Mean Curvature Flow and Related Topics

Tom Begley

December 3, 2015

The following notes are based on a lecture course given by Neshan Wickramasekera at the University of Cambridge in spring 2015. Any errors are due to me, please do tell me about them if you find some (T.Begley@maths.cam.ac.uk). Suggested background reading for this course is [6] [16] [22]

Contents

1 Simplest setting: Curves in $\mathbb{R}^2$

We consider an immersed curve $\gamma : [a, b] \to \mathbb{R}^2$ (i.e. $\gamma'(u) \neq 0$ for all $u \in [a, b]$). Since γ is immersed we can define the arc length function

$$s(u) := \int_a^u |\gamma'(r)| dr$$

for each $u \in [a,b]$. This is a diffeomorphism $[a,b] \to [0,L]$ where $L = s(b)$ is the length of γ. Consequently we can reparametrise γ by arc length via the composition

$$\sigma = \gamma \circ s^{-1} : [0,L] \to \mathbb{R}^2.$$

It is easy to check that $|\sigma'(u)| \equiv 1$ on $[0,L]$. Indeed from the definition of s we see that

$$s'(u) = |\gamma'(u)|,$$

so it follows that

$$\sigma'(u) = \gamma'(s^{-1}(u))(s^{-1})'(u) = \frac{\gamma'(s^{-1}(u))}{s'(s^{-1}(u))} = \frac{\gamma'(s^{-1}(u))}{|\gamma'(s^{-1}(u))|}.$$

Taking norms we get precisely the desired result. Differentiating the identity $|\sigma'(u)| \equiv 1$ we get $\langle \sigma''(u), \sigma'(u) \rangle = 0$, or in other words $\sigma'' \perp \sigma'$. In particular since σ' is tangent to the curve, σ'' is normal to the curve. We define $\vec{k} = \sigma''(u)$ to be the curvature vector at u. If we choose a continuous unit normal $N(u)$ to the curve, we may define the (signed) scalar curvature $k(u)$ by

$$\vec{k}(u) = k(u)N(u)$$

or equivalently

$$k(u) = \langle \vec{k}(u), N(u) \rangle.$$

Notice that the sign of k will depend on the choice of N, but $\vec{k}$ is well defined regardless of the choice we make. As a concrete example of the above, suppose that $\gamma = (\gamma_1, \gamma_2)$ and choose

$$N = \frac{-\gamma_2' e_1 + \gamma_1' e_2}{|\gamma_1' e_2 - y_2' e_1|},$$

then

$$k(u) = \left. \frac{\gamma_1'\gamma_2'' - \gamma_2'\gamma_1''}{((\gamma_1')^2 + (\gamma_2')^2)^{3/2}} \right|_u$$

We can now define mean curvature flow for curves in the plane, which is also called curve shortening flow.

Definition. A one parameter family of curves

$$\Gamma : [a,b] \times [0,T) \to \mathbb{R}^2 \qquad \gamma_t(\cdot) = \Gamma(\cdot, t)$$

is a curve shortening flow starting at the initial curve γ_0 if

$$\begin{aligned} \tfrac{\partial \Gamma}{\partial t}(r,t) &= k_{\gamma_t}(r) N_{\gamma_t}(r) \\ \Gamma(r,0) &= \gamma_0(r). \end{aligned} \tag{1.1}$$

Henceforth we will suppress the subscript γ_t which was included here only for emphasis of the time dependence on the right hand side. We will also assume that γ_0 is smooth, closed and embedded. That is to say that $\gamma_0(a) = \gamma_0(b)$ and γ has no self-intersections (other than at the

endpoints).
Note that the PDE in the definition of curve shortening flow could also be stated as

$$\frac{\partial \gamma}{\partial t} = \frac{\partial^2 \gamma}{\partial s^2}$$

where on the right hand side we use the notation $\partial^2/\partial s^2$ to denote the second derivative with respect to arc length. Superficially this looks like the heat equation which is of course very well understood. However it is important to note that because arc length is not preserved under the flow, the right hand side is in fact not linear. To see this in more detail we rewrite the equation (1.1) in local (extrinsic) coordinates. Specifically we write the flow as the one parameter family of curves $\gamma : [a,b] \times [0,T) \to \mathbb{R}^2 : (u,t) \mapsto \gamma(u,t)$ and compute the evolution

$$\frac{\partial \gamma}{\partial t} = \frac{\partial}{\partial s}\left(\frac{\partial \gamma}{\partial u}\frac{\partial u}{\partial s}\right) = \frac{\partial}{\partial s}\left(\frac{1}{v}\frac{\partial \gamma}{\partial u}\right)$$

where $v = |\partial\gamma/\partial u|$. We therefore have $\partial/\partial s = v^{-1}\partial/\partial u$. Continuing the computation

$$\begin{aligned}\frac{\partial \gamma}{\partial t} &= \frac{1}{v}\frac{\partial}{\partial u}\left(\frac{1}{v}\frac{\partial \gamma}{\partial u}\right)\\ &= \frac{1}{v^2}\frac{\partial^2 \gamma}{\partial u^2} - \frac{1}{v^4}\left(\frac{\partial \gamma}{\partial u}\cdot\frac{\partial^2 \gamma}{\partial u^2}\right)\frac{\partial \gamma}{\partial u},\end{aligned}$$

or in coordinates, writing $\gamma = (\gamma_1, \gamma_2)$ we have for $j = 1, 2$

$$\begin{aligned}\frac{\partial \gamma_j}{\partial t} &= \frac{1}{v^2}\frac{\partial^2 \gamma_j}{\partial u^2} - \frac{1}{v^4}\left(\sum_{i=1}^{2}\frac{\partial \gamma_i}{\partial u}\frac{\partial^2 \gamma_i}{\partial u^2}\right)\frac{\partial \gamma_j}{\partial u}\\ &= \sum_{i=1}^{2}\frac{1}{v^2}\left(\delta_{ij} - \frac{1}{v^2}\frac{\partial \gamma_i}{\partial u}\frac{\partial \gamma_j}{\partial u}\right)\frac{\partial^2 \gamma_i}{\partial u^2}\\ &= \sum_{i=1}^{2} a_{ij}\frac{\partial^2 \gamma_i}{\partial u^2}.\end{aligned}$$

From the above representation the non-linearity is much clearer. The equation is however still at least quasi-linear (linear in its highest order derivatives), and parabolic. We can compute the trace and determinant of the matrix a_{ij}

$$\det(a_{ij}) = \frac{1}{v^4}\left(\left(1 - \frac{1}{v^2}\left(\frac{\partial \gamma_1}{\partial u}\right)^2\right)\left(1 - \frac{1}{v^2}\left(\frac{\partial \gamma_2}{\partial u}\right)^2\right) - \frac{1}{v^4}\left(\frac{\partial \gamma_1}{\partial u}\right)^2\left(\frac{\partial \gamma_2}{\partial u}\right)^2\right) = 0$$

which follows because $v^{-2}\left((\partial\gamma_1/\partial u)^2 + (\partial\gamma_2/\partial u)^2\right) = 1$. Similarly one can check that

$$\operatorname{trace}(a_{ij}) = \frac{1}{v^2} > 0.$$

Hence the eigenvalues of a_{ij} are everywhere non-negative, but one of them is always zero. This is the sense in which the equation is degenerate. Degeneracy is typical of geometric flows.

Theorem 1.1 (Short time existence)**.** *Given* γ_0 *as above, there exists* $T > 0$ *such that* (1.1) *has a unique smooth solution valid for* $t \in [0, T)$*.*

We will not discuss the proof here, instead see [16].

Example. Suppose that $\gamma(\theta) = (\cos\theta, \sin\theta)$ parametrises the unit circle in $\mathbb{R}^2$ centred at the origin. We consider the curve shortening flow of the circle radius R_0 which is parametrised by $\gamma_0 = R_0\gamma$. By symmetry one expects the curve shortening flow of the circle to remain a circle with a possibly different radius. Hence we expect a solution of the form $\gamma_t = R(t)\gamma$ with $R(0) = R_0$. One can compute easily that the curvature vector of γ_t points inward and has magnitude $R(t)^{-1}$, which we may write in terms of γ as $-R(t)^{-1}\gamma$. Hence if γ_t is a curve shortening flow we must have

$$\frac{\partial}{\partial t}\gamma_t = R'(t)\gamma = -\frac{1}{R(t)}\gamma$$

which implies that $R'(t) = -(R(t))^{-1}$. It is easily seen that the solution is given by $R(t) = \sqrt{R_0^2 - 2t}$. This tells us that the flow can only exist smoothly for a finite time, as at time $t = R_0^2/2$ the curve will collapse down to a point.

We now state some facts about curve shortening flow, some with proof.

(1) For a curve shortening flow as above we have the identity

$$\frac{d}{dt}(\text{length}(\gamma_t)) = -\int |\vec{k}|^2 ds,$$

where ds denotes the integral with respect to arc length. In particular this tells us that the length of the curve under curve shortening flow is decreasing.

Proof. Recall that the length of γ_t is given explicitly by

$$\text{length}(\gamma_t) := \int_a^b |\gamma_t'(u)| du.$$

We define unit tangent vectors and speed of the parametrisation of γ_t as follows

$$T_t(u) := \frac{\gamma_t'(u)}{|\gamma_t'(u)|} \qquad v_t(u) := |\gamma_t'(u)|.$$

We now compute dv_t/dt.

$$\begin{aligned}\frac{d}{dt}(v_t)^2 &= \frac{d}{dt}\langle\gamma_t', \gamma_t'\rangle = 2\left\langle \gamma_t', \frac{d}{dt}\gamma_t \right\rangle \\ &= 2\left\langle \gamma_t', \frac{d}{du}(kN) \right\rangle = 2\left\langle \gamma_t', k\frac{dN}{du} + N\frac{dk}{du} \right\rangle \\ &= 2k\left\langle \gamma_t', \frac{dN}{du} \right\rangle.\end{aligned}$$

The well known Frenet formulas tell us that

$$\frac{dT}{du} = vkN \qquad \frac{dN}{du} = -vkT$$

so

$$\frac{d}{dt}(v_t)^2 = 2k\langle vT, -vkT\rangle = -2k^2v^2,$$

from which is follows that

$$\frac{dv}{dt} = -k^2v.$$

The main claim now easily follows by differentiating under the integral. □

(2) The flow satisfies an avoidance principle. That is to say if γ_t^1 and γ_t^2 are both curve shortening flows existing smoothly on an interval $[0, T)$ and such that $\gamma_0^1 \cap \gamma_0^2 = \emptyset$. Then $\gamma_t^1 \cap \gamma_t^2 = \emptyset$ for all $t \in [0, T)$. This is also true of mean curvature flow of hypersurfaces in higher dimensions. In the higher codimension it is no longer true.
The rough idea is as follows. Suppose that the claim were not true. Since the flows are compact, evolving smoothly, and initially disjoint, there is a positive first time $t_1 > 0$ where they touch. At the point where they touch, the curvature vectors must point in the same direction. Indeed, the intersection point cannot be transverse, since otherwise at a slightly earlier time the curves would still intersect. Hence the tangent lines coincide and so the curvature vectors lie on the same line. Moreover the curvature vectors must point in the same direction (else once again, there must have been an intersection at a slightly earlier time). Finally the curvature of the "inner" curve must be greater than or equal to the curvature of the outer curve at the intersection point. If we have a strict inequality that actually then implies once again that there must have been a previous intersection. Thus we arrive at a contradiction. We will later make this picture rigorous using the maximum principle for parabolic PDEs.

(3) The flow exists smoothly for a finite time. This follows immediately from the avoidance principle. Indeed since the initial curve is compact, it is bounded. Hence we can enclose it in a suitably large circle. The circle we know shrinks to a point in finite time ($T = R^2/2$ where R is the radius of the initial circle), hence by the avoidance principle the evolution of the inner curve cannot exist smoothly for this entire time.

Theorem 1.2 (Gage-Hamilton)**.** *A closed embedded convex curve remains convex and shrinks to a round point in finite time under curve shortening flow.*

The meaning of shrinking to a round point is that if we rescale the flow so that enclosed area is constant then the curve converges smoothly to a circle.

Theorem 1.3 (Grayson)**.** *Any embedded closed curve under curve shortening flow becomes convex in finite time, and hence shrinks to a round point.*

We will prove these later in the course, but the proof is long and requires more machinery than we have so far developed. Hence we will now simply state some further results about curve shortening flow.

Lemma 1.4. *If $A(t)$ denotes the area enclosed by the moving curve at time t, then $A(t) = A(0) - 2\pi t$.*

Corollary 1.5. *The flow γ_t becomes extinct at time $t = A(0)/2\pi$.*

Proof. This follows immediately from Grayson's theorem and the preceding lemma. □

Lemma 1.6. *If $T = \frac{\partial\gamma}{\partial u}/\left|\frac{\partial\gamma}{\partial u}\right|$ is the unit tangent in the direction of the parametrisation and N is the inward pointing unit normal then*

$$\frac{\partial T}{\partial t} = \frac{\partial k}{\partial s}N \qquad \frac{\partial N}{\partial t} = -\frac{\partial k}{\partial s}T$$

where s is the arc length.

