TOPICS IN DIFFERENTIAL GEOMETRY
MEAN CURVATURE FLOW
MATH 258, WINTER 2016-2017
OR HERSHKOVITS

NOTES BY CHAO LI, EVANGELIE ZACHOS

Contents

Date: August 30, 2017.

These are the notes of Or Hershkovits' course on mean curvature flow taught at Stanford University in the Winter of 2016-2017. We would like to thank Or Hershkovits for an excellent class. Please be aware that it is likely that we have introduced numerous typos and mistakes in our compilation process, and would appreciate it if these are brought to our attention.

This course will focus on the theory of mean curvature flow and its applications. Topics covered include curve shortening flows in $\mathbb{R}^2$, mean convex flows, mean curvature flows with surgery, and their applications to various geometric and topological questions such as the Riemannian Penrose inequality, and the path-connectedness of the space of all 2-convex embedded spheres. The course will assume the reader to know basic submanifold geometry. Although some familiarity of heat equation and the maximum principle will be helpful.

1. Introduction and overview

1.1. Geometric background.

Definition 1.1. Let M^n be a compact smooth manifold. A family of embeddings of M^n into $\mathbb{R}^{n+1}$, $\varphi : M \times [0,T] \to \mathbb{R}^{n+1}$ is said to be evolved by mean curvature flow, if

$$\frac{d\varphi}{dt} = \vec{H}(\varphi(x,t)),$$

where $\vec{H}$ is the mean curvature vector.

Remark 1.2. For a submanifold $M \subset \mathbb{R}^{n+1}$, let us recall that

$$\nabla_X^{\mathbb{R}^{n+1}} Y = \nabla_X^M Y + A(X,Y), \forall X, Y \in TM.$$

Here A is the vector valued second fundamental form. The mean curvature vector is defined by $\vec{H} = \operatorname{tr} A = \sum_{i=1}^n A(E_i, E_i)$, $\{E_i\}$ is an orthonormal frame.

Remark 1.3. Let us view the mean curvature vector in a different way. By definition,

$$\begin{aligned}
\operatorname{Hess}^M \varphi(X,Y) &= XY\varphi - (\nabla_X^M Y)\varphi \\
&= XY\varphi - (\nabla_X^{\mathbb{R}^{n+1}} Y)\varphi + A(X,Y)\varphi \\
&= A(X,Y)\varphi
\end{aligned}$$

Therefore $\vec{H} = \operatorname{tr} \operatorname{Hess}^M \varphi = \Delta\varphi$. So the mean curvature flow can be written as:

$$\frac{d\varphi}{dt} = \Delta_{g(t)}\varphi.$$

Hence it is a geometric heat equation.

Remark 1.4. Another way to understand the mean curvature flow is as follows. Let X be a compactly supported vector field in $\mathbb{R}^{n+1}$, ϕ_s is the one-parameter family of diffeomorphisms it generates. Then we may derive without much difficulty that

$$\left.\frac{d}{ds}\right|_{s=0} \operatorname{Vol}(\phi_s(M)) = -\int_M \left\langle \vec{H}, X \right\rangle d\operatorname{Vol}_M.$$

Hence the mean curvature vector is the direction where the volume of the submanifold decreases the fastest. In other words, the mean curvature flow is the gradient flow of the area functional.

The first question in the study of mean curvature flow is its existence. By the basic theory of general heat equation, if the initial data is C^3 then there exists smooth mean curvature flow starting from it. In general, one may ask how rough can the initial surface be so that there is short time existence, and what regularity properties the flow has. Let us mention some results in this direction.

The first result is due to Ecker and Huisken, who proved that if the initial surface is locally Lipschitz, then the mean curvature flow has short time existence. Here we call a hypersurface (R, C)- locally Lipschitz, if for any point p, $M \cap B(p, R)$ is a Lipschitz graph over some hyperplane with Lipschitz constant bounded by C. Their results also enable one to control the Lipschitz constant of the short time solution by the Lipschitz constant of the initial data. Note that such a gradient estimate does not hold for the usual heat equation. For instance, consider the parabolic square $P = [0, 1] \times [0, T)$. Let a C^2 function u solves $u_t = \Delta u$ in P with initial data $u = 0$ on $[0, 1] \times \{0\}$. Then at a given point $(\xi, t) \in P$, there is no control on $|u'(\xi, t)|$. For the mean curvature flow, by Ecker and Huisken's result, the Lipschitz constant on an open ball controls the Lipschitz constant in short time on a smaller ball.

Another result worth mentioning here is due to Laver, who proved that for a continuous curve $\gamma : S^1 \to \mathbb{R}^2$, such that the 2-Hausdorff measure of the image of γ is 0, there exists a unique mean curvature flow starting from it. This shows that the initial data of a curve shortening flow can be very rough (imagine a square-filling curve whose 2-Hausdorff measure is 0).

1.2. **Useful tools.** The first useful tool in the study of mean curvature is the avoidance principle. It is an application of the maximum principle of the heat equation. Roughly speaking, if two hypersurfaces do not touch initially, then under the mean curvature flow they won't touch.

Theorem 1.5. *Let $\{M_t\}_{t\in[a,b]}$ and $\{M'_t\}_{t\in[a,b]}$ be two family of embedded hypersurfaces evolved by the mean curvature flows, and suppose $M_a \cap M'_a = \phi$. Then $M_t \cap M'_t = \phi$, for all $t \in [a, b]$.*

To get an immediate corollary, let us look the evolution of a sphere of radius R_0. It is easy to see that under the mean curvature flow, $R'(t) = -\frac{n}{R}$. Hence $R(t) = \sqrt{R_0^2 - 2nt}$. Since any compact embedded hypersurface is enclosed by a sphere, we conclude that the mean curvature flow extincts in finite time (in other words, the flow becomes singular in finite time).

Another useful method that we will discuss is the blow-up analysis. To study the nature of singularities of the mean curvature flow, let us distinguish two cases. Let T be the first singular time of the mean curvature flow, A is the second fundamental form of the embedding.

- There is some constant C such that $|A|(\cdot, t) \leq \frac{C}{\sqrt{T-t}}$.
- $\max_{M_t} \sqrt{T-t}|A|$ is unbounded.

In the first case, by rescaling the flow we are able to take a subsequencial limit to get a tangent flow. Using Huisken's monotonicity formula, we are able to get classification in some situations.

In the second case, by taking a different limit we may get a convex limit. Then using analytic results specifically for convex flows, we are able to make conclusions.

As an example, we will see that for the curve shortening flow, by using the Hamilton-Li-Yau Harnack inequality and the two-point maximum principle, we are going to argue that the second case does not appear. And in the first case, the monotonicity formula for curves will give us a complete classification of the tangent flow. Combining these ideas, we will prove:

Theorem 1.6 (Grayson)**.** *Let γ_0 be a closed connected curve in $\mathbb{R}^2$. Then there exist $x_0 \in \mathbb{R}^2$ and $T < \infty$, such that the mean curvature flow γ_t emenating from γ_0 converges to a round point around x_0:*

$$\frac{\gamma_t - x_0}{\sqrt{T-t}} \to S^1(c).$$

Let us note here that in higher dimensions, other singularities may form. One of them is the famous dumpbell hypersurface in $\mathbb{R}^n$, $n \geq 3$. We will see that the singularity it form will not be round.

A third method to study the flow beyond singularity is through weak solutions- for instance, the level set flow and the Brakke flow. The main issue here is the regularity of these weak solutions. Toward that end we have the following theorem:

Theorem 1.7 (Ilmanen). *A generic embedded weak solution of the mean curvature flow of hypersurfaces in $\mathbb{R}^{n+1}$ has a singular set with parabolic dimension less than n.*

It is conjectured that the genericity is not necessary, and that the singular set should be at most $(n-1)$ dimensional.

For the special case with mean convex initial hypersurfaces, we know a little bit more:

- The singularity is at most $(n-1)$ dimensional at all time, and
- The singularity is at most $(n-3)$ dimensional at almost all time.

2. The avoidance principle

In this section we discuss the maximum principle of parabolic equations.

Theorem 2.1. *Suppose M is a compact manifold, $g(t)$ is a time dependent family of metrics, X is a time dependent vector field. Let $F : \mathbb{R} \times [0,T] \to \mathbb{R}$, $u : M \times [0,T] \to \mathbb{R}$ be two C^2 functions, satisfying*

$$\frac{du}{dt} \geq \Delta_{g(t)} u + \langle \nabla_{g(t)} u, X \rangle + F(u(x,t),t).$$

Further, suppose that $u(\cdot,0) \geq \alpha$ for some positive α.

Let $h : [0,T] \to \mathbb{R}$ be a comparasion function which solves the ODE

$$\begin{cases} \frac{dh}{dt} & = F(h(t),t) \\ h(0) & = \alpha. \end{cases}$$

Then $u(x,t) \geq h(t)$ for all $t \in [0,T], x \in M$.

Proof. Let $\epsilon > 0$. Consider h^ϵ to be the solution to

$$\begin{cases} \frac{dh^\epsilon}{dt} & = F(h^\epsilon(t),t) - \epsilon \\ h^\epsilon(0) = \alpha - \epsilon. \end{cases}$$

Claim: $u(\cdot,t) > h^\epsilon(t)$.

If not, let t_0 be the first time that the claim is contradicted. By assupmtion $t_0 > 0$, and there exists x_0 such that $u(x_0,t_0) = h^\epsilon(t_0)$. Further, $u(x,t_0) \geq u(x_0,t_0)$ for all $x \in M$. Therefore

$$\Delta_{g(t)} \geq 0, \qquad \nabla_{g(t)} u = 0.$$

Therefore

$$\frac{du}{dt}(x_0,t_0) \geq F(u(x_0,t_0)) = F(h^\epsilon(t_0),t_0) = \frac{dh^\epsilon}{dt} + \epsilon.$$

Contradiction. The claim is proved.

Since F is Lipschitz, $h^\epsilon \to h$ as $\epsilon \to 0$. Hence we conclude that $u(x,t) \geq h(t)$.

$\square$

Proposition 2.2. *Suppose that M is as above, $F : M \times [0,T] \to \mathbb{R}$ is a smooth function, and let $\varphi(t) = \min_{x \in M} F(x,t)$. Then:*

1. *φ is Lipschitz.*
2. *Let t be a time where φ is differentiable, and let $x \in M$ be such that $F(x,t) = \varphi(t)$. Then $\frac{\partial}{\partial t} F(x,t) = \frac{d}{dt}\varphi(t)$.*

Proof. Choose x_0 such that $\varphi(t_0) = F(x_0,t_0)$. Then for any t_1,

$$\varphi(t_1) - \varphi(t_0) \leq F(x_0,t_1) - F(x_0,t_0).$$

Therefore φ is Lipschitz.

Suppose further that t_0 is a time of differentiability of φ. Let $h > 0$. Then

$$\frac{1}{h}(\varphi(t_0+h) - \varphi(t_0)) \leq \frac{1}{h}(F(x_0,t_0+h) - F(x_0,t_0)).$$

Therefore

$$\left.\frac{d\varphi}{dt}\right|_{t_0} \leq \frac{\partial F}{\partial t}(x_0, t_0).$$

For similar reasons,

$$\left.\frac{d\varphi}{dt}\right|_{t_0} \geq \frac{\partial F}{\partial t}(x_0, t_0).$$

Therefore we conclude that they are equal. □

As a corollary, we prove the avoidance principle.

Corollary 2.3. *Let $\{M_t^1\}, \{M_t^2\}, t \in [a, b]$ be two family of compact embedded hypersurfaces evolving by the mean curvature flow. Suppose $M_a^1 \cap M_a^2 = \phi$. Then*

$$\frac{d}{dt}\operatorname{dist}(M_t^1, M_t^2) \geq 0$$

at points of differentiability. In particular, their distance is increasing.

Proof. Let t be a time of differentiability of the distance function. Consider the function $F : M^1 \times M^2 \times [a, b] \to \mathbb{R}^+$,

$$F(x_1, x_2, t) = \operatorname{dist}(\xi^1(x_1, t), \xi^2(x_2, t)).$$

Then F is a Lipschitz function. Suppose $x_1 \in M^1, x_2 \in M^2$ be the closest points. Then by the previous proposition,

$$\frac{d}{dt}\operatorname{dist}(M_t^1, M_t^2) = \frac{d}{dt}\operatorname{dist}(x_1, x_2).$$

Since x_1, x_2 are the closest points, their tangent planes must be parallel. Let ν be the unit normal vector.

Suppose, by contradiction, that $\frac{d}{dt}\operatorname{dist}(x_1, x_2) < 0$.

Now since the surfaces evolve by the mean curvature flow, we have

$$\frac{d}{dt}\operatorname{dist}(x_1, x_2) = \left\langle \vec{H}_1 - \vec{H}_2, \nu \right\rangle.$$

Let $E_1, \ldots, E_n$ be an orthonormal frame of $T_{x_1}M^1$. Then

$$\sum_i \langle A_2(E_i, E_i) - A_1(E_i, E_i), \nu \rangle < 0.$$

Therefore for some i we have $\langle A_1(E_i, E_i), \nu\rangle > \langle A_2(E_i, E_i), \nu\rangle$. Hence along the unit speed geodesics along E_i direction, the distance between M_1 and M_2 decreases, contradictory to the choice of x_1, x_2. □

3. Evolution of geometric quantities

3.1. **Submanifold geometry.** Recall that $\nabla_X^{\mathbb{R}^{n+1}} Y = \nabla_X^M Y + A(X, Y)$. Let $\varphi : M^n \to \mathbb{R}^{n+1}$ be the embedding, ν be the outward unit normal vector. Denote $h_{ij} = -\left\langle A(\frac{\partial\varphi}{\partial x_i}, \frac{\partial\varphi}{\partial x_j}), \nu \right\rangle$ and $\operatorname{Hess}^M \varphi(X, Y) = A(X, Y)$. Note that we have chosen the sign convention such that the second fundamental form of the sphere is positive definite.

The Christoffel simbol Γ_{ij}^k is defined via

$$\frac{\partial^2\varphi}{\partial x_i \partial x_j} - \Gamma_{ij}^k \frac{\partial\varphi}{\partial x_k} = A(\frac{\partial\varphi}{\partial x_i}, \frac{\partial\varphi}{\partial x_j}) = -h_{ij}\nu.$$

Hence

$$\frac{\partial^2\varphi}{\partial x_i \partial x_j} = \Gamma_{ij}^k \frac{\partial\varphi}{\partial x_j} - h_{ij}\nu,$$

$$\left\langle \frac{\partial \nu}{\partial x_i}, \frac{\partial \varphi}{\partial x_j} \right\rangle = -\left\langle \nu, \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \right\rangle = h_{ij}.$$

And we define the Riemannian curvature tensor to be

$$\langle R(X,Y)Z, W\rangle = \left\langle \nabla^2_{X,Y} Z - \nabla^2_{Y,X} Z, W \right\rangle,$$

so that the Gauss equation becomes

$$\langle R(X,Y)Z, W\rangle = A(X,W)A(Y,Z) - A(X,Z)A(Y,W),$$

or in local coordinates, $R_{ijkl} = h_{il}h_{jk} - h_{ik}h_{jl}$.

The Codazzi equation for hypersurfaces in the Euclidean space is

$$\nabla_k h_{ij} = \nabla_i h_{kj},$$

or ∇A is fully symmetric.

