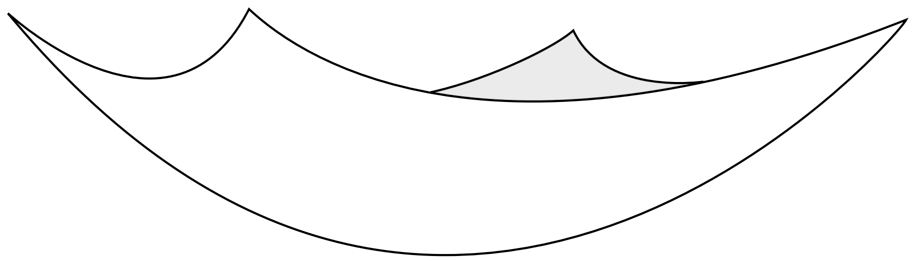
UNDERGRADUATE HONORS THESIS

Regularity of hypersurfaces in $\mathbb{R}^{n+1}$ moving by mean curvature flow



Christos A. Mantoulidis

Thesis Adviser: Professor Leon M. Simon   Submitted On: June 1, 2011 · Stanford, CA

Acknowledgments

I would like to extend my deepest gratitude to my honors thesis and undergraduate adviser, Professor Leon Simon. He has been my guide in mathematics from the day I set foot in his “Honors Multivariable Mathematics” lectures as a freshman to the day I submitted my thesis as a senior. Through classes and independent reading courses he exposed me to more mathematics than I ever would have imagined. He gave me mathematical maturity by helping me understand that I do my best work when I work slowly and meticulously, without getting ahead of myself or glowing over details. In Shakespeare’s words: “Wisely and slow; they stumble that run fast.” Thank you.

I would like to thank Professor Brian White for benevolently sharing his expertise on Mean Curvature Flows with me on multiple occasions throughout the course of my research. I certainly owe thanks to Professor Richard Schoen for introducing me to the Curve Shortening Flow problem during my first course in Partial Differential Equations, igniting my interest in geometric flows. I want to thank all three Professors for being selflessly supportive of me, especially as the deadline of my thesis was approaching.

A big thank you goes to Professor Michael Eichmair, who was the biggest catalyst for my mathematical journey. I wish you only the best.

Contents

CHAPTER 1

Introduction

The aim of this thesis is to study the behavior of surfaces of unit codimension that are deformed in the normal direction with a speed equal to their mean curvature at each point. More concretely, we wish to study the problem of starting with an initial hypersurface M_0 and proceeding to evolve it in such a way that we end up with a family of hypersurfaces M_t where, for a small time increase Δt, the surface $M_{t+\Delta t}$ is approximately obtained by translating each point x in M_t by Δt times the mean curvature vector of M_t at x.

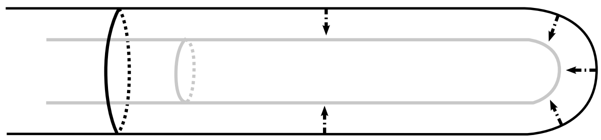


Figure 1: Two-dimensional cylindrical cap moving by mean curvature.

If we denote mean curvature by H, mathematically the mean curvature problem is described as:

$$\frac{\partial}{\partial t}x = H(x)$$

A plethora of approaches have been used to study this geometric problem. In this chapter we will outline the most fruitful of them and the results that they have given rise to. In particular, our overview will include:

(1) "Curve shortening flow," which is the smooth one-dimensional analog of mean curvature flow. This problem has been essentially completely solved.

(2) Smooth mean curvature flow in dimensions higher than one, which is usually studied using techniques from the field of Partial Differential Equations.

(3) "Brakke flow," which drops the smoothness assumption in favor of studying measures moving by mean curvature. This is the oldest approach and is named after Kenneth Brakke, who first tackled this problem. This falls under the category of Geometric Measure Theory.

(4) "Level-set flows and viscosity solutions," a very general approach studying sets moving in a mean-curvature-like way.

Our approach to the problem will feature techniques from both Partial Differential Equations and Geometric Measure Theory, and will lead to a proof of Brakke's regularity theorem for the first singular time of a family of smooth hypersurfaces flowing by mean curvature.

