij}. Assume that, for some vector v,

$$h_{ij} v^j = \varepsilon H v_i \,.$$

Then we derive

$$b_{ij} v^i v^j = 2\varepsilon H h_{ij} v^i v^j - 2h_{im} g^{mI} h_{lj} v^i v^j = 2\varepsilon^2 H^2 |v|^2 - 2\varepsilon^2 H^2 |v|^2 = 0 \,.$$

That the second inequality remains true follows in the same way after reversing signs. □

Theorem 7.4 (Huisken [Hui84])**.** *Let $n \geq 2$ and $M_0 \subset \mathbb{R}^{n+1}$ be closed, convex and embedded. Then the mean curvature flow $(M_t)_{t \in [0,T)}$ starting at M_0 converges to a round point.*

Proof. See [Man11, Theorem 3.4.10]. Let T be the maximal time of smooth existence of the mean curvature flow of an n-dimensional convex hypersurface. By Theorems 1.10, 7.1 and 7.3, we have that after any positive time $H > 0$ and there exists $\varepsilon > 0$, independent of time, such that $h_{ij} \succeq \varepsilon H g_{ij}$. If at time T we have a type-II singularity, we get an unbounded, eternal convex blow-up limit flow with $H \geq 0$, using Hamilton's procedure. By the strong maximum principle, actually $H > 0$ for every time (otherwise $H \equiv 0$, but this and the convexity would imply that the limit flow is simply a fixed hyperplane) and the condition $h_{ij} \succeq \varepsilon H g_{ij}$ passes to the limit. Then, by Theorem 3.6, all the hypersurfaces of the limit flow are compact, in contradiction with the unboundedness, hence type-II singularities cannot develop. Dealing with type-I singularities, any blow-up limit is embedded, strictly convex and compact, again by this theorem. Hence, by Theorem 5.7 it can be only the sphere $\mathbb{S}^n$. This implies that the full sequence of rescaled hypersurfaces converges in C^∞ to such sphere. Finally, as the blow-up limit is unique and compact, the original hypersurface shrinks to a point in finite time. □

Remark 7.5 (Exponential convergence, [Hui84, Lemma 10.6])**.** Consider the normalized flow

$$\tilde{X}(\cdot,t) = \psi(t)X(\cdot,t)$$

where ψ is chosen so that

$$\int_{\tilde{M}_t} d\tilde{\mu} = |M_0|$$

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

$$\tilde{t}(t) = \int_0^t \psi^2(\tau)\, d\tau\,,$$

we get that $\tilde{g}_{ij} = \psi^2 g_{ij}$, $\tilde{H} = \psi^{-1}H$,

$$\psi^{-1}\partial_t\psi = \frac{\int_{\tilde{M}_t} H^2\, d\tilde{\mu}}{n\int_{\tilde{M}_t} d\tilde{\mu}} =: \frac{h}{n} = \psi^{-2}\frac{\tilde{h}}{n}$$

and

$$\partial_{\tilde{t}}\tilde{X} = \psi^{-2}\partial_t\tilde{X} = -\tilde{H}\tilde{\boldsymbol{\nu}} + \frac{\tilde{h}}{n}\tilde{X}$$

for $\tilde{t} \in [0,\infty)$. Then there exist constants $\delta > 0$ and $C, C_m < \infty$ such that

$$\tilde{H}_{\max} - \tilde{H}_{\min} \leq Ce^{-\delta\tilde{t}}\,,$$

$$\left|\tilde{h}_{ij}\tilde{H} - \frac{\tilde{h}}{n}\tilde{g}_{ij}\right| \leq Ce^{-\delta\tilde{t}}\,,$$

$$\max_{\tilde{M}}\left|\nabla^m\tilde{A}\right| \leq C_m e^{-\delta\tilde{t}}$$

for all $m > 0$.

8. Hamilton's Harnack Inequality

We follow [Urb91, Section 2], [And94] and [Sch17d, Chapter 4]. For convex hypersurfaces, the initial value problem (MCF) can be reduced to an initial value problem for the support function. Let M be a smooth, closed, stricly convex hypersurface (A is positive definite everywhere). Recall the Gauss map $\boldsymbol{\nu} : M^n \to \mathbb{S}^n$, unit normal $\bar{\boldsymbol{\nu}} : M^n \to \mathbb{R}^{n+1}$ and the Weingarten map $S : TM^n \to TM^n$ which gives the rate of change in the direction of the normal along the surface with

$$S(v) := dX^{-1}\big(D_{dX(v)}\bar{\boldsymbol{\nu}}\big) = dX^{-1}(d_v\boldsymbol{\nu})\,.$$

The second fundamental form A is the symmetric tensor given by the normal component of the connection on $\mathbb{R}^{n+1}$.

$$\begin{aligned} A(u,v) &= -\langle d^2X(v,w), \bar{\boldsymbol{\nu}}\rangle = -\langle D_{dX(v)}dX(w), \bar{\boldsymbol{\nu}}\rangle \\ &= \langle dX(w), D_{dX(v)}\bar{\boldsymbol{\nu}}\rangle = g(w, S(v)) \end{aligned}$$

for all $v, w \in TM^n$, where $dX : TM^n \to \mathbb{R}^{n+1}$. The eigenvalues $\lambda_1 \ldots \lambda_n$ of S are called the principal curvatures. Without loss of generality, we may assume that M encloses the origin. All information about the hypersurface is contained in the support function $s : M^n \to \mathbb{R}$ where

$$s(p) := \langle \bar{\boldsymbol{\nu}}(p), X(p)\rangle\,.$$

For strictly convex hypersurfaces $\boldsymbol{\nu}$ is a global diffeomorphism, and we can parametrise the hypersurface by $\tilde{X} : \boldsymbol{\nu}(M^n) \subset \mathbb{S}^n \to \mathbb{R}^{n+1}$ where

$$\tilde{X}(z) := X(\boldsymbol{\nu}^{-1}(z))$$

for all $z \in \boldsymbol{\nu}(M^n)$. We will consider the support function

$$s(z) := \langle \bar{z}, \tilde{X}(z)\rangle\,. \tag{8.1}$$

In the following, Indetify $\bar{z}$ with z. If the support function is known, the hypersurface is given as the boundary of the convex region

$$\bigcap_{z\in\mathbb{S}^n} \left\{y \in \mathbb{R}^{n+1} \,|\, \langle y, z\rangle \leq s(z)\right\} .$$

Let σ_{ij} be the metric and $\tilde{\nabla}$ be the gradient on $\mathbb{S}^n$. Differentiating (8.1) we obtain

$$\tilde{\nabla}_i s = \langle \tilde{\nabla}_i \tilde{X}, z\rangle + \langle \tilde{X}, \tilde{\nabla}_i z\rangle = \langle \tilde{X}, \tilde{\nabla}_i z\rangle ,$$

since $\tilde{\nabla}_i \tilde{X}(z)$ is tangential to M at $\tilde{X}(z)$, and z is the normal to M at $\tilde{X}(z)$. Since $\langle z, z\rangle = 1$, we obtain

$$\langle z, \tilde{\nabla}_i z\rangle = 0$$

and writing $\tilde{\nabla}_{ij} := \tilde{\nabla}_i \tilde{\nabla}_j$, we obtain

$$\langle z, \tilde{\nabla}_{ij} z\rangle = -\langle \tilde{\nabla}_j z, \tilde{\nabla}_i z\rangle = -\sigma_{ij} ,$$

Hence,

$$\begin{aligned}\tilde{X} &= \langle \tilde{X}, z\rangle z + \sigma^{ij} \langle \tilde{X}, \tilde{\nabla}_i z\rangle \tilde{\nabla}_j z \\ &= sz + \sigma^{ij} \tilde{\nabla}_i s \nabla_j z = sz + \tilde{\nabla} s\end{aligned}$$

From this, we conclude at a fixed point

$$\begin{aligned}\tilde{\nabla}_i \tilde{X} &= \tilde{\nabla}_i s z + s\tilde{\nabla}_i z + \tilde{\nabla}_{ki} s \sigma^{kl} \tilde{\nabla}_l z + \tilde{\nabla}_k s \sigma^{kl} \tilde{\nabla}_{li} z \\ &= \tilde{\nabla}_i s z + s\tilde{\nabla}_i z + \tilde{\nabla}_{ki} s \sigma^{kl} \tilde{\nabla}_l z - \tilde{\nabla}_k s \sigma^{kl} \sigma_{li} z \\ &= s\tilde{\nabla}_i z + \tilde{\nabla}_{ki} s \sigma^{kl} \tilde{\nabla}_l z\end{aligned}$$

and

$$\begin{aligned}\tilde{\nabla}_{ij} \tilde{X} &= \tilde{\nabla}_j s \tilde{\nabla}_i z - s\sigma_{ij} z + \tilde{\nabla}_{kij} s \sigma^{kl} \tilde{\nabla}_l z - \tilde{\nabla}_{ki} s \sigma^{kl} \sigma_{lj} z \\ &= \tilde{\nabla}_j s \tilde{\nabla}_i z - s\sigma_{ij} z + \tilde{\nabla}_{kij} s \sigma^{kl} \tilde{\nabla}_l z - \tilde{\nabla}_{ij} s z\end{aligned}$$

so that

$$\tilde{h}_{ij} = -\langle \tilde{\nabla}_{ij} \tilde{X}, z\rangle = s\sigma_{ij} + \tilde{\nabla}_{ij} s$$

and

$$\tilde{g}_{ij} = s^2 \sigma_{ij} + 2s\tilde{\nabla}_{ij} s + \tilde{\nabla}_{ik} s \sigma^{kl} \tilde{\nabla}_{jl} s = \tilde{h}_{ik} \sigma^{kl} \tilde{h}_{lj}$$

as well as

$$\tilde{h}_i^j = \tilde{g}^{jk} \tilde{h}_{ik} = \tilde{a}^{jl} \sigma_{lm} \tilde{a}^{mk} \tilde{h}_{ik} = \sigma_{il} \tilde{a}^{lj}$$

where here $(\tilde{a}^{ij})_{ij} = \big((\tilde{h}_{ij})_{ij}\big)^{-1}$ and

$$\tilde{H} = \tilde{h}_i^i = \sigma_{ij} \tilde{a}^{ij}$$

We consider the Weingarten map $\tilde{S} : T\mathbb{S}^n \to T\mathbb{S}^n$ with

$$\tilde{S}(v) := d\tilde{X}^{-1}(d_v \tilde{\boldsymbol{\nu}}) .$$

Since $d\tilde{\boldsymbol{\nu}} = \mathrm{id}$, we have $\tilde{S}^{-1} = d\tilde{X}$. We define

$$\tilde{S}^{-1}(v) = (\sigma^* \tilde{\nabla}^2 s + s\,\mathrm{id})(v) = \tilde{\nabla}_v(\tilde{\nabla} s) + s\,\mathrm{id}(v) =: \mathcal{A}(v) \tag{8.2}$$

so that

$$\begin{aligned}\tilde{g}(u, v) &= \tilde{g}_{ij} v^i w^j = \tilde{h}_{ik} \sigma^{kl} \tilde{h}_{lj} v^i w^j = \sigma_{km} \tilde{a}_i^m \sigma^{kl} \sigma_{ln} \tilde{a}_j^n v^i w^j \\ &= \sigma_{km} \tilde{a}_i^m \tilde{a}_j^k v^i w^j = \sigma(\mathcal{A}(u), \mathcal{A}(v)) .\end{aligned}$$

The great advantage of the support function is that it allows us to consider a family of convex hypersurfaces simply as an evolving scalar function defined on the sphere. This makes things much simpler than the more abstract framework allowing arbitrary parametrizations, since we no longer have different descriptions of the same hypersurface. Furthermore, the identification with the sphere provides a time-independent metric and connection, which vastly simplifies many calculations, including especially those presented here for the proof of the Harnack inequalities.

For the remainder of this section, we consider a familiy of embeddigns $X : M^n \times [0,T) \to \mathbb{R}^{n+1}$ that solve the initial value problem

$$\begin{cases} \partial_t X(p,t) = -F\big(S(p,t), \boldsymbol{\nu}(p,t)\big)\boldsymbol{\nu}(p,t) & \text{for } (p,t) \in M^n \times [0,T) \\ X(\cdot,0) = X_0 & \text{on } M^n\,. \end{cases} \tag{8.3}$$

where F is such that the equation is parabolic and invariant under diffeomorphisms of M^n and translations in space and time. We want to reduce (8.3) to an initial value problem for the support function. Let X be a solution of (8.3), and suppose that for each $t \in [0,T)$, $X(\cdot,t)$ is a parametrization of a smooth, closed, uniformly convex hypersurface M_t. We define a new parametrization $\tilde{X}(\cdot,t)$ by

$$\tilde{X}(z,t) = X\big(\boldsymbol{\nu}_t^{-1}(z),t\big)\ .$$

Then

$$\partial_t \tilde{X} = \partial_i X \partial_t \big(\boldsymbol{\nu}_t^{-1}\big)^i + \partial_t X = \partial_i X \partial_t \big(\boldsymbol{\nu}_t^{-1}\big)^i - \tilde{F} z$$

so that

$$\partial_t s = \langle \partial_t \tilde{X}, z\rangle = -\tilde{F}$$

since $\partial_i X$ is tangential. This proves the following theorem:

Theorem 8.1 (Andrews, [And94, Theorem 3.1])**.** *Let $X : M^n \times [0,T) \to \mathbb{R}^{n+1}$ be a family of strictly convex immersions satisfying* (8.3)*. Then*

$$\begin{cases} \partial_t s(z,t) = \Phi(\mathcal{A}[s(z,t)], z) & \textit{on } \mathbb{S}^n \times [0,T) \\ s(\cdot,0) = s_0 & \textit{on } \mathbb{S}\,. \end{cases} \tag{8.4}$$

where id *is the identity matrix, s_0 is the support function of M_0,*

$$\Phi(\mathcal{A}) = -\operatorname{tr}_\sigma \mathcal{A}^{-1} \qquad \textit{and} \qquad \mathcal{A} = \sigma^* \tilde{\nabla}^2 s + \mathrm{id}\, s\,.$$

The expression (8.2) allows us to use the support function to calculate functions of the curvature of a hypersurface. We can define $\Phi : U \subset T^*\mathbb{S}^n \to \mathbb{R}$ in terms of X by

$$\Phi(X) = -\tilde{F}(\tilde{X}^{-1})$$

for all positive definite maps X. Furthermore, $\dot{\Phi}(\mathcal{A}) : T^*\mathbb{S}^n \to T\mathbb{S}^n$ is given by

$$\dot{\Phi}(\mathcal{A})(\mathcal{B}) = \partial_r|_{r=0}\Phi(\mathcal{A} + r\mathcal{B})$$

and $\ddot{\Phi}(\mathcal{A}) : T\mathbb{S}^n \otimes T^*\mathbb{S}^n \to T\mathbb{S}^n \otimes T^*\mathbb{S}^n$ by

$$\ddot{\Phi}(\mathcal{A})(\mathcal{B},\mathcal{C}) = \partial_r|_{r=0}\dot{\Phi}(\mathcal{A} + r\mathcal{C})(\mathcal{B})\,.$$

We call Φ concave (convex), if

$$\ddot{\Phi}(\mathcal{A})(\mathcal{B},\mathcal{B}) \le (\ge)\ 0$$

for all $\mathcal{A},\mathcal{B} \in T^*\mathbb{S}^n$. We call Φ α-concave (α-convex), if

$$\Phi = \operatorname{sign} \alpha B^\alpha\,,$$

where B is positive and concave (convex), $\alpha \in \mathbb{R}$. α-concavity (-convexity) is equivalent to

$$\ddot{\Phi} = \alpha(\alpha-1)B^{\alpha-2}\dot{B} \otimes \dot{B} + \alpha B^{\alpha-1}\ddot{B} \preceq (\succeq)\ \frac{\alpha-1}{\alpha\Phi}\dot{\Phi} \otimes \dot{\Phi}\,. \tag{8.5}$$

(These conditions become considerably more complicated when written in terms of the principal curvatures and a speed function F. For example, concavity of Φ, becomes $\ddot{F}(X,X) + 2\dot{F}(X \circ S^{-1} \circ X) \ge 0$.)