For a tensor T, commuting two covariant derivatives produces terms involves the curvature:

$$\begin{aligned}
&\nabla^2_{X,Y} T(Z_1, \dots, Z_n) - \nabla^2_{Y,X} T(Z_1, \dots, Z_n) \\
&\qquad = -T(R(X,Y)Z_1, \dots, Z_n) - \dots - T(Z_1, \dots, R(X,Y)Z_n).
\end{aligned}$$

We are now ready to derive the Simons' equations.

$$\begin{aligned}
\nabla^2_{kk} h_{ij} &= \nabla^2_{ki} h_{kj} \qquad \text{(by the Codazzi equation)} \\
&= \nabla^2_{ik} h_{kj} - R_{kikl} h_{lj} - R_{kijl} h_{lk} \\
&= \nabla^2_{ij} h_{kk} - (h_{kl}h_{ik} - h_{kk}h_{il})h_{lj} - (h_{kl}h_{ij} - h_{kj}h_{il})h_{lk} \qquad \text{(by the Gauss equation)} \\
&= \nabla^2_{ij} H + H h_{il} h_{lj} - |A|^2 h_{ij}
\end{aligned}$$

3.2. **Evolution of geometric quantities.** Recall that ν is the outward unit normal vector of the hypersurface. We first calculate the evolution of ν under the mean curvature flow.

$$\begin{aligned}
\left\langle \frac{\partial}{\partial t}\nu, \frac{\partial \varphi}{\partial x_i} \right\rangle &= -\left\langle \nu, \frac{\partial^2 \varphi}{\partial x_i \partial t} \right\rangle \\
&= -\left\langle \nu, \frac{\partial}{\partial x_i}(-H\nu) \right\rangle \\
&= \left\langle \nu, \frac{\partial H}{\partial x_i}\nu \right\rangle \\
&= \frac{\partial}{\partial x_i} H
\end{aligned}$$

Therefore $\frac{\partial}{\partial t}\nu = \nabla H$.

Next, the evolution of the metric $g_{ij} = \left\langle \frac{\partial \varphi}{\partial x_i}, \frac{\partial \varphi}{\partial x_j} \right\rangle$.

$$\begin{aligned}
\frac{\partial}{\partial t} g_{ij} &= \left\langle \frac{\partial^2 \varphi}{\partial x_i \partial t}, \frac{\partial \varphi}{\partial x_j} \right\rangle + \left\langle \frac{\partial^2 \varphi}{\partial x_j \partial t}, \frac{\partial \varphi}{\partial x_i} \right\rangle \\
&= \left\langle \frac{\partial}{\partial x_i}(-H\nu), \frac{\partial \varphi}{\partial x_j} \right\rangle + \left\langle \frac{\partial}{\partial x_j}(-H\nu), \frac{\partial \varphi}{\partial x_i} \right\rangle \\
&= 2H \left\langle \nu, \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \right\rangle \\
&= -2H h_{ij}
\end{aligned}$$

Thus we conclude

Proposition 3.1. *Suppose M_t^n is a family of embedded hypersurfaces evolving by the mean curvature flow. Then*

$$\begin{cases} \frac{\partial X}{\partial t} & = -H\nu \\ \frac{\partial}{\partial t}\nu & = \nabla H \\ \frac{\partial}{\partial t} g_{ij} & = -2Hh_{ij}. \end{cases}$$

Our goal in this section is to use the maximum principle to conclude things about how curvature changes under MCF. To this end, we note

$$\begin{aligned} \frac{\partial}{\partial t} h_{ij} = -\frac{\partial}{\partial t}\left\langle \frac{\partial^2 \phi}{\partial x_i x_j}, \nu \right\rangle &= \left\langle \frac{\partial^2}{\partial x_i x_j}(H\nu), \nu \right\rangle - \left\langle \frac{\partial^2 \phi}{\partial x_i x_j}, \nabla H \right\rangle \\ &= \frac{\partial^2}{\partial x_i x_j} H + H\left\langle \frac{\partial^2}{\partial x_i x_j}\nu, \nu \right\rangle - \Gamma_{ij}^k \frac{\partial \phi}{\partial x_k}\frac{\partial H}{\partial x_j} \\ &= \nabla_{ij}^2 H - H\left\langle h_{kj}\frac{\partial \phi}{\partial x_k}, h_{ki}\frac{\partial \phi}{\partial x_k} \right\rangle = \nabla_{ij}^2 H - Hh_{kj}h_{ki} \end{aligned}$$

where to go from the first to the second line, we used the fact that

$$\frac{\partial^2 \phi}{\partial x_i x_j} = \Gamma_{ij}^k \frac{\partial \phi}{\partial x_k} - h_{ij}\nu$$

and repeatedly used the fact that $\nabla H = \frac{\partial}{\partial t}\nu$ is tangential and. To go to the third line, we use that H is normal and so $\nabla_{ij}^2 H = \frac{\partial^2 H}{\partial x_i x_j} - \Gamma_{ij}^k \frac{\partial \phi}{\partial x_k}\frac{\partial H}{\partial x_j}$.

We are almost done, once we use Simon's equation

$$\Delta h_{ij} = \nabla_{ij}^2 H + h_{i\ell}h_{\ell j}H - |A|^2 h_{ij}$$

to conclude that

$$\frac{\partial}{\partial t} h_{ij} = \Delta h_{ij} - 2h_{kj}h_{ki} + |A|^2 h_{ij}$$

Note the similarity between this equation and the heat equation! We're almost there. Now we change perspective a little. We write things down in a coordinate invariant way, and then we exploit the fact that we are using conormal coordinates at a point (so that $g_{ik}(0) = \delta_{ik}$).

$$\begin{aligned} \frac{\partial}{\partial t} h_j^i = \frac{\partial}{\partial t}(g^{ik}h_{kj}) = -\left(\frac{\partial}{\partial t} g_{ik}\right) h_{kj} + g^{ik}\frac{\partial}{\partial t} h_{kj} &= 2Hh_{ik}h_{kj} + \Delta h_j^i - 2Hh_{ik}h_{kj} + |A|^2 h_j^i \\ &= \Delta h_j^i + |A| h_j^i \end{aligned}$$

Taking the trace of this formula gives us

$$\frac{\partial}{\partial t} H = \Delta H + |A|^2 H$$

Corollary 3.2. *By the maximum principle, if $H(0) > 0$, then $H(t) > 0$ for all times t where the flow is defined.*

Remark 3.3. $H = \langle A, I \rangle \implies |H| \leq \sqrt{n}|A| \implies$

$$\frac{\partial}{\partial t} H \geq \frac{H^3}{n} + \Delta H$$

We can apply the maximum principle to this statement, which gives us a different way of seeing that finite time blowup must occur if $H(0) > 0$. But ... at what rate does blow-up occur? To answer that question, we first see how the second fundamental form and its derivatives change with time.

$$\frac{\partial}{\partial t}|A|^2 = \frac{\partial}{\partial t}(h_j^i h_i^j) = (\Delta h_j^i + |A|^2 h_j^i)h_i^j + h_j^i(\Delta h_i^j + |A|^2 h_i^j) = 2|A|^4 + \Delta(|A|^2) - 2|\nabla A|^2$$

The last, negative, term in this expression is especially useful, as we will see.

Now, $\frac{\partial}{\partial t}|A|^2_{\max} \leq 2|A|^4_{\max}$, so $\frac{\partial}{\partial t}\left(\frac{1}{|A|^2_{\max}}\right) \geq -2$. Therefore, if $t_1 < t_2$, we have

$$\frac{1}{|A|^2_{\max}(t_2)} - \frac{1}{|A|^2_{\max}(t_1)} \geq 2(t_1 - t_2)$$

Suppose that the mean curvature flow $(M_t)_{t\in[0,T)}$ (with M_t compact) has $\sup|A|$ unbounded. Then there exist $t_k \to T$ so that $|A|_{\max}(t_k) \xrightarrow{k\to\infty} \infty$, and that

$$0 - \frac{1}{|A|^2_{\max}(t_1)} \geq 2(t_1 - T) \iff |A|_{\max}(t) \geq \frac{1}{\sqrt{2(T-t)}}$$

Now, to move on to answering how derivatives of A change with time, we note that in normal coordinates

$$\frac{\partial}{\partial t}\Gamma^k_{ij} = \frac{\partial}{\partial t}\left(g^{k\ell}(\partial_i g_{jk} + \partial_j g_{k\ell} - \partial_\ell g_{ij})\right) = (\frac{\partial}{\partial t}g^{k\ell})(0) + g^{k\ell}\left(\frac{\partial}{\partial t}\partial_i g_{j\ell} + \frac{\partial}{\partial t}\partial_j g_{kl} - \frac{\partial}{\partial t}\partial_\ell g_{ij}\right)$$

But we don't care too precisely what these remaining terms are. The point is that they are in the class $A \times \nabla A$, that is, you can write bounds on these quantities in terms of bounds on A and bounds on the derivatives of A.

In general, given a tensor T

$$\frac{\partial}{\partial t}(\nabla T) = \nabla\left(\frac{\partial}{\partial t}T\right) + Err, \quad Err \in A \times \nabla A \times T$$

and, more specifically,

$$\frac{\partial}{\partial t}|\nabla A|^2 = \Delta|\nabla A|^2 - 2|\nabla^2 A|^2 + Err, \quad Err \in A \times A \times \nabla A \times \nabla A$$

With that out of the way, we continue to "actually interesting things":

Claim 3.4. *Suppose that $\{M_t\}_{t\in[0,T)}$ is a MCF and that $|A| \leq C$ for $t \in [0,T)$. Then, $|\nabla A| \leq C_1$ over $\left[\frac{T}{2}, T\right]$.*

Proof. We do a trick that allows for an easy application of the maximum principle, using the good (negative) term from the expansion of $\frac{\partial}{\partial t}|\nabla A|^2$ above. Define $f(t) := \alpha|A|^2 + t|\nabla A|^2$ for some $\alpha > 0$ to be determined. This function has several nice properties: for example, $f(0) \in A \times A$. Also, the real key, it behaves well under $\frac{\partial}{\partial t} - \Delta$:

$$\left(\frac{\partial}{\partial t} - \Delta\right) f = \left(\frac{\partial}{\partial t} - \Delta\right)(\alpha|A|^2 + t|\nabla A|^2) =$$

$$\alpha\frac{\partial}{\partial t}|A|^2 + |\nabla A|^2 + t\frac{\partial}{\partial t}|\nabla A|^2 - \alpha\Delta(|A|^2) - t\Delta|\nabla A|^2 =$$

$$\alpha\left[2|A|^4 + \Delta(|A|^2) - 2|\nabla A|^2\right] + |\nabla A|^2 + t\Delta|\nabla A|^2 - 2t|\nabla^2 A|^2 - \alpha\Delta|A|^2 - t\Delta|\nabla A|^2 + t\, Err_{A\times A\times\nabla A\times\nabla A}$$

$$= 2\alpha|A|^4 - 2\alpha|\nabla A|^2 + |\nabla A|^2 - 2t|\nabla^2 A|^2 + t\, Err_{A\times A\times\nabla A\times\nabla A}$$

$$\leq 2\alpha C^4 - 2\alpha|\nabla A|^2 + |\nabla A|^2 + TDC^2|\nabla A|^2$$

where in the last line transition we drop a negative term, multiple terms cancel, and D is a constant representing the error. If α is large enough relative to T, D, and C, then $-2\alpha|\nabla A|^2$ dominates the positive $|\nabla A|^2$ terms, and them we have, for $M := \alpha C^4$,

$$\left(\frac{\partial}{\partial t} - \Delta\right) f \leq M \implies \text{(maximum principle)} \implies f \leq \alpha C^2 + MT$$

This means that for $t \geq T/2$,

$$|\nabla A| \leq \sqrt{\frac{\alpha C^2 + MT}{t}} \leq \frac{\sqrt{2\alpha C^2 + 2MT}}{\sqrt{T}} \leq C_1$$

$\square$

Remark 3.5. Similar estimates with similar proofs hold for $|\nabla^k A|^2$.

Theorem 3.6. *Let $(M_t)_{t\in[0,T)}$ be a MCF defined on a maximal time interval of existence. Then $|A|(t)$ is unbounded, which, as we have shown, implies that $|A|_{\max}(t) \geq \frac{1}{\sqrt{2(T-t)}}$.*

Proof. Suppose $|A| \leq C$, moving by MCF. Then for $s < t$,

$$|\phi(x,t) - \phi(x,s)| \leq \int_s^t |H| \leq C(t-s)$$

and we can pass to a continuous limit of ϕ as $t \to T$, call it $\phi_T : M \to \mathbb{R}^{n+1}$. We will show that this limit is actually a smooth limit.

First pick $v \in T_xM$ nonzero and write it as $v = v^i \frac{\partial \phi}{\partial x_i}$. Then

$$\frac{\partial}{\partial t} \log g(v,v) = \frac{1}{g(v,v)} \left(-2Hh_{ij}v_iv_j\right) \leq C$$

so that

$$\log \frac{g^t(v,v)}{g^s(v,v)} \leq C(t-s) \iff \frac{g^t(v,v)}{g^s(v,v)} \leq e^{c(t-s)} \text{ for } s<t$$

and therefore we can take a continuous limit metric. Now, if $|A| \leq C$, then as many derivatives as we want are bounded for $t \geq T/2$. Furthermore, this means that the time derivatives of Christoffel symbols, which live in $A \times \nabla A$, are also bounded for these times. So, then for any bounded tensors T, $\left|\frac{\partial}{\partial x_i}T - \nabla_i T\right|$ is also bounded, meaning for example that $\frac{\partial^2\phi}{\partial x_i x_j} = \Gamma^k_{ij}\frac{\partial \phi}{\partial x_k} - h_{ij}\nu$ is also bounded and so is $\partial_k \phi$, as high as you wish to go. We can then use Arzela Ascoli to get a sublimit that is ℓ-differentiable for any ℓ. Then take a diagonal subsequence to show that the sublimit is smooth and, by short time existence, we have a contradiction of the maximality of the time T. $\square$

4. A simple proof that neckpinches can happen

This proof will be a clever application of the maximum principle. Let $0 \leq \beta \leq n$ and define $f : \mathbb{R}^{n+1}_{x_1,\ldots,x_{n+1}} \times \mathbb{R}_t \to \mathbb{R}$ by

$$f(x,t) := \sum_{i=1}^n x_i^2 - (n-1-\beta)x_{n+1}^2 + 2\beta t$$

Then, for a fixed y,

$$\frac{d}{dt} f(\phi(y),t) = 2\vec{H} \cdot (x_1, \ldots, x_n, -(n-1-\beta)x_{n+1}) + 2\beta$$

$$\Delta_{M_t} f(\phi(y),t) = 2\vec{H} \cdot (x_1, \ldots, x_n, -(n-1-\beta)x_{n+1}) + 2\sum_{i=1}^{n+1} |\nabla_{M_t} x_i|^2 - 2(n-\beta)|\nabla_{M_t} x_{n+1}|^2$$

so that

$$\frac{d}{dt} f - \Delta_{M_t} f = 2\beta - 2n + 2(n-\beta)|\nabla_{M_t} x_{n+1}|^2$$

But of course $|\nabla_{M_t} x_{n+1}| \leq 1$, so $\frac{\partial}{\partial t} f - \Delta_{M_t} f \leq 0$ and by the maximum principle, $f_{\max}(t)$ is decreasing, for any choice of β. What geometric conclusions can we draw from this fact?