The original objective of this thesis was to fill in the missing steps in Ecker's survey text **[Eck04]** on regularity theory for mean curvature flow. In the process of streamlining and completing Ecker's arguments, a decision was made to adopt a more heavily measure theoretic approach to allow for the simultaneous treatment of Brakke flows wherever possible. An extensive appendix has been included and is required both for justifying Ecker's original results and also for extending them; most important are the chapters on non-compact manifolds and also on local graph representation theorems with a local Arzelà-Ascoli submanifold corollary–topics that seem not to have been discussed elsewhere in the literature. To the best of the author's knowledge original arguments were used throughout the appendix, and also in the main text whenever completing Ecker's presentation and giving it a more measure theoretic twist while deviating as little as possible from his overall structure.

The approach followed by this thesis has the advantage of simultaneous clarity and universality: very general Brakke flow techniques have been used to study smooth–i.e. intuitively and geometrically simple–solutions to mean curvature flow. This point of view allows one to ponder Brakke flow questions from a different perspective and perhaps lead to an eventual strengthening of Brakke's regularity theorem, which is believed to be suboptimal but has not been significantly improved upon in full generality in decades. A list of interesting questions that occur when one works with mean curvature flow from the point of view of this thesis has been included in the Afterword.

1.1. Curve Shortening Flow

The simplest example of mean curvature flow is the one-dimensional case of a smooth family of curves $\gamma(\cdot, t) : I \to \mathbb{R}^2$ (where $I = \mathbb{S}^1$ or an open interval in $\mathbb{R}$) moving per their curvature vector, κ, and starting out at an initial curve γ_0:

$$\left\{ \begin{array}{rcl} \frac{\partial}{\partial t}\gamma & = & \kappa \\ \gamma(\cdot, 0) & = & \gamma_0 \end{array} \right\}$$

This simpler problem is known as the "curve shortening flow." Gage and Hamilton [**GH86**] studied this equation when the initial curve γ_0 is closed ($I = \mathbb{S}^1$) and convex, and concluded under this hypothesis that the solution curves will actually shrink to a point and that if we zoom in sufficiently fast then the curves will converge exponentially to a circle. Later on Grayson [**Gra87**] gave a delicate argument by studying the different types of singularities that could occur *a priori*, and concluded that any initial closed embedded curve will evolve to become convex, and hence by Gage and Hamilton's result will shrink to a round point. This essentially settled the curve shortening problem.

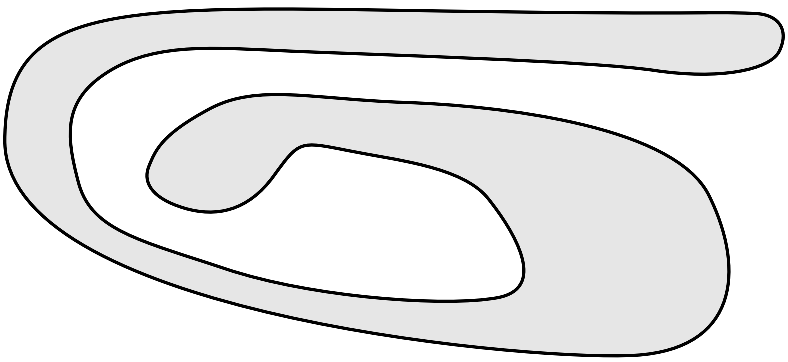


Figure 2: Even complicated closed embedded curves shrink and become asymptotically round.

Later on Hamilton [**Ham95**] and Huisken [**Hui98**] gave geometric arguments that somewhat simplified the proof of Grayson's theorem. Hamilton related the ratio of the isoperimetric profile to that of a circle of the same area, while Huisken studied the ratio of the length of the chord between points on the curve to the length of the actual arc that the curve traces between them, and showed that it converges to that of a circle. Unfortunately, all of these proofs rely on singularity classification results. By giving even stronger isoperimetric estimates, Ben Andrews and Paul Bryan [**AB11**] recently greatly simplified the proof of Grayson's theorem.