Lemma 8.2 (Andrews, [And94, Theorem 3.6 and Lemma 5.1])**.** *The following evolution equations hold under the Gauss map parametrization of the flow* (8.3)*:*

$$\begin{aligned}\partial_t(\tilde\nabla^2 s + s\sigma) &= \tilde\nabla^2\Phi + \Phi\sigma\\ \partial_t\mathcal{A} &= \sigma^*\tilde\nabla^2\Phi + \Phi\,\mathrm{id}\\ \partial_t\Phi(\mathcal{A}) &= \dot\Phi(\mathcal{A})(\sigma^*\tilde\nabla^2\Phi) + \dot\Phi(\mathcal{A})(\mathrm{id})\Phi \qquad (8.6)\\ \partial_t^2\Phi(\mathcal{A}) &= \ddot\Phi(\mathcal{A})(\partial_t\mathcal{A},\partial_t\mathcal{A}) + \dot\Phi(\mathcal{A})(\sigma^*\tilde\nabla^2\partial_t\Phi) + \dot\Phi(\mathcal{A})(\mathrm{id})\partial_t\Phi\,. \qquad (8.7)\end{aligned}$$

Proof. The first equation follows simply by differentiating (8.4), since the metric σ and connection $\tilde\nabla$ are independent of time. The second follows immediately from this. Since Φ depends only on $\mathcal{A}$, we have $\partial_t\Phi = \dot\Phi(\partial_t\mathcal{A})$ which implies the third equation. By (8.6),

$$\begin{aligned}\partial_t^2\Phi &= \partial_t\Big(\dot\Phi(\sigma^*\tilde\nabla^2\Phi) + \dot\Phi(\mathrm{id})\Phi\Big)\\ &= \ddot\Phi(\partial_t\mathcal{A},\sigma^*\tilde\nabla^2\Phi) + \dot\Phi(\sigma^*\tilde\nabla^2\partial_t\Phi) + \ddot\Phi(\partial_t\mathcal{A},\mathrm{id})\Phi + \dot\Phi(\mathrm{id})\partial_t\Phi\\ &= \ddot\Phi(\partial_t\mathcal{A},\partial_t\mathcal{A}) + \dot\Phi(\sigma^*\tilde\nabla^2\partial_t\Phi) + \dot\Phi(\mathrm{id})\partial_t\Phi\,.\end{aligned}$$

□

Lemma 8.3 (Andrews, [And94, Lemma 3.10])**.** *Let* $f : M^n \times [0,T) \to \mathbb{R}$ *and* $\tilde f : \mathbb{S}^n \times [0,T) \to \mathbb{R}$ *be related by*

$$\tilde f(\boldsymbol{\nu}(p,t),t) = f(p,t)$$

for all $p \in M^n$ *and* $t \in [0,T)$*. Then*

$$\partial_t f = \partial_t\tilde f + A^{-1}(\nabla F, \nabla f)\,.$$

Proof. Differentiating yields

$$\begin{aligned}\partial_t f &= \partial_t\tilde f + \partial_{z_i}\tilde f\partial_t\boldsymbol{\nu}^i = \partial_t\tilde f + \partial_{p_j} f\partial_{z_i}(\boldsymbol{\nu}^{-1})^j\partial_{p^i}F\\ &= \partial_t\tilde f + g_{jk}\partial_{p^k} f a_i^j\partial_{p^i}F = \partial_t\tilde f + a_{ij}\partial_{p^i} f\partial_{p^i}F\,,\end{aligned}$$

where $(a^{ij})_{ij} = \big((h_{ij})_{ij}\big)^{-1}$. □

Theorem 8.4 (Andrews, [And94, Theorem 5.6])**.** *Let* X *be a strictly convex solution to* (8.3)*.*

(i) If Φ *is* α*-concave for* $0 < \alpha < 1$ *(*α*-convex for* $\alpha > 1$*), then*

$$\partial_t\Phi + \frac{\alpha\Phi}{(\alpha-1)t} \le (\ge)\, 0\,.$$

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

(ii) If Φ*, is positive and concave (convex), then*

$$\sup_{\mathbb{S}^n}(\partial_t\log\Phi) \qquad \textit{is decreasing (increasing).}$$

Proof. We prove the concave cases. For claim (ii), let Φ be concave and set $R := \partial_t\log\Phi$. Then

$$\partial_t R = \partial_t\left(\frac{\partial_t\Phi}{\Phi}\right) = \frac{\partial_t^2\Phi}{\Phi} - \frac{(\partial_t\Phi)^2}{\Phi^2}$$

as well as

$$\tilde\nabla R = \tilde\nabla\left(\frac{\partial_t\Phi}{\Phi}\right) = \frac{\tilde\nabla\partial_t\Phi}{\Phi} - \frac{\partial_t\Phi\tilde\nabla\Phi}{\Phi^2}$$

and

$$\begin{aligned}\tilde{\nabla}^2 R &= \tilde{\nabla}\left(\frac{\tilde{\nabla}\partial_t\Phi}{\Phi} - \frac{\partial_t\Phi\tilde{\nabla}\Phi}{\Phi^2}\right)\\ &= \frac{\tilde{\nabla}^2\partial_t\Phi}{\Phi} - 2\frac{\tilde{\nabla}\partial_t\Phi\otimes\tilde{\nabla}\Phi}{\Phi^2} - \frac{\partial_t\Phi\tilde{\nabla}^2\Phi}{\Phi^2} + 2\frac{\partial_t\Phi(\tilde{\nabla}\Phi)^2}{\Phi^3}\\ &= \frac{\tilde{\nabla}^2\partial_t\Phi}{\Phi} - 2\frac{\tilde{\nabla}R\otimes\tilde{\nabla}\Phi}{\Phi} - \frac{\partial_t\Phi\tilde{\nabla}^2\Phi}{\Phi^2}\,.\end{aligned}$$

By (8.6) and (8.7),

$$\begin{aligned}\partial_t R &= \frac{1}{\Phi}\left(\dot{\Phi}(\sigma^*\tilde{\nabla}^2\partial_t\Phi) + \dot{\Phi}(\mathrm{id})\partial_t\Phi + \ddot{\Phi}(\partial_t\mathcal{A},\partial_t\mathcal{A})\right) - \frac{R}{\Phi}\left(\dot{\Phi}(\sigma^*\tilde{\nabla}^2\Phi) + \dot{\Phi}(\mathrm{id})\Phi\right)\\ &\le \dot{\Phi}(\sigma^*\tilde{\nabla}^2 R) + \frac{2}{\Phi}\dot{\Phi}\Big(\sigma^*\Big(\tilde{\nabla}\Phi\otimes\tilde{\nabla}R\Big)\Big) + \frac{1}{\Phi^2}\dot{\Phi}(\sigma^*(\partial_t\Phi\tilde{\nabla}^2\Phi))\\ &\quad + \frac{1}{\Phi}\dot{\Phi}(\mathrm{id})\partial_t\Phi - \frac{R}{\Phi}\left(\dot{\Phi}(\sigma^*\tilde{\nabla}^2\Phi) + \dot{\Phi}(\mathrm{id})\Phi\right)\\ &= \dot{\Phi}(\sigma^*\tilde{\nabla}^2 R) + \frac{2}{\Phi}\dot{\Phi}\Big(\sigma^*\Big(\tilde{\nabla}\Phi\otimes\tilde{\nabla}R\Big)\Big)\,.\end{aligned}$$

The strong parabolic maximum principle, Theorem D.3, implies (ii), since the first term is an elliptic operator, and the second a gradient term. For claim (i), let Φ be α-concave with $\alpha < 1$ and set

$$R := t\partial_t\Phi + \frac{\alpha\Phi}{\alpha-1}\,,$$

which is negative at $t = 0$. Then

$$\partial_t R = t\partial_t^2\Phi + \frac{2\alpha-1}{\alpha-1}\partial_t\Phi$$

as well as

$$\tilde{\nabla}R = t\tilde{\nabla}\partial_t\Phi + \frac{\alpha}{\alpha-1}\tilde{\nabla}\Phi$$

and

$$\tilde{\nabla}^2 R = t\tilde{\nabla}^2\partial_t\Phi + \frac{\alpha}{\alpha-1}\tilde{\nabla}^2\Phi\,.$$

By (8.6), (8.7) and (8.5),

$$\begin{aligned}\partial_t R &= t\left(\dot{\Phi}(\sigma^*\tilde{\nabla}^2\partial_t\Phi) + \dot{\Phi}(\mathrm{id})\partial_t\Phi + \ddot{\Phi}(\partial_t\mathcal{A},\partial_t\mathcal{A})\right) + \frac{2\alpha-1}{\alpha-1}\partial_t\Phi\\ &\le \dot{\Phi}(\sigma^*\tilde{\nabla}^2 R) - \frac{\alpha}{\alpha-1}\dot{\Phi}(\sigma^*\tilde{\nabla}^2\Phi) + t\dot{\Phi}(\mathrm{id})\partial_t\Phi\\ &\quad + t\frac{\alpha-1}{\alpha\Phi}(\dot{\Phi}(\partial_t\mathcal{A}))^2 + \frac{2\alpha-1}{\alpha-1}\partial_t\Phi\\ &= \dot{\Phi}(\sigma^*\tilde{\nabla}^2 R) + \frac{\alpha}{\alpha-1}\left(\dot{\Phi}(\mathrm{id})\Phi - \partial_t\Phi\right) + t\dot{\Phi}(\mathrm{id})\partial_t\Phi\\ &\quad + t\frac{\alpha-1}{\alpha\Phi}(\partial_t\Phi)^2 + \frac{2\alpha-1}{\alpha-1}\partial_t\Phi\\ &= \dot{\Phi}(\sigma^*\tilde{\nabla}^2 R) + \frac{\alpha}{\alpha-1}\dot{\Phi}(\mathrm{id})\Phi + t\dot{\Phi}(\mathrm{id})\partial_t\Phi + t\frac{\alpha-1}{\alpha\Phi}(\partial_t\Phi)^2 + \partial_t\Phi\\ &= \dot{\Phi}(\sigma^*\tilde{\nabla}^2 R) + \left(\frac{\alpha-1}{\alpha\Phi}\partial_t\Phi + \dot{\Phi}(\mathrm{id})\right)\left(t\partial_t\Phi + \frac{\alpha\Phi}{\alpha-1}\right)\\ &= \dot{\Phi}(\sigma^*\tilde{\nabla}^2 R) + \left(\frac{\alpha-1}{\alpha\Phi}\partial_t\Phi + \dot{\Phi}(\mathrm{id})\right)R\,.\end{aligned}$$

The weak parabolic maximum principle, Theorem D.2, implies that R stays negative as long as the solution exists. ☐

This calculation can easily be transferred to the standard parametrization, by writing the various quantities in terms of the metric and connection on the hypersurface. This is most easily done by considering the change in the evolution equations coming from the modified parametrization. Here we denote by A^{-1} the map inverse to A.

Corollary 8.5 (Andrews, [And94, Corollary 5.11])**.** *Let X be a strictly convex solution of $\partial_t X = -F\boldsymbol{\nu}$.*

(i) If Φ is α-concave for $\alpha < 1$ (α-convex for $\alpha > 1$), then

$$\partial_t F - A^{-1}(\nabla F, \nabla F) + \frac{\alpha F}{(\alpha - 1)t} \geq (\leq)\, 0\,.$$

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

(ii) If Φ is positive and concave (convex), then

$$\sup_{M^n} \left(\partial_t \log|F| - FA^{-1}(\nabla \log|F|, \nabla \log|F|)\right) \qquad \textit{is decreasing (increasing).}$$

Proof. The claim results from Lemma 8.3 and Theorem 8.4. □

Theorem 8.6 (Andrews, [And94, Theorem 5.17])**.** *Let X be a strictly convex solution of $\partial_t X = -F\boldsymbol{\nu}$. The following inequalities apply in the standard parametrization for the cases described, for any points $p_1, p_2 \in M^n$, any times $0 < t_1 < t_2 < T$, and any curve γ between (p_1, t_1) and (p_2, t_2).*

(i) If Φ is α-concave, $\alpha < 0$, then

$$\frac{F(p_2, t_2)}{F(p_1, t_1)} \geq \left(\frac{t_1}{t_2}\right)^{\alpha/(\alpha-1)} \exp\left(-\frac{1}{4}\int_\gamma F^{-1} A(\dot\gamma, \dot\gamma)\, dt\right)\,.$$

(ii) If Φ is α-convex, $\alpha > 1$, then

$$\frac{F(p_2, t_2)}{F(p_1, t_1)} \geq \left(\frac{t_1}{t_2}\right)^{\alpha/(\alpha-1)} \exp\left(-\frac{1}{4}\int_\gamma |F|^{-1} A(\dot\gamma, \dot\gamma)\, dt\right)\,.$$

(iii) If Φ is convex and positive, then

$$\frac{F(p_2, t_2)}{F(p_1, t_1)} \geq \exp(-C(t_2 - t_1)) \exp\left(-\frac{1}{4}\int_\gamma |F|^{-1} A(\dot\gamma, \dot\gamma)\, dt\right)\,,$$

where $C = \lim_{t \searrow 0} \sup_{M^n} \left(\partial_t \log|F| - FA^{-1}(\nabla \log|F|, \nabla \log|F|)\right),$.

Proof. Along a curve γ,

$$D_{\dot\gamma} \log F = \partial_t \log F + \langle \dot\gamma, \nabla \log F\rangle\,.$$

Furthermore,

$$\langle \dot\gamma, \nabla F\rangle \leq A^{-1}(\nabla F, \nabla F) + \frac{1}{4} A(\dot\gamma, \dot\gamma)$$

so that, by Corollary 8.5(i),

$$\begin{aligned} D_{\dot\gamma} \log F &\geq FA^{-1}(\nabla \log F, \nabla \log F) + \langle \dot\gamma, \nabla \log F\rangle - \frac{\alpha}{(\alpha - 1)t} \\ &\geq -\frac{1}{4} F^{-1} A(\dot\gamma, \dot\gamma) - \frac{\alpha}{(\alpha - 1)t}\,. \end{aligned}$$

Integrating along γ yields claim (i). For claim (ii),

$$D_{\dot\gamma} \log F \geq C - \frac{1}{4} F^{-1} A(\dot\gamma, \dot\gamma) \geq -C - \frac{1}{4} F^{-1} A(\dot\gamma, \dot\gamma)$$

respectively,

$$D_{\dot\gamma} \log F \leq C + \frac{1}{4} F^{-1} A(\dot\gamma, \dot\gamma)\,.$$

□

Remark 8.7. For the mean curvature flow, we have

$$\Phi(\mathcal{A}) = -H(S) = -H(\mathcal{A}^{-1}) = \operatorname{sign}(-1)\left(H^{-1}(\mathcal{A}^{-1})\right)^{-1}.$$

Since $h_{ik}a^{jk} = \delta_i^j$ and $a^{kl} = g^{km}g^{ls}a_{ms}$, we have

$$\partial_{a_{\alpha\beta}} h_{ij} = -h_{ik}h_{jl}\partial_{a_{\alpha\beta}} a^{kl} = -h_{ik}h_{jl}g^{km}g^{ls}\delta_m^\alpha \delta_s^\beta = -h_i^\alpha h_j^\beta$$

and thus

$$\partial_{a_{\alpha\beta}} H = g^{ij}h_i^\alpha h_j^\beta$$

and

$$\partial_{a_{\alpha\beta}}\partial_{a_{\delta\gamma}} H = g^{ij}h_i^\delta h_k^\gamma g^{k\alpha}h_j^\beta + g^{ij}h_i^\alpha h_j^\delta h_k^\gamma g^{k\beta}.$$

This yields

$$\partial_{a_{\alpha\beta}} H^{-1} = -H^{-2}g^{ij}h_i^\alpha h_j^\beta.$$

Since

$$\partial_{a_{\alpha\beta}}\partial_{a_{\delta\gamma}} H^{-1} = 2H^{-2}\partial_{a_{\alpha\beta}} H\partial_{a_{\delta\gamma}} H - H^{-2}\partial_{a_{\alpha\beta}}\partial_{a_{\delta\gamma}} H,$$

the eigenvectors $\{v_i\}$ of $\nabla^2_{\mathcal{A}} H^{-1}$ are the eigenvectors of the Weingarten map S, and

$$\begin{aligned}\nabla^2_{\mathcal{A}} H^{-1}(v_i, v_i, v_i, v_i) &= 2H^{-3}g(S(v_i), S(v_i))\left(g(S(v_i), S(v_i)) - Hg(S(v_i), v_i)\right)\\ &= 2H^{-3}\kappa_i^3\left(\kappa_i - H\right)\end{aligned}$$

is negativ for convex flows. Hence, Φ is (-1)-concave.

Theorem 8.8 (Hamilton [Ham95b, Theorem 1.3])**.** *Let $X : M^n \times (-\infty, T) \to \mathbb{R}^{n+1}$ be an ancient mean curvature flow of a complete, strictly convex hypersurface with bounded second fundamental form at every time and such that H takes its maximum in space and time. Then, X is a translating flow.*

Proof. Define

$$Z := \partial_t H + \frac{H}{2(t-t_0)} - A^{-1}(\nabla H, \nabla H)$$

then

$$(\partial_t - \Delta)Z = 2g^{ij}a^{kl}J_{ik}J_{jl} + \left(|A|^2 - \frac{2}{t-t_0}\right)Z \geq \left(|A|^2 - \frac{2}{t-t_0}\right)Z$$

where

$$J_{ik} = \nabla^2_{ik}H + Hh^2_{ik} - a^{sr}\nabla_s H\nabla_r h_{ik} + \frac{h_{ik}}{2(t-t_0)}.$$

By Corollary 8.5 and Remark 8.7, $Z \geq 0$. On an eternal solution where H attains its maximum in space and time, we can send $t_0 \to -\infty$ and obtain $Z = 0$ at the maximum. By the strong maximum principle, $Z \equiv 0$ so that

$$\partial_t H = A^{-1}(\nabla H, \nabla H).$$

Since

$$g^{ik} = g^{kl}\delta_l^i = g^{kl}h_{jl}a^{ij} = h_j^k a^{ij}$$

and, by Codazzi and $a^{ij}h_{jk} = \delta_k^i$,

$$\begin{aligned}a^{il}\nabla_l H &= a^{il}g^{km}\nabla_l h_{km} = a^{il}g^{km}\nabla_k h_{lm}\\ &= -a^{il}g^{km}h_{ls}h_{mj}\nabla_k a^{sj} = -h_j^k\nabla_k a^{ij},\end{aligned}$$

we obtain

$$\begin{aligned}0 &= -a^{il}\nabla_l H\nabla_i H + \Delta H + H|A|^2\\ &= \left(\nabla_k a^{ij}\nabla_i H + a^{ij}\nabla_k\nabla_i H + Hh_k^j\right)h_j^k.\end{aligned}$$

Consider the vector

$$V = a^{ij}\nabla_i H\nabla_j X + H\boldsymbol{\nu}.$$

Since

$$\nabla_k \nabla_j X = \langle \partial_k \partial_j X, \boldsymbol{\nu} \rangle \boldsymbol{\nu} = -h_{jk} \boldsymbol{\nu}$$

and

$$\nabla_k \boldsymbol{\nu} = h_k^j \nabla_j X = g^{ij} h_{ik} \nabla_j X$$

as well as $a^{ij} h_{jk} = \delta_k^i$, we obtain

$$\begin{aligned}\nabla_k V &= \nabla_k a^{ij} \nabla_i H \nabla_j X + a^{ij} \nabla_k \nabla_i H \nabla_j X + a^{ij} \nabla_i H \nabla_k \nabla_j X + \nabla_k H \boldsymbol{\nu} + H \nabla_k \boldsymbol{\nu} \\ &= \left(\nabla_k a^{ij} \nabla_i H + a^{ij} \nabla_k \nabla_i H + H h_k^j \right) \nabla_j X + \left(\nabla_k H - a^{ij} \nabla_i H h_{jk} \right) \boldsymbol{\nu} = 0 \,.\end{aligned}$$

On the other hand, at a fixed point so that the Christoffel symbols vanish,

$$\begin{aligned}\partial_t a^{ij} &= -a^{ik} a^{jl} \partial_t h_{kl} = -a^{ik} a^{jl} \left(\nabla_k \nabla_l H - H g^{ms} h_{lm} h_{ks} \right) \\ &= -a^{ik} a^{jl} \nabla_k \nabla_l H + H g^{ij} \,.\end{aligned}$$

and

$$\partial_t \partial_i H = \partial_i \left(a^{kl} \partial_k H \partial_l H \right) = -H h_i^l \partial_l H + a^{kl} \partial_k H \partial_i \partial_l H$$

as well as

$$\partial_t \partial_j X = -\partial_j (H \boldsymbol{\nu}) = -\partial_j H \boldsymbol{\nu} - H h_j^k \partial_k X$$

and

$$\partial_t \boldsymbol{\nu} = g^{ij} \partial_i H \partial_j X$$

Together, we obtain,

$$\begin{aligned}\partial_t V &= \partial_t a^{ij} \partial_i H \partial_j X + a^{ij} \partial_t \partial_i H \partial_j X + a^{ij} \partial_i H \partial_t \partial_j X + \partial_t H \boldsymbol{\nu} + H \partial_t \boldsymbol{\nu} \\ &= \big(H g^{ij} \partial_i H - a^{ik} a^{jl} \nabla_k \nabla_l H \partial_i H - a^{ij} H h_i^l \partial_l H + a^{ij} a^{kl} \partial_k H \partial_i \partial_l H \\ &\qquad + H g^{ij} \partial_i H \big) \partial_j X - a^{ij} \partial_i H H h_j^k \partial_k X + \left(a^{kl} \partial_k H \partial_l H - a^{ij} \partial_i H \partial_j H \right) \boldsymbol{\nu} = 0 \,.\end{aligned}$$

Hence V is a constant vectorfield in space and time. Let $t_1 \in (-\infty, T)$ and $\phi : M^n \to M^n$ be a diffeomorphism with $\phi(\cdot, t_1) = \mathrm{id}$ and

$$\partial_t \phi = -a^{ij} \nabla_i H \nabla_j X$$

and $\tilde{X}(p,t) = X(\phi(p,t),t)$. By Theorem 1.3, $\tilde{X}(M^n,t) = X(\phi(M^n,t),t) = M_t$ and

$$\begin{aligned}\tilde{X}(p,t) - \tilde{X}(p,t_1) = X(\phi(p,t),t) - X(p,t_1) &= \int_{t_1}^t \langle DX, \partial_t \phi \rangle + \partial_t X \, d\tau \\ &= -\int_{t_1}^t a^{ij} \nabla_i H \nabla_j X + H \boldsymbol{\nu} \, d\tau = -(t - t_1) V\end{aligned}$$

so that $M_t = M_{t_1} - (t - t_1)V$ and the surfaces move by translation in direction of $-V$. □

9. Noncollapsing

We follow the lines of [And12].