Proof. First we observe that

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\partial}{\partial s} &= \frac{\partial}{\partial t}\left(\frac{1}{v}\frac{\partial}{\partial u}\right) = \frac{1}{v}\frac{\partial}{\partial t}\frac{\partial}{\partial u} - \frac{1}{v^2}(-k^2 v)\frac{\partial}{\partial u}\\ &= \frac{1}{v}\frac{\partial}{\partial u}\frac{\partial}{\partial t} + k^2\frac{\partial}{\partial s}\end{aligned}$$

Hence

$$\begin{aligned}\frac{\partial}{\partial t}T &= \frac{\partial}{\partial t}\frac{\partial\gamma}{\partial s} = \frac{\partial(kN)}{\partial s} + k^2\frac{\partial\gamma}{\partial s}\\ &= k\frac{\partial N}{\partial s} + \frac{\partial k}{\partial s}N + k^2 T\\ &= -k^2 T + \frac{\partial k}{\partial s}N + k^2 T = \frac{\partial k}{\partial s}N.\end{aligned}$$

To calculate the formula for $\partial N/\partial t$ we use the fact that $\langle N, T\rangle \equiv 0$. □

Proof of Lemma 1.4. By the divergence theorem

$$\int_\Omega \mathrm{div} X dxdy = -\int_\gamma \langle X, N\rangle ds.$$

Choose $X(x,y) = (x,y)$, so that $\mathrm{div}X \equiv 2$. Then

$$2A(t) = -\int_{\gamma_t}\langle\gamma, N\rangle ds = -\int_a^b v\langle\gamma, N\rangle du.$$

Differentiating in t we have

$$\begin{aligned}2A'(t) &= -\int_a^b\left(\left\langle\frac{\partial\gamma}{\partial t}, N\right\rangle v + \left\langle\gamma, \frac{\partial N}{\partial t}\right\rangle v + \langle\gamma, N\rangle\frac{\partial v}{\partial t}\right)du\\ &= -\int_a^b\left(kv - \left\langle\gamma, \frac{\partial(kT)}{\partial u}\right\rangle + \left\langle\gamma, \frac{\partial T}{\partial u}k\right\rangle - k^2\langle\gamma, N\rangle v\right)du\\ &= -\int_a^b\left(kv + \left\langle\frac{\partial\gamma}{\partial u}, kT\right\rangle\right)du\\ &= -2\int_{\gamma_t} kds = -4\pi\end{aligned}$$

Hence, by integrating we have

$$A(t) = A(0) - 2\pi t$$

□

1.1 Towards the proof of Grayson's theorem

The proof we will present in this course is not the original proof of Gage-Hamilton-Grayson theorem (which is interesting in its own right, and can be found in the original papers of Gage-Hamilton [12] and Grayson [13] but rather one which utilises some more modern ideas which weren't known at the time of the original proofs. One of the key ingredients is an idea of Huisken's. The idea is to rule out possibilities such as the following picture by considering the ratio of extrinsic distance to intrinsic distance and showing it is bounded below. What Huisken was able to do [15] is show that a quantity related to this ratio is monotone under curve shortening flow.

Let us be more precise. We suppose that $\gamma(\cdot, t) : S^1 \to \mathbb{R}^2$ is an embedded closed curve moving by curve shortening flow. We define the extrinsic distance

$$D_t : S^1 \times S^1 \to \mathbb{R} : (p, q) \mapsto |\gamma_t(p) - \gamma_t(q)|$$

and the intrinsic distance

$$l_t : S^1 \times S^1 \to \mathbb{R} = \min\left\{\left|\int_p^q |\gamma_t'(u)| du\right|, L - \left|\int_p^q |\gamma_t'(u)| du\right|\right\} \tag{1.2}$$

where $L = L(t)$ is the length of γ_t. Notice that l_t is not smooth at points (p, q) where $l_t(p, q) = L/2$. For this reason we define the related quantity

$$\psi_t(p, q) := \frac{L}{\pi} \sin\left(\frac{\pi l}{L}\right)$$

Since $\sin(\pi s/L) = \sin(\pi(L - s)/L)$, it follows that ψ_t is smooth. Now the ratio D_t/ψ_t is well defined away from the diagonal of $S^1 \times S^1$, and extends continuously to the diagonal if we impose $D_t/\psi_t \equiv 1$ on the diagonal. We can now prove the monotonicity result of Huisken, originally proved in [15]

Theorem 1.7. *The minimum of D_t/ψ_t is non-decreasing in t if (γ_t) is a curve shortening flow.*

Proof. Because of out definition of D_t/ψ_t on the diagonal, we always have that

$$\min_{S^1 \times S^1} \frac{D_t}{\psi_t} \leq 1.$$

It will suffice to show that at any time $t_0 \in (0, T)$, if the minimum is strictly less than 1, we have at any point (p, q) where the minimum is attained that

$$\left(\frac{\partial}{\partial t}\bigg|_{t=t_0} \frac{D_t}{\psi_t}\right)(p, q) > 0.$$

It will then follow that $\min_{S^1\times S^1} D_t/\psi_t$ is increasing locally near t_0. Moreover, if the minimum ever becomes 1, then it must remain 1 under the flow. We calculate using the fact that $\partial v/\partial t = -k^2 v$

$$\begin{aligned}
\left.\frac{\partial}{\partial t}\right|_{t=t_0}\left(\frac{D_t}{\psi_t}\right)(p,q) &= \left.\frac{\partial}{\partial t}\right|_{t=t_0} \frac{|\gamma(p,t)-\gamma(q,t)|}{\frac{L}{\pi}\sin\left(\frac{\pi l}{L}\right)} \\
&= \frac{1}{\frac{L}{\pi}\sin\left(\frac{\pi l}{L}\right)}\left\langle \frac{\gamma(p,t)-\gamma(q,t)}{|\gamma(p,t)-\gamma(q,t)|}, \vec{k}(p,t)-\vec{k}(q,t)\right\rangle \\
&\quad - \frac{D_t}{\left(\frac{L}{\pi}\sin\left(\frac{\pi l}{L}\right)\right)^2}\left(\frac{\sin\left(\frac{\pi l}{L}\right)}{\pi}\int_{\gamma_t} k^2 ds + \frac{l}{\pi}\cos\left(\frac{\pi l}{L}\right)\frac{\pi}{L}\int_{[p,q]} k^s ds\right. \\
&\quad \left. -\frac{L}{\pi}\cos\left(\frac{\pi l}{L}\right)\frac{\pi l}{L^2}\int_{\gamma_t} k^2 ds\right) \\
&= \frac{1}{\psi_{t_0}}\left\langle \omega_{t_0}, \vec{k}(q,t_0)-\vec{k}(p,t_0)\right\rangle + \frac{D_{t_0}}{\pi\psi_{t_0}^2}\sin\alpha\int_{\gamma_{t_0}} k^2 ds + \frac{D_{t_0}}{\psi_{t_0}^2}\cos\alpha\int_{[p,q]} k^2 ds \\
&\quad - \frac{D_{t_0} l}{\psi_{t_0}^2 L}\cos\alpha\int_{\gamma_{t_0}} k^2 ds
\end{aligned}$$

where we defined

$$\alpha := \frac{l\pi}{L} \qquad \omega_{t_0} := \frac{\gamma(q,t_0)-\gamma(p,t_0)}{D_{t_0}(p,q)} \quad (\text{so } |\omega_{t_0}| = 1)$$

so

$$\begin{aligned}
\left.\frac{\partial}{\partial t}\right|_{t=t_0}\left(\frac{D_t}{\psi_t}\right)(p,q) &= \frac{1}{\psi_{t_0}}\left\langle \omega_{t_0}, \vec{k}(q,t_0)-\vec{k}(p,t_0)\right\rangle + \frac{D_{t_0}}{\psi_{t_0}^2}\cos\alpha\int_{[p,q]} k^2 ds \\
&\quad + \frac{D_{t_0}}{\psi_{t_0}^2\pi}\sin\alpha\left(1-\frac{\alpha}{\tan\alpha}\right)\int_{[p,q]} k^2 ds.
\end{aligned}$$

Notice that since $0<\alpha\le\pi/2$, $1-\alpha/\tan\alpha>0$.
Claim: Let $e_1 := \partial\gamma(p,t_0)/\partial s$ and $e_2 := \partial\gamma(q,t_0)/\partial s$. Then

(i) $\langle\omega_{t_0}, e_1\rangle = \langle\omega_{t_0}, e_1\rangle = \frac{D_{t_0}}{\psi_{t_0}}\cos\alpha$

(ii) If $e_1 = e_2$ then

$$\langle\omega_{t_0}, \vec{k}(q,t_0)-\vec{k}(p,t_0)\rangle \ge 0$$

(iii) If $e_1 \ne e_2$ then

$$\langle\omega_{t_0}, \vec{k}(q,t_0)-\vec{k}(p,t_0)\rangle \ge -\frac{4\pi^2}{L^2}D_{t_0}.$$

Given the claim, the proof now follows because if $e_1 = e_2$ then it immediately follows that

$$\left.\frac{\partial}{\partial t}\right|_{t=t_0}\left(\frac{D_t}{\psi_t}\right)(p,q) > 0.$$

On the other hand if $e_1 \ne e_2$ then

$$\left.\frac{\partial}{\partial t}\right|_{t=t_0}\left(\frac{D_t}{\psi_t}\right)(p,q) \ge \frac{4\pi^2 D_{t_0}}{L^2\psi_{t_0}} + \frac{D_{t_0}\sin\alpha}{\psi_{t_0}^2\pi}\left(1-\frac{\alpha}{\tan\alpha}\right)\int_{\gamma_t} k^2 ds + \frac{D_{t_0}}{\psi_{t_0}^2}\cos\alpha\int_{[p,q]} k^2 ds.$$

Now by Gauss-Bonnet we have

$$2\pi = \int_{\gamma_t} kds \leq L^{1/2}\left(\int_{\gamma_t} k^2 ds\right)^{1/2} \quad \Rightarrow \quad \int_{\gamma_t} k^2 ds \geq \frac{4\pi^2}{L}$$

and moreover if the angle between e_1 and e_2 is β then

$$\beta^2 = \left(\int_{[p,q]} kds\right)^2 \leq l \int_{[p,q]} k^2 ds$$

and so

$$\int_{[p,q]} k^2 ds \geq \frac{\beta}{l}.$$

By (i) of the claim we have

$$\cos\left(\frac{\beta}{2}\right) = \frac{D_{t_0}}{\psi_{t_0}} \cos\alpha < \cos\alpha$$

and so $\beta/2 > \alpha$ or equivalently $\beta^2 > 4\alpha^2$. Therefore

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{D_t}{\psi_t}\right)\bigg|_{t=t_0}(p,q) &> \frac{-4\pi^2 D_{t_0}}{L^2\psi_{t_0}} + \frac{D_{t_0}}{\psi_{t_0}^2 \pi}\sin\alpha\left(1 - \frac{\alpha}{\tan\alpha}\right)\frac{4\pi^2}{L} + \frac{D_{t_0}}{\psi_{t_0}^2}\cos\alpha\frac{4\alpha^2}{L} \\
&= -\frac{4\pi^2 D_{t_0}}{L^2\psi_{t_0}} + \frac{D_{t_0}}{\psi_{t_0}^2 \pi}\sin\alpha\frac{\pi^2}{L} - \frac{4\pi^2}{L}\frac{D_{t_0}}{\psi_{t_0}^2}\alpha\cos\alpha + \frac{4D_{t_0}\alpha^2\cos\alpha}{\psi_{t_0}^2 l} \\
&= \frac{4D_{t_0}}{\psi_{t_0}^2}\alpha\cos\alpha\left(\frac{\alpha}{l} - \frac{\pi}{L}\right) = 0
\end{aligned}$$

It now only remains to check the claim. **Proof of claim:** Let σ_1 and σ_2 be values of arclength corresponding to the points p and q and assume without loss of generality that $\sigma_2 > \sigma_1$. Then we define (in a neighbourhood of (σ_1, σ_2) the function

$$F(s_1, s_2) := \frac{D_{t_0}(s_1, s_2)}{\psi_{t_0}(s_1, s_2)}.$$

We have abused notation slightly here, since D_{t_0} and ψ_{t_0} were not originally defined as functions of arclength, however we simply do the natural thing and identify s_i with the corresponding points p and q. Parts (i), (ii) and (iii) of the claim now follows from calculating first and second variation of F at (σ_1, σ_2) and using the fact that F attains a minimum here. Specifically

(i) We have

$$\frac{d}{ds}\bigg|_{s=0} F(\sigma_1 + s, \sigma_2) = 0 = \frac{d}{ds}\bigg|_{s=0} F(\sigma_1, \sigma_2 + s)$$

from which it follows that $\langle \omega_{t_0}, e_1\rangle = \langle \omega_{t_0}, e_2\rangle = \frac{D_{t_0}}{\psi_{t_0}}\cos\alpha$

(ii) If $e_1 = e_2$ then we compute second variation in the direction (e_1, e_2) to get

$$\frac{d^2}{ds^2}\bigg|_{s=0} F(\sigma_1 + s, \sigma_2 + s) \geq 0 \quad \Rightarrow \quad \langle \omega_{t_0}, \vec{k}(\sigma_2, t_0) - \vec{k}(\sigma_1, t_0)\rangle \geq 0.$$

(iii) If on the other hand $e_1 \neq e_2$ then computing second variation in the direction $(e_1, -e_2)$

$$\left.\frac{d^2}{ds^2}\right|_{s=0} F(\sigma_1 + s, \sigma_2 - s) \geq 0 \quad \Rightarrow \quad \langle \omega_{t_0}, \vec{k}(\sigma_2, t_0) - \vec{k}(\sigma_1, t_0) \rangle \geq \frac{-4\pi^2 D_{t_0}}{L^2} \tag{1.3}$$

□

We will return to the proof of the Gage-Hamilton-Grayson result later in the course after we have developed more machinery. The remaining ingredients are valid in higher dimensions also, so we will now move on to introduce mean curvature flow in general dimensions.

2 Minimal surfaces

In this section we cover some basic facts about minimal surfaces, which are static solutions of mean curvature flow, i.e. critical for area. There are many parallels between the the study of minimal surfaces and mean curvature flow, in particular we will cover the monotonicity formula and Allard regularity, both of which have analogous versions for mean curvature flow.