1.2. Smooth Mean Curvature Flow (Classical approach)

This is the direct generalization of curve shortening to higher dimensions. In place of the one-dimensional domain I in curve shortening flows, which was either $\mathbb{S}^1$ or an open subinterval of $\mathbb{R}$, we now have a fixed background n-dimensional, smooth, abstract manifold M^n. The mean curvature flow problem is the initial value problem of a smooth family of embeddings $F(\cdot,t) : M^n \to \mathbb{R}^{n+1}$ starting out as a given initial embedding F_0 of M^n into $\mathbb{R}^{n+1}$ (or perhaps some Riemannian manifold (N^{n+1}, g) instead of $\mathbb{R}^{n+1}$), and flowing in the direction of their mean curvature vector, H:

$$\begin{Bmatrix} \frac{\partial}{\partial t}F & = & H \\ F(\cdot,0) & = & F_0 \end{Bmatrix}$$

To save the reader some suspense we note right away that there is no analogue of Grayson's theorem in any dimension above one, as evidenced by the collapsing and singularity formation of the following dumbbell in $\mathbb{R}^3$.

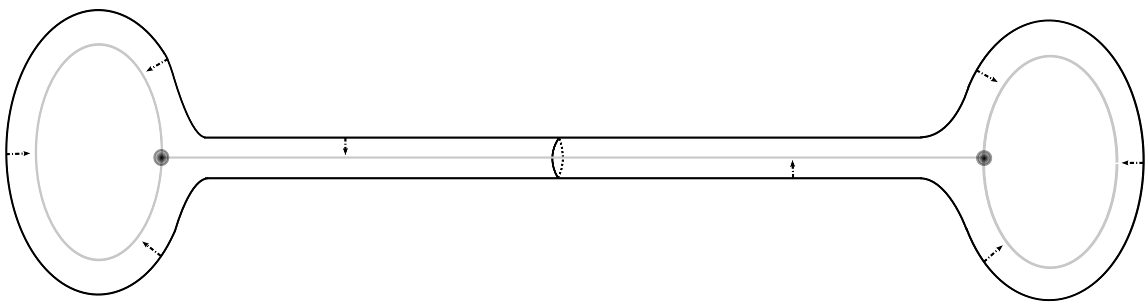


Figure 3: An initially smooth surface forming a collapsing neck and two corner points.

The issue with singularities is that our solutions cease to be smooth and develop corners, so we are no longer in the position to flow by mean curvature in the classical way. The next thing to understand, then, is the size of the singular set that has formed and the nature of these singularities. In the previous figure, the singular set is the line joining the two spheres. Before proceeding to address partial results in this direction, let us make a couple of notes about cases that have been proven to steer clear of singularities. A good reference for the classical approach is **[Zhu02]**.

The first is the case of smooth, compact, convex hypersurfaces, generalizing Gage and Hamilton's dimension one theorem. In the higher dimensional case, Huisken **[Hui84]** confirmed that no singularities will arise since initially compact and convex hypersurfaces will shrink to a point and become asymptotically round. He did this by controlling their curvature norm in terms of mean curvature and giving strong $\mathscr{L}^p$ bounds on mean curvature of convex flows. Later on Andrews [**And94**] simplified the proof by using a pinching estimate, and rather strongly capitalizing on the geometric fact that convex surfaces will always become "rounder" with the flow.

Another case that is largely understood with the use of classical PDE techniques is that of entire graphs flowing by mean curvature in $\mathbb{R}^{n+1}$. Ecker and Huisken [**EH91**] showed that an initially merely locally Lipschitz graph will never develop a singularity and will in fact produce an eternal mean curvature flow solution, i.e. one that exists for all times.

There is very little known about other initial data that will result in flows without singularities, and so the focus has been shifted to understanding the nature of singularities when they do occur. A significant step forward was made in [**Hui90**] with the discovery of Huisken's monotonicity formula which allows one to blow up a singularity and study its geometric structure. In particular, after a blow-up the neighborhood surrounding the singularity is going to turn into a shrinking self-similar solution of mean curvature flow [**Ilm95**], although possibly not a smooth one at that–our cue to begin the section on Brakke flows.