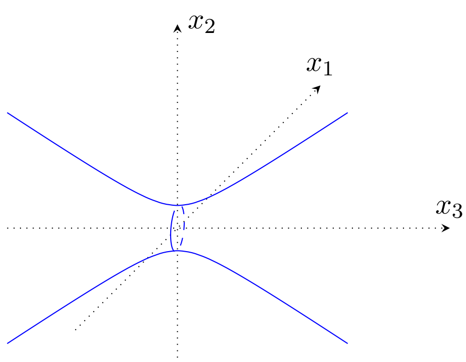
$\beta = n$:

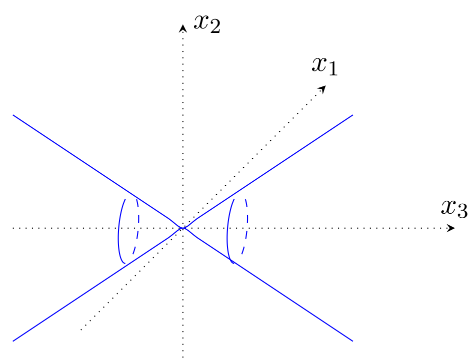
$$f(x,t) = \sum_{i=1}^{n+1} x_i^2 + 2\beta t$$
$$f(x,0) \le R^2 \iff M_0 \subseteq B(0,R)$$
$$\implies f(x,t) \le R^2 \iff M_t \subseteq B(0, \sqrt{R^2 - 2nt})$$

$\beta = n-1$:

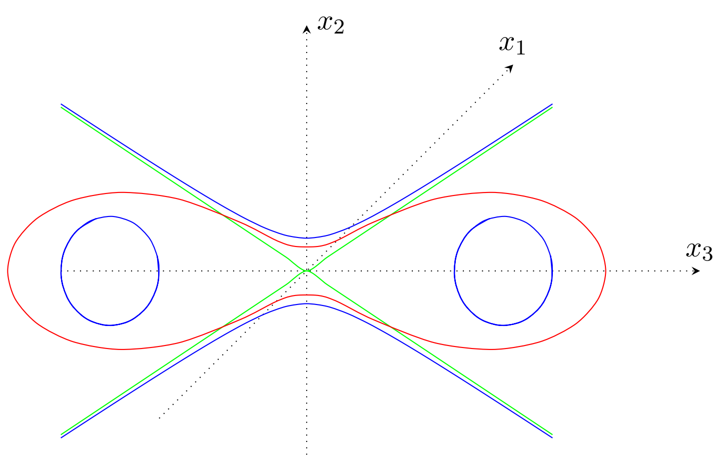
$$f(x,t) = \sum_{i=1}^{n} x_i^2 + 2\beta t$$
$$f(x,0) \le R^2 \iff M_0 \subseteq \text{ the cylinder } D^n(0,R) \times \mathbb{R}$$
$$\implies f(x,t) \le R^2 \iff M_0 \subseteq D^n(0, \sqrt{R - 2(n-1)t}) \times \mathbb{R}$$

$0 < \beta < n-1$: As before, $f(x,0) \le R^2 \implies f(x,t) \le R^2$ are equivalent to M_0 and M_t, respectively, being contained in very particular spaces. Here the maximal time is $\tau = \frac{R^2}{2\beta}$. They are a bit harder to draw, but here are my sketched attempts, where M_0 and M_τ respectively, must be contained within





We can now construct a dumbbell which must pinch off in finite time. To do so, we pick a smooth surface (pictured below in red), which is contained in a hyperboloid with radius at the origin R_2, and where each end contains a sphere of radius R_1. We pick R_2 small and R_1 large so that the function corresponding to $0 < \beta < n-1$ collapses before the R_1-spheres disappear, ensuring that a neckpinch singularity occurs in finite time.



5. A gradient estimate for MCF

If ϕ is a graph over $B^n(0,R)$, so that we can locally express $x_{n+1} = u(x_1, \dots, x_n)$, then we define ν as the upward facing normal. Beware! This is *different* from our usual convention, and it will hold only for this section. Because we are using a different convention, let us just rewrite here our typical ingredients in terms of our new convention:

$$h_{ij} = \left\langle \frac{\partial^2 \phi}{\partial x_i x_j}, \nu \right\rangle$$
$$\frac{\partial}{\partial t}\nu = -\nabla H$$
$$\frac{\partial}{\partial x_i}\nu = -h_i^j \frac{\partial \phi}{\partial x_j}$$

Now define $v := \frac{1}{\nu \cdot e_{n+1}}$ and note that in the graphical region this is well-defined. Now,

$$\frac{\partial \phi}{\partial x_i} = (0, \dots, 0, 1_i, 0, \dots, 0, \frac{\partial u}{\partial x_i})$$
$$\nu = \frac{\left(-\frac{\partial u}{\partial x_1}, \dots, -\frac{\partial u}{\partial x_n}, 1\right)}{\sqrt{1 + |\nabla u|^2}}$$
$$\implies v = \sqrt{1 + |\nabla u|^2}$$

With v nicely defined, let's observe that

$$\frac{\partial}{\partial t} v = \frac{-1}{(\nu \cdot e_{n+1})^2} \cdot \frac{\partial}{\partial t}\nu \cdot e_{n+1} = -v^2 \left\langle -\nabla H, e_{n+1}\right\rangle = v^2 \langle \nabla H, e_{n+1}\rangle$$
$$\Delta v = -\Delta(\nu \cdot e_{n+1})v^2 + 2v^3(\nabla \nu \cdot e_{n+1})^3$$

To evaluate this, we note that $\nabla_i(\nu \cdot e_{n+1}) = -h_{ij}\left\langle \frac{\partial \phi}{\partial x_j}, e_{n+1}\right\rangle$ and thus

$$\Delta \left\langle \nu, e_{n+1}\right\rangle = -(\nabla_i h_{ij})\left\langle \frac{\partial \phi}{\partial x_n}, e_{n+1}\right\rangle = -h_{ij}^2 \langle \nu, e_{n+1}\rangle = -\left\langle \nabla H, e_{n+1}\right\rangle - |A|^2 \frac{1}{v}$$

and we can plug this result into our earlier calculation to observe that

$$\Delta v = -\Delta(v \cdot e_{n+1})v^2 + 2v^3 |\nabla(v e_{n+1})|^2 = -\Delta(v e_{n+1})v^2 + \frac{2|\nabla v|^2}{v}$$
$$\left(\frac{\partial}{\partial t} - \Delta\right) v = v^2 \langle \nabla H, e_{n+1}\rangle - v^2 \langle \nabla H, e_{n+1}\rangle - |A|^2 v - \frac{2|\nabla v|^2}{v} = -|A|^2 v - \frac{2|\nabla v|^2}{v}$$

In this final expression, the first term should remind you of the negative term we used to prove Claim 3.4, but the second term is what we will use right now to prove the following:

Theorem 5.1 (Ecker-Huisken Gradient Estimate)**.** *Let $(M_t)_{t>0}$ be a smooth embedded solution to MCF. Let $\rho > 0, x_0 \in \mathbb{R}^{n+1}$. Then, so long as v is defined on $B(x_0, \sqrt{\rho^2 - 2nt})$,*

$$\sup_{B(x_0, \sqrt{\rho^2 - 2nt}) \cap M_t} \left(1 - \frac{|x - x_0|^2 + 2nt}{\rho^2}\right) \leq \sup_{B(x_0, \rho) \cap M_0} v(x)$$

Proof. We have $\rho > 0$ fixed. Define

$$\mu(r) = (\rho^2 - r)^2$$
$$r(x,t) = |x|^2 + 2nt$$
$$\phi(x,t) = \mu(r(x,t))$$

Now, we want to know what $\left(\frac{\partial}{\partial t} - \Delta\right)(\phi v^2)$ is. We start by collecting some useful facts.

(1)
$$\left(\frac{\partial}{\partial t} - \Delta\right) r(x,t) = 0$$

(2)
$$\left(\frac{\partial}{\partial t} - \Delta\right) \phi(x,t) = \mu'(r)\left(\left[\frac{\partial}{\partial t} - \Delta\right] r\right) - \mu''(r) \cdot |\nabla r|^2 = -2|\nabla r|^2$$
because $\mu'(r) = -2(\rho^2 - r)$ and $\mu''(r) = 2$.

(3) $\nabla\phi = -2(\rho^2 - r)\nabla r$

(4) Combining the previous two observations:
$$\left(\frac{\partial}{\partial t} - \Delta\right)\phi = \frac{-1}{2(\rho^2 - r)^2}|\nabla\phi|^2 = \frac{-1}{2\phi}|\nabla\phi|^2$$

Now we are in a position to expand our actual goal:
$$\begin{aligned}\left(\frac{\partial}{\partial t} - \Delta\right)(\phi v^2) &= \left[\left(\frac{\partial}{\partial t} - \Delta\right)\phi\right] v^2 + \phi\left[\left(\frac{\partial}{\partial t} - \Delta\right)(v^2)\right] + \text{ cross-gradient terms} \\ &= \frac{-|\nabla\phi|^2}{2\phi}v^2 + \phi\left[-|A|^2 v - \frac{2|\nabla v|^2}{v}\right](2v) + \phi\left[-2|\nabla v|^2\right] - 2\nabla(\phi)\cdot\nabla(v^2)\end{aligned}$$

The $-\phi|A|^2(2v^2)$ term is ≤ 0 and can be discarded, so that we are left with
$$\left(\frac{\partial}{\partial t} - \Delta\right)(\phi v^2) \leq \frac{-|\nabla\phi|^2}{2\phi}v^2 - 6\phi|\nabla v|^2 - 2\nabla(\phi)\cdot\nabla(v^2)$$

We want to use the absorbing inequality (Cauchy-Schwarz) to rewrite this inequality to obtain something of the form
$$\left(\frac{\partial}{\partial t} - \Delta\right)(\phi v^2) \leq \nabla[\phi v^2]\cdot X = (\nabla\phi)v^2\cdot X + 2v\phi\nabla v\cdot X$$
so that we can apply the maximum principle to conclude that $\max(\phi v^2)$ is nonincreasing with time. We can see that the only term we have to worry about in the above expression is $-2\nabla\phi\cdot\nabla(v^2)$, because it could be positive. We could attempt to solve this issue by noting that
$$|4\nabla\phi\cdot v\nabla v| = \left|\frac{v\nabla\phi}{\phi^{1/2}}\cdot 2\nabla v\phi^{1/2}\right| \leq \frac{|\nabla\phi|^2 v^2}{2\phi} + 8|\nabla v|^2\phi$$
But this isn't good enough, because $8|\nabla v|^2\phi - 6|\nabla v|^2\phi \geq 0$. Instead, we rewrite the inequality to have terms of the form $\nabla[\phi v^2]\cdot X$, and we hope that the remaining stuff can be absorbed. First, note that
$$-2\nabla(v^2)\cdot\nabla(\phi) = -3\nabla(v^2)\cdot\nabla(\phi) + 2\nabla(v^2)\cdot\nabla(\phi) = -6v\nabla v\cdot\nabla\phi + \frac{\nabla(v^2\phi)\cdot\nabla\phi}{\phi} - \frac{v^2|\nabla\phi|^2}{\phi}$$
The inequality transforms into
$$\left(\frac{\partial}{\partial t} - \Delta\right)(\phi v^2) \leq \frac{-|3\nabla\phi|^2}{2\phi}v^2 - 6\phi|\nabla v|^2 - 6v\nabla v\cdot\nabla\phi + \frac{\nabla(v^2\phi)\cdot\nabla\phi}{\phi}$$
The only bad term in the above inequality is the $6v\nabla v\cdot\nabla\phi$ term because it could be positive and doesn't vanish at $(v^2\phi)_{\max}$. But we note that by Cauchy-Schwarz,
$$|6v\nabla v\nabla\phi| = \left|\left(\sqrt{3}\frac{v\nabla\phi}{\phi^{1/2}}\right)\cdot\left(2\sqrt{3}\phi^{1/2}\nabla v\right)\right| \leq \frac{3}{2}\frac{|v\nabla\phi|^2}{\phi} + 6\phi|\nabla v|^2$$

so that
$$\left(\frac{\partial}{\partial t}-\Delta\right)(\phi v^2)\leq\frac{\nabla(v^2\phi)\cdot\nabla\phi}{\phi}$$
as we desired.

Now we use the maximum principle to conclude that $(\phi v^2)_{\max}$ is nonincreasing on the ball $B(0,\rho)$. $\square$

A side note to keep in mind is that this trick doesn't work in the heat equation, so that here we see one example of how mean curvature flow works better.

Remark 5.2. Recall that $r(x)=|x|^2+2nt$ and $\varphi=(\rho^2-r(x,t))^2$, so φ is zero on the boundary of the ball $B(x_0,\sqrt{\rho^2-2nt})$, and the maximum of φV^2 is attained inside the parabolic ball.

Theorem 5.3. *For any V_0, there exists some C_0 such that, suppose $(M_t)_{t\in[a,b]}$ is an embedded, proper mean curvature flow inside the parabolic ball $P(\bar{x},r)$, and $V\leq V_0$ in $P(\bar{x},r)$. Then $|A|\leq\frac{C_0}{r}$ in $P(\bar{x},r/2)$.*

The proof just uses the fact
$$(\frac{\partial}{\partial t}-\Delta)V=-\frac{2|\nabla V|^2}{V}-|A|^2V,$$
together with the maximum principle. Keep in mind that $|A|$ is behaves like the $|V|$. For the same reason, we also have higher derivatives estimates.

Theorem 5.4 (Higher derivatives estimates)**.** *For any $A_0\in\mathbb{R},m\in\mathbb{Z}$, there exists C_0 such that if $|A|\leq A_0$ in $P(\bar{x},r)$ then $|\nabla^m A|\leq C_0$ in $P(\bar{x},r/2)$.*

6. Integral estimates and the monotonicity formula

6.1. **Integral estimates.** We now derive the monotonicity formula and its various consequences. Recall that under the mean curvature flow, the volume form evolves by
$$\frac{d}{dt}\operatorname{Vol}=-H^2d\operatorname{Vol}.$$

Let $f:\mathbb{R}^{n+1}\times\mathbb{R}\to\mathbb{R}$ be a compactly support $C^{1,\alpha}$ function. Then under a smooth mean curvature flow (M_t),
$$\begin{aligned}\frac{d}{dt}\int_{M_t}f\,d\operatorname{Vol}&=\int(\frac{\partial f}{\partial t}+\nabla f\cdot\vec{H}-fH^2)d\operatorname{Vol}\\&=\int(\frac{df}{dt}-H^2f)d\operatorname{Vol}\\&=\int(\frac{dt}{df}\pm\Delta_{M_t}f)-H^2f\,d\operatorname{Vol}.\end{aligned}$$
Where in the last equality we have used the fact $\int\Delta f=0$ for any smooth function f.

Define
$$\varphi^\rho(x,t)=\left(1-\frac{|x|^2+2nt}{\rho^2}\right)_+^3.$$

Since $(\frac{\partial}{\partial t}-\Delta_{M_t})(\frac{|x|^2+2nt}{\rho^2})=0$ we deduce that (taking minus sign in the above equality)
$$(\frac{\partial}{\partial t}-\Delta_{M_t})\varphi^\rho=-3\left(1-\frac{|x|^2+2nt}{\rho^2}\right)_+^2\leq 0.$$

Therefore we conclude that $\frac{d}{dt}\int \varphi^\rho \leq 0$. We are going to utilize this fact to recover some of the results we got earlier with different approaches.