2.1 Geometry of hypersurfaces in $\mathbb{R}^{n+1}$

This section is based on chapter 7 of [19]. Suppose that $M^n \subset \mathbb{R}^{n+1}$ is a C^k hypersurface. This means that for any $x \in M$, there are open sets U, $V \subset \mathbb{R}^{n+1}$ such that $x \in U$, $0 \in V$ and there is a C^k diffeomorphism $\phi : V \to U$ such that $\phi(0) = x$ and

$$\phi(V \cap \{x^{n+1} = 0\}) = M \cap U.$$

We call ϕ a local representation of M at x. The tangent space to M at $y \in M$ is defined

$$T_yM := \{\gamma'(0) \mid \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^{n+1} \text{ for some } \varepsilon > 0 \text{ is } C^1,\ \gamma((-\varepsilon, \varepsilon)) \subset M \ \ \gamma(0) = y\}.$$

One can check that given any local representation ϕ of M at y, the set

$$\left\{\frac{\partial \phi}{\partial x^1}(0), \dots, \frac{\partial \phi}{\partial x^n}(0)\right\}$$

forms a basis for T_yM.
Given a function $f : M \to \mathbb{R}^N$, we say f is C^l if there is an open set $U \subset \mathbb{R}^{n+1}$ such that $M \subset U$, and a C^l function $\bar{f} : U \to \mathbb{R}^N$ such that $\bar{f}|_M = f$.
Given a tangent vector $\tau \in T_yM$ and a C^1 function $f : M \to \mathbb{R}^N$ we define the directional derivative of f at y in the direction τ by

$$D_\tau f(y) = \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)) \in \mathbb{R}^N$$

where γ is any C^1 map $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^{n+1}$ such that $\gamma((-\varepsilon, \varepsilon)) \subset M$, $\gamma(0) = y$ and $\gamma'(0) = \tau$. One can easily verify that this definition is indeed independent of the choice of γ. The directional derivatives induce a linear map, the differential. Indeed given a C^1 function $f : M \to \mathbb{R}^N$ and $y \in M$ we define

$$df_y : T_yM \to \mathbb{R}^N : df_y(\tau) = D_\tau f(y).$$

2.2 Metric concepts

The following concepts depend on the choice of metric, i.e. the inner product we choose on our tangent spaces. We will always simply consider the standard Euclidean inner product on the ambient space restricted to the tangent planes. Given $f : M \to \mathbb{R}$ a C^1 map, we define the gradient of f at $y \in M$ by

$$\nabla^M f(y) := \sum_{i=1}^{n} D_{t_i} f(y)\tau_i \in T_y M,$$

where $\{\tau_1, \ldots, \tau_n\}$ is any orthonormal basis for $T_y M$. This is simply the projection of the ambient gradient of any extension of f at y projected onto the tangent space, i.e.

$$\nabla^M f(y) = (D\bar{f}(y))^T = D\bar{f}(y) - (D\bar{f}(y) \cdot \nu)\nu.$$

In a similar manner, given a C^1 vectorfield $X : M \to \mathbb{R}^{n+1}$ (which is not necessarily tangential) we define the divergence of X at y by

$$\mathrm{div}_M X(y) := \sum_{i=1}^{n} \tau_i \cdot D_{\tau_i} X(y),$$

where $\{\tau_1, \ldots, \tau_n\}$ is any orthonormal basis for $T_y M$. Equivalently, if $\{e_1, \ldots, e_{n+1}\}$ is an orthonormal basis for $\mathbb{R}^{n+1}$, then

$$\mathrm{div}_M X(y) = \sum_{i=1}^{n+1} e_j \cdot \nabla^M X^j(y),$$

where $X^j - X \cdot e_j$.

Theorem 2.1. *If $X : M \to \mathbb{R}^{n+1}$ is a C^1 tangential vectorfield (i.e. $X(y) \in T_y M$ for every $y \in M$) and if the closure $\overline{M}$ of M is a hypersurface with boundary, then*

$$\int_M \mathrm{div}_M X d\mathcal{H}^n = -\int_{\partial M} X \cdot \eta d\mathcal{H}^{n-1},$$

where η is the inward pointing conormal to ∂M (i.e. normal to ∂M, and tangential to M).

2.3 Curvature

We assume now that M is at least C^2. We define the second fundamental form at $y \in M$ as follows

$$B_y : T_y M \times T_y M \to (T_y M)^\perp : B_y(\tau, \eta) := -(\tau \cdot D_\eta \nu)\nu,$$

where ν is a choice of unit normal. Notice that since ν appears twice, B_y is independent of the choice of orientation. The second fundamental form measures normal curvature relative to the ambient (in our case Euclidean) space. This is illustrated by the fact that $B_y(\tau, \tau) = (\gamma''(0))^\perp = (\gamma''(0) \cdot \nu)\nu$ for any C^2 curve $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = y$, $\gamma'(0) = \tau$. Indeed this follows since $\nu(\gamma(t)) \cdot \gamma'(t) \equiv 0$, so differentiating both sides with respect to t at 0 yields

$$D_\tau \nu(y) \cdot \tau + \nu(y) \cdot \gamma''(0) = 0$$

which implies

$$B_y(\tau,\tau)\cdot\nu = -D_\tau\nu(y)\cdot\tau = \nu\cdot\gamma''(0).$$

More generally if $\tau,\eta\in T_yM$, $\phi: V\to M$ where $V\subset\mathbb{R}^2$ is open and $0\in V$, $\partial\phi(0)/\partial x^1=\tau$, and $\partial\phi(0)/\partial x^2=\eta$ and $\phi(0)=y$, then it follows that

$$B_y(\tau,\eta) = \left(\frac{\partial^2\phi}{\partial x^1\partial x^2}(0)\right)^\perp.$$

Notice it immediately follows from this that B_y is symmetric.
We now define the mean curvature to be the trace of the second fundamental form. That is for each $y\in M$ we define

$$\vec{H}(y) := \text{trace}B_y = \sum_{i=1}^n B_y(\tau_i,\tau_i) = -\sum_{i=1}^n(\tau_i\cdot D_{\tau_i}\nu)\nu = -(\text{div}_M\nu)\nu.$$

If $X: M\to\mathbb{R}^{n+1}$ is a C^1 vectorfield (not necessarily tangential), then we can decompose X into tangential and normal parts $X(y)=X^T(y)+X^\perp(y)$ (note that $X^\perp(y)=(X(y)\cdot\nu)\nu$ and so

$$\text{div}_M X^\perp = (\nabla^M(X\cdot\nu))\cdot\nu + (X\cdot\nu)\text{div}_M\nu = -X\cdot\vec{H},$$

where the first term vanishes because the gradient is tangential. Consequently, using the divergence theorem and the above calculation we have

$$\int_M \text{div}_M X d\mathcal{H}^n = -\int_M X\cdot\vec{H}d\mathcal{H}^n - \int_{\partial M}(X\cdot\eta)d\mathcal{H}^{n-1}.$$

In particular if X is compactly supported away from ∂M we simple have

$$\int_M \text{div}_M X d\mathcal{H}^n = -\int_M X\cdot\vec{H}d\mathcal{H}^n.$$

2.4 First variation formula

Consider $M\subset U$ with $\partial M\cap U=\emptyset$ and a vectorfield $X\in C_c^1(U;\mathbb{R}^{n+1})$. Let

$$\phi_t: U\to\mathbb{R}^{n+1}\qquad \phi_t(x)=x+tX(x),$$

which is one to one for small t since X is compactly supported and C^1. Then we have the following theorem.

Theorem 2.2 (First variation formula)**.** *With the above set-up*

$$\left.\frac{d}{dt}\right|_{t=0}\mathcal{H}^n(\phi_t(M)) = \int_M \text{div}_M X d\mathcal{H}^n = -\int_M X\cdot\vec{H}d\mathcal{H}^n$$

Remark. Notice that this formula in particular implies that mean curvature flow is like the L^2-gradient flow for area, since by the first variation formula

$$\left.\frac{d}{dt}\right|_{t=0}\mathcal{H}^n(\phi_t(M)) \geq -\left(\int_M |X|^2 d\mathcal{H}^n\right)^{1/2}\left(\int_M |H|^2 d\mathcal{H}^n\right)^{1/2}$$

with equality if and only if X is a scalar multiple of $\vec{H}$. Hence to decrease area as quickly as possible we flow in the direction of mean curvature.

Proof of first variation formula. We denote $M_t := \phi_t(M)$. We wish to show that

$$\delta M(X) := \frac{d}{dt}\bigg|_{t=0} \mathcal{H}^n(M_t \cap \mathrm{spt} X) = \int_M \mathrm{div}_M X d\mathcal{H}^n.$$

By the area formula (see for example [8]) we have

$$\mathcal{H}^n(M_t \cap \mathrm{spt} X) = \int_M J\psi_t d\mathcal{H}^n$$

where $\psi_t = \phi_t|_M$ and $J\psi_t = \sqrt{\det(d\psi_t^* \circ d\psi_t)}$ is the Jacobian. Now

$$d\psi_t|_x(\tau) = D_\tau \psi_t = \tau + tD\tau X.$$

We fix an orthonormal basis $\{\tau_1, \dots, \tau_n\}$ of $T_x M$ and an orthonormal basis $\{e_1, \dots, e_{n+1}\}$ for $\mathbb{R}^{n+1}$, then $d\psi_t|_x$ has matrix representation

$$a_{ij} := e_i \cdot \tau_j + tD_{\tau_j} X^i$$

for $1 \leq i \leq n+1$ and $1 \leq j \leq n$. Therefore $d\psi_t|_x^* \circ d\psi_t|_x$ has the matrix

$$\begin{aligned} b_{ij} &= \int_{l=1}^{n+1} a_{li} a_{lj} = \sum_{l=1}^{n+1} (e_l \cdot \tau_i + tD_{\tau_i} X^l)(e_l \cdot \tau_j + tD_{\tau_j} X^l \\ &= \delta_{ij} + t \underbrace{(\tau_j \cdot D_{\tau_i} X + \tau_i \cdot D_{\tau_j} X)}_{=:p_{ij}} + t^2 \underbrace{(D_{\tau_i} X \cdot D_{\tau_j} X)}_{=:q_{ij}}. \end{aligned}$$

Taylor expanding the determinant,

$$\det(I + tP + t^2 Q) = 1 + t\mathrm{trace}P + t^2 \left(\mathrm{trace}Q + \frac{1}{2}(\mathrm{trace}P)^2 - \frac{1}{2}\mathrm{trace}(P^2) \right) + O(t^3).$$

Moreover we have that

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + O(x^3),$$

hence

$$J\psi_t = 1 + \frac{1}{2}t\mathrm{trace}P + \frac{1}{2}t^2 \left(\mathrm{trace}Q + \frac{1}{2}(\mathrm{trace}P)^2 - \frac{1}{2}\mathrm{trace}(P^2) \right) - \frac{1}{8}t^2(\mathrm{trace}P)^2 + O(t^3).$$

Notice that $\mathrm{trace}P = 2\mathrm{div}_M X$ and so it immediately follows from differentiating under the integral that

$$\delta M(X) = \int_M \mathrm{div}_M X d\mathcal{H}^n.$$

□

We can also now compute the second variation under the following additional assumptions

(1) M is orientable, and

(2) X is a normal vectorfield.

Indeed we can then write $X(x) = \xi(x)\nu_x$ for some $\xi \in C^1_c(M;\mathbb{R})$, where ν is a continuous choice of unit normal. Then with $X = \xi\nu$ we have

$$\text{trace}P = 2\text{div}_M(\xi\nu) = 2\xi\text{div}_M\nu = -2\xi(\vec{H}\cdot\nu),$$

also

$$\text{trace}Q = \sum_{i=1}^n D_{\tau_i}(\xi\nu)\cdot D_{\tau_j}(\xi\nu) = \sum_{i=1}^n \xi^2|D_{\tau_i}\nu|^2 + (D_{\tau_i}\xi)^2 = \xi^2|B_x|^2 + |\nabla^M\xi|^2,$$

and finally

$$\begin{aligned}\text{trace}(P^2) = \sum_{i,l=1}^n p_{il}p_{li} = \sum_{i,l=1}^n(\tau_l\cdot D_{\tau_i}(\xi\nu) + \tau_i\cdot D_{\tau_l}(\xi\nu))^2 \\ &= 4\sum_{i,l=1}^n \xi^2 h_{li}^2 \\ &= 4\xi^2|B_x|^2.\end{aligned}$$

Therefore we can calculate the second variation, with $X = \xi\nu$ as

$$\begin{aligned}\delta^2 M(X,X) :&= \frac{d^2}{dt^2}\bigg|_{t=0}\mathcal{H}^n(M_t\cap\text{spt}X) \\ &= \int_M \xi|B|^2 + |\nabla\xi|^2 + 2\xi^2H^2 - \xi^2|B|^2 - H^2\xi^2 d\mathcal{H}^n \\ &= 2\int_M |\nabla\xi|^2 - \xi^2|B|^2 + \xi^2H^2 d\mathcal{H}^n.\end{aligned}$$

Definition. We say that M is stationary in U if $\delta M(X) = 0$ for all $X \in C^1_c(M;\mathbb{R}^{n+1})$. We say that M is stable in U if $\delta^2 M(X,X) \geq 0$ for all $X \in C^1_c(U;\mathbb{R}^{n+1})$.

Suppose that M is stationary in some open set $U \subset \mathbb{R}^{n+1}$. Then let $X = \xi e_j$ where $\xi \in C^1_c(U)$, and e_j is the jth basis vector for $\mathbb{R}^{n+1}$. Then by the first variation formula

$$\int_M \text{div}_M(\xi e_j)d\mathcal{H}^n = \int_M \nabla^M\xi\cdot e_j d\mathcal{H}^n = \int_M \nabla^M\xi\cdot\nabla^M x^j d\mathcal{H}^n = 0$$

for each $j = 1,\ldots,n+1$. This implies $\Delta_M x^j = 0$ for each $j = 1,\ldots,n+1$. A similar argument establishes that, without the assumption of stationarity, we always have $\vec{H}(x) = \Delta_m x$.

Corollary 2.3. *There are no compact, stationary hypersurfaces (in fact no compact stationary submanifolds) in* $\mathbb{R}^{n+1}$.

Proof. Apply the maximum principle to the coordinated functions restricted to M to conclude that each coordinate function is constant. Alternatively, since M is compact we can let $\xi = x^j$ in the formula

$$\int_M \nabla^M x^j\cdot\nabla^M\xi d\mathcal{H}^n$$

for each $j = 1,\ldots,n+1$, and again we get a contradiction. □

2.5 Monotonicity formula

The monotonicity formula is a very powerful consequence of the first variation formula.