1.3. Brakke Flow

One of the biggest setbacks of the classical PDE approach is that we are forced to stop our study of the evolution at the first singular time, in view of the fact that our manifolds cease to be smooth surfaces with well-defined mean curvature (in the classical sense of the term). In his monograph **[Bra78]**, Brakke pioneered the study of mean curvature flow in the context of more general non-smooth objects evolving in a mean-curvature-like way and obtained an impressive regularity theory. The study of non-smooth "manifolds" falls under the general category of Geometric Measure Theory. For a reference on Geometric Measure Theory, see [**Sim83**].

For the purposes of this introduction, a Brakke flow in $\mathbb{R}^{n+1}$ is taken to be a one-parameter family of n-rectifiable Radon measures μ_t such that:

$$\frac{d}{dt}\int \varphi \, d\mu_t \leq \int \left(\nabla\varphi \cdot H - |H|^2\varphi\right) \, d\mu_t \quad \text{for all } \varphi \in \mathcal{C}_c^1\left(\mathbb{R}^{n+1}\right)$$

where H denotes the generalized mean curvature vector of μ_t.

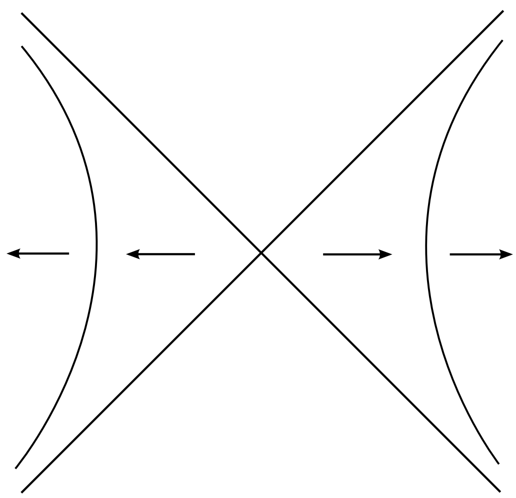


Figure 4: An evolution captured by Brakke flow because of the inequality.

This was taken to be the definition of Brakke motion in particular because this identity (with an equality sign) holds in the case of smooth compact hypersurfaces flowing by mean curvature in $\mathbb{R}^{n+1}$, as will be checked during the course of this thesis. The choice of using an inequality sign in defining Brakke motion was made to increase the scope of mean curvature flow to a larger class of flows (see the figure above), seeing as to how all arguments made only require inequality to hold rather than equality. Furthermore, Brakke flows actually allow for the derivative above to be merely an upper derivative to account for the case of discontinuities and non-differentiability that frequently occur in the non-smooth context.

We immediately see that Brakke flows help us overcome our previous limitation in studying the flow only up to the first singular time, simply because it is not required that our surfaces be smooth anymore. Brakke flows have helped produce a number of important results in the study of the size and the nature of the singular set. For instance, in the previous section it was noted that we may blow up the area surrounding a singularity and obtain a shrinking self-similar solution to mean curvature, albeit not smooth. Ilmanen [**Ilm95**] showed exactly that the blow up will comprise a Brakke motion.

Brakke's regularity theorem [**Bra78**] proves that in the case of a unit-density Brakke flow in $\mathbb{R}^{n+1}$ that experiences no loss in mass at time t, our surface at time t will actually be a smooth manifold away from a closed set of $\mathscr{H}^n$-measure zero. Despite the fact that the assumptions in Brakke's theorem (unit density and area continuity) are deemed too strong and his result too weak (people expect a stronger Hausdorff dimension bound), Brakke's result remains unsurpassed to this date. Consult the Afterword of this thesis for a discussion of open questions in this direction.

Very recently Ecker [**Eck11**] and Han, Sun [**HS11**] independently improved Brakke's theorem in the two-dimensional case by dropping the unit density and area continuity hypotheses in the case of smooth manifolds flowing by mean curvature; in other words, they showed that the singular set at the first singular time has $\mathscr{H}^2$-measure zero. In the general higher dimensional case, Ilmanen [**Ilm94**] made partial progress in improving Brakke's theorem by proving that for almost every initial surface, and almost every time t, our surface at time t will be smooth almost everywhere. His proof did not consist entirely of Brakke flow techniques, but also made substantial use of level set flows–the subject of our next section.