Example 6.1 (Avoidance of balls). *Suppose* $M_0 \cap B(p,\rho) = \phi$. *Then* $\int_{M_0} \varphi^\rho = 0$. *Hence by the monotonicity,* $\int_{M_t} \varphi^\rho = 0$, *and therefore* $1 - \frac{|x|^2+2nt}{\rho^2} \leq 0$ *on* M_t, *or in other words,* $M_t \cap B(0, \sqrt{\rho^2 - 2nt}) = \phi$.

Example 6.2 (Volume control). *Take* ϵ *small, we see that*

$$\int_{M_{\epsilon\rho^2}} \varphi^\rho \leq \int_{M_0} \varphi^\rho \leq \mathrm{Vol}(M_0 \cap B(0,\rho)).$$

Observe that

$$\int_{M_{\epsilon\rho^2}} \varphi^\rho \geq \frac{1}{8}\mathrm{Vol}(M_{\epsilon\rho^2} \cap B(0,\rho/2)),$$

for $\epsilon \leq \frac{1}{8n}$. *Therefore*

$$\frac{1}{8}\mathrm{Vol}(M_{\frac{1}{8n}\rho^2} \cap B(0,\rho/2)) \leq \mathrm{Vol}(M_0 \cap B(0,\rho)).$$

We therefore have:

Corollary 6.3 (Polynomial growth is preserved under MCF). *If* M_0 *has polynomial volume growth, that is, if* $\mathrm{Vol}(M_0 \cap B(p,R)) \leq A_0 R^k$ *for some* A_0, k *and all* R *sufficiently large, then* M_t *also has polynomial volume growth.*

6.2. **Huisken's monotonicity formula.** In this section we describe Huisken's monotonicity formula for the mean curvature flow. A key observation here is to plug the backwards heat kernel into the general monotonicity formula obtained in the previous section.

For $t < 0, x \in \mathbb{R}^{n+1}$, denote the backwards heat kernel

$$\Phi(x,t) = \frac{1}{(-4\pi t)^{n/2}} e^{|x|^2/4t}.$$

And

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

Take plus sign in the general formula, we have

$$\frac{d}{dt}\int_{M_t} \Phi = \int \left(\frac{\partial}{\partial t} + \Delta_{M_t}\right)\Phi - H^2\Phi.$$

Therefore we deduce

$$\begin{aligned}&(\frac{\partial}{\partial t} + \Delta_{M_t})\Phi - H^2\Phi \\ &\qquad = \frac{\partial \Phi}{\partial t} + 2\nabla\Phi \cdot \vec{H} + \mathrm{div}_{M_t}\nabla\Phi - H^2\Phi.\end{aligned}$$

Since

$$\begin{aligned}\mathrm{div}_{M_t}\nabla\Phi &= \mathrm{div}_{M_t}\nabla^{M_t}\Phi + \mathrm{div}_{M_t}(\nabla\Phi^\perp) \\ &= \Delta_{M_t}\Phi + \vec{H}\cdot\nabla\Phi,\end{aligned}$$

Above can be simplified as

$$\frac{\partial\Phi}{\partial t} + \mathrm{div}_{M_t}\nabla\Phi - |\vec{H} - \frac{\nabla^\perp\Phi}{\Phi}|^2\Phi + \frac{\nabla^\perp\Phi}{\Phi}.$$

Claim: $(*) = \frac{\partial\Phi}{\partial t} + \mathrm{div}_{M_t}\nabla\Phi + \frac{\nabla^\perp\Phi}{\Phi} = 0$.

Proof. Note that $u = \frac{1}{\sqrt{-t}}\Phi$ solves the backwards heat equation. Then

$$(\frac{\partial}{\partial t} + \Delta_{\mathbb{R}^{n+1}})\Phi = u \cdot \frac{\partial}{\partial t}\sqrt{-t} = u \cdot \frac{-1}{2\sqrt{-t}} = \frac{\Phi}{2t}.$$

Therefore

$$\begin{aligned}(*) &= \frac{\Phi}{2t} - \mathrm{Hess}(\Phi)(\nu,\nu) + \frac{|\nabla^\perp \Phi|^2}{\Phi} \\ &= \Phi\left[\frac{1}{2t} - \mathrm{Hess}(\log\Phi)(\nu,\nu)\right].\end{aligned}$$

Since Φ is the heat kernel, we deduce that

$$\mathrm{Hess}(\log\Phi) = \mathrm{Hess}(\frac{|x|^2}{4t})(\nu,\nu) = \frac{1}{2t}$$

Therefore $(*) = 0$ and the claim is proved. □

Now we conclude that

$$\frac{\partial}{\partial t}\int_{M_t}\Phi = -\int |\vec{H} - \frac{\nabla^\perp\Phi}{\Phi}|^2 \Phi d\,\mathrm{Vol}\,. \tag{6.1}$$

Let us look at various consequences of the monotonicity formula.

6.3. **Rescaling and tangent flow.** Let $(M_t)_{t\in(a,b)}$ be a mean curvature flow and let $t_0 \in (a,b)$, $x_0 \in \mathbb{R}^{n+1}$. For $\lambda > 0$, define

$$\tilde{M}_{(x_0,t_0),s} = \lambda(M_{t_0+s\lambda^{-2}} - x_0).$$

Then $\tilde{M}$ is also a mean curvature flow. Let

$$\psi(x,s) = \lambda[\varphi(x, t_0 + \lambda^{-2}s) - x_0],$$

then

$$\frac{d\psi}{ds} = \lambda\frac{d\varphi}{ds} = \frac{1}{\lambda}\vec{H}_\varphi(x,,t_0+\lambda^{-2}s) = \vec{H}_\psi(x,s).$$

Recall that

$$\Phi_{(x_0,t_0)}(x,t) = \frac{1}{[4\pi(t_0-t)^{n/2}]}e^{-|x-x_0|/4(t_0-t)^2}$$

is the backwards heat kernel in $\mathbb{R}^n$, and the monotonicity formula

$$\frac{d}{dt}\int_{M_t}\Phi_{x_0,t_0} = -\int|\vec{H} - \frac{\nabla^\perp\Phi}{\Phi}|^2 d\,\mathrm{Vol}\,.$$

The fundamental property of the quantity $\int_{M_t}\Phi_{(x_0,t_0)}$ is that it is scaling invariant.

Proposition 6.4. $\int_{M_t}\Phi_{(x_0,t_0)}$ *is invariant under parabolic scaling.*

Proof. Let $\tilde{M}_{(x_0,t_0),s}$ be the rescaled flow defined above. For simplicity we suppose the density function is taken at a point in space time $(\rho, 0)$. Then we have

$$\begin{aligned}\int_{\tilde{M}_s}\Phi_{(\rho,0)}d\,\mathrm{Vol}_{\tilde{M}_s} &= \int_{\tilde{M}_s}\frac{1}{(-4\pi s)^{n/2}}e^{|x|^2/4s}d\,\mathrm{Vol} \\ (\text{set } t_0+s\lambda^{-2} = t \text{ and } x = \lambda(y-y_0)) \quad &= \lambda^n\int_{M_t}\frac{1}{[-4\pi(t-t_0)\lambda^2]^{n/2}}e^{|\lambda(y-y_0)|^2/4(t-t_0)\lambda^2}d\,\mathrm{Vol}_{M_t} \\ &= \int_{M_t}\frac{1}{4\pi(t_0-t)^{n/2}}e^{-|y-y_0|^2/4(t_0-t)}d\,\mathrm{Vol}_{M_t}\,.\end{aligned}$$

This finishes the proof. □

Recall that if we rescale a manifold locally by larger and larger scale, then the rescaled manifold becomes its tangent plane. We are going to use rescaling to study the local property of a mean curvature. To do so, let us first study the limit case, that is, a mean curvature flow that is invariant under parabolic rescaling. Suppose $(M_t)_{t\in[T,0)}$ is a mean curvature flow that moves by scaling, that is,

$$M_t = \sqrt{-t}M_{-1}.$$

Since $\int_{M_t}\Phi_{(0,0)}$ is invariant under parabolic rescaling, we have that, by the monotonicity formula,

$$\vec{H} - \frac{\nabla^\perp\Phi}{\Phi} = 0.$$

Also,

$$\frac{\nabla^\perp\Phi}{\Phi} = \nabla^\perp(\log\Phi) = \nabla^\perp(\frac{|x|^2}{4t}) = \frac{1}{2t}\langle x,\nu\rangle\,\nu.$$

Therefore we obtain that

$$H = -\frac{1}{2t}\langle x,\nu\rangle\,.$$

Proposition 6.5. *If at some time t_0, M_{t_0} satisfies $H = -\frac{1}{2t_0}\langle x,\nu\rangle$, then the mean curvature flow emenating from M_{t_0} evolves by scalings, and satisfies $H = -\frac{1}{2t}\langle x,\nu\rangle$.*

Proof. Since the condition $H = -\frac{1}{2t}\langle x,\nu\rangle$ is invariant under parabolic scaling, it suffices to check the statement for $t_0 = -1$.

Now define $M_t = \sqrt{-t}M_{-1}$. Then we check that

$$\left\langle\frac{\partial X}{\partial t},\nu\right\rangle = -\frac{1}{2\sqrt{-t}}\langle X_{-1},\nu\rangle = -\frac{1}{\sqrt{-t}}H(X_{-1}) = -H(X_{-t}).$$

By the uniqueness of the solution of the mean curvature flow with initial condition, (M_t) is the unique mean curvature flow emenating from M_{-1}. □

Definition 6.6. The Gaussian density pf a mean curvature flow $(M_t)t\in[a,b)$ at a point (x_0,t_0), where $x\in\mathbb{R}^{n+1}, t_0\in(a,b]$, is defined to be

$$\Theta(M,x_0,t_0) = \lim_{t\to t_0}\int_{M_t}\Phi_{(x_0,t_0)}d\operatorname{Vol}.$$

Fix (x_0,t_0) and take a sequence λ_k converging to infinity. For a mean curvature flow (M_t), consider the rescaled flows

$$M_s^k = \lambda_k(M_{t_0+\lambda_k^{-2}s} - x_0).$$

For a parabolic ball $P(0,R)$, points in $M_s^k\cap P(0,R)$ come from points that are very close to (x_0,t_0). Suppose one can take a smooth, multiplicity 1 limit M_s^k, that is,

$$M_s^k \to M_s, \quad \text{in } B(0,R)\times[-R^2,-R^{-2}]\cap M_s^k \text{ smoothly graphically.}$$

Then by the scaling invariance of the Gaussian density, we have

$$\int_{M_s^k}\Phi - \Theta(M_{t_0},0,0) = \int_{M_{\lambda_k^{-2}s}}\Phi_{(0,0)}d\operatorname{Vol}_{M_{\lambda_k^{-2}s}} - \Theta(M_{t_0,0,0}) \to 0.$$

Therefore the limit flow $\tilde{M}_s$ satisfies

$$\int_{\tilde{M}_s}\Phi_{(0,0)} = \Theta(M_t,0,0), \quad \text{for any } s$$

hence the flow $\tilde{M}_s$ moves by dilation, namely, $\tilde{M}_s = \sqrt{-s}\tilde{M}_{-1}$. Moreover, the flow $\tilde{M}_s$ is defined on $\mathbb{R}^{n+1}\times(-\infty,0)$.

Next we are going to see under which conditions the above limit flow construction works. Let us point out here that the major difficulty is to guarantee that the convergence is of multiplicity 1.

Suppose $(M_t)_{t\in[T,0)}$ is a smooth embedded mean curvature flow and suppose also that (x_0,t_0) is a singular point. Then $\sup_{t\to 0}|A(t)|=\infty$, hence $|A(t)|>\frac{C}{\sqrt{-t}}$ for some constant C.

Definition 6.7. A singular point (x_0,t_0) is called type I singualar point, if there is another constant C_1 such that $|A(t)|\leq\frac{C_1}{\sqrt{-t}}$.

Type I singularity is important, since it guarantees that the tangent flow at it is well-defined. Namely, we have:

Theorem 6.8. *If $(x_0,0)$ is a type I singularity of a mean curvature flow $(M_t)_{t\in[-T,0)}$, then for any sequence $\lambda_k\to\infty$, there is a subsequence λ_{k_n}, such that*

$$M^n_s=\lambda_{k_n}M_{\lambda_{k_n}^{-2}s}$$

converges smoothly with multiplicity 1 to a limit flow $\tilde{M}_s$.

COMMENTS AT BEGINNING OF WEEK 4 LEFT OUT FOR NOW

In order to prove this theorem, we need one more definition.

Definition 6.9. Given a smooth, embedded MCF, a point $x_0\in\mathbb{R}^{n+1}$ is said to be reached by the flow at time T if there exist $x_{t_n}\in M_{t_n}$ so that $t_n\to t$ and $x_{t_n}\to x_0$.
Such an x_0 is called singular if there exist such a sequence with $|A(x_{t_n})|\to\infty$.

Proof. By the monotonicity formula, it suffices to just prove the first part, that $\widetilde{M}^k_s\to\widetilde{M}_s$ smoothly with multiplicity one. We pick $s<0$ and we note that the type I condition implies that $|A^k(s')|\leq\frac{C}{\sqrt{-s}}$ for $s'\leq s$. In fact, Furthermore, for small time interval ending at s, we have derivative bounds $|\nabla^j A^k(s')|\leq\frac{C}{\sqrt{-s}}$ as well. Now define d_t to be the intrinsic distance in M_t. There exist ϵ,c so that $d_t(x,y)\leq\epsilon\sqrt{-t}$ implies that

$$C^{-1}\leq\frac{d_t(x,y)}{|x-y|}\leq C$$

Note that the condition we need for this comparability of distance is an upper bound of the form $\epsilon\frac{1}{\sup A}$ and the type I condition allows us to write this in terms of t. On the points where we have this bound, we have smooth convergence with multiplicity one. What about other points?

Claim 6.10. *Define the set $A_\epsilon:=\{(x,y,t)$ where $x,y\in M_t,d_t(x,y)\geq\epsilon\sqrt{-t}\}$. Pick x,y to minimize $\frac{|x-y|}{\sqrt{-t}}$ over A_ϵ at time t. Then either*

1. *$d_t(x,y)=\epsilon\sqrt{-t}$ or*
2. *We have a picture that looks like the following: ADD PICTURE*

□

Note that the above also gives a proof of the following nice fact: if you have a MCF starting from an embedded hypersurface, up until the time of singularity, you stay embedded.

There is one more concern that we have: if x_0 is obtained at a type I singularity, why is $\widetilde{M}_s$ nonempty? To show this, we need to compare the rate of convergence at a point and the scaling rates. The point is that if $x_0=0$ is obtained at time 0, then for all $t<0$, $\overline{M_t\cap B(0,\sqrt{-2nt})}$ is nonempty (this follows from the avoidance principle), and so there exists a point in $\widetilde{M}_s\cap\overline{B(0,\sqrt{-2ns})}$

6.4. **Examples of shrinkers.** Recall that the definition of shrinkers is that they solve the equation $H + \frac{\langle x,\nu\rangle}{2t} = 0$. We provide a list of shrinkers defined on the time interval $(-\infty, 0)$.