Theorem 2.4 (Monotonicity formula)**.** *Suppose that $M^n \subset U$ where $U \subset \mathbb{R}^{n+1}$ is open (more generally we can let $U \subset \mathbb{R}^{n+k}$). Suppose that M is stationary in U, then for all $y \in U$ (note we don't impose $y \in M$) and for $0 < \sigma < \rho < \operatorname{dist}(y, \partial U)$ we have*

$$\frac{\mathcal{H}^n(M \cap B_\rho(y)}{\rho^n} - \frac{\mathcal{H}^n(M \cap B_\sigma(y)}{\sigma^n} = \int_{M\cap(B_\rho(y)\setminus B_\sigma(y))} \frac{|(x-y)^\perp|^2}{|x-y|^{n+2}} d\mathcal{H}^n(x). \tag{2.1}$$

Here $B_\rho(y)$ and $B_\sigma(y)$ are ambient balls in $\mathbb{R}^{n+1}$ (or $\mathbb{R}^{n+k}$), $(x-y)^\perp$ is the projection of $(x-y)$ onto $(T_xM)^\perp$.

Remark. In particular we have that for all fixed y, the mass ratios $\rho^{-n}\mathcal{H}^n(M \cap B_\rho(y))$ are monotone, so we can define the density at y as

$$\Theta_M(y) := \lim_{\rho\to 0} \frac{\mathcal{H}^n(M \cap B_\rho(y))}{\omega_n \rho^n}.$$

Passing to the limit $\sigma \searrow 0$ in (2.1) we have

$$\frac{\mathcal{H}^n(M \cap B_\rho(y)}{\rho^n} - \Theta_M(y) = \int_{M\cap B_\rho(y)} \frac{|(x-y)^\perp|^2}{|x-y|^{n+2}} d\mathcal{H}^n(x).$$

Proof of the monotonicity formula. Fix $y \in U$, $0 < \sigma < \rho < \operatorname{dist}(y, \partial U)$ and let $X(x) := (x-y)\eta(x)$ where $\eta \in C_c^1(U;\mathbb{R})$. Then

$$\begin{aligned}
\operatorname{div}_M X &= \eta(x) \sum_{j=1}^{n+1} e_j \cdot (p_{T_xM} e_j) + (x-y)\cdot \nabla^M \eta(x) \\
&= n\eta(x) + (x-y)\cdot\nabla^M\eta(x).
\end{aligned}$$

Thus, by stationarity and the first variation formula we have

$$n\int_M \eta(x) d\mathcal{H}^n + \int_M (x-y)\cdot\nabla^M\eta(x) d\mathcal{H}^n = 0.$$

We now fix $\tau \in (\sigma, \rho)$, and $\delta > 0$ such that $\tau + \delta < \rho$. We take $\gamma : [0,\infty) \to [0,\infty)$ to be a smooth approximation function with $\gamma(t) \equiv 1$ for $t \le \tau$, $\gamma(t) \equiv 0$ for $t \ge \tau+\delta$, and γ linear on $[\tau, \tau+\delta]$. For simplicity we will pretend we can choose precisely this γ, but a simple approximation argument will validate this choice. We then choose $\eta(x) := \gamma(|x-y|)$ in the above formula and get

$$\begin{aligned}
0 &= n\int_M \eta(x) d\mathcal{H}^n + \int_{M\cap(B_{\tau+\delta}\setminus B_\tau)} (x-y)\cdot\left(\gamma'(|x-y|)\frac{p_{T_xM}(x-y)}{|x-y|}\right) d\mathcal{H}^n \\
&= n\int_M \eta(x) d\mathcal{H}^n - \frac{1}{\delta}\int_{M\cap(B_{\tau+\delta}\setminus B_\tau)} \frac{|p_{T_xM}(x-y)|^2}{|x-y|} d\mathcal{H}^n.
\end{aligned}$$

Let $\delta \to 0$, then

$$n\mathcal{H}^n(M \cap B_\tau(y)) - \frac{d}{d\tau}\left(\int_{M\cap B_\tau(y)} \frac{|p_{T_xM}(x-y)|^2}{|x-y|} d\mathcal{H}^n\right) = 0$$

which implies, by the coarea formula, that

$$\begin{aligned} 0 &= n\mathcal{H}^n(M \cap B_\tau(y)) - \int_{M\cap\partial B_\tau(y)} \frac{|p_{T_xM}(x-y)|^2}{|x-y|} d\mathcal{H}^{n-1} \\ &= n\mathcal{H}^n(M \cap B_\tau(y)) - \int_{M\cap\partial B_\tau(y)} \tau - \frac{|(x-y)^\perp|^2}{|x-y|} d\mathcal{H}^{n-1}. \end{aligned}$$

Rearranging (and dividing through by τ) we find

$$\frac{d}{d\tau}\left(\frac{\mathcal{H}^n(M\cap B_\tau(y))}{\tau^n}\right) = -\frac{n}{\tau}\mathcal{H}^n(M\cap B_\tau(y)) + \frac{d}{d\tau}\mathcal{H}^n(M\cap B_\tau(y)) = \int_{M\cap\partial B_\tau(y)} \frac{|(x-y)^\perp|^2}{|x-y|^2}$$

Integrating over $[\sigma, \rho]$ will evidently yield the desired result. □

An immediate consequence is the following.

Corollary 2.5. *The density $\Theta_M(\cdot)$ is upper semi-continuous on U, i.e., if $y_j, y \in U$ and $y_j \to y$ then*

$$\Theta_M(y) \geq \limsup_{j\to\infty} \Theta_M(y_j)$$

Proof. Let $\varepsilon > 0$, then for small ρ we have

$$\begin{aligned} \Theta_M(y) + \varepsilon &\geq \frac{\mathcal{H}^n(M\cap B_\rho(y))}{\omega_n\rho^n} \geq \frac{\mathcal{H}^n(M\cap B_{\rho-|y_j-y|}(y_j)}{\omega_n\rho^n} \\ &= \frac{(\rho-|y_j-y|)^n}{\rho^n}\frac{\mathcal{H}^n(M\cap B_{\rho-|y_j-y|}(y_j))}{\omega_n(\rho-|y_j-y|)^n} \\ &\geq \left(1-\frac{|y_j-y|}{\rho}\right)^n \Theta_M(y_j). \end{aligned}$$

The result now follows. □

We will from now on assume that M has no removable singularities in U, i.e. if $y \in U \cap \overline{M}$ and there exists $\sigma > 0$ such that $\overline{M} \cap \overline{B}_\sigma(y)$ is a smooth, compact, n-dimensional manifold with boundary contained in $\overline{B}_\sigma(y)$, then $y \in M$. we then define the singular set of M as

$$\mathrm{sing}M := (\overline{M} \setminus M) \cap U.$$

Corollary 2.6. *If M is stationary in U, then $\mathcal{H}^n(\mathrm{sing}M) = 0$.*

Proof. Note that $\Theta_M(y) = 1$ for all $y \in \mathrm{reg}M$, and hence by upper semi-continuity we have $\Theta_M(y) \geq 1$ everywhere. By some measure theory nonsense we know that $\mathcal{H}^n$-almost every $y \in U \setminus M$ satisfies

$$\Theta^*_M(y) = \Theta_M(y) = 0,$$

and so the claim follows. □

2.6 Allard's Theorem

Let us quickly recap what we know so far. We assume that $M^n \subset U$ is a C^1 submanifold of $U \subset \mathbb{R}^{n+k}$ open. We furthermore assume that M has no removable singularities and is stationary in U, which is to say that

$$\int_M \mathrm{div}_M X d\mathcal{H}^n = 0$$

for every $X \in C_c^1(U, \mathbb{R}^{n+k})$. It then follows that

(i) $\mathcal{H}^n((\overline{M} \setminus M) \cap U) = 0$,

(ii) M has no manifold boundary in U (which follows from the first variation formula and the divergence theorem),

(iii) M is C^2 (in fact smooth, even analytic) and $\vec{H} = 0$ on M.

There are no known examples showing that (i) is sharp. Indeed all known examples satisfy

$$\dim_{\mathcal{H}}((\overline{M} \setminus M) \cap U) \leq n - 1$$

or equivalently

$$\mathcal{H}^{n-1+\delta}((\overline{M} \setminus M) \cap U) = 0$$

for every $\delta > 0$. The following theorem due to Allard is also a consequence of the monotonicity formula, albeit a highly non-trivial one.

Theorem 2.7 (Allard [2] [19]).

(i) There is a fixed $\varepsilon = \varepsilon(n, k) > 0$ such that if $y \in \mathrm{sing}M$, then $\Theta_M(y) \geq 1 + \varepsilon$. In other words, in order for a singularity to form, there has to be some sort of concentration of area.

(ii) Suppose that $U = B_1^{n+k}$. There is a $\sigma = \sigma(n, k) \in (0, 1)$ such that if $0 \in \overline{M}$, $\mathcal{H}^n(M \cap B_1^{n+k})/\omega_n < 1 + \varepsilon$, then $\mathrm{sing}M \cap B_\sigma^{n+k} = \emptyset$ (so $M \cap B_\sigma^{n+k}$ is C^∞), and

$$\sup_{M \cap B_\sigma^{n+k}} |B| \leq C,$$

where $C = C(n, k) < \infty$.

Remark. Scaling and translating (ii) implies the following. Suppose that $U = B_R^{n+k}(z)$ and $M \subset U$ is stationary with $z \in \overline{M}$. If $\mathcal{H}^n(M \cap B_R^{n+k}(z))/(\omega_n R^n) < 1 + \varepsilon$, then

$$\sup_{M \cap B_{\sigma R}^{n+k}(z)} |B| \leq \frac{C}{R},$$

where ε, σ and C are fixed constants independent of M and R.

We list some examples of singular stationary hypersurfaces. The simplest example is that of three lines in the plane meeting at angles of 120°. More generally one can consider any number of lines meeting at a point, provided the unit direction vectors pointing away from the meeting

point sum to 0.
Moving to higher dimensions, if we take the Clifford torus

$$\Sigma := \frac{1}{\sqrt{2}} S^1 \times \frac{1}{\sqrt{2}} S^1 \subset S^3 \subset \mathbb{R}^4,$$

then the cone $C(\Sigma)$ over Σ

$$\begin{aligned} C(\Sigma) &= \{\lambda x | \lambda > 0 \quad x \in \Sigma\} \\ &= \{x \in \mathbb{R}^4 | x_1^2 + x_2^2 = x_3^2 + x_4^2\} \end{aligned}$$

is stationary in $\mathbb{R}^4$, with an isolated singularity at the origin. More generally we can define

$$\Sigma_{pq} := \sqrt{\frac{p}{p+q}} S^p \times \sqrt{\frac{q}{p+q}} S^q \subset S^{p+q+1} \subset \mathbb{R}^{p+q+2}.$$

Then $C(\Sigma_{pq})$ is a stationary cone of dimension $p+q+1$ in $\mathbb{R}^{p+q+2}$. If p and q are large enough, then this cones are in fact stable, and some of them are even locally area minimising! (That is, roughly speaking, to say that any compact piece of the cone has smaller area than anything else with the same boundary).

Theorem 2.8 (Simons [20])**.** *If $2 \leq n \leq 6$ and M is a hypersurface in $\mathbb{R}^{n+1}$ with $\overline{M} \setminus M \subset \{0\}$, and if M is stationary and stable in $\mathbb{R}^{n+1}$, and conical (i.e. invariant under rescalings, so if $\eta_\lambda(x) := x/\lambda$ then $\eta_\lambda(M) = M$); then M is a hyperplane.*

Remark. Note that the theorem is false in the case $n = 1$, as we can consider two transverse lines crossing at the origin. The theorem is also false if $n = 7$, as we can take

$$M = C\left(\frac{1}{\sqrt{2}} S^3 \times \frac{1}{\sqrt{2}} S^3\right)$$

We can check that M is stable. Indeed one can verify by direct computation that $\vec{H}(x) = 0$ and that $|B|(x) = 6|x|^{-2}$ for each $x \in M$. Then of course we have that

$$\int_M \mathrm{div}_M X d\mathcal{H}^n = 0$$

for every $X \in C_c^1(\mathbb{R}^{n+1} \setminus \{0\})$. Furthermore if we let $X(x) := \xi(x)x/|x|^2$ where $\xi \in C_c^1(M)$ then we have that

$$\mathrm{div}_M X = \frac{n\xi(x)}{|x|^2} + x \cdot \left(\frac{\nabla \xi(x)}{|x|^2} - \frac{2\xi(x) p_{T_x M}(x)}{|x|^4}\right).$$

But on a cone we have $x^\perp \equiv 0$ so $p_{T_x M}(x) \equiv x$, and so

$$\mathrm{div}_M X = \frac{(n-2)\xi(x)}{|x|^2} + \frac{x \cdot \nabla \xi(x)}{|x|^2},$$

so by minimality

$$(n-2)\int_M \frac{\xi(x)}{|x|^2} d\mathcal{H}^n = -\int_M \frac{x \cdot \nabla \xi(x)}{|x|^2} d\mathcal{H}^n.$$

If we now replace ξ with ξ^2 we find

$$(n-2)\int_M \frac{\xi^2(x)}{|x|^2}d\mathcal{H}^n = -2\int_M \frac{\xi(x)x\cdot\nabla\xi(x)}{|x|^2}d\mathcal{H}^n \le 2\left(\int_M \frac{\xi^2(x)}{|x|^2}d\mathcal{H}^n\right)^{1/2}\left(\int_M |\nabla\xi|^2 d\mathcal{H}^n\right)^{1/2}$$

and so

$$\frac{(n-2)^2}{4}\int_M \frac{\xi^2(x)}{|x|^2}d\mathcal{H}^n \le \int_M |\nabla\xi|^2 d\mathcal{H}^n$$

for every $\xi \in C_c^1(M)$. So if $|B|^2(x) \le (n-2)^2/(4|x|^2)$ then we recover the stability inequality

$$\int_M |B|^2\xi^2 d\mathcal{H}^n \le \int_M |\nabla\xi|^2 d\mathcal{H}^n$$

3 Minimal Graphs

We now restrict our attention to minimal graphs in particular. That is to say we have the following set-up. $\Omega \subset \mathbb{R}^n$ is open, $u : \Omega \to \mathbb{R}$ is C^1 (or indeed Lipschitz will suffice). We denote by G the graph of u

$$G = \text{graph}(u) := \{(x, u(x)) | x \in \Omega\}.$$

If G is stationary, that is

$$\int_G \text{div}_G X d\mathcal{H}^n = 0$$

for every $X \in C_c^1(\Omega \times \mathbb{R}; \mathbb{R}^{n+1})$, then by choosing $X(x) := \tilde{\xi}(x)e_{n+1}$ where $\tilde{\xi}(\tilde{x}, x^{n+1}) = \xi(\tilde{x})$ in a neighbourhood of G, we get

$$\sum_{i=1}^n \int_\Omega \frac{D_i u}{\sqrt{1+|Du|^2}} D_i\xi d\mathcal{H}^n = 0,$$

which is the minimal surface equation in weak form. Alternatively, or indeed equivalently, one can derive the same equation by considering compactly supported variations of the graph function itself, i.e. we say the graph of u is stationary if and only if

$$\left.\frac{d}{dt}\right|_{t=0} \text{area}(\text{graph}(u + t\xi)) = 0$$

for any $\xi \in C_c^1(\Omega)$. It turns out that solutions of the minimal surface equation are in fact smooth.