Definition 9.1 (α-noncollapsed)**.** A mean convex hypersurface M bounding an open region Ω in $\mathbb{R}^n$ is α-noncollapsed (on the scale of the mean curvature) if for every $x \in M$ there is an open ball B of radius $\alpha / H(x)$ contained in Ω with $x \in \partial B$.

Note that every compact, smooth, strictly mean convex domain is α-Andrews for some $\alpha > 0$.

Given a hypersurface $M = X(M^n)$, define $Z : M^n \times M^n \to \mathbb{R}$ by

$$Z(p,q) = \frac{H(p)}{2} |X(q) - X(q)|^2 + \alpha \langle X(q) - X(p), \boldsymbol{\nu}(p) \rangle$$

Then we have the following characterization:

Lemma 9.2 (Andrews [And12, Proposition 2])**.** *M is α-noncollapsed if and only if $Z(p,q) \geq 0$ for all $p, q \in M^n$.*

Proof. A ball in Ω of radius $\alpha/H(p)$ with $X(p)$ as a boundary point must have centre at the point

$$z(p) = X(p) - \frac{\alpha}{H(p)}\boldsymbol{\nu}(p)\,.$$

The statement that this ball is contained in Ω is equivalent to the statement that no points of M are of distance less than $\alpha/H(p)$ from z, that is

$$0 \leq |X(q) - z(p)|^2 - \left(\frac{\alpha}{H(p)}\right)^2 = 2\frac{Z(p,q)}{H(p)}$$

for all $p, q \in M^n$. Since $H > 0$ this is equivalent to the statement that $Z \geq 0$ everywhere. If $Z \geq 0$, then by the same equation as above, yields the claim. $\square$

Theorem 9.3 (Andrews [And12, Theorem 3])**.** *Let M^n be a compact manifold, and $X : M^n \times [0,T) \to \mathbb{R}^{n+1}$ evolve by* (MCF) *with $H > 0$. If M_0 is α-noncollapsed for some $\alpha > 0$, then M_t α-noncollapsed for every $t \in [0,T)$.*

Proof. By PLemma 9.2, the claim is equivalent to the statement that the function $Z : M^n \times M^n \times [0,T) \to \mathbb{R}$ with

$$Z(p,q,t) = \frac{H(p,t)}{2}|X(q,t) - X(q,t)|^2 + \alpha\langle X(q,t) - X(p,t), \boldsymbol{\nu}(p,t)\rangle$$

is nonnegative everywhere provided that it is nonnegative on $M^n \times M^n \times \{0\}$. We prove this using the maximum principle. For convenience we denote $H_p = H(p,t)$ and $\boldsymbol{\nu}_p = \boldsymbol{\nu}(p,t)$ an define

$$d = |X(q,t) - X(p,t)| \quad \text{and} \quad w = \frac{X(q,t) - X(p,t)}{d}$$

so that

$$Z = d^2\frac{H_p}{2} + \alpha d\langle w, \boldsymbol{\nu}_p\rangle\,.$$

We compute the first and second derivatives of Z, with respect to some choices of local normal coordinates $\{p^i\}$ near p and $\{q^i\}$ near p. Then

$$\partial_{q_i} Z = dH_p\langle w, \partial_{q_i}\rangle + \alpha\langle\partial_{q_i}, \boldsymbol{\nu}_p\rangle \tag{9.1}$$

$$\partial_{p_i} Z = -dH_p\langle w, \partial_{p_i}\rangle + \frac{d^2}{2}\nabla_{p_i} H_p + \alpha d h_{ij}^p g_p^{jk}\langle w, \partial_{p_k}\rangle \tag{9.2}$$

$$\partial_{q_i}\partial_{q_j} Z = H_p\langle\partial_{q_i}, \partial_{q_j}\rangle - dH_p h_{ij}^q\langle w, \boldsymbol{\nu}_q\rangle - \alpha h_{ij}^q\langle\boldsymbol{\nu}_q, \boldsymbol{\nu}_p\rangle \tag{9.3}$$

$$\partial_{q_i}\partial_{p_j} Z = -H_p\langle\partial_{q_i}, \partial_{p_j}\rangle + d\nabla_{p_j} H_p\langle w, \partial_{q_i}\rangle + \alpha h_{jk}^p g_p^{kl}\langle\partial_{q_i}, \partial_{p_l}\rangle \tag{9.4}$$

$$\begin{aligned}\partial_{p_i}\partial_{p_j} Z = {} & H_p\langle\partial_{p_i}, \partial_{p_j}\rangle - d\nabla_{p_j} H_p\langle w, \partial_{p_i}\rangle + dH_p h_{ij}^p\langle w, \boldsymbol{\nu}_p\rangle \\ & - d\nabla_{p_i} H_p\langle w, \partial_{p_j}\rangle + \frac{d^2}{2}\nabla_{p_i}\nabla_{p_j} H_p \\ & + \alpha d\nabla_{p_j} h_{ik}^p g_p^{kl}\langle w, \partial_{p_l}\rangle - \alpha h_{ij}^p - \alpha d h_{ik}^p g^{kl} h_{jl}^p\langle w, \boldsymbol{\nu}_p\rangle\end{aligned} \tag{9.5}$$

$$\begin{aligned}\partial_t Z = {} & dH_p\langle w, -H_q\boldsymbol{\nu}_q + H_p\boldsymbol{\nu}_p\rangle + \frac{d^2}{2}\left(\Delta H_p + H_p|A_p|^2\right) \\ & + \alpha\langle -H_q\boldsymbol{\nu}_q + H_p\boldsymbol{\nu}_p, \boldsymbol{\nu}_p\rangle + \alpha d\langle w, \nabla H_p\rangle\,.\end{aligned} \tag{9.6}$$

Equation (9.1) yields

$$0 = \left\langle \partial_{q_i}, \boldsymbol{\nu}_p + \frac{dH_p}{\alpha}w\right\rangle - \frac{1}{\alpha}\partial_{q_i} Z = \left\langle \partial_{q_i}, \boldsymbol{\nu}_p + \frac{dH_p}{\alpha}w - \frac{1}{\alpha}\nabla_q Z\right\rangle\,.$$

Thus, the vector $\boldsymbol{\nu}_p+(dH_p/\alpha)w-(1/\alpha)\nabla_q Z$ is normal to the hypersurface at $X(q)$, and is a multiple of $\boldsymbol{\nu}_q$. Furthermore,

$$\begin{aligned}
&\left|\boldsymbol{\nu}_p+\frac{dH_p}{\alpha}w-\frac{1}{\alpha}\nabla_q Z\right|^2\\
&\quad=1+\left(\frac{dH_p}{\alpha}\right)^2+2\frac{dH_p}{\alpha}\langle\boldsymbol{\nu}_p,w\rangle+\frac{1}{\alpha^2}|\nabla_q Z|^2-\frac{2}{\alpha}\left\langle\nabla_q Z,\boldsymbol{\nu}_p+\frac{dH_p}{\alpha}w\right\rangle\\
&\quad=1+\left(\frac{dH_p}{\alpha}\right)^2+2\frac{dH_p}{\alpha}\left(Z-d^2\frac{H_p}{2}\right)+\frac{1}{\alpha^2}|\nabla_q Z|^2\\
&\quad\quad-\frac{2}{\alpha}\left\langle\nabla_q Z,\boldsymbol{\nu}_p+\frac{dH_p}{\alpha}w-\frac{1}{\alpha}\nabla_q Z\right\rangle-\frac{2}{\alpha^2}|\nabla_q Z|^2\\
&\quad=1+2\frac{H_p}{\alpha^2}Z-\frac{1}{\alpha^2}|\nabla_q Z|^2\,.
\end{aligned}$$

where we used the fact that $\nabla_q Z$ is in the tangent space at $X(q)$, hence orthogonal to $\boldsymbol{\nu}_p+(dH_p/\alpha)w-(1/\alpha)\nabla_q Z$. This yields

$$\boldsymbol{\nu}_p+\frac{dH_p}{\alpha}w-\frac{1}{\alpha}\nabla_q Z=\boldsymbol{\nu}_q\sqrt{1+2\frac{H_p}{\alpha^2}Z-\frac{1}{\alpha^2}|\nabla_q Z|^2}\,. \tag{9.7}$$

We compute at a point (p,q), $p\neq q$. Choose local coordinates so that $\{\partial_{p_i}\}$ are othronormal, $\{\partial_{q_i}\}$ are othronormal and $\partial_{p_i}=\partial_{q_i}$ for $i=1,\dots,n-1$. Thus ∂_{p_n} and ∂_{q_n} are coplanar with $\boldsymbol{\nu}_p$ and $\boldsymbol{\nu}_q$. With (9.3), (9.4), (9.5) and (9.6),

$$\begin{aligned}
LZ:=&\left(\partial_t-g_q^{ij}\partial_{q_i}\partial_{q_j}-g_p^{ij}\partial_{p_i}\partial_{p_j}-2g_p^{ik}g_q^{jl}\langle\partial_{p_k},\partial_{q_l}\rangle\partial_{p_i}\partial_{q_j}\right)Z\\
&=dH_p\langle w,-H_q\boldsymbol{\nu}_q+H_p\boldsymbol{\nu}_p\rangle+\frac{d^2}{2}\left(\Delta H_p+H_p|A_p|^2\right)\\
&\quad+\alpha\langle -H_q\boldsymbol{\nu}_q+H_p\boldsymbol{\nu}_p,\boldsymbol{\nu}_p\rangle+\alpha d\langle w,\nabla H_p\rangle\\
&\quad-nH_p+dH_pH_q\langle w,\boldsymbol{\nu}_q\rangle+\alpha H_q\langle\boldsymbol{\nu}_q,\boldsymbol{\nu}_p\rangle\\
&\quad-nH_p+2d\langle w,\nabla H_p\rangle-dH_p^2\langle w,\boldsymbol{\nu}_p\rangle-\frac{d^2}{2}\Delta_p H_p-\alpha d\langle w,\nabla H_p\rangle\\
&\quad+\alpha H_p+\alpha d\langle w,\boldsymbol{\nu}_p\rangle|A_p|^2\\
&\quad+2(n-1)H_p+2\langle\partial_{p_n},\partial_{q_n}\rangle^2H_p-2dg_p^{ik}g_q^{jl}\langle\partial_{p_k},\partial_{q_l}\rangle\langle w,\partial_{q_j}\rangle\nabla_{p_i}H_p\\
&\quad-2\alpha\left(H_p-h^p_{nn}+\langle\partial_{p_n},\partial_{q_n}\rangle^2h^p_{nn}\right)\\
&=Z|A_p|^2+2d\left\langle w,\partial_{p_k}-\langle\partial_{p_k},\partial_{q_l}\rangle g_q^{lj}\partial_{q_j}\right\rangle g_p^{ki}\nabla_{p_i}H_p\\
&\quad-2(H_p-\alpha h^p_{nn})\left(1-\langle\partial_{p_n},\partial_{q_n}\rangle^2\right)\,.
\end{aligned}$$

We observe that

$$\partial_{p_n}=\langle\partial_{p_n},\partial_{q_n}\rangle\partial_{q_n}+\langle\partial_{p_n},\boldsymbol{\nu}_q\rangle\boldsymbol{\nu}_q\,,$$

so that

$$\langle\partial_{p_n},\boldsymbol{\nu}_q\rangle=\langle\partial_{p_n},\partial_{q_n}\rangle\langle\partial_{q_n},\boldsymbol{\nu}_q\rangle+\langle\partial_{p_n},\boldsymbol{\nu}_q\rangle$$

and

$$1=\langle\partial_{p_n},\partial_{q_n}\rangle^2+\langle\partial_{p_n},\boldsymbol{\nu}_q\rangle^2\,.$$

At a critical point of Z, we obtain fron (9.1) that $\langle w,\partial_{q_i}\rangle=\alpha/(dH_p)\langle\partial_{q_i},\boldsymbol{\nu}_p\rangle$. Hence,

$$\langle w,\partial_{p_i}\rangle=\langle w,\partial_{q_i}\rangle=0$$

for $i=1,\dots,n-1$ and

$$\langle w,\partial_{q_n}\rangle=\frac{\alpha}{dH_p}\langle\partial_{q_n},\boldsymbol{\nu}_p\rangle$$

Furthermore, by (9.2),

$$\nabla_{p_i}H_p=\frac{2}{d}\langle w,H_p\partial_{p_i}-\alpha h^p_{im}g_p^{ms}\partial_{p_s}\rangle$$

and by (9.7),

$$\boldsymbol{\nu}_p + \frac{dH_p}{\alpha} w = \boldsymbol{\nu}_q \sqrt{1 + 2\frac{H_p}{\alpha^2} Z} =: \rho \boldsymbol{\nu}_q \,,$$

so that

$$\frac{dH_p}{\alpha} \langle w, \partial_{p_n} \rangle = \rho \langle \boldsymbol{\nu}_q, \partial_{p_n} \rangle \,.$$

Hence,

$$\begin{aligned} &2d \left\langle w, \partial_{p_k} - \langle \partial_{p_k}, \partial_{q_l} \rangle g_q^{lj} \partial_{q_j} \right\rangle g_p^{ki} \nabla_{p_i} H_p \\ &\quad = 4(H_p - \alpha h^p_{nn}) \left\langle w, \partial_{p_n} - \langle \partial_{p_n}, \partial_{q_n} \rangle \partial_{q_n} \right\rangle \langle w, \partial_{p_n} \rangle \,. \end{aligned}$$

so that

$$LZ = |A_p|^2 Z + 2(H_p - \alpha h^p_{nn}) Q$$

where

$$\begin{aligned} Q &= 2 \left\langle w, \partial_{p_n} - \langle \partial_{p_n}, \partial_{q_n} \rangle \partial_{q_n} \right\rangle \langle w, \partial_{p_n} \rangle - \langle \partial_{p_n}, \boldsymbol{\nu}_q \rangle^2 \\ &= 2 \langle \partial_{p_n}, \boldsymbol{\nu}_q \rangle \langle w, \boldsymbol{\nu}_q \rangle \langle w, \partial_{p_n} \rangle - \langle \partial_{p_n}, \boldsymbol{\nu}_q \rangle^2 \\ &= \langle \partial_{p_n}, \boldsymbol{\nu}_q \rangle \langle 2 \langle w, \partial_{p_n} \rangle w - \partial_{p_n}, \boldsymbol{\nu}_q \rangle \\ &= \frac{1}{\rho} \langle \partial_{p_n}, \boldsymbol{\nu}_q \rangle \left\langle 2 \langle w, \partial_{p_n} \rangle w - \partial_{p_n}, \boldsymbol{\nu}_p + \frac{dH_p}{\alpha} w \right\rangle \\ &= \frac{1}{\rho} \langle \partial_{p_n}, \boldsymbol{\nu}_q \rangle \langle \partial_{p_n}, w \rangle \left(2 \langle \boldsymbol{\nu}_p, w \rangle + \frac{2dH_p}{\alpha} - \frac{dH_p}{\alpha} \right) \\ &= \frac{2}{\alpha d \rho} \langle \partial_{p_n}, \boldsymbol{\nu}_q \rangle \langle \partial_{p_n}, w \rangle \left(\alpha d \langle \boldsymbol{\nu}_p, w \rangle + \frac{d^2 H_p}{2} \right) \\ &= \frac{2H_p}{\alpha^2 \rho^2} \langle \partial_{p_n}, w \rangle^2 Z \,. \end{aligned}$$

Since the coefficient of Z is a smooth function which is bounded on $(M \times M) \setminus \{p = q\}$, the maximum principle implies that Z remains nonnegative if initially nonnegative (Z is zero on the diagonal $\{p = q\}$). □

Remark 9.4 (Andrews [And12, Remark])**.** We made no use of the sign assumption on α, so the result also holds for negative α. This proves "exterior noncollapsing", ie the hypersurface remains outside the ball of radius $|\alpha|/H_p$ which touches the tangent plane at p on the exterior.