1.4. Level Set Flows/Viscosity Solutions

Much like Brakke flows removed the differentiability constraint from mean curvature flow solutions, we can go one step further and even remove Brakke's measure theoretic constraints too in view of a more general theory of set-theoretic solutions to mean curvature flow. This approach was introduced independently by Evans, Spruck [**ES91, ES92a, ES92b**], and Chen, Giga, Goto [**CGG91**], who coined the terms "level set flow" and "viscosity solution," respectively.

Even though level set flows/viscosity solutions (evidently) study the evolution of the zero-level set of a time dependent function satisfying a particular differential equation, they are better understood in terms of "set-theoretic subsolutions" to mean curvature flow, as presented by Ilmanen [**Ilm95**]. A family of sets Γ_t is a set theoretic subsolution to mean curvature if any compact, smooth mean curvature flow M_t that is originally disjoint from Γ_t remains disjoint at all future times. This approach is motivated by the fact that two smooth, compact, initially disjoint surfaces flowing by mean curvature will remain disjoint for as long as they are smooth. This is a simple application of the weak maximum principle. In any case, a level set flow is just a maximal family of set theoretic subsolutions.

It is worth noting where the level set flow stands in terms of advantages and disadvantages. On one hand, it shares the ability of Brakke flow to study mean curvature evolution past what used to qualify as a singular time in the smooth case–since there are no singular times anymore. On the other hand, it differs from Brakke flows substantially in that it has built-in uniqueness, almost by construction. (Brakke flows are not necessarily unique, as one can see by rotating the cross diagram from the previous section by ninety degrees.)

Combinations of the level set flow and the Brakke flow have made significant contributions to the theory of mean curvature flow. As noted, with the use of Brakke flows and level set flows Ilmanen [**Ilm95**] showed that Brakke's unit density and area continuity hypothesis are almost always satisfied, for almost any initial surface. White [**Whi00**] improved Brakke's regularity theorem in surfaces with non-negative mean curvature ("mean convex surfaces"), both by discarding the unit density and area continuity hypotheses and by establishing a lower Hausdorff dimension bound on the singular set. In particular, he showed that with initial data comprising of a mean convex surface, at each time t the singular set has $\mathscr{H}^{n-1}$-measure zero, and at almost every time it actually has $\mathscr{H}^{n-3}$-measure zero. In particular, in the two dimensional case our surfaces are smooth manifolds for almost all times. Later, White [**Whi03**] classified the nature of singularities that can occur in the mean convex case using a blow up argument.

1.5. Foreword

As explained in the introduction, the end goal of this thesis is to conclude Brakke's regularity theorem for the first singular time of a family of smooth hypersurfaces moving by mean curvature in $\mathbb{R}^{n+1}$. Extra care has been taken to distinguish "classical" from "geometric measure theoretical" arguments in the way the presentation is structured. When there was a choice to be had between classical arguments or measure theoretic arguments, the latter were always favored.

The following assumptions are made throughout the course of this thesis:

(1) Our manifolds are compact and at least two-dimensional. In other words the background manifold M^n is a smooth, compact manifold of dimension $n \geq 2$.

(2) Our manifolds are all (properly) embedded in $\mathbb{R}^{n+1}$. That is, for all times t the map $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$ is an embedding.

(3) Our manifolds have no boundary in the open sets of the ambient space in which we are studying them. In particular, when looking at global mean curvature flows we assume that our manifolds are without boundary.

The assumption $n \geq 2$ is immaterial in view of the fact that the case $n = 1$ (curve shortening flow) has been completely settled. The compactness assumption guarantees a certain degree of equivalence with Brakke flows and makes for an overall much cleaner presentation of the results. A list of topics that might go wrong in the non-compact case can be found in Chapter D of the appendix. The embeddedness and lack-of-boundary assumptions are crucial to the classical PDE arguments involving limits, which will break in the case of merely immersed manifolds, or manifolds with boundary.

We commence our work in Chapter 2 where we use classical PDE tools to begin to understand our mean curvature evolution equation from a geometric point of view. We establish short-time existence, and then proceed to study the evolution of various geometric quantities associated with our surface (e.g. the metric, the unit normal, the second fundamental form, and all its derivatives). We introduce the notion of parabolic rescaling which is used in blow up arguments, and then we conclude the chapter with a section on a localized weak maximum principle for mean curvature flow.