Proposition 3.1. *If $u \in C^1(\Omega)$ solves the minimal surface equation, then $u \in C^\infty(\Omega)$.*

Sketch of Proof. **Step 1.** Since $u \in C^1(\Omega)$, it immediately follows that $u \in W^{1,2}_{loc}(\Omega)$. We can improve this to $u \in W^{2,2}_{loc}(\Omega)$ using difference quotients, and the fact that u is a solution of the MSE. This is by now a standard PDE argument, and can be found in, for example, [7].
Step 2. Differentiate the equation to find an equation satisfied by the derivatives of u. More specifically we choose $\xi \in C_c^2(\Omega)$ and use $D_l\xi$ as a test function for some l. Then

$$0 = \int \frac{D_i u}{\sqrt{1+|Du|^2}} D_i u D_i D_l \xi d\mathcal{H}^n = -\int D_l\left(\frac{D_i u}{\sqrt{1+|Du|^2}}\right) D_i\xi d\mathcal{H}^n.$$

Defining $w = D_l u$ this implies

$$\begin{aligned} 0 &= \int_\Omega \left(\frac{D_i w}{\sqrt{1+|Du|^2}} - \frac{D_i u}{(1+|Du|^2)^{3/2}} D_j u D_j w \right) D_i \xi d\mathcal{H}^n \\ \Rightarrow \quad 0 &= \int_\Omega \frac{1}{\sqrt{1+|Du|^2}} \left(\delta_{ij} - \frac{D_i u D_j u}{1+|Du|^2} \right) D_j w D_i \xi d\mathcal{H}^n \\ &= \int_\Omega a_{ij}(x) D_j w D_i \xi d\mathcal{H}^n, \end{aligned}$$

where we defined

$$a_{ij}(x) = \frac{1}{\sqrt{1+|Du|^2}} \left(\delta_{ij} - \frac{D_i u D_j u}{1+|Du|^2} \right).$$

The matrix a_{ij} is positive definite, indeed

$$a_{ij}(x)\xi^i \xi^j = \frac{1}{\sqrt{1+|Du|^2}} \left(|\xi|^2 - \frac{(\xi \cdot Du)^2}{1+|Du|^2} \right) \geq \frac{|xi|^2}{(1+|Du|^2)^{3/2}}.$$

In fact we see that the equation satisfied by w is uniformly elliptic on compact subsets $\tilde{\Omega}$ of Ω, and the a^{ij} are also locally bounded. DeGiorgi-Nash-Moser theory then tells us that $w \in C^\alpha(\tilde{\Omega})$, i.e. $u \in C^{1,\alpha}(\tilde{\Omega})$

Step 3. Now $w = D_l u$ satisfies a divergence form, uniformly elliptic equation in $\tilde{\Omega}$ with coefficients in $C^\alpha(\tilde{\Omega})$. From Schauder theory we therefore get that $w \in C^{1,\alpha}(\tilde{\Omega})$ so $u \in C^{2,\alpha}(\tilde{\Omega})$.

Step 4. Now we can use Schauder theory for classical solutions and bootstrap to get smoothness. □

Remark. The fact that C^1 solutions to the minimal surface equation are smooth extends to higher codimension, however it is no longer possible to prove it in the same way as outlined above because the DeGiorgi-Nash-Moser part of the proof breaks down for elliptic systems. Instead we can use Allard regularity to replace this step. This is of course very specific to the minimal surface equation, whereas DeGiorgi-Nash-Moser theory in codimension one is very general.
In codimension 1, one case start with a Lipschitz or even BV solution, and still get full regularity. In higher codimension this is no longer the case, as there are Lipschitz solutions to the minimal surface equation that are not smooth.

We want to prove a special case of Allard's regularity theorem, mentioned in the above remark and which was stated in a certain form in a previous section. For the proof we will need the following compactness theorem.

Theorem 3.2. *Suppose that $U \subset \mathbb{R}^{n+k}$ is open, that $M_i \subset U$ are stationary, n-dimensional submanifolds with $\operatorname{sing} M_i \cap U = \emptyset$, $p_i \in M_i$ with $p_i \to p \in U$ and $|B_{M_i}(p_i)| \to \infty$ and such that there exists $\Lambda > 0$ such that*

$$\frac{\mathcal{H}^n(M_i \cap B_\rho^{n+k}(x))}{\omega_n \rho^n} \leq \Lambda$$

for all $x \in M_i \cap U$ and $\rho \in (0, \operatorname{dist}(x, \partial U))$. Then there exist $q_i \in M_i$ with $q_i \to p$ such that $\lambda_i := |B_{M_i}(q_i)| \to \infty$, and if we define $\tilde{M}_i := \lambda_i(M_i - q_i)$ then (after passing to a subsequence) $\tilde{M}_i \to M$

smoothly locally where M is smooth, stationary in $\mathbb{R}^{n+k}$, $|B_M(x)| \leq 1$ for all x, $|B_M(0)| = 1$ and

$$\frac{\mathcal{H}^n(M \cap B_R^{n+k}(0))}{\omega_n R^n} \leq \Lambda,$$

for all $R > 0$.

Remark. $\tilde{M}_i \to M$ smoothly locally means that for all $p \in M$ and $\rho > 0$ such that there exists a smooth function

$$u : (p + T_pM) \cap B_\rho^{n+k}(p) \to T_pM^\perp$$

with $M \cap B_{\rho/2}^{n+k}(p) = (\mathrm{graph} u) \cap B_{\rho/2}^{n+k}(p)$, (note that this is always possible at any $p \in M$ for some $\rho(p) > 0$), we have that for sufficiently large i, there are smooth functions

$$u_i^l : (p + T_pM) \cap B_\rho^{n+k}(p) \to T_pM^\perp$$

for $1 \leq l \leq m$ such that

$$\tilde{M}_i \cap B_{\rho/2}^{n+k}(p) = \bigcup_{l=1}^m \mathrm{graph}(u_i^l) \cap B_{\rho/2}^{n+k}(p),$$

and $u_i^l \to u$ in C^j-norm for all j on $(p + T_pM) \cap B_\rho^{n+k}(p)$.

Example. If we take M to be the catenoid, which is parametrised by $z = \cosh^{-1}(r)$ where $r = \sqrt{x^2 + y^2}$, and $M_i = \mu_i M$ where $\mu_i \searrow 0$ then we get smooth convergence with multiplicity 2 away from the origin, but not at/near the origin. This is no contradiction, as there is no bound on the second fundamental form at the origin.

Proof. Choose $\rho_i \to 0$ such that $\rho_i |B_{M_i}(p_i)| \to \infty$ (e.g. $\rho_i |B_{M_i}(p_i)|^{-1/2}$). Now let $q_i \in M_i \cap B_{\rho_i}^{n+k}(p_i)$ be such that

$$|B_{M_i}(q_i)| \mathrm{dist}(q_i, \partial B_{\rho_i}(p_i)) \geq |B_{M_i}(x) \mathrm{dist}(x, \partial B_{\rho_i}(p_i))$$

for all $x \in B_{\rho_i}(p_i)$. Note that, provided i is large enough, we will have that $B_{\rho_i}(p_i) \subset\subset U$ as $p_i \to p \in U$ and $\rho_i \to 0$. Now let $\sigma_i := \mathrm{dist}(q_i, \partial B_{\rho_i}(p_i))$, then

$$\begin{aligned} |B_{M_i}(q_i)| \sigma_i &\geq |B_{M_i}(x)| \mathrm{dist}(x, \partial B_{\rho_i}(p_i)) \\ &\geq |B_{M_i}(x)| \mathrm{dist}(x, \partial B_{\sigma_i}(q_i)), \end{aligned}$$

for every $x \in M_i \cap B_{\sigma_i}^{n+k}(q_i)$. Therefore

$$\lambda_i = |B_{M_i}(q_i)| \geq |B_{M_i}(p_i)| \frac{\rho_i}{\sigma_i} \to \infty.$$

In particular we have

$$\lambda_i \sigma_i = |B_{M_i}(q_i)| \geq |B_{M_i}(p_i)| \rho_i \to \infty.$$

Let $\tilde{M}_i := \lambda_i(M_i - q_i)$. Then $\tilde{M}_i$ is stationary in $B_{\lambda_i\sigma_i}^{n+k}(0)$ with $\lambda_i\sigma_i \to \infty$. Moreover $0 \in \tilde{M}_i$, and for all $x \in \tilde{M}_i \cap B_{\lambda_i\sigma_i}^{n+k}(0)$ we have

$$|B_{\tilde{M}_i}(x)| = \frac{1}{\lambda_i} \left| B_{M_i}\left(q_i + \frac{x}{\lambda_i}\right) \right| \leq \frac{1}{\lambda_i} \frac{\lambda_i\sigma_i}{\mathrm{dist}(q_i + \lambda_i^{-1}x, \partial B_{\sigma_i}(q_i))} = \frac{\sigma_i}{\sigma_i - \lambda_i^{-1}|x|} \to 1,$$

if x is fixed. Moreover

$$\frac{\mathcal{H}^n(\tilde{M}_i \cap B^{n+k}_{\lambda_i\sigma_i}(0))}{\omega_n(\lambda_i\sigma_i)^n} = \frac{\lambda_i^n \mathcal{H}^n(M_i \cap B_{\sigma_i}(q_i))}{\omega_n(\lambda_i\sigma_i)^n} = \frac{\mathcal{H}^n(M_i \cap B_{\sigma_i}(q_i))}{\omega_n\sigma_i^n} \leq \Lambda$$

so by monotonicity, for all $R > 0$ we have

$$\frac{\mathcal{H}^n(\tilde{M}_i \cap B^{n+k}_R(0))}{\omega_n R^n} \leq \frac{\mathcal{H}^n(\tilde{M}_i \cap B^{n+k}_{\lambda_i\sigma_i}(0))}{\omega_n(\lambda_i\sigma_i)^n} \leq \Lambda.$$

In summary we have for all $R > 0$ and i large

(i) $|B_{\tilde{M}_i}(x)| \leq 2$ for all $x \in \tilde{M}_i \cap B^{n+k}_R(0)$,

(ii) $(\omega_n R^n)^{-1}\mathcal{H}^n(\tilde{M}_i \cap B^{n+k}_R(0)) \leq \Lambda$ and $\tilde{M}_i$ is stationary in $B^{n+k}_R(0)$.

Claim: From (i) and (ii) it follows that there exists a smooth staionard n-dimensional submanifold $M_R \subset B^{n+k}_R(0)$ such that $\tilde{M}_i \to M$ (up to subsequences) smoothly on compact sets in $B^{n+k}_R(0)$. The claim can be formulated as the following. Given $M_i^n \subset U \subset \mathbb{R}^{n+k}$, where U is open; $0 \in U$; each M_i is C^2, stationary and embedded; $\mathrm{sing} M_i \cap U = \emptyset$ and

(i) $\sup_{M_i \cap K} |B_{M_i}| \leq C_K$ for all i.

(ii) $\mathcal{H}^n(M_i \cap K) \leq A_K$ for all i.

Then there exists a smooth, stationary, (embedded in codimension 1) $M \subset U$ such that, after passing to a subsequence, we have $M_i \to M$ smoothly on compact sets. We now prove this reformulated claim.
Fix K compact. By the uniform curvature bound (i), for every $p \in M_i \cap K$ we have

$$\tilde{M}_i \cap B^{n+k}_{\rho_0/2}(p) \subset \mathrm{graph} u_i \subset \tilde{M}_i$$

where $\tilde{M}_i$ is the connected component of M_i containing p, and

$$u_i : (p + T_p M_i) \cap B^{n+k}_{\rho_0}(p) \to (T_p M_i)^\perp.$$

By (ii), only a bounded number of points $\{p_1, \ldots, p_m\}$ are needed to cover all of $M_i \cap K$ by such graphs. Passing to a subsequence we have $p_j + T_{p_j} M_i \to T_j$ as $i \to \infty$ for each j. Write M_j as the union of C^2 graphs v_j^i for $j = 1, \ldots, m$ defined on balls of fixed size in T_j. These v_j^i solve the minimal surface equation, and hence we have uniform C^2 on the v_j^i, so by Arzelà-Ascoli we get subsequential $C^{1,\alpha}$ convergence to some limit $v_j^i \to v_j$ for each j. The v_j are weak solutions to the minimal surface equation, so the regularity theory implies they are in fact smooth. Let $M := \bigcup_{j=1}^m \mathrm{graph} v_j$. One can then show that M is smooth, and in codimension 1 the maximum principle implies that it must also be embedded. □

With the above compactness theorem in hand, we now state the following special case of Allard's regularity theorem.