10. Convexity estimates

We follow the lines of [HK17]. In this chapter, we will also work with the evolving family $\{\Omega_t\}_{t \in I}$ where $\partial \Omega_t = M_t$. We will also consider families of possibly noncompact closed domains $\{\Omega_t \subset U\}_{t \in I}$ in an open set $U \subset \mathbb{R}^{n+1}$. For the mean curvature flow, time scales like distance squared.

Definition 10.1 (α-Andrews condition)**.** A smooth mean curvature flow $\mathcal{M}$ is α-Andrews if every time slice is α-noncollapsed.

Remark 10.2. By Theorem 9.3, if the initial set M_0 is compact and α-Andrews, then so is the whole flow $\mathcal{M}$.

Theorem 10.3 (Half-space convergence, Haslhofer–Kleiner [HK17, Theorem 2.1])**.** *Let $T_0 \geq 0$ and $\{\mathcal{M}^j\}$ be a sequence of α-Andrews flows such that:*

(i) For every $r < \infty$, the flow $\mathcal{M}^j$ is defined in $P(0, T_0, r)$ and there exists t_j so that $B_r(0) \subset \Omega^j_{t_j}$ for j sufficiently large.
(ii) The origin $0 \in \mathbb{R}^{n+1}$ lies in M^j_0 for every j.
(iii) Let $K \subset \{x^{n+1} < 0\}$ be compact, then $K \subset \Omega^j_0$ for j sufficiently large.

Then $\mathcal{O}^j$ converges smoothly on compact subsets of $\mathbb{R}^{n+1} \times (-\infty, T_0]$ to the static plane $\{x^{n+1} = 0\} \times (-\infty, T_0]$.

Remark 10.4. (1) Assumption (i) can be weakend by: For every $r < \infty$, the flow $\mathcal{M}^j$ is defined in $P(0, T_0, r)$ for j sufficiently large. For a proof, see [HK17, Appendix D].
(2) Assumption (1) is satisfied for every blowup sequence.
(3) The case $t_j \leq T_0 - R^2$ is of course allowed. In fact, it follows from the assertion of the theorem that $t_j \to -\infty$.

Proof of Theorem 10.3. We begin by proving convergence to a half-space in a weak sense. For $R \in (0, \infty)$ and $d \in \mathbb{R}$, let

$$B_R^d := B_R((-R + d)e_{n+1})$$

be the closed ball of radius R tangent to the horizontal hyperplane $\{x^{n+1} = d\}$ at the point de_{n+1}. If we evolve ∂B_R^d under (MCF) and start at time

$$t_0 = -\frac{dR}{n} + \frac{d^2}{2n} + \varepsilon\,,$$

for $\varepsilon > 0$, then $R(t) = \sqrt{R^2 - 2n(t - t_0)}$ (see Example 1.1(i)) and $\partial B_{R(t)}^d$ has left the upper half-space $\{x^{n+1} > 0\}$ at $t = \varepsilon$. Since $0 \in M_0^j$ for all j, $\overline{B}_R^d$ is not contained in Ω_0^j. Furthermore, the comparison principle, Theorem 1.8, yields that $\overline{B}_R^d$ cannot be contained in the interior of Ω_t^j for any $t \in [t_0, 0]$. Let By assumption (i) and (iii), By condition (iii), for large j we can find $d_j \leq d$ such that $\overline{B}_R^{d_j}$ has first interior contact with M_t^j at some point x_j, where

$$\langle x_j, e_{n+1}\rangle < d\,, \quad |x_j|^2 \leq t_0 \quad \text{and} \quad \liminf_{j\to\infty} \langle x_j, e_{n+1}\rangle \geq 0\,.$$

Hence the mean curvature satisfies

$$H(x_j, t) \leq \frac{n}{R}\,.$$

Since M_t^j satisfies the α-Andrews condition, there is a closed ball $\overline{B}_{R_j}$ with radius $R_j \geq \alpha R/n$ making exterior contact with M_0^j at x_j. As d and R are arbitrary, this implies that for any $t_1 < 0$ and any compact subset $V \subset \{x^{n+1} > 0\}$, for large j the time slice M_t^j is disjoint from V for all $t \geq t_1$. Likewise, for any $t_2 < 0$ and any compact subset $W \subset \{x^{n+1} < 0\}$, the time slice M_t^j contains W for all $t \in [t_2, T_0]$ and large j because $M_{t_2}^j$ will contain a ball whose forward evolution under (MCF) contains W at any time $t \in [t_2, T_0]$. This means that the sequence of mean curvature flows $\{\mathcal{M}^j\}$ converges in the pointed Hausdorff topology to a static plane in $\mathbb{R}^{n+1} \times (-\infty, T_0]$.

In general, let $U \subset \mathbb{R}^{n+1}$ be an open set and $\{K_\tau \subset U\}_{\tau \geq t}$ is a smooth family of mean convex domains such that $\{\partial K_\tau\}$ foliates $U \setminus \operatorname{int}(K_t)$. Let $K' \supset K_t$ be a closed domain that agrees with K_t outside a compact smooth domain $V \subset U$. Let $\boldsymbol{\nu}$ be the vectorfield in $U \setminus \operatorname{int}(K_t)$ defined by the outward unit normals of the foliation. Since $\operatorname{div} \boldsymbol{\nu} = H \geq 0$ we obtain with the area formula, Theorem A.1,

$$\begin{aligned}\mu^n(\partial K' \cap V) - \mu^n(\partial K_t \cap V) &= \int_t^{t_0} \partial_\tau \mu^n(\partial K_\tau \cap V)\, d\tau \\ &= \int_t^{t_0} \int_{\partial K_\tau \cap V} \operatorname{div} \boldsymbol{\nu}\, d\mu^n\, d\tau = \int_{(K' \setminus K_t)\cap V} H\, d\mu^n \geq 0\,.\end{aligned}$$

Hence, K_t has the following one-sided minimization property:

$$|\partial K_t \cap V| \leq |\partial K' \cap V|\,.$$

Now in our situation, one can take as a comparison domain

$$K' = \Omega_t^j \cup \big(\overline{B}_r(x) \cap \{x_{n+1} \leq \delta\}\big)$$

for $\delta > 0$ small. Hence, we get for every $\varepsilon > 0$, every time $t \leq T_0$, and every ball $B_r(x)$ centered on the hyperplane $\{x^{n+1} = 0\}$ that

$$\begin{aligned}\mu^n(M_t^j \cap B_r(x)) &\leq \mu^n(B_r(x) \cap \{x_{n+1} = \delta\}) + \mu^n(\partial B_r(x) \cap \{0 \leq x_{n+1} \leq \delta\}) \\ &\leq (1+\varepsilon)\omega_n r^n\end{aligned}$$

for j large enough. Let $(x,t) \in P(x_0, t_0, r)$. Then

$$\begin{aligned}&\int_{M^j_{t-r^2} \cap B_r(x)} \Phi_{(x,t)}(y, t-r^2)\, d\mu^n_{t-r^2} \\ &= \frac{1}{(4\pi(t-(t-r^2)))^{n/2}} \int_{M^j_{t-r^2} \cap B_r(x)} \exp\left(-\frac{|x-y|^2}{4(t-(t-r^2))}\right) d\mu^n_{t-r^2} \\ &\leq \frac{\mu^n(M_t^j \cap B_r(x))}{(4\pi r^2)^{n/2}} \leq \frac{(1+\varepsilon)\omega_n}{(4\pi)^{n/2}} = \frac{(1+\varepsilon)}{\Gamma(n+1/2)4^{n/2}} < (1+\varepsilon)\,.\end{aligned}$$

By Thorem 5.10 with $r \to \infty$, we have smooth convergence to a plane. □

The next theorem ensures that sequences of α-Andrews flows have subsequences that converge locally to smooth mean curvature flows provided we normalize the mean curvature at a single point.

Theorem 10.5 (Curvature estimate, Haslhofer–Kleiner [HK17, Theorem 1.8])**.** *For all $\alpha > 0$ there exist $\rho = \rho(\alpha) > 0$ and $C_l = C_l(\alpha) < \infty$, $l \in \mathbb{N} \cup \{0\}$, with the following property: If $\mathcal{M}$ is an α-Andrews flow in a parabolic ball $P(x,t,r)$ centered at $x \in M_t$ with $H(x,t) \leq 1/r$, then $\mathcal{M}$ is smooth in the parabolic ball $P(x,t,\rho r)$ and*

$$\sup_{P(x,t,\rho r)} \left|\nabla^l A\right| \leq \frac{C_l}{r^{l+1}}\,.$$

Proof. We will first show that there exists a $\rho' > 0$ such that the estimate holds for $l = 0$ with $C_0 = 1/\rho'$. Suppose this does not hold. Then there are sequences of α-Andrews flows $\{\mathcal{M}^j\}_{j\in\mathbb{N}}$, points $\{p_j \in M_{t_j}\}_{j\in\mathbb{N}}$ and scales $\{r_j\}_{j\in\mathbb{N}}$ such that $\mathcal{M}^j$ is defined in $P(x_j,t_j,r_j)$, some time slice contains $B_{r_j}(x_j)$ and $H(x_j,t_j) \leq 1/r_j$, but

$$\sup_{P(x_j,t_j,r_j/j)} \left|\nabla^l A\right| \geq \frac{j}{r_j}$$

for every $j \in \mathbb{N}$. After parabolically rescaling according to

$$(x,t) \mapsto \left(\frac{j}{r_j}(x-x_j), \frac{j^2}{r_j^2}(t-t_j)\right)$$

and applying an isometry, we obtain a new sequence $\{\hat{\mathcal{M}}^j\}$ of α-Andrews flows such that:

(a) $\hat{\mathcal{M}}^j$ is defined in $P(0,0,j)$ and some time slice contains $B_j(0)$.
(b) $0 \in \hat{M}_0^j$ and the outward unit normal of $\hat{M}_0^j$ at $(0,0)$ is e_{n+1}.
(c) $H_{\hat{M}_0^j}(0,0) \leq 1/j \to 0$ as $j \to \infty$.
(d) $\sup_{P(0,0,1)} |A| \geq 1$.

By (a), (b), (c) and the α-Andrews condition, $\{\hat{\mathcal{M}}^j\}$ satisfies assumptions (i), (ii) and (iii) of Theorem 10.3, and hence it converges smoothly on compact subsets of spacetime to a static half-space; this contradicts (d). Finally, by Ecker–Huisken [EH91], see also [Eck04, Propositions 3.21 and 3.22], we get uniform bounds on all scale-invariant derivatives of A in $P(x,t,\rho' r/2)$. By setting $\rho = \rho'/2$ the claim follows. □

Corollary 10.6 (Huisken–Sinestrari [HS09, Theorem 1.6], see also [HK17, Corollary 2.6])**.** *Let $\mathcal{M}$ be a mean convex flow where the initial time slice is compact. Then*

$$|\nabla A| \leq CH^2$$

for a constant $C < \infty$ depending only on the initial time slice.

Proposition 10.7 (White, [Whi03, Proposition A.4])**.** *Let $\mathcal{M}$ be mean convex. If κ_1/H attains a minimum value γ at (p, b), then κ_1/H is a nonnegative constant in a spacetime neighborhood of (p, b).*

Proof. Let $v = v^i\partial_i$ be a time-parallel vectorfield, that is

$$\partial_t v^i = -\frac{1}{2}g^{ij}(\partial_t g_{jk})v^k = Hg^{ij}h_{jk}v^k = Hh^i_k v^k\,.$$

Since $\partial_t(g_{ij}v^iv^j) = 0$, the length of v is constant in time. Then

$$\begin{aligned}\partial_t(A(v,v)) &= \partial_t(h_{ij}v^iv^j) = (\partial_t h_{ij})v^iv^j + 2h_{ij}(\partial_t v^i)v^j \\ &= (\Delta h_{ij} + |A|^2 h_{ij} - 2Hh^k_i h_{jk})v^iv^j - 2h_{ij}Hh^i_l v^l v^j \\ &= ((\Delta + |A|^2)h_{ij})v^iv^j\end{aligned}$$

and

$$\begin{aligned}\partial_t(Hg(v,v)) &= (\partial_t H)g(v,v) = (\Delta H + |A|^2 H)g(v,v) \\ &= ((\Delta + |A|^2)(Hg))(v,v)\,.\end{aligned}$$

Define the tensor $m := A - \gamma Hg$, which is positive semidefinite (by choice of γ) and satisfies

$$\partial_t(m(v,v)) = (\Delta m)(v,v) + |A|^2 m(v,v) \geq (\Delta m)(v,v)\,.$$

Note that the first eigenvalue $\lambda = \kappa_1 - \gamma H$ of m is everywhere nonnegative and is 0 at (p, b). Thus by Theorem D.8, λ is identically 0. Fix a time t. Then M^n is locally a metric product $N_1 \times N_2$. Let v_1 and v_n be unit eigenvectors of A (at some given point) with eigenvalues κ_1 and κ_n, respectively, and assume that $\kappa_1 \leq 0$. Then $\kappa_n > 0$ since $H > 0$. Thus v_1 and v_n will be horizontal and vertical, respectively, with respect to the product structure $N_1 \times N_2$. Moreover, by Theorem D.8, $\langle v_1, \nabla_w v_n\rangle = 0$ for every vector field w. The sectional curvature determined by v_1 and v_n is given by

$$\begin{aligned}\kappa_1\kappa_n = K(v_1, v_n) &= \frac{\langle R(v_1, v_n)v_n, v_1\rangle}{g(v_1, v_1)g(v_n, v_n) - g(v_1, v_n)^2} \\ &= \big\langle \nabla_{v_1}\nabla_{v_n}v_n - \nabla_{v_n}\nabla_{v_1}v_n - \nabla_{[v_1,v_n]}v_n, v_1\big\rangle = 0\,.\end{aligned}$$

Since κ_n is positive, κ_1 must vanish. □

The next theorem says that a boundary point (x, t) in an α-Andrews flow has almost positive definite second fundamental form as long as the flow has had a chance to evolve over a portion of spacetime that is large compared with the scale given by $H(x, t)$.

Theorem 10.8 (Convexity estimate, Haslhofer–Kleiner [HK17, Theorem 1.9])**.** *For all $\alpha, \varepsilon > 0$ there exists $\eta = \eta(\varepsilon, \alpha) < \infty$ such that if $\mathcal{M}$ is an α-Andrews flow in a parabolic ball $P(x, t, \eta r)$ centered at $x \in M_t$ with $H(x, t) \leq 1/r$, then*

$$\kappa_1(x, t) \geq -\frac{\varepsilon}{r}\,.$$

Proof. Fix $\alpha > 0$ and let $r_{\text{out}}(x, t)$ be the radius of the ball touching M_t at x from the outside. The α-Andrews condition implies

$$\frac{\alpha}{r_{\text{out}}(x,t)} = H(x, t) \leq \frac{1}{r}\,.$$

Hence

$$\kappa_1(x,t) \geq -\frac{1}{r_{\mathrm{out}}(x,t)} = -\frac{H(x,t)}{\alpha} \geq -\frac{1}{\alpha r}\,,$$

so that the assertion holds for $\varepsilon = 1/\alpha$. Let $\varepsilon_0 \leq 1/\alpha$ be the infimum of the ε's for which it holds, and suppose $\varepsilon_0 > 0$. It follows that there is a sequence $\{\mathcal{M}^j\}$ of α-Andrews flows, where for all j,

$$(0,0) \in \mathcal{M}^j\,, \quad H(0,0) \leq 1 \quad \text{and} \quad \mathcal{M}^j \text{ is defined in } P(0,0,j)$$

but

$$\kappa_1 \to -\varepsilon_0 \quad \text{for} \quad j \to \infty\,.$$

After passing to a subsequence, $\{\mathcal{M}^j\}$ converges smoothly to a mean curvature flow $\mathcal{M}^\infty$ in the parabolic ball $P(0,0,\rho)$, where $\rho = \rho(\alpha)$ is the quantity from Theorem 10.5. Then for $\mathcal{M}^\infty$ we have $\kappa_1(0,0) = -\varepsilon_0$ and thus $H(0,0) = 1$. By continuity, $H > 1/2$ in $P(0,0,r)$ for some $r \in (0,\rho)$. Furthermore, we have $\kappa_1/H \geq -\varepsilon_0$ everywhere in $P(0,0,r)$. This is because every $(x,t) \in \mathcal{M}^\infty \cap P(0,0,r)$ is a limit of a sequence $\{(x_j,t_j) \in \mathcal{M}^j\}$ and for every $\varepsilon > \varepsilon_0$, if $\eta = \eta(\varepsilon,\alpha)$, then $\mathcal{M}^j$ is defined in $P(x_j,t_j,\eta/H(x_j,t_j))$ for large j, which implies that the ratio $\kappa_1/H(x_j,t_j)$ is bounded below by $-\varepsilon$. Thus, in the parabolic ball $P(0,0,r)$, the ratio κ_1/H attains a negative minimum ε_0 at $(0,0)$. This contradicts Proposition 10.7. □

As an immediate consequence of Theorem 10.8, we obtain the original versions of the convexity estimate:

Corollary 10.9 (Huisken–Sinestrari [HS99a, Theorem 1.4], see also [HK17, Corollary 2.10])**.** *Let $\mathcal{M}$ be a smooth mean convex flow, where the initial time slice is compact. Then for all $\varepsilon > 0$ there is an $H_0 < \infty$ such that if $H(x,t) \geq H_0$ then $\kappa_1/H(x,t) \geq -\varepsilon$.*

Proposition 10.10 (Huisken–Sinestrari, [HS99b, Theorem 4.1])**.** *If M_0 has nonnegative mean curvature, then any limiting flow of a type-II singularity has convex surfaces M^∞_τ, $\tau \in \mathbb{R}$.. Furthermore, either M^∞_τ is a strictly convex translating soliton or (up to rigid motion) $M^\infty_\tau = \mathbb{R}^{n-k} \times N_\tau$, where N_τ is a k-dimensional strictly convex translating soliton in $\mathbb{R}^{k+1}$.*