We proceed to Chapter 3, which studies the evolution of integral quantities of carefully chosen test functions in the ambient space. This allows us to establish area, mean curvature, and Hausdorff density estimates. We introduce Huisken's monotonicity formula, Brakke's clearing out lemma, and the concept of Gauss/parabolic density. This approach is largely influenced by Brakke flows and special care has been taken to have arguments that require merely the Brakke motion inequality to work (albeit, we need an actual derivative rather than just an upper derivative). Chapter C of the appendix is devoted to showing that Gauss density equals one at smooth times of the flow, a computation that is not required for Brakke's regularity theorem but nevertheless has been included for reasons of completeness. It is presented separately from Chapter 3 because it does not work for non-smooth Brakke flows, unlike the other arguments in the chapter.

Up next is Chapter 4, which uses both classical and measure theoretic tools to establish sufficient conditions for non-singularity within an open set. In particular, via measure theory we establish a relationship between $\mathscr{L}^2$ and $\mathscr{L}^\infty$ flatness of our surfaces, and then via a blow up argument we show that the latter implies curvature bounds. In view of Chapter E of the appendix, this guarantees regularity.

Finally, we conclude our work with Chapter 5 which contains the proof of Brakke's regularity theorem. Assuming unit density and area continuity for the first singular time, we pick out a set of points on the limiting surface around which integral approximations are bound to work, and show that around these points our limiting surface is almost $\mathscr{L}^2$-flat. The result will follow from the fact that these points are of full $\mathscr{H}^n$-measure within our limiting surface.

CHAPTER 2

Evolution under Flow

We will now delve into studying mean curvature flow of smooth n-dimensional surfaces in $\mathbb{R}^{n+1}$. As explained, we begin with a fixed abstract, smooth, n-dimensional manifold M^n and we wish to study the existence and behavior of a smooth family of embeddings $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$ moving by mean curvature for $t \geq 0$. In other words, in this chapter we will study the initial value problem:

$$\begin{Bmatrix} \frac{\partial}{\partial t} F(p,t) & = & H(p,t) \\ F(p,0) & = & F_0(p) \end{Bmatrix} \tag{MCF-IVP}$$

where $H(p,t)$ denotes the mean curvature vector at the point $F(p,t)$ of the embedded n-dimensional manifold $F(M^n, t)$, and $F_0 : M^n \to \mathbb{R}^{n+1}$ is a given embedding.

The outline of this chapter is as follows. In section 2.1 we discuss issues of notation and vocabulary as will be used in the remainder of this work. In section 2.2 we establish a short-time existence theorem, in view of the fact that our background manifold is assumed to be compact. In section 2.3 we study the basic evolution of geometric quantities (metric, second fundamental form, curvature norms, etc.) under mean curvature flow. In section 2.4 we introduce the technique of parabolic rescaling a family of submanifolds moving by mean curvature flow. Finally, in section 2.5 we establish a localized weak maximum principle for our flows.

2.1. Notation and Nomenclature

So far we have defined what it means for a family of embeddings $F : M^n \times I \to \mathbb{R}^{n+1}$ to move by mean curvature:

$$\frac{\partial}{\partial t}F(p,t) = H(p,t) \quad \text{for all } (p,t) \in M^n \times I \tag{MCF}$$

In certain cases we may not need to require that there is motion by mean curvature globally. That is, we may only require that (MCF) be true for $(p,t) \in \mathcal{O}$, some (relatively) open subset of $M^n \times I$. In this case we will write that our mappings $F : M^n \times I \to \mathbb{R}^{n+1}$ move by mean curvature only in $\mathcal{O} \subseteq M^n \times I$. The subset $\mathcal{O}$ will often arise by considering surfaces that only satisfy mean curvature movement within some open subset $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ of the ambient space outside of which we have no interest. This will be the case most of the time in Chapter 3. In this case $\mathcal{O} = \{(p,t) \in M^n \times I \;:\; F(p,t) \in \mathcal{U}\} \subseteq M^n \times I$, and we will write that we have movement by mean curvature in the local domain $\mathcal{U} \times I$.