Theorem 3.3 (A priori curvature estimates for smooth, stationary surfaces with appropriate mass bounds)**.** *Let* $M^n \subset B_R^{n+k}(y)$ *be stationary where* $y \in M$*, with* $\text{sing}M \cap B_R^{n+k}(y) = \emptyset$ *and* M *smooth. There exist* $\varepsilon = \varepsilon(n,k) \in (0,1)$*,* $\sigma = \sigma(n,k) \in (0,1)$ *such that if*

$$\frac{\mathcal{H}^n(M \cap B_R^{n+k}(0))}{\omega_n R^n} \leq 1 + \varepsilon$$

then

$$\sup_{M \cap B_{\sigma R}^{n+k}(y)} |B_M| \leq \frac{C}{R}$$

where $C = C(n,k)$ *independent of* M*.*

Remark. (1) The theorem is true without the assumption that $\text{sing}M = \emptyset$. Then part of the conclusion is that there are no singularities in $B_{\sigma R}^{n+k}(y)$.

(2) It is not possible to relax the mass bound to

$$\frac{\mathcal{H}^n(M \cap B_R^{n+k}(y))}{\omega_n R^n} < 2$$

as there are counter examples, consider a sequence of rescaled catenoids for example.

Proof. Without loss of generality we can assume $y = 0$ and $R = 1$. By the monotonicity formula, for all $z \in M \cap B_{1/2}^{n+k}(0)$ for all $0 < \rho < 1 - |z|$. Then

$$\frac{\mathcal{H}^n(M \cap B_\rho^{n+k}(z))}{\omega_n \rho^n} \leq \frac{\mathcal{H}^n(M \cap B_{1-|z|}^{n+k}(z))}{\omega_n (1-|z|)^n} \leq \frac{1+\varepsilon}{(1-|z|)^n}$$

since $\mathcal{H}^n(M \cap B_1^{n+k}(0)) \leq 1 + \varepsilon$. Thus

$$\frac{\mathcal{H}^n(M \cap B_\rho^{n+k}(z))}{\omega_n \rho^n} \leq 1 + \varepsilon/2$$

for any $z \in B_\delta^{n+k}(0) \cap M$ and $0 < \rho < \text{dist}(z, \partial B_\delta^{n+k}(0))$, where $\delta = \delta(\varepsilon) > 0$ is small.
Suppose for contradiction that with $\sigma = \delta/2$ we have M_i stationary in $B_\delta^{n+k}(0)$, $0 \in M_i$ with

$$\frac{\mathcal{H}^n(M_i \cap B_\rho^{n+k}(z))}{\omega_n \rho^n} \leq 1 + \varepsilon/2$$

for all $z \in B_\delta^{n+k}(0) \cap M_i$ and $0 < \rho < \text{dist}(z, \partial B_\delta^{n+k}(0))$, and points $p_i \in M_i \cap B_{\delta/2}^{n+k}(0)$ with $|B_{M_i}(p_i)| \geq i$. Applying the previous theorem we get a smooth, stationary, embedded (in any codimension because of the mass bound), n-dimensional submanifold M of $\mathbb{R}^{n+k}$ with $B_M \leq 1$, $|B_M(0)| = 1$ and $(\omega_n R^n)^{-1}\mathcal{H}^n(M \cap B_R^{n+k}(0)) \leq 1 + \varepsilon/2$ for all $R > 0$.
We claim that there is no M satisfying these conditions if ε is chosen to be small enough. Indeed, if this were not the case then there exist $\tilde{M}_i$ such that $|B_{\tilde{M}_i}| \leq 1$, $|B_{\tilde{M}_i}(0)| = 1$ and such that

$$\frac{\mathcal{H}^n(\tilde{M}_i \cap B_R^{n+k}(0))}{\omega_n R^n} \leq 1 + 1/2^i,$$

for all i and $R > 0$. Passing to a subsequence using the compactness theory again, we find that there is some smooth, stationary $M' \subset \mathbb{R}^{n+k}$, such that $\tilde{M}_i \to M'$ smoothly, and moreover $|B_{M'}| \leq 1$, $|B_{M'}(0)| = 1$ and

$$\frac{\mathcal{H}^n(M' \cap B_R^{n+k}(0))}{\omega_n R^n} \equiv 1$$

for all $R > 0$. By the monotonicity formula, M' is a cone, hence since it is also smooth, it must be a plane. This however contradicts the fact that $|B_{M'}(0)| = 1$. □

3.1 Regularity of stationary surfaces

Suppose that $M \subset U$ is stationary (where $U \subset \mathbb{R}^{n+k}$ is open, and $\dim(M) = n$). Defining $\mathrm{sing}M := (\overline{M} \setminus M) \cap U$. We know that $\mathcal{H}^n(\mathrm{sing}M) = 0$, but not much else...
We can make natural, stronger hypotheses than just stationarity. For example:
1. M is locally area minimising. Roughly, any compact piece of M has area less than or equal to the area of any surface in U having the same boundary as the boundary of that piece. Under this assumption we can say much more about the regularity:

Theorem 3.4 (DeGiorgi, Reifenberg, Federer-Fleming, Fleming, Almgren, J. Simons, Federer [3,5,9–11,17,21])**.** *In codimension* 1*,* $\dim_{\mathcal{H}}(\mathrm{sing}M) \leq n-7$*. In particular* $\mathrm{sing}M = \emptyset$ *if* $1 \leq n \leq 6$*.*

This theorem is sharp, for example the Simons cone

$$C\left(\frac{1}{\sqrt{2}}S^3 \times \frac{1}{\sqrt{2}}S^3\right) \subset \mathbb{R}^8$$

is locally area minimising and has an isolated singularity. In higher codimension we can still say more than in the stationary case, but more exotic singular structures can arise than in the hypersurface case.

Theorem 3.5 (Almgren [4])**.** *For any* $k \geq 2$*, if* M *is locally area minimising then*

$$\dim_{\mathcal{H}}(\mathrm{sing}M) \leq n - 2 \tag{3.1}$$

This is also sharp as the algebraic curve

$$V = \{z^2 = w^3\} \cap (\mathbb{C} \times \mathbb{C})$$

is locally area-minimising and has an isolated singularity at $(0, 0)$.
2. M is stationary and stable, that is to say that for any $X \in C_c^1(U; \mathbb{R}^{n+k})$ we have, defining $\phi_t(x) := x + tX(x)$, the following

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{H}^n(\phi_t(M)) = \int_M \mathrm{div}_M X \, d\mathcal{H}^n = 0$$
$$\left.\frac{d^2}{d^2t}\right|_{t=0} \mathcal{H}^n(\phi_t(M)) \geq 0.$$

If $k \geq 2$, then no sharp dimension estimate for sing$M$ is known. This is ultimately because in higher codimension it is not clear how to leverage the stability assumption. In codimension one however, we can write the stability assumption as

$$\int_M |B|^2 \xi^2 d\mathcal{H}^n \leq \int_M |\nabla \xi|^2 d\mathcal{H}^n \qquad \forall \xi \in C^1_c(M).$$

From the example of two crossing lines, it seems that $\dim_{\mathcal{H}}(\text{sing}M) \leq n-1$ is the best result we could hope for. The following theorem says that in actual fact, if we rule out such singularities, then we have the same regularity as in the area minimising case.

Theorem 3.6 (Wickramasekera [24]). *Suppose that M is stationary, stable and codimension 1. Assume that M has no 'classical singularities' (we say that a point $y \in \overline{M} \cap U$ is a classical singularity if there exists $\sigma > 0$ such that $\overline{M} \cap B^{n+1}_\sigma(y)$ is a union of smooth n-dimensional manifolds with boundary, meeting only along their common boundary). Then*

$$\dim_{\mathcal{H}}(\text{sing}M) \leq n-7,$$

and hence, in particular, is empty if $n \leq 6$.

Recall that Allard's theorem says that if M is stationary in $U = B^{n+k}_1(0)$, with $0 \in \overline{M}$ and

$$\frac{\mathcal{H}^n(M \cap B^{n+k}_1(0))}{\omega_n} \leq 1 + \varepsilon$$

where $\varepsilon = \varepsilon(n,k) \in (0,1)$, then

(1) $\text{sing}M \cap B^{n+k}_\sigma(0) = \emptyset$, and

(2) $\sup_{M \cap B^{n+k}_\sigma} |B| \leq C$.

For stationary, stable hypersurfaces we have a similar estimate.

Theorem 3.7. *Given $\Lambda > 0$ there exists $\varepsilon = \varepsilon(n,\Lambda) \in (0,1)$ such that if M is stationary, stable in $U = B^{n+1}_1(0)$,*

$$\frac{\mathcal{H}^n(M \cap B^{n+1}_1(0))}{\omega_n} \leq \Lambda, \quad \text{and} \quad \int_{M \cap (B_{1/2} \times \mathbb{R})} |x^{n+1}|^2 d\mathcal{H}^n \leq \varepsilon,$$

then

$$\sup_{M \cap (B_{1/4} \times \mathbb{R})} |B| \leq C = C(n,\Lambda).$$

In fact, $M \cap (B_{1/4} \times \mathbb{R}) = \bigcup_{j=1}^{l'} \text{graph}(u_j)$ where $u_j : B_{1/4} \to \mathbb{R}$ are smooth, solve the minimal surface equation and $u_1 < \cdots < u_{l'}$ where l' is bounded in terms of l.

An earlier version of this result was proved by Schoen-Simon [18], which assumed in place of the 'no classical singularities' assumption, that $\mathcal{H}^{n-2}(\text{sing}M) = 0$.

4 Mean Curvature Flow

We will consider only the flow of embedded hypersurfaces. In particular suppose that M^n is a smooth manifold, and

$$F_0 : M \times [0, T) \to \mathbb{R}^{n+1},$$

is a smooth embedding. That is to say that each $F_0(\cdot)$ is a homeomorphism onto its image, and the differential has full rank at every point. We also will further assume that $M_0 := F_0(M)$ is closed (compact without boundary). A mean curvature flow starting at $M_0 = F_0(M)$ is a smooth one parameter family of embeddings

$$F : M \times [0, T) \to \mathbb{R}^{n+1}$$

which solves the initial value problem

$$\begin{cases} \frac{\partial F}{\partial t}(p, t) &= \vec{H}(p, t) \\ F(p, 0) &= F_0(p) \end{cases}$$

where $\vec{H}$ is the mean curvature vector. We will write $F_t(\cdot) := F(\cdot, t)$ and $M_t := F_t(M)$. Much like in the case of curve shortening flow, we have a short time existence theorem for closed initial hypersurfaces.

Theorem 4.1 (Short-time existence)**.** *Given M compact, there exists a smooth solution F satisfying the above initial value problem for some $T = T(M, F_0) > 0$.*

One of the main tools in the analysis of mean curvature flows is the monotonicity formula of Gerhard Huisken [14]. We will state it shortly, but first we introduce some notation. We write

$$\Phi(x, t) := \frac{1}{(-4\pi t)^{n/2}} \exp\left(\frac{|x|^2}{4t}\right) \qquad t < 0, \ \ x \in \mathbb{R}^n$$

or more generally, centred at a point (x_0, t_0)

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

Notice that $\psi(x, t) := (-4\pi t)^{-1/2}\Phi(x, t)$ solves

$$\left(\frac{\partial}{\partial t} + \Delta_{\mathbb{R}^{n+1}}\right)\psi = 0,$$

(in fact ψ is the backwards heat kernel).

Theorem 4.2 (Monotonicity formula [14])**.** *Suppose that $(M_t)_{t\in[0,T)}$ is a mean curvature flow of compact embedded n-dimensional surfaces, then*

$$\frac{d}{dt}\int_{M_t} \Phi_{(x_0,t_0)}(x, t) d\mathcal{H}^n(x) = -\int_{M_t} \left|\vec{H} - \frac{(x_0 - x)^\perp}{2(t_0 - t)}\right|^2 \Phi_{(x_0,t_0)}(x, t) d\mathcal{H}^n(x)$$

for all $0 \le t < t_0$.

Notice in particular that the right hand side is non-positive. This tells us that the integral on the left hand side is monotone in t. Since the scale of the kernel $\Phi_{(x_0,t_0)}(x,t)$ is shrinking as t increases, ultimately becoming closer and closer to a point mass at x_0 as t approaches t_0, the monotonicity formula gives us control over the rate at which mass can concentrate at a point under the flow. It is the parabolic analogue for mean curvature flow of the monotonicity formula for the mass ratios of minimal surfaces. We will see some applications soon, but first we introduce what is like a 'first variation' for the flow.