Proof. We follow the lines of [Man11, Remark 2.5.6 and Proposition 4.2.7]. Around a singularity, we can send $\varepsilon \to 0$ in Corollary 10.9. This yields the convexity of the limit flow. For the splitting, we observe that the Weingarten operator satisfies $h^i_j \succeq 0$ on $(M^\infty_\tau)_{\tau\in\mathbb{R}}$ and

$$\partial_\tau h^i_j = \Delta h^i_j + |A|^2 h^i_j\,.$$

Let $\tau \in \mathbb{R}$. By the strong maximum priciple for 2-tensors, Theorem D.7, there exists $\delta(\tau) > 0$ so that

$$\operatorname{rank} S(\tau) = \operatorname{rank} A(\tau) =: m(\tau) \in \mathbb{N}$$

on $(\tau, \tau+\delta)$ and

$$m(\tau_2) = \inf_{M^\infty_{\tau_2}} \operatorname{rank} A \geq \sup_{M^\infty_{\tau_1}} \operatorname{rank} A = m(\tau_1)$$

for $\tau_2 > \tau_1$. Hence $m(\tau)$ is nondecreasing and there exists $\tau_0 \in \mathbb{R}$, so that the global minimum

$$m := \min_{\tau\in\mathbb{R}} m(\tau)$$

is attained at some point of $M^\infty_{\tau_0}$, that is,

$$m(\tau) = m$$

for all $\tau \leq \tau_0$. Assume that $m < n$, then

$$\ker A_x(\tau) \subset T_x M^\infty_\tau$$

is $(n-m)$-dimensional at every point $x \in M_\tau^\infty$. Let $v \in \ker A_x$ and γ be a geodesic in M_τ^∞ starting at x in direction of v. Then

$$\nabla^{\mathbb{R}^{n+1}}_{\dot\gamma}\dot\gamma = \nabla^M_{\dot\gamma}\dot\gamma + A(\dot\gamma,\dot\gamma)\boldsymbol{\nu} = 0$$

so that γ remains always in $\ker A$ and is also a geodesic in $\mathbb{R}^{n+1}$. Hence, for every $\tau \le \tau_0$ the hypersurface M_τ^∞ contains an $(n-m)$-dimensional affine subspace of $\mathbb{R}^{n+1}$. By Theorem D.7, $\ker A(\tau)$ is invariant by parallel transport and time for all $\tau \le \tau_0$, so that is the same affine subspace for all $\tau \le \tau_0$. Thus,

$$M_\tau^\infty = \ker A(\tau) \times N_\tau$$

splits as a product of an $(n-m)$-dimensional flat part and a family of either strictly convex, m-dimensional hypersurfaces $N_\tau \subset \mathbb{R}^{m+1}$ evolving by (MCF). Since A is bounded on $(M_\tau^\infty)_{\tau\in\mathbb{R}}$, the flow is unique (see Remark 1.7) and the above holds also for every $\tau > \tau_0$.

To show that N_τ is a translating solution, by Theorem 7.3, H and $|A|$ are comparable quantities, that is, there exists a time-independent constant ε so that

$$\varepsilon|A| \le H \le \sqrt{n}|A|$$

for $t \in [\delta, T)$. Hence, we can modify the type-II rescaling (see Definition 6.1) by replacing $|A|^2$ with H^2 and get the same estimates on the second fundamental form and its covariant derivatives. We then still get an eternal smooth limit flow, complete with bounded curvature and its covariant derivatives, with the only difference that this time it is the mean curvature H which gets a global maximum equal to one at time zero. Now Theorem 8.8 yields that $\mathcal{M}$ is translating. □

11. Cylindrical estimates

The cylindrical estimate says, roughly speaking, that near a boundary point in a uniformly k-convex flow, either the flow is uniformly $(k-1)$-convex or it is close to a shrinking round $(k-1)$-cylinder $\mathbb{R}^{k-1} \times \mathbb{S}^{n-k}$ provided the flow exists in a subset of backward spacetime that is large compared to the scale given by the mean curvature. To state this precisely, we say that an α-Andrews flow is ε-close to a shrinking round l-cylinder (or cylindrical domain) $\mathbb{R}^l \times \mathbb{S}^{n+1-l}$ near (x_0, t_0) if after applying the parabolic rescaling

$$(x, t) \mapsto (\lambda(x - x_0), \lambda^2(t - t_0)),$$

where $\lambda = H(x_0, t_0)$, and a rotation it becomes ε-close in the $C^{\lfloor 1/\varepsilon \rfloor}$-norm on $P(0, 0, 1/\varepsilon)$ to the standard shrinking l-cylinder with $H(0, 0) = 1$. See Huisken and Sinestrari [HS09, Theorem 1.5].

Theorem 11.1 (Cylindrical estimate, Haslhofer–Kleiner [HK17, Theorem 1.15])**.** *Let $\alpha, \beta, \varepsilon > 0$. Let $\mathcal{M}$ be an α-Andrews flow that is uniformly k-convex in the sense that $\kappa_1 + \cdots + \kappa_k \ge \beta H$. Let $x \in M_t$. Then there exists $\delta = \delta(\varepsilon, \alpha, \beta) > 0$ such that, if $\mathcal{M}$ is defined in $P(x, t, (\delta H(x, t))^{-1})$ and*

$$\frac{\kappa_1 + \cdots + \kappa_{k-1}}{H}(x, t) < \delta$$

then $\mathcal{M}$ is ε-close to a shrinking round $(k-1)$-cylinder $\mathbb{R}^{k-1} \times \mathbb{S}^{n-k}$ near (x, t).

Appendix A. Hypersurfaces in $\mathbb{R}^{n+1}$

A topological space is called *Hausdorff space* if for any two distinct points there exists a neighbourhood of each which is disjoint from the neighbourhood of the other. A topological space M^n is called locally *Euclidean* of dimension n, if M^n can be covered with open sets where every set is homeomorphic to an open subset of $\mathbb{R}^n$. A pair (U, φ), where $U \subset M^n$ is open and $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ is a homeomorphism, is called *chard* of M^n. A collection A of chards is called *atlas* of M^n if

$$M^n \subset \bigcup_{(U,\varphi)\in A} U\,.$$

Two chards (U,φ) and (V,ψ) are called C^k-*compatible*, $k \geq 1$, if

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$$

is a C^k-diffeomorphism. An atlas is called of class C^k, if each of its chards are C^k-compatible. If A is a C^k-atlas, there exists exactly one maximal C^k-atlas A_0 with $A \subset A_0$; it contains all chards which are C^k compatible with the chards of A. A differentiable (C^k-)structure on M^n is a maximal C^k-atlas on M^n. A local Euclidean Hausdorff space with a differentiable structure is called *differentiable manifold*.

Let N^{n+m} be a differentiable manifold. A subset $M^n \subset N^{n+m}$, $n, m \geq 1$, is called n-dimensional C^k-*submanifold* of N^{n+m} if for every $x \in M^n$ there exists an open neighbourhood $U \subset N^{n+m}$ and a C^k diffeomorphism $\varphi : U \to \varphi(U) \subset \mathbb{R}^{n+m}$ with

$$\varphi(U \cap M) = \varphi(U) \cap (\mathbb{R}^n \times \{0_{\mathbb{R}^m}\})\,.$$

Such an M^n owns a C^k-atlas, that is

$$A := \{(U \cap M, \varphi|_{U\cap M}) \mid \text{ where } (U,\varphi) \text{ as above}\}\,.$$

Then, M^n is locally Euclidean of dimension m and

$$(\psi|_{V\cap M}) \circ (\varphi|_{U\cap M})^{-1} = \psi \circ \varphi^{-1}|_{(\mathbb{R}^n\times\{0\})\cap\varphi(U\cap V)} \in C^k$$

for two diffeomorphisms ψ and φ.

A topological *manifold with boundary* is a Hausdorff space in which every point has a neighborhood homeomorphic to an open subset of the Euclidean half-space $\mathbb{R}^n_+ = \{(x_1,\ldots,x_n) \in \mathbb{R}^n \,|\, x_n \geq 0\}$. The boundary ∂M^n of M^n is the set of all points $p \in M^n$ such that $(\varphi(p))^n = 0$ for all chards (U,φ) of M^n. If M^n is a manifold with boundary, then the interior $\operatorname{int} M^n = M^n \setminus \partial M^n$ is a manifold (without boundary) of dimension n and boundary ∂M^n is a manifold (without boundary) of dimension $n-1$.

Let M^n be an abstract, smooth, compact, n-dimensional manifold without boundary and X a smooth immersion ($\operatorname{rank} DX \equiv n$) with

$$X : M^n \to \mathbb{R}^{n+m}\,.$$

We call $M := X(M^n)$ a *hypersurface* in $\mathbb{R}^{n+m}$. For all $p \in M^n$ and $v, w \in T_pM^n$, the embedding X induces an isomorphism

$$dX_p : T_pM^n \to T_{X(p)}M\,,$$

and the *first fundamental form* or *metric* $g_p : T_pM^n \times T_pM^n \to \mathbb{R}$ with

$$g_p(v,w) := \langle dX_p(v), dX_p(w)\rangle_{\mathbb{R}^{n+m}}\,.$$

Let $(U_i,\varphi_i)_{i\in I}$ be an atlas of M^n and

$$\partial_i = \frac{\partial}{\partial p_i} = d\varphi^{-1}(e_i) \in TM^n$$

then the matrix entries of the metric are

$$g_{ij} = g(\partial_i,\partial_j) = \langle dX(\partial_i), dX(\partial_j)\rangle_{\mathbb{R}^{n+m}} = \langle \partial_i X, \partial_j X\rangle_{\mathbb{R}^{n+m}} = \delta_{\alpha\beta}\partial_i X^\alpha \partial_j X^\beta$$

for $1 \leq \alpha, \beta \leq n+m$. We define by $(g^{ij})_{ij}$ the coordinate dependent inverse of the matrix $(g_{ij})_{ij}$ and the measure

$$d\mu^n = \sqrt{\det(g_{ij})}\, dp\,.$$

Observe that

$$\partial_k g_{ij} = \langle \partial_k \partial_i X, \partial_j X\rangle + \langle \partial_i X, \partial_k \partial_j X\rangle$$

and

$$\partial_k g^{ij} = -g^{pi} g^{qj} \partial_k g_{pq}\,.$$

The corresponding *Levi–Cevita connection* on M^n is given by

$$\nabla_v w = dX^{-1}\Big(\big(D_{dX(v)} dX(w)\big)^{\top}\Big)\,.$$

Here D is the standard connection in $\mathbb{R}^{n+m}$, and $^{\top}$ denotes the tangential component with respect to M, that is the orthogonal projection onto $dX(p)(T_pM^n) = T_{X(p)}M$. The connection can be evaluated in coordinates in terms of the *Christoffel symbols* Γ_{ij}^k defined by

$$\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k \partial_k\,,$$

where Γ_{ij}^k is explicitly given by We define the Christoffel symbols by

$$\Gamma_{ij}^k := g^{kl}\langle \partial_i\partial_j X, \partial_l X\rangle\,.$$

Here and in the following, we sum over repeated indices. Then,

$$\Gamma_{ij}^k \partial_k X = \langle \partial_i \partial_j X, \partial_l X\rangle \partial_l X\,.$$

At a fixed point, we can choose a coordinate system such that $\Gamma_{ij}^k = 0$. We calculate

$$0 = \partial_k \delta_j^i = \partial_k(g^{il} g_{jl}) = g^{il}\partial_k g_{jl} + g_{jl}\partial_k g^{il}\,,$$

so that

$$\begin{aligned}\partial_k g^{ij} &= -g^{il} g^{jm}\partial_k g_{lm} = -g^{il}g^{jm}\partial_k\langle \partial_l X, \partial_m X\rangle\\ &= -g^{il}g^{jm}(\langle \partial_k\partial_l X, \partial_m X\rangle + \langle \partial_l X, \partial_k \partial_m X\rangle) = -g^{il}\Gamma_{kl}^j - g^{jm}\Gamma_{km}^i\,.\end{aligned}$$

Being in a Levi–Cevita connection the *Lie bracket* $[\,\cdot\,,\,\cdot\,]$ is given by

$$[v,w] = \nabla_v w - \nabla_w v = \big(v\big(\mu^k\big) - w\big(\lambda^k\big)\big)\,\partial_k\,.$$

The tangential gradient of a function $f \in C^1(M)$ is given by

$$\nabla^M f = g^{ij}\partial_i f \partial_j\,.$$

The tangential divergence $\operatorname{div}_M : T_pM^n \to \mathbb{R}$ is given by

$$\operatorname{div}_M v = g^{ij}\langle \partial_i v, \partial_j X\rangle_{\mathbb{R}^{n+m}}\,.$$

For the embedding vector X, we therefore have

$$\operatorname{div}_M X = g^{ij}\langle \partial_i X, \partial_j X\rangle_{\mathbb{R}^{n+m}} = g^{ij} g_{ij} = n\,.$$

For $\omega = df = \frac{\partial f}{\partial p_i} dp^i$, we obtain the Hessian of the function f

$$(\operatorname{Hess}_M f)(v,w) := \left(\nabla^2 f\right)(v,w)\,,$$

or in coordinates

$$\nabla_i \nabla_j f = (\operatorname{Hess}_M f)(\partial_i, \partial_j) = \partial_i \partial_j f - \Gamma_{ij}^k \partial_k f\,.$$

The *Laplace–Beltrami operator* $\Delta_M : C^2(M^n) \to C^0(M^n)$ is defined as

$$\Delta_M f := \frac{1}{\sqrt{\det g_{kl}}}\partial_j\Big(\sqrt{\det g_{kl}}\, g^{ij}\partial_j f\Big) = \operatorname{div}_M(\nabla^M f) = g^{ij}\nabla_i\nabla_j f\,.$$

We define the *second fundamental form* $\mathbf{A}_p : T_pM^n \times T_pM^n \to \big(T_{X(p)}M\big)^\perp$ by

$$\begin{aligned}\mathbf{A}_p(v,w) &:= -\sum_{k=1}^m \big\langle D_{dX_p(v)}dX_p(w), \boldsymbol{\nu}_k(p)\big\rangle\, \boldsymbol{\nu}_k(p)\\ &= \sum_{k=1}^m \big\langle dX_p(w), D_{dX_p(v)}\boldsymbol{\nu}_k(p)\big\rangle\, \boldsymbol{\nu}_k(p)\,,\end{aligned}$$

where $\{\boldsymbol{\nu}_k\}_{1\le k\le m}$ is an orthonormal frame for $(TM)^\perp$. In coordinates $\{p_i\}_{1\le i\le n}$,

$$\mathbf{A}_{ij} := \mathbf{A}_p(\partial_i,\partial_j) = \sum_{k=1}^m \langle \partial_i X, \partial_j \boldsymbol{\nu}_k\rangle\, \boldsymbol{\nu}_k\,.$$

The *mean curvature vector* $\mathbf{H} : M \to (TM)^\perp$ is the trace of the second fundamental form

$$\mathbf{H} := -g^{ij}\mathbf{A}_{ij} = -g^{ij}\sum_{k=1}^m \langle \partial_i X, \partial_j \boldsymbol{\nu}_k\rangle\, \boldsymbol{\nu}_k = -\sum_{k=1}^m \operatorname{div}(\boldsymbol{\nu}_k)\boldsymbol{\nu}_k\,.$$

We calculate that

$$\begin{aligned}\Delta_M X &= g^{ij}\big(\partial_i\partial_j X - \Gamma_{ij}^k\partial_k X\big) = g^{ij}\sum_{k=1}^m \langle \partial_i\partial_j X, \boldsymbol{\nu}_k\rangle\, \boldsymbol{\nu}_k\\ &= -g^{ij}\sum_{k=1}^m \langle \partial_i X, \partial_j\boldsymbol{\nu}_k\rangle\, \boldsymbol{\nu}_k = \mathbf{H}\,.\end{aligned}$$

For a submanifold Σ of M, the mean curvature vector is given by

$$\mathbf{H}_\Sigma = -\sum_{k=1}^m \operatorname{div}_\Sigma(\boldsymbol{\nu}_k)\boldsymbol{\nu}_k - \operatorname{div}_\Sigma(\boldsymbol{\nu}_\Sigma)\boldsymbol{\nu}_\Sigma\,,$$

where $\boldsymbol{\nu}_\Sigma$ is the unit co-normal of Σ. Since $\boldsymbol{\nu}_\Sigma$ tangential to M,

$$\langle \mathbf{H}_\Sigma, \boldsymbol{\nu}_\Sigma\rangle = -\operatorname{div}_\Sigma \boldsymbol{\nu}_\Sigma$$

and on Σ,

$$\begin{aligned}\Delta_\Sigma X &= g_\Sigma^{ij}\big(\partial_i\partial_j X - \Gamma_{ij}^k\partial_k X\big)\\ &= \sum_{k=1}^m g_\Sigma^{ij}\langle \partial_i\partial_j X, \boldsymbol{\nu}_k\rangle\, \boldsymbol{\nu}_k + g_\Sigma^{ij}\langle \partial_i\partial_j X, \boldsymbol{\nu}_\Sigma\rangle\, \boldsymbol{\nu}_\Sigma = \mathbf{H}_\Sigma\,.\end{aligned}$$

For $m=1$,

$$\mathbf{A}(v,w) = A(v,w)\boldsymbol{\nu}\,,$$

where $\boldsymbol{\nu}$ is the outward pointing unit normal to M and $A : TM^n \times TM^n \to \mathbb{R}$ is given by

$$A(v,w) = -\big\langle D_{dX(v)}dX(w), \boldsymbol{\nu}\big\rangle = \big\langle dX(w), D_{dX(v)}\boldsymbol{\nu}\big\rangle\,.$$

where $\boldsymbol{\nu}$ is the outward pointing unit normal to M. In coordinates,

$$h_{ij} := A(\partial_i,\partial_j) = -\langle \partial_i\partial_j X, \boldsymbol{\nu}\rangle = \langle \partial_i X, \partial_j \boldsymbol{\nu}\rangle\,.$$