- the plane through the origin
- spheres of radius $\sqrt{-2nt}$
- cylinders $S^k(\sqrt{-2kt}) \times \mathbb{R}^{n-k}$
- the Argenent torus, a large but narrow torus

As an aside, the Argenent torus turns out to be very useful for comparision reasons – for example, it can be used to provide an alternate proof that neckpinches exist.

6.5. **Applications of monotonicity.** First, observe that

$$\frac{\partial}{\partial t}\left(\int_{M_t} f\Phi\right) = \int_{M_t}\left[\frac{\partial}{\partial t}f\Phi + f\frac{\partial}{\partial t}\Phi - H^2 f\Phi\right] dvol$$

$$= \int_{M_t}\left(\frac{\partial}{\partial f} - \Delta_{M_t} f\right)\Phi + \left(\frac{\partial}{\partial \Phi} - \Delta_{M_t}\Phi\right) f - H^2 f\Phi \, dvol =: \star$$

Now if $\left(\frac{\partial}{\partial t}f - \Delta_{M_t} f\right) \leq 0$, $f > 0$, then

$$\star \leq \int_{M_t}\left(\frac{\partial}{\partial t}\Phi + \Delta_{M_t}\Phi\right) f - H^2 f\Phi = -\int_{M_t}\left|\vec{H} + \frac{\nabla^\perp \Phi}{\Phi}\right|^2 \Phi f \, dvol$$

Now, our favorite f are functions like

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

These functions let you show that densities on small balls are controlled by quantities at a fixed earlier time. The main trick is to use the backwards heat kernel at $(x_0, t_0) = (0, r^2)$ to get something like the density $\frac{\text{Vol}(M_t \cap B(0,r))}{r^n}$ and then to pick the correct cut-off. The details are left as an exercise.

Here are some examples of what we can derive about densities:

First, suppose M_t is a MCF that does not reach 0 at time 0. What is $\Theta(\{M_t\}, (0,0))$? For t sufficiently small, $M_t \cap \overline{B(0, \sqrt{-2nt})} = \varnothing$, so that in the scaling limit, we get 0. This density is preserved under scaling, so that $\theta(\{M_t\}, (0,0)) = 0$.

As another example, if x_0 is a smooth point of flow at time 0, then $\theta(\{M_t\}, (0,0)) = 1$. Again, this is obtained by looking at the scaling limit, which is a plane. In both of these examples we are using, in a critical way, that we have uniform control of densities.

As an (easy) exercise, consider a MCF which is smooth at times $t_1 < t_2$ and $x_2 \in M_{t_2}$. What can you say about $\int_{M_{t_1}} \Phi(x_2, t_2)$? If the flow is not a shrinker, then it should be > 1.

To compute the density, it doesn't matter how the flow looks away from the point. If $f \in C_c^\infty(N_\epsilon(x_0, t))$, then

$$\lim_{t \to t_0} \int_{M_t} f\Phi_{(x_0,t_0)} = \theta(M, (x_0, t_0)) f(x_0, t_0)$$

Fact 6.11. *θ is upper semicontinuous.*

Proof. Let $(x_j, t_j) \to (x_0, t_0)$ for a time t_0. Then

$$\int_{M_t} \Phi_{(x_j,t_j)} \xrightarrow{j\to\infty} \int_{M_t} \Phi_{(x_0,t_0)}$$

and then by monotonicity

$$\int_{M_t} \Phi_{(x_j,t_j)} \geq \theta(M, (x_j, t_j))$$

so that we can conclude that

$$\limsup_{j\to\infty} \theta(M,(x_j,t_j)) \leq \int_{M_t} \Phi_{(x_0,t_0)}$$

Taking the limit on the right hand side as $t \to t_0$ gives us the desired inequality:

$$\limsup_{j\to\infty} \theta(M,(x_j,t_j)) \leq \theta(M,(x_0,t_0))$$

$\square$

Corollary 6.12. *For x_0 obtained at time t_0 where $\{M_t\}_{t\in(T,t_0)}$, $\theta(\{M_t\},(x_0,t_0)) \geq 1$.*

6.5.1. INCLUDE PICTURE

Lemma 6.13 (Clearing out Lemma)**.** *For all $\alpha < 1$, there exists $\theta(n)$ so that if M_t reaches $0 = x_0$ at time 0,*

$$\frac{Vol(M_{-\alpha r^2/2n} \cap B(0,r))}{r^n} \geq \theta$$

This lemma is often used in situations where the left hand side is $< \theta$ for some r, and thus 0 is not reached by the flow.

6.5.2. *Noncompact maximum principle.*

Fact 6.14. *If M_t is properly embedded, smooth, with polynomial growth and $\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq 0$, then $f(x,t) \leq f_{\max}(t=0) = M$.*

Proof. We use monotonicity. By picking $p > 2$, $(f-M)_+^p$ is a C^2 nonnegative function with $(f-M)_+^p \leq 0$. Then $\int_{M_t} \Phi_{(x_0,t_0)}(f-M)_+^p$ is nonincreasing which lets us conclude that $(f-M)_+^p \equiv 0$. $\square$

6.5.3. *Extinction estimates.* If T is a singular time of the flow $\{M_t\}_{[0,T)}$, and we pick regular points $(x_n,t_n) \to (x,T)$, a singular point. Then

$$\int_{M_0} \frac{1}{(4\pi T)^{n/2}} e^{|x-x_n|^2/4T} \geq 1$$

implies that $(4\pi T)^{n/2} \leq \mathrm{Vol}(M_0)$ and thus

$$T \leq \left(\frac{\mathrm{Vol}(M_0)}{4\pi}\right)^{2/n}$$

Remark 6.15. We could also consider MCF in higher codimensions, and we would still have a k-dimensional monotonicity formula.

Remark 6.16. This theorem is also true for weak flows, as Ilmanen showed when he constructed the *track*under MCF, so S such that $\partial S = M_0^k$ and then noted that

$$\mathrm{Vol}\left(\int_{[0,T)}\right) \leq (\mathrm{Vol}(M_0))T^{1/2}$$

$$\mathrm{Vol}(S) \leq \mathrm{Vol}(M_0)^{2/n-1}/(4\pi)^{1/n}$$

This is much better than the isomperimetric inequality, $\frac{1}{\sqrt{4\pi}}\mathrm{Vol}(M_0)\frac{k+1}{k}$.

6.6. **Brian White's Regularity for Type 1 Singular Flows.** We return to our favorite functions, where for a point $X = (x_0, t_0)$ and $\rho > 0$,

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

These functions, again, have the benefit that $\left(\frac{\partial}{\partial t} - \Delta\right)\phi^{X,\rho} \le 0$ and $\phi^{X,\rho} \ge 0$. Then we define

$$\theta^\rho(M, X, r) = \int_{M_{t-r^2}} \Phi_X \phi^{X,\rho}$$

and as $r \to 0$, the function decreases and converges to density, as we defined earlier.

Theorem 6.17 (Local Regularity Theorem, White 2005)**.** *There exist $\epsilon > 0, c > 0$ so that for each spacetime point X_0 and $\rho > 0$, if M_t is a MCF in $P(X_0, 2n\rho)$ and for some $r < \rho$,*

$$\sup_{X \in P(X_0, r)} \theta^\rho(M, X, r) < 1 + \epsilon,$$

then

$$\sup_{P(X_0, r/2)} |A| \le \frac{c}{r}$$

Proof. Suppose not. Then rescale so that $r = 1$ and $X_0 = (0,0)$. Then there exist MCFs M^j in $P(0, 2n\rho_j)$ for $\rho_j \ge 2n$ so that $\sup_{P(0,1)} \theta^{\rho_j}(M, X, r) < 1 + \frac{1}{j}$ but there exists $X \in P(0, \frac{1}{2})$ so that $|A| > j$. It would be nice if we could pick a point with maximal $|A|$ in $P(0, \frac{1}{2})$, rescale so that $|A| \to 1$, and derive a contradiction, where we have a point with curvature on what must be a plane. But we can't finish the proof that easily – the "maximal" point might be on the boundary. But if we a bit more careful, we can derive this contradiction with the following claim.

Claim 6.18 (Point Selection)**.** *We can pick points p_j in $P(0, \frac{3}{4})$ so that $|A(p_j)| \ge j$ and $|A(p)| \le 2|A(p_j)|$ for $p \in P\left(p^j, \frac{j}{10|A|(p_j)}\right)$.*

Proof of Claim. Fix j. Look for $q = p_j$. There exists $q^0 \in P(0, \frac{1}{2})$ so that $|A|(q^0) = Q_0 > j$. Maybe q^0 satisfies the other condition, in which case we are done. If q^0 does not, then there exists some $q^1 \in P\left(q^0, \frac{j}{10Q_0}\right)$ so that $|A|(q^1) = Q_1 \ge 2Q_0 > 2j$. If q^1 satisfies the other condition, then we stop. If not, then we find a q^2 similarly and continue the process, and so on. If we must continue the process until we find a q^n, note that we have moved at most $\frac{j}{10Q_0}\left[1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n}\right] \le \frac{1}{5}$, and so we have found a point $q \in P(0, \frac{3}{4})$. On the other hand, note that $|A^j|$ is bounded on $P(0, \frac{3}{4})$ because we have a smooth compact flow, and so this process must halt. □

To use this claim, we move p_j to the spacetime origin and rescale by $P_j = |A|(p_j)$. Then $\widetilde{M}^j$ is defined in $P(0, 2n\rho_j P_j)$ so that $\theta^{\rho_j P_j}(\widetilde{M}^j, 0, P_j) < 1 + \frac{1}{j}$, $|A| \le 2$ on $P(0, \frac{j}{10})$, and $|A|(0) = 1$. We take the limit $\widetilde{M}$ on $\mathbb{R}^{n+1} \times (-\infty, 0]$ so that $\int_{\widetilde{M}_t} \Phi_{(0,0)} = 1$ and $|A(0)| = 1$ and $|A| \le 2$ always. This provides a contradiction, as $|A(0)| = 1$ means that $\widetilde{M}$ is not a plane, but using the monotonicity formula we have that $\widetilde{M}_{-1}$ is a plane and therefore as $\widetilde{M}_t = \sqrt{-t}\widetilde{M}_1 \to \widetilde{M}_0$ smoothly (because the curvatures are bounded), $\widetilde{M}_0$ is a plane. □

Corollary 6.19. *If $\{M_t\}_{t \in [a,0)}$ satisfies $\theta(\{M_t\}, (0,0)) = \lim_{t \to 0} \int_{M_t} \Phi_{(0,0)} = 1$, then 0 is not a singular point of the flow at time 0.*

This corollary follows because for $t_0 < 0$, with $\int_{M_{t_0}} \Phi_{(0,0)} < 1 + \frac{\epsilon}{2}$, $\int_{M_t} \Phi_{(x_0,t_0)} < 1 + \epsilon$ in some neighborhood of $(0,0)$. By the local regularity theorem just proven, that means in a slightly smaller neighborhood around $(0,0)$, the curvature is bounded.

7. Classification of singularities

In this section we derive some results for the classification of hypersurfaces evolved by the mean curvature flow. The idea is to analyze the singularity of the flow, and hence the study of the tangent flow at the singularity is most important. By White's regularity theorem, the tangent flow at any singular point cannot be a hyperplane. Therefore the question is two fold: one is to study when type I singularity occurs, and the other is to classify all self shrinkers. As we will see, these can be done under some special cases, namely for the curve shortening flow, and for mean convex initial surfaces.

7.1. Classification of self shrinkers.

For the curve shortening flow, the following theorem gives a complete classification of the self shrinkers.

Theorem 7.1 (Gage-Hamilton). *If γ is an embedded curve in $\mathbb{R}^2$ satisfying $\kappa = H = \langle x, \nu\rangle$, then γ is a circle or a line.*

Proof. Let $p_0 \in \gamma$ and choose length parametrization $\gamma(s)$. Now $\kappa = \langle x, \nu\rangle$ and $\ddot{\gamma}(s) = -\kappa\nu$, so

$$\langle \nu, \nu\rangle = 1 \implies \langle \nu, \nu_s\rangle = 0$$
$$\langle \gamma, \dot{\gamma}(s)\rangle = 0 \implies \langle \nu_s, \dot{\gamma}(s)\rangle = \kappa(s)$$

so that $\nu_s = \kappa(s)\dot{\gamma}(s)$, and we have the equations

$$\kappa_s = \langle \dot{\gamma}(s), \nu\rangle + \langle \gamma(s), \kappa\dot{\gamma}(s)\rangle = \kappa\langle x, \dot{\gamma}(s)\rangle$$
$$\kappa_{ss} = \kappa_s\langle x, \dot{\gamma}(s)\rangle + \kappa - \kappa\langle x, \kappa\nu\rangle = \kappa\langle x, \dot{\gamma}(s)\rangle^2 + \kappa - \kappa^3.$$

From the equation for κ_s, we note that if $\kappa = 0$ at time t_0, then $\kappa = 0$ at all times, by the uniqueness of the solution of the ODE, and so we have a line through the origin. Assume that $\kappa \neq 0$ for all times, say $\kappa > 0$. We want κ to be a closed curve.

$$\partial_s|x|^2 = 2\langle \dot{\gamma}, x\rangle = 2\frac{\kappa_s}{\kappa} = 2\partial_s \log\kappa$$

The solution $\kappa = ce^{|x|^2/2}$ lets us conclude that κ is bounded from above and below.

Now we take the Gauss map $N : \gamma \to S^1$ and its inverse $\theta : S^1 \to N$. Note that $\frac{\partial N}{\partial s} = \kappa$ so that $\frac{\partial s}{\partial \theta} = \frac{1}{\kappa}$. Then

$$\kappa_\theta = \frac{1}{\kappa}\cdot\kappa_s = \langle x, \dot{\gamma}(s)\rangle$$
$$\kappa_{\theta\theta} = \frac{1}{\kappa}\left[1 - \kappa\langle x, \nu\rangle\right] = \frac{1}{\kappa} - \kappa$$

Then, if we multiply this last equation by $2\kappa_\theta$, it becomes

$$2\left[\kappa_\theta\kappa_{\theta\theta} + \kappa_\theta\kappa - \frac{\kappa_\theta}{\kappa}\right] = \partial_\theta\left[\kappa_\theta^2 + \kappa^2 - 2\log\kappa\right] = \partial_\theta[E] = 0$$

Thus $E(\theta)$ is a conserved quantity along the curve. Now, notice that $\kappa^2 - \log(\kappa^2) \geq 1$ (which can be seen by looking at the Taylor series), and $\kappa^2 - 2\log\kappa = 1$ iff $\kappa = 1$. If $E = 1$ somewhere, then $E = 1$ always by this equation. But $E(\theta) = 1$ implies that $\kappa(\theta) = 1$, and $\kappa \equiv 1$ implies that γ is a circle.