Lemma 4.3. *Let $U \subset \mathbb{R}^{n+1}$ be open, then $\phi : U \times [0,T) \to \mathbb{R}$, with $\phi(\cdot,t) \in C_c^2(U)$ and $\frac{\partial \phi}{\partial t}(\cdot,t) \in C_c^0(U)$ then*

$$\frac{d}{dt}\int_{M_t} \phi(x,t)d\mathcal{H}^n(x) = \int_{M_t} \frac{\partial \phi}{\partial t}(x,t) + \vec{H}(x)\cdot D\phi(x,t) - H^2(x)\phi(x,t)d\mathcal{H}^n(x).$$

Proof. Fix $t \in (0,T)$, then

$$\begin{aligned} \frac{d}{dt}\int_{M_t} \phi(x,t)d\mathcal{H}^n(x) &= \frac{d}{ds}\bigg|_{s=0} \int_{M_{t+s}} \phi(x,t+s)d\mathcal{H}^n(x) \\ &= \frac{d}{ds}\bigg|_{s=0} \int_{M_t} \phi(g_s(y),t+s)Jg_s(y)d\mathcal{H}^n(y), \end{aligned}$$

by the area formula, where $g_s(y) = F_{t+s}\circ F_t^{-1}(y)$. Then for $y \in M_t$ we have

$$g_s(y) = y + s\,\frac{d}{ds}\bigg|_{s=0} g_s(y) + O(s^2) = y + s\vec{H}(y) + O(s^2).$$

Moreover,

$$\begin{aligned} \frac{d}{ds}\bigg|_{s=0} Jg_s(y) &= \operatorname{div}_{M_t}\vec{H}(y) = \operatorname{div}_{M_t}((\vec{H}\cdot\nu)\nu) \\ &= \nabla^{M_t}(\vec{H}\cdot\nu)\cdot\nu + (\vec{H}\cdot\nu)\operatorname{div}_{M_t}\nu = -H^2, \end{aligned}$$

which follows because $\vec{H} = -(\operatorname{div}_{M_t}\nu)\nu$, and hence we see that the main result follows also. □

Corollary 4.4. *If (M_t) is a compact mean curvature flow in U then*

$$\frac{d}{dt}\mathcal{H}^n(M_t) = -\int_{M_t} |\vec{H}|^2 d\mathcal{H}^n.$$

Proof. As M_0 is compact, we can choose ϕ in such a way that $\phi \equiv 1$ in a neighbourhood of M_t. □

Corollary 4.5. *If ϕ in the lemma satisfies*

$$\left(\frac{d}{dt} - \Delta_{M_t}\right)\phi \le 0$$

and if M_t is compact then

$$\frac{d}{dt}\int_{M_t} \phi d\mathcal{H}^n \le -\int_{M_t} |\vec{H}|^2\phi d\mathcal{H}^n$$

Proof. By the chain rule we have

$$\left(\frac{d}{dt} - \Delta_{M_t}\right)\phi(x,t) = \frac{\partial \phi}{\partial t} + D\phi \vec{H} - \Delta_{M_t}\phi$$

which implies

$$\frac{\partial \phi}{\partial t} + D\phi \cdot \vec{H} \leq \Delta_{M_t}\phi,$$

and so

$$\frac{d}{dt}\int_{M_t}\phi \leq -\int_{M_t}|\vec{H}|^2\phi + \int_{M_t}\Delta_{M_t}\phi = -\int_{M_t}|\vec{H}|^2\phi d\mathcal{H}^n,$$

by the divergence theorem. □

Suppose that $M_0 \subset B_\rho^{n+1}(0)$ and take

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

Then we can check by direct calculation that

$$\left(\frac{d}{dt} - \Delta_{M_t}\right) \leq 0.$$

Hence

$$\int_{M_t}\phi_\rho(x,t)d\mathcal{H}^n \leq \int_{M_0}\phi_\rho(x,0)d\mathcal{H}^n.$$

Now $M_0 \subset B_\rho^{n+1}(0)$ implies that

$$\int_{M_0}\phi_\rho(x,0)d\mathcal{H}^n = 0,$$

and so by the corollary we have

$$\int_{M_t}\phi_\rho(x,t)d\mathcal{H}^n = 0$$

for all $t \in (0,T)$ which implies

$$M_t \subset B^{n+1}_{\sqrt{\rho^2-2nt}}(0).$$

Consequently we see that any compact mean curvature flow must become singular in finite time. We now prove the monotonicity formula.

Proof of monotonicity formula. We assume without loss of generality that $(x_0,t_0) = (0,0)$. Then by direct computation we have that

$$\frac{\partial \Phi}{\partial t} + \mathrm{div}_{M_t}D\Phi + \frac{|D^\perp\Phi|^2}{\Phi} = 0.$$

Hence we compute

$$\left(\frac{d}{dt}+\Delta_{M_t}\right)\Phi = \frac{\partial\Phi}{\partial t} + D\Phi\cdot\vec{H} + \operatorname{div}_{M_t}(D\Phi-(D\Phi\cdot\nu)\nu)$$
$$= \frac{-|D^{\perp}\Phi|^2}{\phi} + D\Phi\cdot\vec{H} + (D^{\perp}\Phi\cdot\nu)(\vec{H}\cdot\nu)$$
$$= \frac{-|D^{\perp}\Phi|^2}{\phi} + 2D^{\perp}\Phi\cdot\vec{H}$$
$$= -\left|\vec{H}-\frac{D^{\perp}\Phi}{\Phi}\right|^2\Phi + |\vec{H}|^2\Phi.$$

Hence

$$\frac{d}{dt}\int_{M_t}\Phi d\mathcal{H}^n = \int_{M_t}\left(\frac{d}{dt}+\Delta_{M_t}\right)\Phi - |\vec{H}|^2\Phi d\mathcal{H}^n = -\int_{M_t}\left|\vec{H}-\frac{D^{\perp}\Phi}{\Phi}\right|^2\Phi d\mathcal{H}^n$$

□

.

4.1 Local curvature estimates for smooth mean curvature flows

In this section we will discuss the local regularity theorem of Brian White, which serves as an analogue of Allard's theorem for mean curvature flows. To do so we first introduce some notation that will prove convenient in what follows. The notation and terminology is based on White's [23]. We denote spacetime as $\mathbb{R}^{n+1,1} := \mathbb{R}^{n+1}\times\mathbb{R}$. A general point in spacetime will be denoted $X=(x,t)\in\mathbb{R}^{n+1,1}$. We also introduce the parabolic norm

$$\|X\| := \max\{|x|,|t|^{1/2}\}$$

and parabolic dilation for $\lambda>0$

$$\mathcal{D}_\lambda : \mathbb{R}^{n+1,1}\to\mathbb{R}^{n+1,1} : (x,t)\mapsto(\lambda x,\lambda^2 t).$$

This is the natural scaling for parabolic problems, as rescaled solutions are themselves solutions. Moreover, notice that the parabolic norm behaves nicely under this rescaling

$$\|\mathcal{D}_\lambda(X)\| = |\lambda|\|X\|.$$

Under the parabolic norm, open balls take the form

$$B^{n+1,1}_\rho((x_0,t_0)) = B^{n+1}_\rho(x_0)\times(t_0-\rho^2,t_0+\rho^2).$$

We will denote the projection onto the time axis by $\tau:\mathbb{R}^{n+1,1}\to\mathbb{R}$. Finally we will adopt a slightly different view of mean curvature flows, identifying them with their spacetime track, rather than considering them as an evolving hypersurface. More specifically we say that $\mathcal{M}$ is a smooth n-dimensional mean curvature flow in an open set $U \subset \mathbb{R}^{n+1,1}$ if $\mathcal{M} \subset U$ and for all $X_0 = (x_0, t_0) \in \mathcal{M}$ there exists $\rho > 0$ such that

$$\mathcal{M} \cap B_\rho^{n+1,1}(X_0) = \big\{(F(p,t),t)|p \in M \;\; t \in (t_0 - \rho^2, t_0 + \rho^2)\big\}$$

for some n-dimensional smooth manifold M and a smooth map

$$F : M \times (t_0 - \rho^2, t_0 + \rho^2) \to \mathbb{R}^{n+1}$$

such that if $M_t := F_t(M)$, then $\overline{M}_t$ is a compact smooth embedded hypersurface with boundary in $B_\rho^{n+1}(x_0)$ and $\partial \overline{M}_t \subset \partial B_\rho^{n+1}(x_0)$ and

$$\frac{\partial F}{\partial t}(p,t) = \vec{H}(F(p,t)),$$

where $\vec{H}$ is the mean curvature vector of M_t. We write $\mathcal{M}(t) := \{x|(x,t) \in \mathcal{M}\}$. The final ingredient before we can state the local regularity theorem is a suitable norm.

Definition. Suppose that $0 \in \mathcal{M}$. If $\mathcal{M}$ can be rotated to get $\mathcal{M}'$ such that

$$\mathcal{M}' \cap B_1^{n+1,1}(0) \subset \operatorname{graph} u$$

where $u : B^{n,1}, 1(0) \to \mathbb{R}$ and $\|u\|_{C^{2,\alpha}(B_1^{n,1}(0))} \leq 1$ then we say that $K_{2,\alpha}(\mathcal{M}, 0) \leq 1$. More generally, if $X \in \mathcal{M}$ we define

$$K_{2,\alpha}(\mathcal{M}, X) := \inf\{\lambda > 0 | K_{2,\alpha}(\mathcal{D}_\lambda(\mathcal{M} - X), 0) \leq 1\}$$

Notice that according to this definition we have

$$K_{2,\alpha}(\mathcal{D}_\lambda \mathcal{M}, \mathcal{D}_\lambda X) = \lambda^{-1} K_{2,\alpha}(\mathcal{M}, X) \qquad \forall \lambda > 0.$$

Finally we let

$$K_{2,\alpha;U}(\mathcal{M}) := \sup_{X \in U \cap \mathcal{M}} d(X,U) K_{2,alpha}(\mathcal{M}, X).$$

This is scale invariant, which is to say that

$$K_{2,\alpha;\mathcal{D}_\lambda U}(\mathcal{D}_\lambda \mathcal{M}) = K_{2,\alpha;U}(\mathcal{M})$$

In the above definition we used the parabolic $C^{2,\alpha}$ norm which is defined

$$\|u\|_{C^{k,\alpha}(\Omega)} := \Sigma_{j+2l \leq k} \|D^j \partial_t^l u\|_{C^{0,\alpha}(\Omega)},$$

where

$$\|w\|_{C^{0,\alpha}(\Omega)} = \sup_{X \in \Omega} |w| + \sup_{X \neq Y} \frac{|w(X) - w(Y)|}{\|X - Y\|^\alpha}.$$

We can now state the local regularity theorem.

Theorem 4.6 (White [23])**.** *There exists $\varepsilon(n) > 0$ such that the following holds. Let $\mathcal{M}$ be a smooth n-dimensional proper mean curvature flow in some open $U \subset \mathbb{R}^{n+1,1}$. By proper we mean that $\mathcal{M} = \overline{\mathcal{M}} \cap U$. Moreover suppose that for all $X = (x,t) \in U$ and for all $r \in (0, d(X,U))$ we have*

$$\Theta(\mathcal{M}, X, r) := \int_{\mathcal{M}(t-r^2)} \frac{1}{(4\pi r^2)^{n/2}} \exp\left(-\frac{|y-x|^2}{4r^2}\right) d\mathcal{H}^n(y) < 1 + \varepsilon,$$

where $d(X,U) = \inf\{\|X - Y\| | Y \in \mathbb{R}^{n+1,1} \setminus U\}$.

$$K_{2,\alpha;U}(\mathcal{M}) \leq C = C(n,\alpha)$$

Our next goal is to prove the theorem, for which we will employ a contradiction argument and use compactness. For this we will need the following appropriate form of the Arzelà-Ascoli theorem.

Theorem 4.7. *If $\mathcal{M}_i$ are n-dimensional smooth mean curvature flows in $U \subset \mathbb{R}^{n+1,1}$ open, such that for every $U' \subset\subset U$ we have*

$$\sup_{U'} K_{2,\alpha}(\mathcal{M}_i, X) \leq C(U') \qquad \forall i,$$

Assume without loss of generality that $0 \in \mathcal{M}_i$ for all i. Then there exists a smooth mean curvature flow $\mathcal{M}$ in U with $0 \in \mathcal{M}$ such that (up to subsequences) we have

$$\mathcal{M}_i \to \mathcal{M}$$

locally in C^2.

Proof of Theorem 4.6. Suppose that the statement fails, then for all $i = 1, 2, \ldots$ there are $U_i \subset \mathbb{R}^{n+1,1}$ open and $\mathcal{M}_i \subset U_i$ proper smooth mean curvature flows with

$$\Theta(\mathcal{M}_i, X, R) < 1 + \frac{1}{i}$$

for all $X \in \mathcal{M}_i$, and $0 < r < d(X, U_i)$, but with

$$K_{2,\alpha;U_i}(\mathcal{M}_i) \to \infty. \tag{4.1}$$

By compactly exhausting the U_i if necessary we can assume without loss of generality that

$$K_{2,\alpha;U_i}(\mathcal{M}_i) =: s_i < \infty$$

and that $s_i \to \infty$. Pick $X_i \in \mathcal{M}_\rangle$ such that

$$d(X_i, U_i) K_{2,\alpha}(\mathcal{M}_i, X_i) \geq \frac{s_i}{2}.$$

Translate X_i to the origin and dilate so that we have $K_{2,\alpha}(\mathcal{M}_i, 0) = 1$. It therefore follows from (4.1) that

$$d(0, U_i) \to \infty \qquad \forall X \in \mathcal{M}_i.$$

We observe

$$
\begin{aligned}
& K_{2,\alpha}(\mathcal{M}_i, X) d(X, U_i) \leq s_i \leq 2d(0, U_i) \\
\Rightarrow \quad & K_{2,\alpha}(\mathcal{M}_i, X) \leq \frac{2d(0, U_i)}{d(X, U_i)} \leq \frac{2d(0, U_i)}{d(0, U_i) - \|X\|}
\end{aligned}
$$

which implies that for suitably large i we have uniform local bounds on $K_{2,\alpha}(\mathcal{M}_i, X)$. Therefore by Theorem 4.7 we deduce that there exists a smooth, proper mean curvature flow $\mathcal{M}$ in $\mathbb{R}^{n+1,1}$ such that

$$\mathcal{M}_i \to \mathcal{M}$$

locally (parabolically) in C^2. Moreover we have that $\Theta(\mathcal{M}, X, r) \leq 1$ for all $X \in \mathcal{M}$ and for all $0 < r < \infty$. Now from the monotonicity formula for $\mathcal{M}$ we see that in fact $\Theta(\mathcal{M}, X, r) \equiv 1$ for all $r > 0$ because $\mathcal{M}$ is smooth. And so (from the monotonicity formula again

$$\int_{\mathcal{M}(t-r^2)} \left| \vec{H} - \frac{(x-y)^{\perp}}{2r^2} \right|^2 \Phi_{(x,t)} d\mathcal{H}^n(y) = 0.$$