Let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be the eigenvalues of A, that is

$$h_{ij}\xi_k^i\xi_k^j = \lambda_k g_{ij}$$

for eigenvectors $\xi_k \in TM$ and $k = 1,\ldots,n$. The *Weingarten operator* $S : TM^n \to TM^n$ is given by

$$S(v) := dX^{-1}\big(D_{dX(v)}\boldsymbol{\nu}\big)$$

so that

$$A(v,w) = g(v, S(w))\,,$$

where in coordinates,

$$h_j^i := g^{ik}h_{kj}$$

and the Weingarten equations by

$$\partial_i \boldsymbol{\nu} = h_i^j \partial_j X \,.$$

The norm of the second fundamental form is given by

$$|A|^2 = g^{ik} g^{lj} h_{kl} h_{ij} = h^{ij} h_{ij} \,,$$

and the mean curvature vector is given by

$$\mathbf{H} = -g^{ij} h_{ij} \boldsymbol{\nu} = -H \boldsymbol{\nu} \,,$$

where we define the *mean curvature* H of M as the trace of the second fundamental form with

$$H = g^{ij} h_{ij} = \operatorname{div}_M \boldsymbol{\nu} \,.$$

The *Gauss curvature* is given by

$$K := \det(h_{ij}) \,.$$

We have the Gauss formula

$$\nabla_i \nabla_j X = \partial_i \partial_j X - \Gamma_{ij}^k \partial_k X = -h_{ij} \boldsymbol{\nu}$$

which as before leads to $\Delta_M X = \mathbf{H}$. More useful identities are the Codazzi equations in $\mathbb{R}^{n+1}$

$$\nabla_k h_{ij} - \nabla_j h_{ik} = \Gamma_{ij}^l h_{lk} - \Gamma_{ik}^l h_{lj}$$

and Simons' identity

$$\Delta h_{ij} = \nabla_i \nabla_j H + H h_{ik} h_j^k - |A|^2 h_{ij} \,. \tag{A.1}$$

We define the *Riemannian curvature tensor* by

$$R(u,v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w \,.$$

In coordinates that is

$$R_{lij}^k := \nabla_i \Gamma_{jl}^k - \nabla_j \Gamma_{il}^k + \Gamma_{im}^k \Gamma_{jl}^m - \Gamma_{jm}^k \Gamma_{il}^m \,.$$

Moreover, we set

$$R_{klij} := g^{kr} R_{lij}^r$$

and define the Ricci tensor by

$$R_{ik} := R_{ijkl} g^{jl}$$

and the *scalar curvature* by

$$R := R_{ij} g^{ij} \,.$$

The Gauss equation are

$$R_{ijkl} = h_{ik} h_{jl} - h_{il} h_{jk} \,.$$

The *sectional curvature* in direction of two linearly independent vectors v and w is given by

$$K(v,w) = \frac{\langle R(v,w)w, v\rangle}{g(v,v)g(w,w) - g(v,w)^2} \,.$$

Theorem A.1 (First variation of the area formula, see [Sim83, p. 51])**.** *Let $M \subset \mathbb{R}^{n+1}$ be a smooth, compact, n-dimensional hypersurface with boundary. Let $U \subset \mathbb{R}^{n+1}$ be a open and bounded such that $M \subset U$. Let $\phi : U \times (-1,1) \to U$ be a one-parameter family of $C^{2;1}$-diffeomorphisms. Set $M_t := \phi(M,t)$ and $v(p) := \partial_t \phi(p,0)$. Then*

$$\partial_t|_{t=0} \mu^n(M_t) = \int_M \operatorname{div}_M v \, d\mu^n \,.$$

Theorem A.2 (Divergence theorem, see [Sim83, p. 43], [DHTK10, p. 304], [Eck04, p. 116])**.** *Let $M \subset \mathbb{R}^{n+1}$ be a smooth, compact, n-dimensional manifold with boundary. Let v be a C^1-vectorfield on M. Then*

$$\int_M \operatorname{div}_M v \, d\mu^n = -\int_M \langle v, \mathbf{H}_M \rangle_{\mathbb{R}^{n+1}} \, d\mu^n + \int_{\partial M} \langle v, \boldsymbol{\nu}_{\partial M} \rangle_{\mathbb{R}^{n+1}} \, d\mu^{n-1} \,.$$

Theorem A.3 (Rademacher's theorem, see [Fed69, Theorem 3.1.6])**.** *Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$ be Lipschitz continuous. Then f is differentiable almost everywhere in U.*

Lemma A.4 (Fatou's lemma, [AE06, Theorem 3.7])**.** *Let $(\Omega, \sigma, d\mu)$ be a measure space and let $(f_i : \Omega \to [0,\infty))_{i\in\mathbb{N}}$ be a sequence of non-negative integrable functions such that $\liminf_{i\to\infty} \int_\Omega f_i \, d\mu < \infty$. Then*

$$\int_\Omega \liminf_{i\to\infty} f_i \, d\mu \le \liminf_{i\to\infty} \int_\Omega f_i \, d\mu \,.$$

Appendix B. Frobenius' theorem

Let M^n be a smooth manifold and v a smooth vector field on M^n. The *integral curve* of v is a curve $\gamma : (a,b) \to M^n$ such that

$$\dot\gamma(t) = v(\gamma(t))$$

for all $t \in (a,b)$. (The existence of γ is given by Picard–Lindelöf.) If v is non-vanishing, then its integral curves are connected, immersed 1-dimensional submanifolds of M^n.

A k-dimensional (tangent) *distribution* D on M^n is a choice of k-dimensional linear subspaces $D_p \subset T_pM^n$ at each point $p \in M^n$, where

$$D = \bigsqcup_{p\in M^n} D_p \subset TM^n \,.$$

If D is a k-dimensional distribution, then we can find a vector field v_1 such that $v_1(p) \in D_p$ for all p in some neighborhood $U \subset M^n$. We can continue (possibly shrinking the neighborhood) until we have vector fields $v_1, \dots, v_k$ such that $v_1(p), \dots, v_k(p)$ form a basis for D_p at each $p \in U$.

An immersed submanifold $N \subset M^n$ is an *integral manifold* of the distribution D if $T_pN = D_p$ for all $p \in N$, and D is *integrable* if each point of M^n there exists an integral manifold of D.

A distribution D is called *involutable* if $[v,w] \in D$ for all $v, w \in D$.

A parametrization $\phi : U \subset M^n \to \mathbb{R}^n$ is *flat* for D if $\phi(U) \subset \mathbb{R}^n$ is a product of connected open sets in $\mathbb{R}^k \times \mathbb{R}^{n-k}$ and for each $p \in U$, D_p is spanned by precisely the first k basis vector fields. A distribution D is *completely integrable* if there exists a flat parametrization for D in a neighborhood of every point of M^n.

Theorem B.1 (Frobenius' theorem)**.** *Let D be a distribution on a smooth manifold M^n. Then, D is completely integrable if and only if D is involutable.*

A k-dimensional *foliation* $\mathcal{F}$ on M^n is a collection of disjoint, connected, immersed k-dimensional submanifolds N of M^n (the *leaves* of the foliation) such that

(i) the union of the leaves is all of M^n, i.e., $M^n = \bigsqcup_{N\in\mathcal{F}} N$, and
(ii) there is a parametrization ϕ around each $p \in U \subset M^n$ such that $\phi(U)$ is a product of connected open sets in $\mathbb{R}^k \times \mathbb{R}^{n-k}$ and each leaf N intersects U in the empty set or a countable union of k-dimensional slices of the form $x_{k+1} = c_{k+1}, \dots, x_m = c_m$.

Theorem B.2 (Alternate Frobenius)**.** *If D is an involutive distribution on M^n, then the collection of all maximal connected integral manifolds N of D forms a foliation of M^n.*

Appendix C. Sard's theorem

Section copied from [Sch05, Section 3]. See also [BJ73].

Definition C.1. Let $f : M \to N$ differentiable. A point $p \in M$ is called *regular*, if the differential of f in p is surjektiv. A point $q \in N$ is called *regular value*, if $f^{-1}(q)$ consists of regular points. Non-regular points or values are called *singular* or *critical*.

We want to prove the following theorem.

Theorem C.2 (Sard's theorem)**.** *Let M^m and N^n be differentiable manifolds with a countable basis of their topology. The critical set S of a C^k function $f : M \to N$ consists of those points at which the differential $df : TM \to TN$ has rank less than n as a linear transformation. If $k \geq \max\{n-m+1, 1\}$, then the image of S has Lebesgue measure zero as a subset of N.*

Corollary C.3. *Let M^m be a differentiable manifold and $f : M^m \to \mathbb{R}^n$ a diffenrentiable. Then $f^{-1}(x) \subset M^m$ is a differentiable submanifold of co-dimension n for almost every $x \in \mathbb{R}^n$.*

Remark C.4. The set $f^{-1}(x)$ can be empty. Sard's theorem also holds for maps $f : \mathbb{R}^n \to \mathbb{R}^p$, $f \in C^k$ with $k > \max\{n-p, 0\}$ and manifolds with according dimensions.

Definition C.5. A subset $C \subset \mathbb{R}^n$ is of *measure zero*, if for every $\varepsilon > 0$ there exists a sequence $(W_i)_{i\in\mathbb{N}}$ of cubes in $\mathbb{R}^n$ with

$$C \subset \bigcup_{i\in\mathbb{N}} W_i \quad \text{and} \quad \sum_{i\in\mathbb{N}} |W_i| < \varepsilon\,.$$

Remark C.6. (i) The countable set of zero sets is again a zero set.
(ii) One obtains an equivalent definition for open oder closed cubes or balls.

Lemma C.7. *Let $U \subset \mathbb{R}^m$ be open and $C \subset U$ of measure zero. Let $f : U \to \mathbb{R}^m$ be Lipschitz. Then $f(C)$ has measure zero.*

Proof. Exercise. □

Definition C.8. A subset C of a differentiable manifold has measure zero, if for every chard $h : U \to U' \subset \mathbb{R}^m$ the set $h(C \cap U) \subset \mathbb{R}^m$ is of measure zero.

Remark C.9. The assumption of differentiability is important here, since zero sets are not necessarily maintained under homeomorphisms. Since a manifold owns a countable basis of the topologie, there exists an atlas with countably many chards. It is sufficient to apply the definition for such chards. Well-definedness follows, since zero sets are maintained under differentiable chard changes and countable unions.

Lemma C.10. *An open covering of the interval $[0,1]$ by subintervals contains a countable cover $[0,1] = \bigcup_{j=1}^k I_j$ with $\sum_{j=1}^k |I_j| \leq 2$.*

Proof. Due to the compactness, there exists a finite subcover. Choose one where no interval can be left out without loosing the covering property. Let the intervals I_j, $j = 1, \ldots, k$ be numbered so that with $I_j = (a_j, b_j)$ always holds $a_j < a_{j+1}$, $j = 1, \ldots, k-1$. Minimality and covering property imply $a_i < a_{i+1} < b_i < a_{i+2}$. So that

$$\begin{aligned}\sum_i (b_i - a_i) &= \sum_i (a_{i+1} - a_i) + \sum_i (b_i - a_{i+1}) \\ &< \sum_i (a_{i+1} - a_i) + \sum_i (a_{i+1} - a_{i+1}) \leq 2\,,\end{aligned}$$

where we used that we have telecope sums in the end. □

Theorem C.11 (Fubini)**.** *Let $\mathbb{R}^{n-1}_t := \{x \in \mathbb{R}^n \,|\, x^n = t\} \subset \mathbb{R}^n$. Let $C \subset \mathbb{R}^n$ be compact and $C_: = C \cap \mathbb{R}^{n-1}_t$ be of measure zero in $\mathbb{R}^{n-1}_t \cong \mathbb{R}^{n-1}$ for all $t \in \mathbb{R}$. Then C is of measure zero in $\mathbb{R}^n$.*

Proof. Since the property of being of measure zero is maintained under countable unions, we can assume that $C \subset \mathbb{R}^{n-1} \times [0,1]$. For $t \in [0,1]$, C_t is of measure zero in $\mathbb{R}^{n-1} \times \{t\}$. Let $\varepsilon > 0$ and W_t^i be a cover of C_t by open cubes with $\sum_i |W_t^i| < \varepsilon$. Define $W_t := \bigcup_i W_t^i$ identify these with subsets of $\mathbb{R}^{n-1}$. The function $|x^n - t|$ is for fixed $t \in [0,1]$ on C continuous, vanishes exactly on C_t und attains a positive minimum in the compact set $C \setminus (W_t \times [0,1])$, which we call α. It follows

$$\{x \in C : |x^n - t| < \alpha\} \subset W_t \times I_t^\alpha \,,$$

where $I_t^\alpha = (t-\alpha, t+\alpha)$ and $\bigcup_t I_t^\alpha = [0,1]$. Choose a subcover of $[0,1]$ among the intervals I_t^α with $\sum_{t_i} |I_{t_i}^\alpha| \leq 2$. Observe that $\alpha = \alpha(t_i)$. It holds

$$C \subset \bigcup_{t_j, i} W_{t_j}^i \times I_{t_j}^\alpha \,,$$

where i is the index of the cube and we take the union over cuboids. Moreover,

$$\sum_{t_j, i} |W_{t_j}^i \times I_{t_j}^\alpha| \leq 2\varepsilon \,.$$

Sending $\varepsilon \to 0$ yields the lemma. □

Remark C.12. The requirement that C is compact, can be weakened as follows: C is a countable union of compact sets, that each suffice the assumptions of the theorem. This is fulfilled by closed and open sets (which cannot be zero sets), for images of these set under continuous maps, countable union und finite intersections of these.

Proof of Theorem C.2. After introducing maps it is sufficient to show: Let $U \subset \mathbb{R}^n$ be open, $f : U \to \mathbb{R}^p$ smooth and $D \subset U$ be the set of critical points of f, then $f(D) \subset \mathbb{R}^p$ has measure zero.

We prove by induction over n. In case $n = 0$, $\mathbb{R}^n$ is a point. So, $f(U)$ is at most a point and has measure zero. Assume the claim is true for the case $n-1$. We proof the case n. Let $D_i \subset U$ be the set of all points points, in which the partial derivative of order $\leq i$ vanish. We obtain the decreasing sequence of relatively closed sets

$$D \supset D_1 \supset D_2 \supset \dots \,.$$

We claim that

(i) $f(D \setminus D_1)$ is of measure zero,
(ii) $f(D_i \setminus D_{i+1})$ is of measure zero,
(iii) for k big enough, $f(D_k)$ is of measure zero.

We observe, that (iii) is neccessary, since also the points, in which all derivatives vanish, can be captured. By Remark C.12, all sets occuring in (i)–(iii) can be used. Moreover, it is sufficient to prove that every point in $D \setminus D_1$ resp. $D_i \setminus D_{i+1}$ resp. D_k has a neighbourhood V, so that $f(V \cap (D \setminus D_1))$ resp. $f(V \cap (D_i \setminus D_{i+1}))$ resp. $f(V \cap D_k)$ are of measure zero. The claim then follows, since the countable union of zero set is again a zero set.

Proof of (i): Assume, that $p \geq 2$, since for $p = 1$ we already have $D = D_1$. Let $x_0 \in D \setminus D_1$. Since $x_0 \notin D_1$, there exists a partial derivative that is not vanishing in x_0, w.l.o.g. $\partial_1 f \neq 0$. Define $h : U \to \mathbb{R}^n$ by

$$h : x = (x^1, \dots, x^n) \mapsto (f^1(x), x^2, \dots, x^n) \,.$$

Then h is not singular in x_0. Hence there exists a neighbourhood V of x_0, so that $h : V \to h(V) = V'$ is a diffeomorphism. Define $g := f \circ h^{-1}$. In a neighbourhood of $h(x)$, g is of the form

$$g : (z^1, \dots, z^n) \mapsto (z^1, g^2(z), \dots, g^n(z)) \,.$$

The hyperplane $\{z\,|\,z^1 = t\}$ is (locally) mapped into the hyperplane $\{y\,|\,y^1 = t\}$. Define

$$g_t : \{t\} \times \mathbb{R}^{n-1} \cap V' \to \{t\} \times \mathbb{R}^{p-1}$$

als restriction of g. We have

$$Dg_t = \begin{pmatrix} 1 & 0 \\ ? & Dg \end{pmatrix} .$$

Hence a point in $(\{t\} \times \mathbb{R}^{n-1}) \cap V'$ is critical for g if and only if it is for g_t. By the induction assumption the set of critical values of g_t is of measure zero in $\{t\} \times \mathbb{R}^{p-1}$. Since g maps entsprechende hyperplanes onto itself, the set of critical values of g also has a intersection of measure zero with the hyperplane $\{y\,|\,y^1 = t\}$. By Fubini, Theorem C.11, the critical values of g have measure zero. Since f and g only differ by an diffeomorphism, also the criticalen values of f have measure zero. This holds locally, as long as $\partial_1 f \neq 0$. This proves (i).

Proof of (ii): We argument similarly as in the proof of (i). Let $x_0 \in D_k \setminus D_{k+1}$. Then there exist a non-vanishing $(k+1)$-st derivative, w.l.o.g.

$$\frac{\partial^{k+1} f^1}{\partial x^1 \partial x^{\nu_1} \dots \partial x^{\nu_k}}(x_0) \notin 0 .$$

Assume, that this holds in a neighbourhood V of x_0. Define $w : V \to \mathbb{R}$ by

$$w := \frac{\partial^k f^1}{\partial x^{\nu_1} \dots \partial x^{\nu_k}}(x_0) \neq 0 .$$

It holds $w(x) = 0$, $\frac{\partial}{\partial x^1} w(x) \neq 0$. The map

$$h : x \to (w(x), x^2, \dots, x^n)$$

defines a diffeomorphism $h : V \to V' = h(V)$. w and therefore all k-th derivatives of f^1 vanish at most for $x = x_0$. Hence

$$h(D_k \cap V) \subset \{0\} \times \mathbb{R}^{n-1} \subset \mathbb{R}^n .$$

Define

$$g : f \circ h^{-1} : V' \to \mathbb{R}^p$$

and

$$g_0 : \{0\} \times \mathbb{R}^{n-1} \cap V' \to \mathbb{R}^p .$$

By the induction assumption, the set of critical values of g_0 has measure zero. Let $x \in h(D_k \cap V)$. Then all derivatives of g up to order k vanish there. Since $h(D_k \cap V) \subset \{0\} \times \mathbb{R}^{k-1}$, g_0 is defined there and has vanishing derivatives up to order k. In particular, all first derivatives vanish there as well and thus we are dealing with critical points of g_0. Hence

$$(g_0 \circ h)(D_k \cap V) = (g \circ h)(D_k \cap V) = f(D_k \cap V)$$

has measure zero.