For the remainder of the proof, let us assume that $E > 1$, meaning that κ varies. $\kappa_{\theta\theta} = \frac{1}{\kappa} - \kappa$. Note that all critical points are nondegenerate, as a degenerate critical point would simply that $\frac{1}{\kappa} = \kappa$ and thus that $\kappa = 1$ and $E = 1$, putting us in the previous case. At the minimum points, $\kappa_{\theta\theta} > 0$ and so $\kappa < 1$. At the maximum points, $\kappa_{\theta\theta} < 0$ and $\kappa > 1$. The equation also tells us that all critical points are global maxima or minima. However, we do not know how many there are. To that end, we use the following elementary geometry theorem (which we will not prove):

Theorem 7.2 (Four Vertices Theorem)**.** *For a convex curve $\gamma \subset \mathbb{R}^2$, its curvature k has at least 4 critical points.*

By moving the θ parameter is necessary, let us assume that k attains its maximum at $\theta = 0$. Let $T/2 > 0$ be the first time where k attains its minimum. If we reflect the graph of k in $[0, T/2]$ accross the line $\theta = T/2$, then the second derivative of k is unchanged at $T/2$, and the first derivative at $T/2$ vanishes. Therefore the reflected graph gives an extension of the solution to the ODE $k_{\theta\theta} = 1/k - k$. By the uniqueness of the solution of ODE, this solution must coincide with the original one. Therefore we have proved that k is periodic with period T. Denote $T = \frac{2\pi}{m}$, where m is the number of periods in $[0, 2\pi]$.

According to the four vertices theorem, m has to be at least 2, since in each period there are only 2 critical points. The trick is to calculate $(k^2)_{\theta\theta\theta}$. We have:

$$\begin{aligned}(k \cdot k)_{\theta\theta\theta} &= 2kk_{\theta\theta\theta} + 6k_\theta k_{\theta\theta} \\ &= 2k(\frac{1}{k} - k)_\theta + 6k_\theta(\frac{1}{k} - k) \\ &= -4(k^2)_\theta - 4\frac{k_\theta}{k}.\end{aligned}$$

Integrate against $\sin(2\theta)$ from 0 to $T/2$, and use that $\frac{k_\theta}{k} < 0$ for $\theta \in [0, T/2]$, we get:

$$\begin{aligned}4\int_0^{T/2} \sin(2\theta)\frac{k_0}{k} &= \int_0^{T/2} [(k^2)_{\theta\theta\theta} + 4(k^2)_\theta] \sin(2\theta) \\ &= (k^2)_{\theta\theta}(T/2)\sin T \\ &= \sin T \cdot (2kk_{\theta\theta}(T/2)) \\ &= \sin T \cdot 2(1 - k_{\min}^2)\end{aligned}$$

That's a contradiction, since $\frac{k_\theta}{k} < 0$ and $k_{\min} < 1$.

□

For mean curvature flows in general dimension, we have:

Theorem 7.3 (Huisken)**.** *If $M \subset \mathbb{R}^{n+1}$ is properly embedded with polynomial volume growth, $H = \langle X, \nu \rangle$ and $|A|$ is bounded, then it is a plane, a sphere of a cylinder.*

Remark 7.4. Note that by previous results, M is the tangent flow of a mean curvature flow with mean convex, embedded and compact initial hypersurface.

Proof. First observe that the mean curvature flow $M_t = \sqrt{-2t}M_{-\frac{1}{2}}$, $t < 0$, satisfies $H(t) = \frac{1}{\sqrt{-2t}H(-\frac{1}{2})}$, and $\frac{\partial H}{\partial t} = \frac{1}{(\sqrt{-2t})^3}H(-\frac{1}{2})$, and therefore $\frac{\partial H}{\partial t}|_{-\frac{1}{2}} = H$.

On the other hand, for a mean curvature flow, we have

$$\frac{\partial H}{\partial t} = \Delta H + |A|^2 H.$$

We want to conclude that $H = \Delta H + |A|^2 H$. This is not entirely correct, since moving by homothety by scale $\sqrt{-2t}$ is not moving in the normal direction, hence we need to modify the evolution equation in accordance to the extra tangential diffeomorphism. In fact, let X be the time dependent tangential vector field given by the homothety. Then

$$\frac{\partial H}{\partial t} = \Delta H + |A|^2 H - \langle X, \nabla H \rangle.$$

Therefore $H = \Delta H + |A|^2 H - \langle X, \nabla H\rangle$, or

$$\nabla H = (1 - |A|^2) + \langle X, \nabla H\rangle\,.$$

And similarly, we have

$$\nabla|A| = (1 - |A|^2)|A| + \langle X, \nabla|A|\rangle + \frac{|\nabla A|^2 - |\nabla||A||^2}{|A|}.$$

Note by Kato's inequality, the last term above is always nonnegative.

By the strong maximum principle if H is zero somewhere, then H is identically zero, and hence the surface is a hyperplane. From now on we assume $H > 0$ everywhere.

We calculate

$$\Delta|A|H - |A|\Delta H = \langle X, \nabla|A|\rangle\, H + \frac{|\nabla A|^2 - |\nabla|A||^2}{|A|} - \langle X, \nabla H\rangle\, |A|.$$

Note that

$$\operatorname{div}(\nabla|A| \cdot He^{-|X|^2/2}) = \Delta|A|He^{|X|^2/2} + \nabla|A| \cdot \nabla He^{-|X|^2/2} - \langle X, \nabla|A|\rangle\, He^{-|X|^2/2}.$$

Therefore above gives

$$\operatorname{div}(\nabla|A|He^{-|X|^2/2}) - \operatorname{div}(\nabla H|A|e^{-|X|^2/2}) = \frac{|\nabla A|^2 - |\nabla|A||^2}{|A|}.$$

Let r be a radius such that $\partial B(0,r) \cap M$ is a codimensional 2 submanifold. Note that such r are almost everythere. Denote ρ to be the outward unit conormal vector of $\partial B(0,r)$ in M. Integrate the above equality in $B(0,r)$ and use Stokes' formula, we obtain that

$$\int_{\partial B(0,r)} \langle \nabla|A|H - \nabla H|A|, \rho\rangle\, e^{-|X|^2/2} = \int_{B(0,r)\cap M} \frac{|\nabla A|^2 - |\nabla|A||^2}{|A|} e^{-|X|^2/2}.$$

Note that by Kato's inequality, the right hand side is always nonnegative, but the left hand side converges to zero, as the hypersurface has Euclidean volume growth. Therefore we conclude $|\nabla A| = |\nabla|A||$ on the whole hypersurface. We will deal with this rigidity case separately.

□

To continue the proof of Huisken's theorem, we recall that we had subdivided into two cases. If $H = 0$ somewhere, then M is a plane. The remaining case is when $H > 0$ everywhere, in which case $|\nabla A| = \nabla|A|$. This is possible iff $\nabla_k A = c_k(x)A$. This case must include the cases both where M is a sphere and where M is a cylinder. To separate these two cases out, we define $\mathrm{Null}(A)_x$ as those $v \in T_xM$ for which $A(v,w) = 0$ for all $w \in T_xM$ and then we note:

Claim 7.5. *If $v \in \mathrm{Null}(A)_x$, then M splits off a line in the direction v. In other words, there eists an orthogonal transformation O so that $O(M^n) = N^{n-1} \times \mathbb{R}$ and $O_*(v) = e_{n+1}$.*

Proof. Let γ be a unit speed geodesic with initial point and direction (x,v). Then the equation $\nabla_{\dot\gamma}A(\dot\gamma,\dot\gamma) = c_k(\gamma(s))A(\dot\gamma,\dot\gamma)$ gives a linear ODE for $A(\dot\gamma,\dot\gamma)$. $A(\dot\gamma(0),\dot\gamma(0)) = 0$ by the choice of v, and so by the uniqueness of solutions for linear ODEs, $A(\dot\gamma(s),\dot\gamma(s)) = 0$ for any time s. So $\gamma(s)$ is a straight line. In order to see that this straight line is really a full factor that can be split off, let y be another point and $\mu(s)$ be a unit speed from x to y. Given $v, w \in T_xM$, let $v(s)$ and $w(s)$ by the parallel transport of these vectors along $\mu(s)$. Then $\nabla_{\dot\mu}A(v(s),w(s)) = c_k(\mu(s))A(v(s),w(s))$ which implies that $v(s) \in \mathrm{Null}(A)_{\mu(s)}$. Then

$$\nabla^{\mathbb{R}^{n+1}}_{\dot\mu} v(s) = \nabla^M_{\dot\mu} v(s) + A(\dot\mu, v(s)) = 0$$

which implies that $v = v(s)$ for any s as vectors in $\mathbb{R}^{n+1}$. Thus there is an $\mathbb{R}$-factor that we can split off. □

This claim means that, once we have split off all factors, we have reduced our last case to the situation where A has no null space and $N^k \subset \mathbb{R}^{k+1}$. Either $k=1$, in which case we have a curve in $\mathbb{R}^2$, which we have already proven is a circle, or $k>1$. In this scenario, $\nabla_\ell h_{ij} = c_\ell(x) h_{ij}$, so that $\nabla_\ell H = c_\ell(x) H$ or, put differently, $c_\ell(x) = \frac{\nabla_\ell H}{H}$. Then using the Codazzi equation, we have

$$\frac{\nabla_\ell H}{H} h_{ij} = \nabla_\ell h_{ij} = \nabla_i h_{j\ell} = \frac{\nabla_i H}{H} h_{j\ell}$$

For a sphere, $\nabla H = 0$. Suppose that at a point x, $\nabla_x H \neq 0$. Pick an orthonormal frame at x: $e_1, \ldots, e_n$ so that $e_1 = \frac{\nabla H}{|\nabla H|}$. Then

$$0 = |\nabla_\ell h_{ij} - \nabla_j h_{i\ell}|^2 = \left| \frac{\nabla_\ell H}{H} h_{ij} - \frac{\nabla_i H}{H} h_{j\ell} \right|^2 = 0 + \sum_{i\neq 1, \ell=1} \frac{|\nabla H|^2}{H^2} h_{ij}^2 + \sum_{i=1,\ell\neq 1} \frac{|\nabla H|^2}{H^2} h_{j\ell^2}$$
$$= 2 \sum_{1\leq j\leq n, 2\leq i\leq n} \frac{|\nabla H|^2}{H^2} h_{ij}^2$$

This means that for $i \neq 1$ and any j, $h_{ij} = 0$ and so $\mathrm{rank}(A)_x = 1$. But this is impossible if $k \geq 2$ and $\mathrm{Null}(A)_x = 0$. Thus we know that $\nabla H = 0$ and $\nabla A = 0$ at all points.

Now we are almost done. We note that $H = \langle x, \nu \rangle$ so

$$\nabla_i H = \langle x, h_{ij} e_j \rangle = 0$$

and so

$$\nabla_{\ell i} H = h_{i\ell} + \langle x, h_{ij}(-h_{j\ell}\nu) \rangle = 0$$

implying that $h_{i\ell} = H h_{ij} h_{j\ell}$, in other words, $A = HA^2$. We diagonalize A with eigenvalues $\lambda_1, \ldots, \lambda_n$, and then the above equation implies that $A = \frac{1}{H}\mathrm{Id}$. However, because $H = \mathrm{tr}(A)$, $H = \frac{n}{H}$ so that $H = \sqrt{n}$ and $A = \frac{1}{\sqrt{n}}\mathrm{Id}$. By Myer's Theorem, M is compact and furthermore

$$\Delta |x|^2 = 2n + 2\langle \nabla x, x \rangle = 2n - 2H\langle x, \nu \rangle = 2n - 2H^2 = 0$$

so that $|x|^2$ is harmonic on the compact manifold M. But then $|x|^2$ is constant, and so M is a sphere.

7.2. **Type II singularities.** Type II singularities are those for which $\sup |A(t)| \cdot \sqrt{T-t} \xrightarrow{t\to T} \infty$. Type II singularities do not happen for curves in $\mathbb{R}^2$ but they do happen to MCF in higher dimensions. One example is as follows: when we constructed the neckpinch, we chose the radius r of the neck small enough, compared to the size of the spherical ends, so that the neck pinches off before the ends shrink. If r is large enough, then the entire manifold is convex. Somewhere in the middle there is a transition radius, r_0. With this radius, the singularity will be "cylindrical" compared to spherical. We will discuss this more thoroughly in the proof.

7.2.1. *Hamilton's blowup.* For every k, pick a spacetime point (p_k, t_k) so that

$$|A(p_k,t_k)|^2 \left(T - \frac{1}{k} - t_k \right) = \sup_{0\leq t\leq T-\frac{1}{k}, x\in M_t} |A(x,t)|^2 \left(T - \frac{1}{k} - t \right)$$

so that $|A(p_k,t_k)|^2 \left(T - \frac{1}{k} - t_k\right) \xrightarrow{k\to\infty} \infty$. Set $Q_k := |A(p_k,t_k)|$ and consider the flows $Q_k M_{t_k + sQ_k^{-2}} =: M_s^k$ defined on the interval $0 < t_k + sQ_k^{-2} < T$. As we increase k, the interval for s, $(-t_k Q_k^2, (T - t_k)Q_k^2)$ goes to $(-\infty, \infty)$. Note how different this behavior is compared to type I singularities.

Fix s and example $|A^k(x,s)|$. $|A^k(x,s)|^2 = Q_k^{-2}|A(x,t_k+sQ_k^{-2})|^2$. Then

$$|A(x,t_k+sQ_k^{-2})|^2\left(T-\frac{1}{k}-\left[t_k+sQ_k^{-2}\right]\right)\leq |A(p_k,t_k)|^2\left[T-\frac{1}{k}-t_k\right]$$

If we choose p_k so that $t_k+sQ_k^{-2}<T-\frac{1}{k}$, i.e. $s<\left(T-\frac{1}{k}-t_k\right)Q_k^2\to\infty$, so that this holds for every k beyond $k_{\min}(s)$. Then $|A^k(x,s)|^2\leq\frac{T-\frac{1}{k}-t_k}{T-\frac{1}{k}-t_k-sQ_k^{-2}}\to 1$. Therefore M_s^k has uniform curvature bounds on times $(-s,s)$ for all $s>0$. This implies higher derivative bounds as well, and we can pass to a limit, once we show embeddedness. (This is similar to showing multiplicity 1 in the type I case.) The properties of the limit include that $|A|\leq 1$ and $|A|=1$ at $(0,0)$. $M_s^k\to\tilde{M}_s$, an eternal solution of bounded curvature.

Let us give an example of type two singularity.

Example 7.6. *An example of type II singularity of the mean curvature flow is provided by Altschuler-Angenent-Giga, described as following.*

Let us consider a transition stage of the dumbbell. The construction should have two properties:

- *The two big spheres are not too far to each other so that the surface converges, under the mean curvature flow, to the origin.*
- *They are not too close to each other so that the tangent flow of the origin is a cylinder.*

We'll not go into details of this construction. However, after assuming the existence of a dumbbell with the above two properties, we can argue that the orgin is a type II singularity of the mean curvature flow. Indeed, suppose the origin is a type I singularity. Then we have $|A(t)|\leq\frac{C}{\sqrt{T-t}}$, here T is the extinction time. Integrating this with respect to t, we obtain that $\operatorname{diam} M_t\leq C\sqrt{T-t}$. *Therefore $\frac{1}{\sqrt{T-t}}M_t$ is compact. This contradicts to the fact that the tangent cone at the origin is cylindrical.*

For a hypersurface of arbitraty dimension, it is hard to say when it develops a type II singularity in general. Suppose that M_t develops a type II singularity at time T. Then by Hamilton's blowup procedure, there exists an eternal flow $\tilde{M}_{s\in(-\infty,\infty)}$, along which $|A|\leq 1$ and $|A(0,0)|=0$. We could use this fact to prove that when $M^n, n\geq 2$ is convex, then no type II singularity will appear. In fact, we have:

Theorem 7.7. *Suppose $n\geq 2$ and M^n is convex. Then the mean curvature flow starting from M^n disappears at a round point.*

In fact, convexity is equivalent to $A\geq\lambda g$ for some positive constant λ- which, by the maximum principle, is a condition that is preserved under the mean curvature.