Hence the translated flow $\mathcal{M}' = \mathcal{M} - X$ satisfies

$$\vec{H}_{\mathcal{M}'(t')}(y') = \frac{y'^{\perp}}{2t'} \qquad t' < 0$$

which implies that $\mathcal{M}'$ is a self-shrinker, that is, evolves by the relation $\mathcal{M}'(t') = \sqrt{-t'}\mathcal{M}'(-1)$. Since $\mathcal{M}$ is smooth at time zero, the curvature can't blow up and so $\mathcal{M}(-1)$ is a plane.
After maybe rotating we can assume that $\mathcal{M} = (\mathbb{R}^n \times \{0\}) \times \mathbb{R}$. Since $\mathcal{M}_i \to \mathcal{M}$ in C^2 locally, there exist $R_i \to \infty$ such that

$$\mathcal{M}_i \cap B^n_{R_i}(0) \times [-R_i, R_i] = \text{graph} u_i$$

where $u_i : B^n_{R_i}(0) \times [-R_i, R_i] \to \mathbb{R}$, with $\|u_i\|_{C^2} \to 0$. We will get a contradiction if we show that in fact, u_i converge to 0 in $C^{2,\alpha}$. To do so, we first note that each of the u_i satisfies the non-parametric mean curvature flow equation

$$\frac{\partial u_i}{\partial t} - \Delta u_i = -\frac{D_j u_i D_l u_i}{1 + |Du|^2} D_{jl} u_i =: f_i.$$

By the interior parabolic Schauder estimates we have

$$\|u_i\|_{2,\alpha;B_R} \leq C \|f_i\|_{0,\alpha;B_{2R}},$$

and since f_i is quadratic in terms bounded in $C^{0,\alpha}$ and converging to zero uniformly we get

$$\|u_i\|_{2,\alpha} \to 0 \qquad \text{as } i \to \infty$$

contradicting the fact that $K_{2,\alpha}(\mathcal{M}_i, 0) \equiv 1$. $\square$

We so far only stated the monotonicity formula for compact flows, but it is possible to extend this to non-compact flows also provided that

$$\int_{\mathcal{M}(t)} \Phi_{(x_0,t_0)}(x,t) d\mathcal{H}^n < \infty \qquad \forall t.$$

Indeed, suppose without loss of generality that $(x_0, t_0) = (0,0)$ and we write

$$\Phi_{(x_0,t_0)}(x,t) = \Phi(x,t) = \frac{1}{(4\pi(-t))^{n/2}} \exp\left(\frac{|x|^2}{4t}\right) \qquad t < 0.$$

Then

$$\left(\frac{d}{dt} + \Delta_{\mathcal{M}(t)}\right)\Phi - |\vec{H}|^2\Phi = -\left|\vec{H} - \frac{D^\perp \Phi}{\Phi}\right|^2 Phi.$$

For $R > 0$ we introduce the cut-off function $\phi_R \in C_c^2(\mathbb{R}^{n+1})$ which satisfies

$$\phi_R \equiv 1 \text{ on } B_R \qquad \phi_R \equiv 0 \text{ on } \mathbb{R}^{n+1} \setminus B_{2R} \qquad |D\phi_R| \leq \frac{C}{R} \qquad |D^2\phi_R| \leq \frac{C}{R^2}$$

Then we can calculate

$$\begin{aligned}
\frac{d}{dt}\int_{\mathcal{M}(t)} \Phi\phi_R d\mathcal{H}^n &= \int_{\mathcal{M}(t)} \frac{\partial}{\partial}(\Phi\phi_R) + \vec{H}\cdot D(\Phi\phi_R) - |\vec{H}|^2\Phi\phi_R d\mathcal{H}^n \\
&= \int_{\mathcal{M}(t)} \frac{d}{dt}(\Phi\phi_R) - |\vec{H}|^2\Phi\phi_R \\
&= \int_{\mathcal{M}(t)} \Phi\frac{d\phi_R}{dt} + \phi_R\frac{d\Phi}{dt} - |\vec{H}|^2\Phi\phi_R d\mathcal{H}^n \\
&= \int_{\mathcal{M}(t)} \Phi\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\phi_R + \left(\left(\frac{d}{dt} + \Delta_{\mathcal{M}(t)}\right)\Phi - |\vec{H}|^2\Phi\right)\phi_R d\mathcal{H}^n,
\end{aligned}$$

since

$$\int_{\mathcal{M}(t)} \Phi\Delta_{\mathcal{M}(t)}\phi_R = \int_{\mathcal{M}(t)} \phi_R\Delta_{\mathcal{M}(t)}\Phi.$$

Hence

$$\frac{d}{dt}\int_{\mathcal{M}(t)} \Phi\phi_R d\mathcal{H}^n = -\underbrace{\int_{\mathcal{M}(t)} \Phi\left(\Delta_{\mathcal{M}(t)}\phi_R\right) d\mathcal{H}^n}_{\leq \frac{C}{R^2}} - int_{\mathcal{M}(t)}\left|\vec{H} - \frac{D^\perp\Phi}{\Phi}\right|^2 \Phi\phi_R d\mathcal{H}^n$$

so we can let $R \to \infty$ and we will recover the same identity as before.

4.2 Proof of Gage-Hamilton-Grayson Theorem

Now that we have more results for mean curvature flow, and hence curve shortening flow in particular, under our belts, we return to the proof of the Gage-Hamilton-Grayson result which we started discussing back in section 1.1. The proof we sketch is based on Huisken's proof of the result which first appeared in [15].

Sketch of proof. Suppose that $F : S^1 \times [0,T) \to \mathbb{R}^2$ is smooth with $F(\cdot,t) : S^1 \to \mathbb{R}^2$ an embedding for all t, and such that

$$\frac{\partial F}{\partial t}(p,t) = \vec{k}(p,t),$$

and that $F(x,0) = F_0(x)$ for some fixed embedding F_0. We denote $\mathcal{M} := \bigcup_{t\in[0,T)} \gamma_t \times \{t\}$. Suppose that T is the first singular time of the flow, so that $\mathcal{M}$ is a proper smooth flow in $\mathbb{R}^2 \times (0,T)$. Furthermore, that T is the first singular time means there is $x_0 \in \mathbb{R}^2$ with the following properties:

(i) (x_0,T) is reached by the flow, i.e. there are x_j, t_j such that $t_j < T$ with $t_j \nearrow T$ and $x_j \in \mathcal{M}(t_j)$,

(ii) there is no spacetime neighbourhood U_0 of (x_0,T) such that $\mathcal{M} \cap U_0$ can be extended to a proper smooth flow in U_0.

By the comparison principle and considering shrinking circles, $T < \infty$. Rescale $\mathcal{M}$ around (x_0,T) by $\lambda_j \to \infty$. Then

$$\begin{aligned}\mathcal{M}^j &:= \mathcal{D}_{\lambda_j}(\mathcal{M} - (x_0,T)) \\ &= \bigcup \lambda_j(\gamma_t - x_0) \times \lambda_j^2\{t - T\},\end{aligned}$$

so

$$\mathcal{M}^j(s) = \bigcup \lambda_j(\gamma_{T+\lambda_j^{-2}s} - x_0) \times \{s\} \qquad s \in [-\lambda_j^2 T, 0).$$

Monotonicity for $\mathcal{M}$ says

$$\frac{d}{dt}\int_{\mathcal{M}(t)} \Phi_{(x_0,T)}(x,t)d\mathcal{H}^n = -\int_{\mathcal{M}(t)} \left|\vec{k} - \frac{(x_0-x)^\perp}{2(T-t_0)}\right|^2 \Phi_{(x_0,t_0)}d\mathcal{H}^n \qquad t < T.$$

Let

$$\Theta(\mathcal{M},(x_0,T)) := \lim_{t\nearrow T}\int_{\mathcal{M}(t)} \Phi_{(x_0,T)}(x,t)d\mathcal{H}^n,$$

which must exist by the monotonicity formula. Therefore integrating we have

$$\int_{\mathcal{M}(t)} \Phi_{(x_0,T)}(x,t)d\mathcal{H}^n - \Theta(\mathcal{M},(x_0,T)) = \int_t^T \int_{\mathcal{M}(t')} \left|\vec{k} - \frac{(x_0-x)^\perp}{2(T-t')}\right|^2 \Phi_{(x_0,T)}(x,t')d\mathcal{H}^n dt'.$$

In terms of the rescaled flow,

$$\int_{\mathcal{M}^j(s)} \Phi_{(0,0)}(y,s)d\mathcal{H}^n(y) - \Theta(\mathcal{M},(x_0,T)) = \int_s^0 \underbrace{\int_{\mathcal{M}^j(s')} \left|\vec{k_j} - \frac{y^\perp}{2s'}\right|^2 \Phi(y,s')d\mathcal{H}^n}_{=:f_j(s')} ds'.$$

Since

$$\int_{\mathcal{M}(T+\lambda_j^{-2}s)} \Phi_{(x_0,T)}(x,T+\lambda_j^{-2}s)d\mathcal{H}^n \to \Theta(\mathcal{M},(x_0,T))$$

as $j \to \infty$, it follows that $f_j \to 0$ in $L^1_{loc}(-\infty, 0]$ this implies (up to subsequences) that $f^j(s) \to 0$ pointwise almost everywhere. Hence for almost every fixed s we have

$$\int_{\mathcal{M}^j(s)\cap B_R(0)} |\vec{k_j}|^2 d\mathcal{H}^n \leq C$$

with C independent of j. Since we are in one dimension, this implies that the $\mathcal{M}^j(s)$ are bounded in $C^{1,1/2}_{loc}$. Hence we get subsequential convergence in $\mathcal{M}^j(s) \to \tilde{\mathcal{M}}(s)$ in $C^{1,\alpha}_{loc}$ for any $\alpha < 1/2$, where $\tilde{\mathcal{M}}(s)$ is an embedded (by Huisken's scale invariant intrinsic/extrinsic distance estimate discussing in section 1.1) $C^{1,1/2}$ curve.. Wince we have uniform $W^{2,2}_{loc}$ bounds, Rellich's theorem (see [7]) implies the limit is in $W^{2,2}_{loc}$. Elliptic regularity then gives smoothness as $\tilde{\mathcal{M}}(s)$ satisfies

$$\vec{k} = \frac{y^\perp}{2s}$$

in a weak sense. It follows from work of Abresch-Langer [1] that $\tilde{\mathcal{M}}(s)$ must be either a line through the origin or a circle of radius $\sqrt{-2s}$ centred at the origin. If $\tilde{\mathcal{M}}(s)$ is a line then we see that $\Theta(\mathcal{M}, (x_0, T)) = 1$ and so by Theorem 4.6 it follows that (x_0, T) is a smooth point, which is a contradiction. On the other hand, if $\tilde{\mathcal{M}}(s)$ is a circle then we can show $C^{1,\alpha}$ convergence of the rescalings to $S^1(\sqrt{-2s})$ which implies smooth convergence for all times. This then implies the result. □

References

[1] U. Abresch and J. Langer. The normalized curve shortening flow and homothetic solutions. *J. Differential Geom.*, 23(2):175–196, 1986.

[2] William K. Allard. On the first variation of a varifold. *Ann. of Math. (2)*, 95:417–491, 1972.

[3] F. J. Almgren, Jr. Some interior regularity theorems for minimal surfaces and an extension of Bernstein's theorem. *Ann. of Math. (2)*, 84:277–292, 1966.

[4] Frederick J. Almgren, Jr. *Almgren's big regularity paper*, volume 1 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000. Q-valued functions minimizing Dirichlet's integral and the regularity of area-minimizing rectifiable currents up to codimension 2, With a preface by Jean E. Taylor and Vladimir Scheffer.

[5] Ennio De Giorgi. *Frontiere orientate di misura minima.* Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61. Editrice Tecnico Scientifica, Pisa, 1961.

[6] Klaus Ecker. *Regularity theory for mean curvature flow.* Progress in Nonlinear Differential Equations and their Applications, 57. Birkhäuser Boston, Inc., Boston, MA, 2004.

[7] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[8] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[9] Herbert Federer. The singular sets of area minimizing rectifiable currents with codimension one and of area minimizing flat chains modulo two with arbitrary codimension. *Bull. Amer. Math. Soc.*, 76:767–771, 1970.

[10] Herbert Federer and Wendell H. Fleming. Normal and integral currents. *Ann. of Math. (2)*, 72:458–520, 1960.

[11] Wendell H. Fleming. On the oriented Plateau problem. *Rend. Circ. Mat. Palermo (2)*, 11:69–90, 1962.

[12] M. Gage and R. S. Hamilton. The heat equation shrinking convex plane curves. *J. Differential Geom.*, 23(1):69–96, 1986.

[13] Matthew A. Grayson. Shortening embedded curves. *Ann. of Math. (2)*, 129(1):71–111, 1989.

[14] Gerhard Huisken. Local and global behaviour of hypersurfaces moving by mean curvature. In *Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990)*, volume 54 of *Proc. Sympos. Pure Math.*, pages 175–191. Amer. Math. Soc., Providence, RI, 1993.

[15] Gerhard Huisken. A distance comparison principle for evolving curves. *Asian J. Math.*, 2(1):127–133, 1998.

[16] Carlo Mantegazza. *Lecture notes on mean curvature flow*, volume 290 of *Progress in Mathematics*. Birkhäuser/Springer Basel AG, Basel, 2011.

[17] E. R. Reifenberg. Solution of the Plateau Problem for m-dimensional surfaces of varying topological type. *Acta Math.*, 104:1–92, 1960.

[18] Richard Schoen and Leon Simon. Regularity of stable minimal hypersurfaces. *Comm. Pure Appl. Math.*, 34(6):741–797, 1981.

[19] Leon Simon. *Lectures on geometric measure theory*, volume 3 of *Proceedings of the Centre for Mathematical Analysis, Australian National University*. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[20] James Simons. Minimal cones, Plateau's problem, and the Bernstein conjecture. *Proc. Nat. Acad. Sci. U.S.A.*, 58:410–411, 1967.

[21] James Simons. Minimal varieties in riemannian manifolds. *Ann. of Math. (2)*, 88:62–105, 1968.

[22] Brian White. Evolution of curves and surfaces by mean curvature. In *Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002)*, pages 525–538. Higher Ed. Press, Beijing, 2002.

[23] Brian White. A local regularity theorem for mean curvature flow. *Ann. of Math. (2)*, 161(3):1487–1519, 2005.

[24] Neshan Wickramasekera. A general regularity theory for stable codimension 1 integral varifolds. *Ann. of Math. (2)*, 179(3):843–1007, 2014.