Proof of (iii): The set U ist countable union of cubes. Let $W \subset U$ be a cube with side length $a \leq 1$ and let $k > n - 1$. It is sufficient to show, that $f(W \cap D_k)$ is of measure zero. By Taylor it holds that

$$f(x + h) = f(x) + R(x, h)$$

with

$$|R(x, h)| \leq c|h|^{k+1}$$

for $x \in D_k \cap W$ and $x + h \in W$, where the constant c only depends on f and W. We devide W in r^n cubes with side length a/r, $r \in \mathbb{N}$. If W_1 is a cube of this partitioning, which contains a point $x \in D_k$, then every other point in W_1 can be described as $x+h$ with $|h| \le \sqrt{n}a/r$. Hence with Taylor

$$|f(x+h) - f(x)| \le c \left(\frac{\sqrt{n}a}{r} \right)^{k+1} .$$

So that $f(W_1)$ is contained in a cube with side length

$$c(n) \left(\frac{\sqrt{n}a}{r} \right)^{k+1} .$$

There are at most r^n such cubes with points in D_k. The summed up volumes of the images of these cubes in $\mathbb{R}^p$ are at most

$$c(n)^p \left(\frac{\sqrt{n}a}{r} \right)^{p(k+1)} r^n = c r^{n-p(k+1)} .$$

Since $n - p(k+1) < 0$, this will get arbitrary small for $r \to \infty$. $\square$

Corollary C.13 (Brown)**.** *Let M and N be (finite dimensional) manifolds. Let $f : M \to N$ be a differentiable (C^∞-)maps. Then all the regular values of f lay dense in N.*

We want to derive Brouwer's fixed point theorem from Sard's theorem.

Definition C.14. Let $A \subset B$. A retraction is a continuous map $f : B \to A$, so that $f|_A = id$, that is, $f(x) = x$ for all $x \in A$.

Theorem C.15. *There exists no retraction of $\overline{B_1(0)} \subset \mathbb{R}^n$ on $\mathbb{S}^{n-1}$.*

Proof. We prove the claim by contradiction. Let $f : \overline{B_1(0)} \to \mathbb{S}^{n-1}$ be a retraction. Show at first, that then there also exists a C^∞-retraction of $\overline{B_1(0)}$ on $\mathbb{S}^{n-1}$: We find a retraction g, that is close to $\partial B_1(0)$ of the class C^∞, e.g.,

$$g(x) = \begin{cases} f\left(\frac{x}{|x|}\right) & \text{for } \frac{1}{2} \le |x| \le 1 \\ f(2x) & \text{for } 0 \le |x| \le \frac{1}{2} . \end{cases}$$

Mollification in the interior gives a C^∞-retraction. Hence we may assume that $f \in C^\infty\big(\overline{B_1(0)}, \mathbb{S}^{n-1}\big)$. By Corollary C.13 there exists a regular value $y \in \mathbb{S}^{n-1}$ of f. Hence the compact set $f^{-1}(y)$ is a one-dimensional submanifold (first in $B_1(0)$, but since we can mollify f, also up to the boundary, since f is after construction constant on radial line segments close to $\mathbb{S}^{n-1}$). Hence $f^{-1}(y)$ is a one-dimensional manifold with boundary in $\overline{B_1(0)}$, whose boundary is a subset of $\mathbb{S}^{n-1} = \partial B_1$. It holds that $y \in f^{-1}(y)$, since f is a retraction. Let V be the component of $f^{-1}(y)$ that contains y. Then V is a one-dimensional compact connected manifold and thus diffeomorph to a closed interval. Then y is the one boundary point of V. Let z be the other, which as well lays on $\partial B_1(0)$. It follows that $z = f(z)$ in contradiction to $y, z \in f^{-1}(y)$. $\square$

Theorem C.16 (Brouwer's fixed point theorem)**.** *Let $f : \overline{B_1(0)} \to \overline{B_1(0)}$ be continuous. Then f has one fixed point, that is, there exists $x \in \overline{B_1(0)}$ with $f(x) = x$.*

Proof. If $f(x) \neq x$ for all $x \in \overline{B_1(0)}$, we define $g(x)$ to be the intersection of a line with $\mathbb{S}^{n-1}$ beginning in $f(x)$ through x . As constructed g is a retraction of $\overline{B_1(0)}$ on $\mathbb{S}^{n-1}$. $\square$

Appendix D. Maximum principles

Theorem D.1 (Strong elliptic maximum principle)**.** *Let M be closed and $f : M \to \mathbb{R}$ satisfy*

$$-\Delta_M f + b^i \nabla_i^M f + cf \leq 0$$

for some smooth funtions b^i and $c \leq 0$. If $f \leq 0$, but $f \not\equiv 0$, then $f < 0$.

Proof. For a proof see [Eva02, §6.4, Theorem 4] or [Sch17b, Theorem 5.5] for $M^n = \mathbb{R}^n$. □

Let M^n be a smooth n-dimensional manifold with boundary whose closure is compact. Let $X : \bar{M}^n \times [0,T) \to \mathbb{R}^{n+m}$ be a family of smooth embeddings and set $M_t := X(M^n, t)$. For $f \in C^{2;1}(M^n \times [0,T))$, we define the parabolic operator

$$L(f) := \partial_t f - a^{ij} \nabla_i \nabla_j f - b^i \nabla_i f - cf\,,$$

where $a_{ij}, b_i, c \in L^\infty$ may depend on p, t, $(g_{kl})_{kl}$, f, ∇f, and $\nabla^2 f$, and where $(a^{ij})_{ij}$ is positive semi-definite, that is,

$$\lambda |\xi|^2 \leq a_{ij} \xi_i \xi_j \leq \Lambda |\xi|^2$$

for all $\xi \in \mathbb{R}^n$. For $R > 0$, $p_0 \in M^n$ and $t_0 \in [0,T)$, define the spatial neighbourhood

$$\begin{aligned} U_R(p_0, t_0) &:= X^{-1}\big(B_R(X(p_0,t_0)) \cap M_{t_0}\big) \\ &= \big\{p \in M^n \,\big|\, |X(p,t_0) - X(p_0,t_0)| < R\big\}\,, \end{aligned}$$

the parabolic neighbourhood

$$\begin{aligned} Q_R(p_0,t_0) &:= \big\{(p,t) \in M^n \times \big(t_0 - R^2, t_0\big] \,\big|\, |X(p,t) - X(p_0,t)| < R\big\} \\ &= \bigcup_{t \in (t_0 - R^2, t_0]} (U_R(p_0,t) \times \{t\}) \end{aligned}$$

and, for an open set $U \subset M^n$ and $[t_1, t_0] \subset [0,T)$, the parabolic boundary

$$\mathcal{P}(U \times [t_1,t_0]) := (U \times \{t_1\}) \cup (\partial U \times (t_1, t_0])\,.$$

Theorem D.2 (Weak parabolic maximum principle)**.** *Let $U \subset M^n$ be open and let $f \in C^{2;1}(Q) \cap C^0(\mathcal{P}Q)$ for $Q := U \times [t_1, t_0]$.*

(i) If $L(f) \geq 0$ on Q and $f \geq 0$ on $\mathcal{P}Q$. Then $f \geq 0$ in Q.
(ii) If $L(f) \leq 0$ on Q and $f \leq 0$ on $\mathcal{P}Q$. Then $f \leq 0$ in Q.

Theorem D.3 (Strong parabolic maximum principle)**.** *Let $U \subset M^n$ be open, $Q := U \times [0,T)$, and $f \in C^{2;1}(Q) \cap C^0(\bar{Q})$.*

(i) Let $L(f) \geq 0$ in Q. If there exists $(p_0,t_0) \in Q \backslash \mathcal{P}Q$ with $f(p_0,t_0) = \min_{\bar{Q}} f$, then f is constant in $\bar{Q}$.
(ii) Let $L(f) \leq 0$ in Q. If there exists $(p_0,t_0) \in Q \backslash \mathcal{P}Q$ with $f(p_0,t_0) = \max_{\bar{Q}} f$, then f is constant in $\bar{Q}$.

D.1. **2-tensors.** We follow the lines of [CCG⁺08, Chapter 12]. Let $T > 0$ and $(M^n, g(t))_{t \in [0,T)}$ a closed manifold with a family of metrics, that depend smoothly on time. Let $m = (m_{ij})_{1 \leq i,j \leq n}$ be symmetric with $m_{ij} \in C^\infty(M^n \times [0,T))$. Let $b = (b_{ij}(m,p,t))_{1 \leq i,j \leq n}$ be symmetric with $b_{ij} \in C^1(M^n \times [0,T))$ and satisfy the null eigenvector condition, that is, if $m_{ij}\xi^j = 0$ for $1 \leq i \leq n$ then also $b_{ij}\xi^i\xi^j \geq 0$. Let $u^k \in L^\infty(M^n \times [0,T))$, $1 \leq k \leq n$.

Theorem D.4 (Weak parabolic maximum principle for 2-tensors)**.** *Let*

$$\partial_t m_{ij} \succeq \Delta_{g(t)} m_{ij} + u^k \nabla_k^{g(t)} m_{ij} + b_{ij}(m_{kl}, \cdot)$$

in $M^n \times (0,T)$ and $m_{ij}(\cdot, 0) \succeq 0$. Then $m_{ij}(\cdot,t) \succeq 0$ for $0 \leq t < T$.

Proof. See e.g. [Sch17c, Theorem 4.2] □

Theorem D.5 (Strong parabolic maximum principle for 2-tensors I, Hamilton [Ham86, Lemma 8.2])**.** *Let b be locally Lipschitz in m. Let*

$$\partial_t m_{ij} = \Delta_{g(t)} m_{ij} + u^k \nabla_k^{g(t)} m_{ij} + b_{ij}(m_{kl}, \cdot)$$

in $M^n \times (0,T)$, $m_{ij}(\cdot, 0) \succeq 0$ for all $t \in [0,T)$ and $m_{ij}(p_0, 0) \succ 0$ for $p_0 \in M^n$. Then $m_{ij}(\cdot, t) \succ 0$ for $0 < t < T$.

Proof. We follow the lines of [CCG+08, Theorem 12.47]. Let $p \in M^n$ and $U \subset M^n$ so that $p, p_0 \in U$ and so that $\overline{U}$ is a compact manifold with smooth boundary. Define $\varphi_1 : \overline{U} \times [0,T) \to \mathbb{R}$ by

$$\begin{aligned} \varphi_1 \leq \lambda_1(\cdot, 0) &\quad \text{in } \overline{U} \\ \varphi_1 \equiv 0 &\quad \text{on } \partial U \\ 2\varphi_1(p_0) \geq \lambda_1(p_0, 0)\,. & \end{aligned}$$

Let $C > 0$ to be chosen later and let $f : \overline{U} \times [0,T) \to \mathbb{R}$ a solution of

$$\begin{aligned} \partial_t f = \Delta_{g(t)} f + u^k \nabla_k^{g(t)} f - Cf &\quad \text{in } U \times (0,T) \\ f \equiv 0 &\quad \text{on } \partial U \times [0,T) \\ f(\cdot, 0) = \varphi_1 &\quad \text{in } U\,. \end{aligned}$$

Since $m_{ij}(p_0, 0) > 0$, we also have $\varphi_1(p_0) > 0$. The strong maximum principle for functions, Theorem D.3, yields that $f > 0$ in $U \times (0,T)$. The weak maximum principle, Theorem D.2, yields

$$f(x,t) \leq \max_{p \in \overline{U}} \varphi_1(x) \leq \max_{p \in \overline{U}} \lambda_1(x, 0)$$

in $U \times (0,T)$. Define the tensor

$$\tilde{m}_{ij} = m_{ij} + (\varepsilon e^{Ct} - f)\delta_{ij}\,,$$

where $\varepsilon > 0$. Then

$$\tilde{m}_{ij} \succeq \lambda_1 \delta_{ij} + (\varepsilon e^{Ct} - \lambda_1)\delta_{ij} \succ 0$$

and

$$\begin{aligned} \partial_t \tilde{m}_{ij} &= \partial_t m_{ij} + \left(\varepsilon C e^{Ct} - \partial_t f\right)\delta_{ij} \\ &= \Delta_{g(t)}(m_{ij} - f\delta_{ij}) + u^k \nabla_k^{g(t)}(m_{ij} - f\delta_{ij}) \\ &\quad + b_{ij}(m_{kl}) + C\left(\varepsilon e^{Ct} + f\right)\delta_{ij} \\ &= \Delta_{g(t)} \tilde{m}_{ij} + u^k \nabla_k^{g(t)} \tilde{m}_{ij} + b_{ij}(\tilde{m}_{kl}) \\ &\quad - (b_{ij}(\tilde{m}_{kl}) - b_{ij}(m_{kl})) + C\left(\varepsilon e^{Ct} + f\right)\delta_{ij}\,. \end{aligned}$$

Since b_{ij} is Lipschitz in m_{ij},

$$b_{ij}(\tilde{m}_{kl}) - b_{ij}(m_{kl}) \preceq \mathrm{Lip}(b_{kl})(\tilde{m}_{ij} - m_{ij}) = \mathrm{Lip}(b_{kl})(\varepsilon e^{Ct} + f)\delta_{ij}\,.$$

By choosing $C \geq \mathrm{Lip}(b_{ij})$ and ε such that $\varepsilon \leq e^{-Ct}$, we obtain

$$\begin{aligned} \partial_t \tilde{m}_{ij} &\succeq \Delta_{g(t)} \tilde{m}_{ij} + u^k \nabla_k^{g(t)} \tilde{m}_{ij} + b_{ij}(\tilde{m}_{kl}) \\ &\qquad + (C - \mathrm{Lip}(b_{ij}))\left(\varepsilon e^{Ct} + f\right)\delta_{ij} \\ &\succeq \Delta_{g(t)} \tilde{m}_{ij} + u^k \nabla_k^{g(t)} \tilde{m}_{ij} + b_{ij}(\tilde{m}_{kl})\,. \end{aligned}$$

The weak maximum principle, Theorem D.4, implies $\tilde{m}_{ij} \succeq 0$ on $\overline{U} \times [0,T)$ for $\varepsilon \in (0, e^{-Ctt}]$. Thus $m_{ij} \succeq (-\varepsilon e^{Ct} + f)\delta_{ij}$ on $\overline{U} \times [0,T)$ for $\varepsilon \in (0, e^{-Ctt}]$. Letting $\varepsilon \to 0$ yields $m_{ij} \succeq f\delta_{ij} \succ 0$ on $\overline{U} \times [0,T)$. $\square$

Theorem D.6 (Strong parabolic maximum principle for 2-tensors II)**.** *Let*

$$\begin{aligned}\phi_k(p,t) &:= \inf_{\{\tau_1,\dots,\tau_k\}\ \textit{orthonormal}} (m(\tau_1,\tau_1)+\cdots+m(\tau_k,\tau_k))\\ &= \lambda_1(p,t)+\cdots+\lambda_k(p,t)\end{aligned}$$

where $k\in\{1,\dots,n\}$*. Let* b *be locally Lipschitz in* m*. Let*

$$\partial_t m_{ij} = \Delta_{g(t)} m_{ij} + u^k \nabla_k^{g(t)} m_{ij} + b_{ij}(m_{kl},\cdot)$$

in $M^n\times(0,T)$*,* $\phi_k(\cdot,0)\geq 0$ *in* M^n *and* $\phi_k(p_0,0)>0$ *for* $k\in\{1,\dots,n\}$ *and* $p_0\in M^n$*. Then* $\phi_k(\cdot,t)>0$ *for* $0<t<T$*.*

Proof. We follow the lines of [CCG+08, Theorem 12.49]. Let $p\in M^n$ and $U\subset M^n$ so that $p,p_0\in U$ and so that $\overline{U}$ is a compact manifold with smooth boundary. Define $\varphi_k:\overline{U}\times[0,T)\to\mathbb{R}$ by

$$\begin{aligned}k\varphi_k &\leq \phi_k(\cdot,0) &&\text{in } \overline{U}\\ \varphi_k &\equiv 0 &&\text{on } \partial U\\ k\varphi_k(p_0) &\geq \lambda_1(p_0,0)\,.\end{aligned}$$

Let $C>0$ to be chosen later and let $f:\overline{U}\times[0,T)\to\mathbb{R}$ a solution of

$$\begin{aligned}\partial_t f = \Delta_{g(t)} f + u^k\nabla_k^{g(t)} f - Cf &\quad \text{in } U\times(0,T)\\ f\equiv 0 &\quad \text{on } \partial U\times[0,T)\\ f(\cdot,0)=\varphi_k &\quad \text{in } U\,.\end{aligned}$$

Since $\phi_k(p_0,0)>0$, we also have $\varphi_k(p_0)>0$. The strong maximum principle for functions, Theorem D.3, yields that $f>0$ in $U\times(0,T)$. The weak maximum principle, Theorem D.2, yields

$$f(x,t)\leq \max_{p\in\overline{U}}\varphi_k(x)\leq \max_{p\in\overline{U}}\phi_k(x,0)$$

in $U\times(0,T)$. Define the tensor

$$\tilde{m}_{ij} = m_{ij} + (\varepsilon e^{Ct}-f)\delta_{ij}\,,$$

for $\varepsilon>0$ and

$$\begin{aligned}\tilde{\phi}_k(p,t) &:= \inf_{\{\tau_1,\dots,\tau_k\}\ \text{orthonormal}} (\tilde{m}(\tau_1,\tau_1)+\cdots+\tilde{m}(\tau_k,\tau_k))\\ &= \phi_k(x,t)+k(\varepsilon e^{Ct}-f(x,t))\,.\end{aligned}$$

We want to show that $\tilde{\phi}_k>0$ on $\overline{U}\times[0,T)$ for $\varepsilon>0$ small enough. Assume the opposite. Since $\tilde{\phi}_k>0$ in $U\times\{0\}$ and $\partial U\times[0,T)$, there exists a point $(p_1,t_1)\in U\times[0,T)$ with

$$\tilde{\phi}_k(p_1.t_1)=0 \qquad\text{and}\qquad \tilde{\phi}_k(p.t)>0 \ \text{ for all } (p,t)\in U\times[0,t_1)\,.$$