Note that the condition $A\geq\lambda g$ is also scaling invariant, hence is also satisfied by the eternal flow $\tilde{M}_s$. However, any hypersurface satisfying $A\geq\lambda g$ must be compact, and thus cannot be satisfied by an eternal flow. Hence we conclude that:

Proposition 7.8. *Any convex hypersurface does not develop type II singularity.*

In the next section we will study the special case when the hypersurface is of dimension 1, namely the case of a curve shortening flow.

8. Curvature shortening flow

In this section we prove that type II singularity never occur for the curve shortening clow. Suppose γ is a compact curve developing a type II singularity at time T. Let $\tilde{\gamma}_s$ be its Hamilton blowup at time T.

Let s be the arclength parameter. We first calculate that

$$\frac{d}{dt}\int_{\gamma_t}|k|ds = \frac{d}{dt}\left(\sum_{P_i}\int kds - \sum_{N_i}\int kds\right),$$

here P_i's and N_i's are all the intervals on which k is positive and negative, respectively. Use the general fact that $\frac{d}{dt}d\,\mathrm{Vol} = -H^2 d\,\mathrm{Vol}$, one calculates that above is

$$\begin{aligned}
&= \sum_{P_i}\int (k_{ss}+K^3)ds - \sum_{N_i}\int (k_{ss}+k^3)ds - \sum_{P_i} k_{ss}ds + \sum_{N_i}\int k_{ss}ds \\
&= \sum_{P_i}\int k_{ss}ds - \sum_{N_i} k_{ss}ds \\
&= -2\sum_{\{x:k(x)=0\}} |k_s|.
\end{aligned}$$

Observe that the total curvature $\int |k|ds$ is scaling invariant. Therefore on the limit flow $\tilde{\gamma}_s$, we conclude that $-2\sum_{\{x:k(x)=0\}}|k_s| = 0$ for almost every time. That is, for almost every t, $k=0$ implies $k_s = 0$.

Proposition 8.1. *$\tilde{\gamma}_s$ is convex, namely $k \neq 0$.*

Sketch. This statement uses essentially that the flow is one dimensional. We will only illustrate the idea by proving the counterpart for the heat equation. Consider a heat equation $u_t = u_{xx}$. Suppose at $(0,0)$, $u = u_t = u_{xx} = 0$ but $u_{xxx} \neq 0$. Then $u_{tx} = 1$. Therefore

$$u(t,x) = \frac{x^3}{3!} + xt + \text{higher order terms.}$$

Now view $u(t,\cdot)$ as a function of x. Then when $t<0$ it has 3 distinct roots. When $t=0$ is has a multiplicity 3 solution. When $t>0$ it has only 1 real solution. For almost every $\epsilon > t > 0$, at the roots of $u(t,\cdot)$ the derivative in x is not zero, contradiction.

A rigirous proof uses the implicit function theorem. We'll not include it here. □

Now that the limit curve is convex, we use a version of the Harnack inequality, namely the Hamilton Li-Yau differential Harnack inequality.

8.1. **The Hamilton Li-Yau Harnack inequality.** Let $M_t \subset \mathbb{R}^{n+1}$, $t \in [0,T)$ be a convex solution of the mean curvature flow with bounded curvature. Let X be a time dependent vector field. Then we have

Theorem 8.2.
- $\frac{\partial H}{\partial t} + \frac{H}{2t} + 2\langle \nabla H, X\rangle + A(X,X) \geq 0,$
- *If H attains a space time maximum, then the flow is a translation soliton.*

Remark 8.3. The usual Harnack inequality can be obtained by integrating the differential Harnack inequality:

Suppose $(x_1,t_1),(x_2,t_2)$ are two points in the space time, $t_1 < t_2$. Let $(\eta(t),t)$ be a curve connecting them. Then

$$\begin{aligned}\log\left(\frac{H(x_1,t_1)}{H(x_2,t_2)}\right) &= \int_{t_1}^{t_2} \frac{\frac{\partial}{\partial t}H(\eta(t),t)}{H(\eta(t),t)} \\ &= \int_{t_1}^{t_2} \frac{\frac{\partial H}{\partial t} + \langle \nabla H, \eta(t)\rangle}{H} \\ &\geq -\int_{t_1}^{t_2} \frac{1}{2t} + \frac{1}{4}\frac{A(\eta',\eta')}{H} \\ &\geq -\int_{t_1}^{t_2} \frac{1}{2t} + |\eta'|^2.\end{aligned}$$

Choose η to be a geodesic in the space time metric, we obtain the usual Harnack inequality.

Remark 8.4. The usual Li-Yau Harnack inequality for the heat equation is the following. Suppose on a n dimensional compact manifold with nonnegative Ricci curvature, u is a positive solution of the heat equation $u_t = \Delta u$. Then

$$\frac{u_t}{u} - \frac{|\nabla u|^2}{u^2} \geq -\frac{n}{2t}.$$

The idea is to define $f = \log u$ and consider the quantity

$$t[|\nabla f|^2 - f_t].$$

We are going to consider a similar quantity and derive a differential inequality of it.

Remark 8.5. Recall that a mean curvature flow is called a translating soliton (or translator), if $\varphi(x,t) = X + tV$, and $\langle V,\nu\rangle = \left\langle \frac{\partial\varphi}{\partial t}, \nu\right\rangle = -H$.

In $\mathbb{R}^2$ the only translator is called the grim-reaper, defined by

$$x = -\log\cos(y), \quad y \in (-\frac{\pi}{2}, \frac{\pi}{2}).$$

We will only prove the Hamilton Li-Yau Harnack inequality for curves in $\mathbb{R}^2$ and for a specific vector field $X = -\frac{\nabla H}{H}$, or

$$\frac{k_t}{k} - \frac{k_s^2}{k^2} + \frac{1}{2t} \geq 0.$$

8.2. **Huisken's two-point function.**

Theorem 8.6. *$R(t)$ is decreasing along the curve shortening flow.*

Proof. Let us suppose the contrary. Then there exsits $t_1 < t_2$ such that $R(t_2) > r > R(t_1)$, here $r > 1$ is any real number between $R(t_1)$ and $R(t_2)$. Consider the function

$$Z(x,y,t) = r|x-y| - \frac{L(t)}{\pi}\sin(\frac{\pi d_t(x,y)}{L(t)}).$$

Then $Z(x,y,t_1) > 0$ for any $x,y \in \gamma$, and $Z(x_2,y_2,t_2) < 0$ for some $x_2,y_2 \in \gamma$.

Let $t_0 \in [t_1,t_2]$ and x_0,y_0 be such that $Z(x,y,t) \geq 0$ for every x,y and for all $t \in [t_1,t_0]$, and $Z(x_0,y_0,t_0) = 0$. Since the function R is scaling invariant, we may also suppose that $L(t_0) = 2\pi$.

Pick arclength parameter of γ and pick the orientation correctly, such that

$$\frac{d}{dx}d_t(x,y) = -1, \quad \frac{d}{dy}d_t(x,y) = 1.$$

Now we are going to apply the maximum principle to the function Z. Unless otherwise indicated, all the above calculations will be derivatives at the point (x_0,y_0,t). Since at (x_0,y_0,t) is a spatial local minimum, we have

$$
\begin{aligned}
0 = \frac{\partial Z}{\partial y} &= r\frac{1}{2|x-y|} \cdot 2\left\langle \gamma'(x), x-y \right\rangle - 2\cos(\frac{d(x,y)}{2}) \cdot \frac{-1}{2} \\
&= r\left\langle T(x), \frac{x-y}{|x-y|} \right\rangle + \cos(\frac{d(x,y)}{2}).
\end{aligned}
$$

here we use $T(x), T(y)$ to denote $\gamma'(x), \gamma'(y)$, respectively.

Similarly,

$$
0 = \frac{Z}{y} = r\left\langle T(y), \frac{y-x}{|x-y|} \right\rangle - \cos(\frac{d(x,y)}{2}).
$$

Denote $V = \frac{y-x}{|x-y|}$. Then the above shows that

$$
0 = -r\left\langle T(x), V \right\rangle + \cos(\frac{d(x,y)}{2}), \quad 0 = r\left\langle T(y), V \right\rangle - \cos(\frac{d(x,y)}{2}).
$$

Therefore we obtain $\langle T(x), V\rangle = \langle T(y), V\rangle = \cos\alpha$. Here α is the signed angle between $T(y)$ and V. Note that $r\cos\alpha = \cos(\frac{d(x,y)}{2})$, and $r > 1$. Therefore we see $\alpha > d(x,y)/2$.

Next, we calculate the t-derivative of Z. Since we are at the first time when Z hits 0, we have

$$
\begin{aligned}
0 &\geq \frac{\partial Z}{\partial t} \\
&= r\frac{1}{2|x-y|}2\left\langle -k(x)\nu(x) + k(y)\nu(y), x-y \right\rangle + \frac{\int_0^{2\pi}}{\pi}\sin(d(x,y)/2) \\
&- 2\cos(d(x,y)/2)\left[\frac{1}{2}(-\int_x^y k^2 ds) + \frac{\pi d_t(x,y)}{(2\pi)^2}\int_0^{2\pi} k^2 ds\right].
\end{aligned}
$$

Denote β to be the signed angle between $T(x)$ and $T(y)$. Then by Cauchy-Schwartz, we have $\int_x^y k^2 \geq (\int_x^y k)^2/d(x,y) = \frac{\beta^2}{d(x,y)}$. Also note that since $T(x), T(y)$ have the same angle α with the vector V, β is either 0 or 2α. By the above, we conclude that

$$
\begin{aligned}
0 &\geq r\left\langle k(x)\nu(x) - k(y)\nu(y), V \right\rangle + \cos(d(x,y)/2)\frac{\beta^2}{d(x,y)} \\
&+ \frac{1}{\pi}\int_0^{2\pi} k^2 ds \cdot \left[\sin(\frac{d(x,y)}{2}) - \frac{d}{2}\cos(\frac{d}{2})\right].
\end{aligned}
$$

Since $d \leq \pi$, we have $\tan(d/2) \geq d/2$. Therefore the last term in the above equation is always nonnegative. We conclude that

$$
0 \geq r\left\langle k(x)\nu(x) - k(y)\nu(y), V \right\rangle + \cos(\frac{d(x,y)}{2}\frac{\beta^2}{d(x,y)}) = \bigstar. \tag{8.1}
$$

Now we calculate the second derivative of Z. We have

$$
\begin{aligned}
0 \leq \frac{\partial^2 Z}{\partial x^2} &= r\left\langle k(x)\nu(x), V \right\rangle + \frac{r}{|x-y|} - \frac{1}{2}r\frac{1}{|x-y|^3}\left\langle T(x), x-y \right\rangle^2 - \sin(\frac{d(x,y)}{2}) \cdot (\frac{-1}{2}) \\
&= r\left\langle k(x)\nu(x), V \right\rangle + \frac{r}{|x-y|}[1 - \left\langle T(x), V \right\rangle^2] + \frac{1}{2}\sin(\frac{d(x,y)}{2}).
\end{aligned}
$$

Similarly,

$$
0 \leq \frac{\partial^2 Z}{\partial y^2} = -r\left\langle k(y)\nu(y), V \right\rangle + \frac{r}{|x-y|}[1 - \left\langle T(y), V \right\rangle^2] + \frac{1}{2}\sin(\frac{d(x,y)}{2}).
$$

To arrive at a contradiction, we will also need to calculate the mixed derivative.

$$\frac{\partial^2 Z}{\partial x \partial y} = \frac{r}{|x-y|}\langle T(x), -T(y)\rangle + r\langle T(x), V\rangle \langle T(y), V\rangle (-\frac{1}{2})\sin(\frac{d(x,y)}{2}).$$

Combine equation 8.1, we see that

$$0 \geq \bigstar = \frac{\partial^2 Z}{\partial x^2} + \frac{\partial^2 Z}{\partial y^2} + 2\frac{\partial^2 Z}{\partial x \partial y} - \frac{r}{|x-y|}[2 - 4\cos^2\alpha + 2\langle T(x), T(y)\rangle] + \cos(\frac{d(x,y)}{2})\frac{\beta^2}{d(x,y)}.$$

By the maximum principle, we conclude that $\frac{\partial^2 Z}{\partial x^2} + \frac{\partial^2 Z}{\partial y^2} + 2\frac{\partial^2 Z}{\partial x \partial y} \geq 0$, since it is the second derivative in direction $x+y$. Now if $\beta = 2\alpha$, then

$$2 - 4\cos^2\alpha + 2\langle T(x), T(y)\rangle = 2 - 4\cos^\alpha + 2\cos\beta = 0,$$

and $\cos(d(x,y)/2)\frac{\beta^2}{d(x,y)} \geq 0$, contradiction.

□

9. Mean convex mean curvature flow

In this section we study the mean curvature flow with mean convex initial data. Recall the evolution equation for H and $|A|$:

$$\frac{\partial H}{\partial t} = \Delta H + |A|^2 H, \quad \frac{\partial |A|}{\partial t} = \Delta |A| - \frac{|\nabla A|^2 - |\nabla |A||^2}{|A| + |A|^3}.$$

Hence by the maximum principle $H > 0$ is preveserved under the mean curvature flow. Moreover, we have:

Proposition 9.1. $\frac{|A|}{H}$ *is decreasing along the mean curvature flow of compact hypersurfaces.*

Proof.

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{|A|}{H}\right) - \Delta\left(\frac{|A|}{H}\right) &= \frac{(\partial_t - \Delta)|A|}{H} - \frac{|A|(\partial_t - \Delta)H}{H^2} \\
&\quad - 2\frac{|\nabla H|^2 |A|^2}{H^3} + 2\frac{\nabla |A| \cdot \nabla H}{H^2} \\
&\leq -\frac{2|\nabla H|^2 |A|}{H^3} + \frac{2\nabla |A| \cdot \nabla H}{H^2} \\
&= \frac{2}{H}\left\langle \nabla(\frac{|A|}{H}), \nabla H \right\rangle.
\end{aligned}$$

Hence by the maximum principle $\frac{|A|}{H}$ is decreasing. □

The new idea in the study of mean convex mean curvature flow is splitting theorem and Hamilton's strong maximum principle for tensors. As an analogue, let us recall Cheeger-Gromoll splitting theorem: if a manifold has nonnegative Ricci curvature and it contains a geodesic line, then it splits, namely, it is $M \times \mathbb{R}$ as Riemannian manifold. In our case we are going to prove a strong maximum principle for tensors, such that if an interior maximum is attained, then the manifold splits.