Let $\boldsymbol{\tau}_1^0,\dots\boldsymbol{\tau}_k^0\in T_{p_1}M^n$ be orthonormal with

$$\tilde{m}(\boldsymbol{\tau}_1^0,\boldsymbol{\tau}_1^0)+\cdots+\tilde{m}(\boldsymbol{\tau}_k^0,\boldsymbol{\tau}_k^0)=0$$

in (p_1,t_1). Extend each $\boldsymbol{\tau}_i^0$ in space and time to a lokal vectorfield $\boldsymbol{\tau}_i$ by parallel translation of $\boldsymbol{\tau}_i^0$ along geodesics starting from p_1 with respect to $\nabla^{g(t_1)}$ and constant in time. Then

$$\nabla\boldsymbol{\tau}_i(p_1,t_1)=0\,,\ \Delta\boldsymbol{\tau}_i(p_1,t_1)=0\,,\ \partial_t\boldsymbol{\tau}_i(p_1,t_1)=0\,.$$

Define in a neighbourhood of (p_1,t_1)

$$\psi_k(p,t):=\tilde{m}(p,t)(\tau_1,\tau_1)+\cdots+\tilde{m}(p,t)(\tau_k,\tau_k)$$

where $\psi_k(p_1,t_1)=0$ and

$$\psi_k(p,t)\geq\tilde{\phi}_k(p,t)\geq 0$$

for all $p \in U$ and $t \in [0, t_1]$. At (p_1, t_1), we have

$$\begin{aligned}
0 &\geq (\partial_t - \Delta - u^l \nabla_l)\psi_k \\
&= \sum_{i=1}^{k} (\partial_t - \Delta - u^l \nabla_l)\tilde{m}(\boldsymbol{\tau}_i^0, \boldsymbol{\tau}_i^0) \\
&= \sum_{i=1}^{k} b(\tilde{m})(\boldsymbol{\tau}_i^0, \boldsymbol{\tau}_i^0) - \sum_{i=1}^{k} \left(b(\tilde{m}) - b(m)\right)(\boldsymbol{\tau}_i^0, \boldsymbol{\tau}_i^0) + C\left(\varepsilon e^{Ct} + f\right) \\
&\geq \left(kC - \sum_{i=0}^{k} \operatorname{Lip}(b)(\boldsymbol{\tau}_i^0, \boldsymbol{\tau}_i^0)\right)\left(\varepsilon e^{Ct} + f\right) > 0
\end{aligned}$$

if we choose $C \geq \operatorname{Lip}(b_{ij})$ and ε such that $\varepsilon \leq e^{-Ct}$. This is a contradiction. Hence, $\tilde{\phi}_k > 0$ on $\overline{U} \times [0, T)$ for $\varepsilon \leq e^{-Ct}$. Thus $\phi_k \geq -k(\varepsilon e^{Ct} - f)$ on $\overline{U} \times [0, T)$ for $\varepsilon \in (0, e^{-Ctt}]$. Letting $\varepsilon \to 0$ yields $\psi_k \geq f > 0$ on $\overline{U} \times [0, T)$. □

Theorem D.7 (Strong parabolic maximum principle for 2-tensors III, Hamilton [Ham86, Section 8])**.** *Let b be locally Lipschitz in m. Let*

$$\partial_t m_{ij} = \Delta_{g(t)} m_{ij} + u^k \nabla_k^{g(t)} m_{ij} + b_{ij}(m_{kl}, \cdot)$$

in $M^n \times (0, T)$ and $m_{ij}(\cdot, 0) \succeq 0$ for all $t \in [0, T)$. Then

(i) If $t_2 > t_1$ in $[0, T)$, then

$$\inf_{p \in M^n} \operatorname{rank} m(p, t_2) \geq \sup_{p \in M^n} \operatorname{rank} m(p, t_1)$$

and there exists $\delta > 0$ so that $\operatorname{rank} m(p, t)$ is constant for all $p \in M^n$ and $t \in (0, \delta)$.

(ii) ($\ker m$ is smooth in space and time). Let $(0, \delta)$ be the time interval from (i). Then, $\ker m(t) \subset TM^n$ is a smooth subspace which depends smoothly on time for $t \in (0, \delta)$.

(iii) ($\ker m$ is parallel in space and time). Let $(0, \delta)$ be the time interval from (i). Then, $\ker m(t)$ is invariant under parallel transport in space and constant in time for $t \in (0, \delta)$.

Proof. See [CCG$^+$08, Theorem 12.50]. □

We also need the following two variants of the previous theorems. A vectorfield $v = v^i \partial_i$ is called time-parallel provided

$$\partial_t v^i = -\frac{1}{2} g^{ij} (\partial_t g_{jk}) v^k \,.$$

Since $\partial_t (g_{ij} v^i v^j) = 0$, the length of v is constant in time.

Theorem D.8 (Stong maximum principle for 2-tensors IV, White [Whi03, Propositions A.2 and A.3])**.** *Let $\Omega \subset \mathbb{R}^n$ be open and connected. Let m_{ij} be a smooth time-dependent symmetric 2-tensorfield such that*

$$\partial_t (m_{ij} v^i v^j) \geq (\Delta m_{ij}) v^i v^j$$

for all time-parallel vectorfields v. Let λ be the smallest eigenvalue of m. If the minimum value of λ on $\Omega \times (a, b]$ occurs at (p, b), then λ is constant on $\Omega \times (a, b]$. Furthermore, at each time $t \in (a, b]$, Ω is locally isometric to a product $N_1 \times N_2$ of two Riemannian manifolds N_1 and N_2, where $v \perp TN_2$ if and only if v is an eigenvector of m with eigenvalue λ. Moreover, let $v \in TN_1$, $w \in TN_2$ and $V \in T\Omega$, then $\nabla_V v \in TN_1$ and $\nabla_V w \in TN_2$.

Proof. Given a spacetime point $x = (p,t)$, let $v = v_x$ be a unit vector such that $m(v,v) = \lambda$. Extend v to a unit vectorfield $v(\cdot, t)$ at time t by parallel translation along geodesics starting from p. This way of extending v guarantees that

$$(\Delta m)(v,v) = \Delta(m(v,v)) \tag{D.1}$$

at (p,t). Now extend v as a time-parallel vectorfield on $\Omega \times (a,b]$. Then v is a unit vectorfield so

$$\lambda \le m(v,v)\,, \tag{D.2}$$

with equality at (p,t). Suppose for the moment that λ is a smooth function on $\Omega \times (a,b]$. Then by (D.1) and (D.2),

$$\partial_t \lambda = \partial_t(m(v,v)) \ge (\Delta m)(v,v) = \Delta(m(v,v)) \ge \Delta\lambda \tag{D.3}$$

at the point (p,t). Thus if λ is smooth, then

$$\partial_t \lambda \ge \Delta\lambda. \tag{D.4}$$

Even if λ is not smooth, the derivation just given shows that (D.4) holds in a viscosity sense. (In the nonsmooth case, one should think of $\partial_t\lambda$ as

$$\lim_{h\to 0} \inf_{h>0} \frac{\lambda(x,t) - \lambda(x,t-h)}{h}\,.$$

Then by (D.2), we will still have $\partial_t\lambda \ge \partial_t(m(v,v))$ at (p,t).) The strict maximum principle, Theorem D.3, then implies that λ is constant. Now consider the point (p,t) and the special vectorfield v defined above. Since λ is constant, the first and last terms in (D.3) vanish. This forces all the terms to vanish, in particular

$$(\Delta m)(v,v)(p,t) = 0\,.$$

(The argument for nonsmooth λ goes as follows. The maximum principle for smooth λ is proved using smooth functions f such that $\partial_t f < \Delta f$ and then observing that it is impossible for $\lambda - f$ to attain a minimum (on certain domains). In the nonsmooth case, note that if $\lambda - f$ attained a minimum at a spacetime point x, then for $v = v_x$, the function $\bar{f} := m(v,v) - f$ would also have a minimum at the spacetime point x, which readily gives a contradiction since $\bar{f}$ is a smooth function with $\partial_t \bar{f} > \Delta \bar{f}$.)

For the last claim, without loss of generality, we may assume that $\lambda = 0$; otherwise replace m by $m - \lambda g$. Fix a time t. It suffices to prove the conclusion on an open dense subset of Ω. Since the nullity (dimension of the nullspace) of m is locally constant on a dense open subset of Ω, we may assume it is constant throughout Ω. Now fix some point (p,t). Let $\{e_i\}$ be a g-orthonormal basis at (p,t), and extend (spatially) by parallel translation along geodesics emanating from p; this guarantees that $\Delta T = \nabla_{e_i}(\nabla_{e_i} T)$ for any tensor field T. Now $m(v,\cdot) = 0$, so

$$\begin{aligned}
0 = \Delta(m(v,v)) &= \nabla_{e_i}(\nabla_{e_i}(m(v,v)))) \\
&= \nabla_{e_i}((\nabla_{e_i} m)(v,v) + 2m(\nabla_{e_i} v, v)) \\
&= (\Delta m)(v,v) + 2(\nabla_{e_i} m)(\nabla_{e_i} v, v) \\
&= 2\nabla_{e_i}(m(\nabla_{e_i} v, v)) - 2m(\nabla_{e_i} v, \nabla_{e_i} v) = -2m(\nabla_{e_i} v, \nabla_{e_i} v)\,.
\end{aligned}$$

Since m is positive semidefinite, this means $\nabla_{e_i} v$ is in the nullspace of m at (p,t) for each i. Thus for any vector V, the vector $\nabla_V v$ is in the nullspace at (p,t). Since (p,t) is arbitrary, in fact this holds everywhere. In other words, if v is a null vectorfield and V is an arbitrary vectorfield, then $\nabla_V v$ is also a null vectorfield. By the Frobenius theorem, Theorem B.2, the nullspaces of m form an integrable distribution. (Note that the leaves of the foliation are totally geodesic.) Now suppose V is an arbitrary vectorfield, v is a nullvectorfield, and that w is a vectorfield everywhere perpendicular to the nullvectors. Then

$$0 = \nabla_V \langle w, v\rangle = \langle \nabla_V w, v\rangle + \langle w, \nabla_V v\rangle = \langle \nabla_V w, v\rangle \,.$$

Thus (again by Frobenius) the orthogonal complements of the nullspaces of m form an integrable distribution, and the leaves are totally geodesic. Thus we can find a coordinate system $\{p^i\}$ such that

$$g = \begin{pmatrix} (g_{ij})_{1\le i,j\le m} & 0 \\ 0 & (g_{\alpha\beta})_{m+1\le\alpha,\beta\le n} \end{pmatrix} .$$

Since $g_{i\alpha} = 0$, the Christoffel symbol simplify to

$$\Gamma_{ij}^{\alpha} = -\frac{1}{2} g^{\alpha\beta}\partial_\beta g_{ij} \,.$$

Since the horizontal leaves are totally geodesic, Γ_{ij}^{α} vanishes for all α, which implies that $\partial_\beta g_{ij} = 0$, so g_{ij} does not depend on p^β. Notice this holds for all i, j and β. Likewise $g_{\alpha\beta}$ does not depend on any of the p^i. Thus g is a product metric. □

References

[AE06] H. Amann and J. Escher, *Analysis II*, Birkhäuser, 2006.

[AL86] U. Abresch and J. Langer, *The normalized curve shortening flow and homothetic solutions*, J. Diff. Geom. **23** (1986), no. 2, 175–196.

[And94] B. Andrews, *Harnack inequalities for evolving hypersurfaces*, Math. Z. **217** (1994), 179–197.

[And12] ———, *Noncollapsing in mean-convex mean curvature flow*, Geometry & Topology **16** (2012), no. 3, 1413–1418.

[Bak10] C. Baker, *The mean curvature flow of submanifolds of high codimension*, Ph.D. thesis, Australian National University, November 2010.

[BJ73] Theodor Bröcker and Klaus Jänich, *Einführung in die Differentialtopologie*, Heidelberger Taschenbücher, vol. 143, Springer-Verlag, Berlin, 1973.

[Bra78] K. A. Brakke, *The motion of a surface by its mean curvature*, Math. Notes, Princeton University Press, 1978.

[CCG+08] B. Chow, S.-C. Chu, D. Glickenstein, C. Guenther, J. Isenberg, T. Ivey, D. Knopf, P. Lu, F. Luo, and L. Ni, *The Ricci Flow: Techniques and Applications: Part II: Analytic Aspects*, Mathematical Surveys and Monographs, vol. 144, American Mathematical Society, 2008.

[CM12] T. H. Colding and W. P. Minicozzi, *Generic mean curvature flow i; generic singularities*, Annals of Mathematics **175** (2012), 755–833, http://dx.doi.org/10.4007/annals.2012.175.2.7.

[Coo11] A. A. Cooper, *A compactness theorem for the second fundamental form*, Preprint: arXiv:1006.5697v4, 2011.

[DHTK10] U. Dierkes, S. Hildebrandt, A. Tromba, and A. Küster, *Regularity of minimal surfaces*, 2nd ed., Grundlehren der mathematischen Wissenschaften, Springer, 2010.

[Eck04] K. Ecker, *Regularity theory for mean curvature flow*, Birkhäuser, 2004.

[EH89] K. Ecker and G. Huisken, *Interior curvature estimates for hypersurfaces of prescribed mean curvature*, Annales de l'Institut Henri Poincaré (C) Analyse non linéaire **6** (1989), 251–260.

[EH91] ———, *Interior estimates for hypersurfaces moving by mean curvature*, Invent. Math. **105** (1991), no. 1, 547–569 (English).

[Eva02] L. C. Evans, *Partial differential equations*, American Mathematical Society, 2002.

[Fed69] H. Federer, *Geometric measure theory*, Grundlehren der mathematischen Wissenschaften, vol. 153, Springer, Berlin, Heidelberg, New York, 1969.

[GH86] M. E. Gage and R. S. Hamilton, *The heat equation shrinking convex plane curves*, J. Diff. Geom. **23** (1986), 69–96.

[Ham86] R. S. Hamilton, *Four-manifolds with positive curvature operator*, J. Diff. Geom. **24** (1986), no. 2, 153–179.

[Ham94] ———, *Convex hypersurfaces with pinched second fundamental form*, Comm. Anal. Geom. **2** (1994), no. 1, 167–172.

[Ham95a] ———, *The formation of singularities in the Ricci flow*, Proceedings of the conference on geometry and topology held at Harvard University April 23–25, 1993 (Cambridge MA) (C. C. Hsiung and S.-T. Yau, eds.), Surveys in Differential Geomerty, vol. 2, International Press of Boston, Inc., 1995, pp. 7–136.

[Ham95b] ———, *Harnack estimate for the mean curvature flow*, J. Diff. Geom. **41** (1995), no. 1, 215–226.

[HK17] R. Haslhofer and B. Kleiner, *Mean curvature flow of mean convex hypersurfaces*, Comm. Pure Appl. Math. **70** (2017), no. 3, 0511–0546.

[HL99] F. Hirsch and G. Lacombe, *Elements of functional analysis*, Graduate Texts in Mathematics, vol. 192, Springer, New York, 1999.

[HS99a] G. Huisken and C. Sinestrari, *Convexity estimates for mean curvature flow and singularities of mean convex surfaces*, Acta Math. **183** (1999), no. 1, 45–70.

[HS99b] ———, *Mean curvature flow singularities for mean convex surfaces*, Calc. Var. **8** (1999), no. 1, 1–14.

[HS09] ———, *Mean curvature flow with surgeries of two–convex hypersurfaces*, Invent. Math. **175** (2009), 137–221.

[Hui84] G. Huisken, *Flow by mean curvature of convex surfaces into spheres*, J. Diff. Geom. **20** (1984), no. 1, 237–266.

[Hui90] ———, *Asymptotic behavior for singularities of the mean curvature flow*, J. Diff. Geom. **31** (1990), no. 1, 285–299.

[Hui93] ———, *Local and global behaviour of hypersurfaces moving by mean curvature*, Differential geometry. Part 1: Partial differential equations on manifolds. Proceedings of a summer research institute, held at the University of California, Los Angeles, CA, USA, July 8–28, 1990 (Providence, RI) (R. Greene et al., ed.), Proc. Symp. Pure Math., vol. 54, American Mathematical Society, 1993, pp. 175–191.

[Lan85] J. Langer, *A compactness theorem for surfaces with l^p-bounded second fundamental form*, Math. Annalen **270** (1985), 223–234.

[Man11] C. Mantegazza, *Lecture notes on mean curvature flow*, Birkhäuser, 2011.

[MB14] E. Mäder-Baumdicker, *The area preserving curve shortening flow with Neumann free boundary conditions*, Doctoral thesis, Albert-Ludwigs-Universität Freiburg, 2014.

[Pih98] D. M. Pihan, *A length preserving geometric heat flow for curves*, Ph.D. thesis, University of Melbourne, September 1998.

[Sch05] O. Schnürer, *Differentialgeometrie ii*, Lecture notes, 2005.

[Sch17a] ———, *Differentialgeometrie i*, Lecture notes, 2017.

[Sch17b] ———, *Partielle Differentialgleichungen 1*, Lecture notes, 2017.

[Sch17c] ———, *Partielle Differentialgleichungen 1a*, Lecture notes, 2017.

[Sch17d] F. Schulze, *Introduction to mean curvature flow*, LSGNT course, 2017.

[Sch18] O. Schnürer, *Graphischer Mittlerer Krümmungsfluss*, Lecture notes, 2018.

[Sim83] L. Simon, *Lectures on geometric measure theory*, Proceedings of the Centre for Mathematical Analysis, vol. 3, Australian National University, 1983.

[Urb91] J. I. E. Urbas, *An expansion of convex hypersurfaces*, J. Diff. Geom. **33** (1991), no. 1, 91–125.

[Whi03] B. White, *The nature of singularities in mean curvature flow of mean-convex sets*, J. Amer. Math. Soc. **16** (2003), no. 1, 123–138.

[Whi05] ———, *A local regularity theorem for mean curvature flow*, Ann. Math. **161** (2005), no. 3, 1487–1519.