9.1. A strong maximum principle for tensors.

Let Ω be a domain. Suppose on $\Omega \times [0, T]$, there is a family of metrics $g(t)$ and $\frac{\partial}{\partial t} g = h$. Then for any time dependent vector field V, we have

$$\frac{d}{dt} g(V, V) = \frac{d}{dt}(g_{ij} V^i V^j) = h_{ij} V^i V^j + g_{ij}(\frac{\partial}{\partial t} V^i) V^j + g_{ij} V^i \frac{\partial}{\partial t} V^j.$$

If we choose V carefully so that $\frac{\partial}{\partial t} V^i = -\frac{1}{2} g^{ik} h_{kj} V^j$, then the above is

$$= h_{ij} V^i V^j - \frac{1}{2} h_{jl} V^l V^j - \frac{1}{2} h_{il} V^l V^i = 0.$$

Hence we have the following

Definition 9.2. A time dependent vector field V is called time-parallel, if

$$\frac{\partial}{\partial t}V^i = -\frac{1}{2}g^{ik}h_{kl}V^l.$$

Note that by basic theory of ODE, time-parallel vector exists for any Lipschitz initial value.

Theorem 9.3 (The strong maximum principle for tensors)**.** *Suppose on $\Omega \times [0,T]$, a family of metrics $g(t)$ satisfies $\partial_t g = h$, and M is a time dependent 2-tensor, such that for any time-parallel vector field V,*

$$\frac{\partial}{\partial t}M(V,V) \geq (\Delta M)(V,V).$$

Let $\lambda_1(M)$ be its smallest eigenvalue. Then λ_1 is non-decreasing. If $\lambda_1(M)$ attains space time minimum at some point (x,t), $t>0$, then $\lambda_1(M)$ is constant on $\Omega \times [0,T]$.

Moreover, in the above case, $(\Delta M)(V,V) = 0$, for any eigenvector of λ_1.

Proof. Let (x,t) be a point where λ_1 attains its space time minimum, and V is a unit eigenvector of λ_1. We first extend V via parallel transport to a neighborhood of x, and thus $(\Delta M)(V,V) = \Delta M(V,V)$. Next we extend V to be a time-parallel vector field.

Then since λ_1 is the smallest eigenvalue and (x,t) attains its space time minimum, we conclude that at a time t where λ_1 is smooth,

$$\frac{\partial}{\partial t}\lambda_1 \geq \frac{\partial}{\partial t}M(V,V) \geq \Delta M(V,V) \geq \Delta\lambda_1 \geq 0.$$

Hence λ_1 is non-decreasing, and whenever it attains a space time minimum, it must be constant by the maximum principle. Also in this case, $\Delta M(V,V) = 0$.

In general λ_1 may only be Lipschitz. Then the above inequality is understood in the viscosity sence. The maximum principle also works in the viscosity sense. We will elaborate this point below. □

Definition 9.4. A Lipschitz function f is said to satisfy $\frac{\partial}{\partial t}f \geq \Delta f$ in the viscosity sense, if for every C^2 function φ in a backwards neighborhood of any point (x,t), such that $\varphi \geq f$ in the neighborhood and $\varphi = f$ at (x,t), we have $\frac{\partial}{\partial t}\varphi \geq \Delta\varphi$.

The maximum principle also holds for viscosity solutions. Indeed, if (x_0,t_0) is a point of f where f attains its space time minimum, then by definition, for any other C^2 function φ, $\varphi \geq f$ in a neighborhood and $\varphi(x_0,t_0) = f(x_0,t_0)$, we see that φ attains its space time minimum also at (x_0,t_0). Therefore

$$\frac{\partial}{\partial t}\varphi \geq \frac{\partial}{\partial t}f \geq \Delta f \geq \Delta\varphi, \quad \text{at } (x_0,t_0).$$

But $\frac{\partial}{\partial t}\varphi \leq 0$, $\Delta\varphi \geq 0$. Then use a standard perturbation argument we conclude that this cannot happen unless f is a constant function.

We are know ready to state and prove the splitting theorem.

Theorem 9.5. *Under the same assumption as before, suppose the space time minimum is attained. Then Ω splits locally to $\Omega = P \times Q$, such that V is tangent to P if and only if $M(V,V) = \lambda_1 g(V,V)$.*

Proof. We will use the Frobenius theorem to produce a local splitting. Without loss of generality let us assume that $\lambda_1 = 0$, otherwise just consider $M - \lambda_1 I$ instead.

Let x be a point where the space time minimum is attained. Take a normal coordinate $\{E_i\}$ around x. Then for any vector field V, $\nabla_{E_i}(\nabla_{E_i}M(V,V)) = \Delta M(V,V)$.

Take V to be an arbitrary section of the null space of M around x. Then we see that

$$\begin{aligned}
0 = M(V,V) &= \nabla_{E_i}(\nabla_{E_i}M(V,V)) \\
&= \nabla_{E_i}\left((\nabla_{E_i}M)(V,V) + 2M(\nabla_{E_i}V,V)\right) \\
&= (\Delta M)(V,V) + 2(\nabla_{E_i}M)(\nabla_{E_i}V,V).
\end{aligned}$$

By the previous theorem, since V is λ_1-eigensection, we have $(\Delta M)(V,V)=0$. Therefore

$$\begin{aligned}0 &= (\nabla_{E_i}M)(\nabla_{E_i}V,V)\\ &= \nabla_{E_i}M(\nabla_{E_i}V,V)-M(\nabla_{E_i}V,\nabla_{E_i}V)-M(\nabla_{E_i}\nabla_{E_i}V,V)\\ &= -M(\nabla_{E_i}V,\nabla_{E_i}V).\end{aligned}$$

Since $M\geq 0$, we conclude that $\nabla_{E_i}V$ is in the null space of M. Therefore we conclude that $\mathrm{Null}(M)$ is a integrable distribution.

For any vector $W\in\mathrm{Null}(M)^{\perp}$, we prove $\nabla_{E_i}W\in\mathrm{Null}(M)^{\perp}$. In fact, for any vector $V\in\mathrm{Null}(M)$ and any vector field X, we have

$$0=X(W\cdot V)=(\nabla_X W)\cdot V+W\cdot\nabla_X V.$$

Since $\nabla_X V=0$, we conclude that $\nabla_X W\cdot V=0$. That is, $\nabla_X W\in\mathrm{Null}(M)^{\perp}$.

By the Frobenius theorem, both $\mathrm{Null}(M)$ and $\mathrm{Null}(M)^{\perp}$ are integrable distributions. Let P,Q be the submanifolds they define. We have, as differentiable manifolds, $M=P\times Q$ locally. It remains to prove this splitting is also in the sence of Riamannian manifold, i.e., the metric is also a product metric locally.

□

9.2. **Local curvature estimate.** In this section we derive a local curvature estimate from α-noncollapsedness of the mean curvarure flow.

Definition 9.6. A smooth α-noncollapsed flow on an open set $U\subset\mathbb{R}^{n+1}$ is a family of closed sets $K_t\subset U$ such that ∂K_t is mean convex and flows by the mean curvature flow, and is α-noncollapsed.

Theorem 9.7. *For any $\alpha>0$, there exists a $\rho=\rho(\alpha)$ and $C_i=C_i(\alpha)$, $i=0,1,\dots$ such that if K is an α-noncollapsed flow in the parabolic ball $P(p,t,r)$ centered at $p\in\partial K_t$ with $H(p,t)\leq r^{-1}$, then*

$$\sup_{P(p,t,\rho r)}|\nabla^i A|\leq C_i r^{-(1+i)}.$$

Proof. By the standard curvature estimate we only need to prove the statement for $|A|$, namely $i=0$. We argue by contradiction. Suppose the contrary. Then by rescaling at the point where the curvature is maximized, we may assume that we can take a sequence of closed sets representing the mean curvature flow, K^j, defined on $P(0,0,j)$ such that $H(0,0)\leq j^{-1}$ and $\sup_{P(0,0,1)}|A|\geq j$. Assume $\nu(0,0)=e_{n+1}$.

We first prove the following property of the rescaled surfaces K^j: K^j converges in the strong Hausdorff sense to the static half space $\langle e_{n+1},X\rangle\leq 0$. Namely, K_j converges in Hausdorff sense to $\langle e_{n+1},X\rangle\leq 0$ and the complement of K^j converges to $\langle e_{n+1},X\rangle>0$.

First let us prove that K^j converges in the Hausdorff sense to the lower half space. Since K^j is α-noncollapsed and $H(0,0)=j^{-1}$, K^j contains a ball of radius αj in the lower half space which is tangent at $(0,0)$ to the $e_{n+1}=0$ coordinate plane. Also since the surfaces involves as mean convex mean curvature, for $t_2<t_1$, $K^j_{t_1}\subset K^j_{t_2}$. Therefore every point in the lower half space is contained in K^j_t, as $t\to 0$ and $j\to\infty$.

Now let us prove that the complement of K^j converges in the Hausdorff sense to the upper half space. Pick a sequence of real numbers $R_j\to\infty$ and $d_j\to 0$ such that $Rd\to\infty$. Let $B_{R,d}=B(-Re_{n+1},R+d)$. Its mean curvature is $H=\frac{1}{R+d}$. Under the mean curvature flow, the time for $B_{R,d}$ to leave the upper half space is $T=(R+d)d$. By the avoidance principle, for $t\in[-Rd,0]$, K^j_t cannot contain $B_{R,d}$ (otherwise K^j_t will contain an open set of the upper half space, contradicting to the fact that they converge to the lower half space).

Take the largest possible d' such that $B_{R,d'}$ and K^j_t have a contact point. Let p_j be the contact point. We observe that p_j must satisfy the following three properties:

(1) $\langle p_j, e_{n+1}\rangle \leq d$.
(2) $\liminf_{j\to\infty} \langle p_j, e_{n+1}\rangle \geq 0$.
(3) $\|p_j\| \leq 2\sqrt{2Rd}$.

Property 1 follows easily from the choice of d'. Property 2 follows from the fact that K_t^j converges to the lower half space (otherwise there will be an open set of the lower half space that is not contained in every K_t^j.) We now elaborate on property 3. At the point p_j, K_t^j has a tangent ball of radius R. Therefore the mean curvature at p_j is at most $\frac{n}{R}$. By the nocollapsing condition, there exists an exterior tangent ball of radius at least $\frac{\alpha}{n}R$ at p_j. Now that $\operatorname{dist}(p_j, 0) \leq \sqrt{Rd}$ and p_j is on $\partial B(R, d)$, the tangent plane at of k_t^j at p_j is sufficiently close to the coordinate plane $e_{n+1} = 0$. Therefore the ball of radius $\frac{\alpha}{n}R$ tangent to p_j converges to the upper half space as $j \to \infty$. This proves that the complement of K_t^j converges in Hausdorff sense to the upper half space.

Now we finish the proof of the theorem under the extra assumption that $\forall R > 0$, K_t^j contains $B(0, R)$, for j sufficiently large and $t = t(j)$. To do so, we first observe the following

Lemma 9.8 (One-sided minimization). *For $t_2 > t_1$, and any open subset V such that $K_{t_2} \subset V \subset K_{t_1}$, $\operatorname{Vol}(\partial K_{t_2}) \leq \operatorname{Vol}(\partial V)$.*

The proof of lemma is a calibration argument. The mean curvature flow from time t_1 to t_2 creates a foliation of $K_{t_1} - K_{t_2}$ by mean convex hypersurfaces. Define the following vector field v in $K_{t_1} - K_{t_2}$: for $x \in \partial K_t$, let $v(x) = \nu_{\partial K_t}(x)$, $t \in [t_1, t_2]$. Then by the Stokes' formula and the mean convexity of each leaf of the foliation,

$$0 \leq \int_V \operatorname{div}(v) = \int_{\partial V} \langle v, \nu_V\rangle - \int_{\partial K_{t_2}} \left\langle v, \nu_{\partial K_{t_2}}\right\rangle \leq \operatorname{Vol}(\partial V) - \operatorname{Vol}(\partial K_{t_2}),$$

the lemma is proved.

To finish the proof, we see that as j approaches infinity, K_t^j is contained in an arbitrarily thin layer near the coordinate plane $e_{n+1} = 0$. Use the above lemma, we see that for any $\epsilon > 0$,

$$\operatorname{Vol}(B(p, R) \cap \partial K_t^j) \leq (1 + \epsilon)\omega_n R^n,$$

for sufficiently large j. Therefore by B. White's epsilon regularity theorem, we conclude the local curvature estimate $\sup_{B(p,R/2)} |A|$ is uniformly bounded. Therefore, the flow K_t^j converges smoothly to its limite, contradicting $\sup_{P(0,0,1)} \leq j$.

Finally we are going to remove the extra assumption that $B(0, R) \subset K_t^j$ for large j, $t = t(j)$.

Claim 9.9. *There exists μ such that if the flow is defined on $P(p, t, \mu r)$ and $H \leq r^{-1}$, then $\sup |A| \leq C_0 r^{-1}$ on $P(p, t, \rho r)$.*

Assuming the claim, the theorem is proved with a different ρ. Now let us prove the claim.

If the statement of the claim is not true, there exist counter examples for $\mu_j = j$. Then we can assume:

(1) K^j is defined on $P(0, 0, j)$ and $\nu(0, 0) = e_{n+1}$.
(2) $H(0, 0) \to H_\infty \leq 1$.
(3) $\sup |A| > C_0$ in $P(0, 0, \rho)$.
(4) If for $(p, t) \in P(0, 0, \frac{j}{5})$ we have $H(p, r) \leq r^{-1}$ for $r \leq \frac{1}{2}$, then $\sup_{P(p,t,\rho r)} \leq C_0 r^{-1}$.

Note that property (4) is similar to the point selection process we have seen earlier. It naturally appeared because we only assume that the flow is defined in an open ball. We separate two cases.

Case 1 $H_\infty = 0$. We know that K^j converges in strong Hausdorff sense to the lower half space. Use property (4) and cover the whole flow by balls of radius $\frac{1}{2}$, we see that $|A|$ is uniformly bounded for the entire flow. Therefore the flow has to converge smoothly to its limit. Therefore the limit has to be a plane, contradicting property (3).

Case 2 $H_\infty > 0$. In this case, there is some neighborhood of fixed size around $(0,0)$ in which $H \le \frac{H_\infty}{2}$. Therefore, there exist r_0, t_1 such that $B(0, r_0) \subset K_{t_1}$. By the avoidance principle, we deduce that for any radius R, there is some t_2 such that $B(0, R) \subset K_{t_2}$.

Pick a point p_t to be a point on ∂K_t^j such that $|p_t| = \mathrm{dist}(\partial K_t, 0)$. We prove that $H(p_t, t) \ge \frac{C}{|t|}$, $C = C(\alpha, r_0)$ is some constant. Otherwise if $H \le \frac{C}{|t|}$, by α-noncollapsing there exists an inscribed ball of radius comparable to $|t|$. But for $|t|$ large, such a ball cannot reach the origin at time 0. (Note that a ball of radius R becomes a ball of radius $\sqrt{R^2 - 2ns}$ after time s under the mean curvature flow. Hence after time $|t|$, the radius $|t|$ becomes $\sqrt{t^2 - 2n|t|}$, which is positive for $|t|$ large)

Note that $\mathrm{dist}(\partial K_t, 0)$ is a Lipschitz function of t, and at a point (p_t, t) where the function is differentiable, its derivative is given by $H(p_t, t)$. By above we see that $\frac{d}{dt}\mathrm{dist}(\partial K_t, 0) \ge \frac{C}{|t|}$ for some constant C. Since $\int \frac{C}{|t|} dt$ diverges, we deduce that $\mathrm{dist}(\partial K_t, 0) \to \infty$ as t approaches to 0. That is, K_t^j contains a ball of arbitrarily large radius. The theorem follows.

$\square$