Our definition of movement by mean curvature continues to be–visually–cumbersome, primarily because we always have to make explicit reference to the underlying mappings $F(\cdot,t) : M^n \times \mathbb{R}^{n+1}$. In reality, we wish to follow our intuition and talk about surfaces $M_t := F(M^n,t)$ moving by mean curvature. We adopt the following definition:

DEFINITION 2.1.1. We say that the family of submanifolds (M_t) moves by mean curvature in $\mathcal{U} \times I$, where $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ is open, if there exists a smooth family of embeddings $F(\cdot,t) : M^n \to \mathbb{R}^{n+1}$ which parametrizes M_t (that is, $M_t = F(M^n,t)$ for all $t \in I$) and in turn moves by mean curvature in $\mathcal{U} \times I$, i.e. (MCF) holds for every $(p,t) \in M^n \times I$ such that $F(p,t) \in \mathcal{U}$. □

The reader must observe that *not* all parametrizations $G(\cdot,t) : M^n \to \mathbb{R}^{n+1}$ of our M_t will flow by mean curvature per (MCF). As a matter of fact, for each family (M_t) moving by mean curvature flow there is a *unique* choice of parametrization $F(\cdot,t) : M^n \to \mathbb{R}^{n+1}$ satisfying (MCF), up to diffeomorphisms of the background manifold M^n.

The parametrization independent version of (MCF) is:

$$\left(\frac{\partial}{\partial t}G(p,t)\right)^{\perp} = H(p,t) \quad \text{for all } (p,t) \in M^n \times I$$

and allows one to use arbitrary smoothly varying parametrizations $G(\cdot,t) : M^n \to \mathbb{R}^{n+1}$ of (M_t). Furthermore, given any solution G to the parametrization independent equation above one can find a smooth family of diffeomorphisms $\varphi(\cdot,t) : M^n \to M^n$ that allows us to recover the solution of the original equation (MCF) by reparametrizing per $F(\cdot,t) := G(\varphi(\cdot,t),t)$.

2.2. Short-time Existence

Naturally, we want to establish some kind of existence theorem for our flow, otherwise it would be a rather uninteresting geometric flow to study. In this section, we will prove that we do in fact have short-time existence, i.e. for any given embedding $F_0 : M^n \to \mathbb{R}^{n+1}$ there exists a $\varepsilon > 0$ and a smooth family of embeddings $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$ such that (MCF-IVP) holds for all $t \in [0, T)$. Our compactness assumption is crucial here.

There are a few ways to establish short-time existence for mean curvature flow. One possibility is to work locally and express each surface as a graph of a function over an affine space, which satisfies a PDE whose short-time existence one can show with a contraction mapping argument. We would then have to patch those graph maps together. Alternatively, we could avoid having to patch things together by writing each surface entirely as a graph over its preceding surfaces using the unit normal field. Two other ways come from tools that were initially developed to establish short-time existence for Ricci flow. Hamilton **[Ham82]** put together a rather lengthy and technical proof of existence for a general class of nonlinear problems, which he initially applied to Ricci flow, and shortly thereafter Gage and Hamilton **[GH86]** applied the same idea to establish short-time existence for mean curvature flow. The final way was originally developed by DeTurck **[DeT83]** to simplify Hamilton's proof for short-time existence of Ricci flow, and was subsequently adopted to mean curvature flows.

We will present "DeTurck's trick," which is the most elegant of the approaches since it simply reduces the problem to a quasilinear strictly parabolic problem. Notice that this works equally well for immersions, injective immersions, and embeddings, since all three of these classes of mappings are stable on compact manifolds (see below).

THEOREM 2.2.1. Let $F_0 : M^n \to \mathbb{R}^{n+1}$ be a given initial embedding. Then there exists a $T > 0$ and a smooth family of embeddings $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$, $t \in [0, T)$, such that $F(\cdot, 0) = F_0$ and:

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

for all $(p, t) \in M^n \times [0, T)$, where $H(p, t)$ denotes the mean curvature vector of $F(M^n, t)$ at $F(p, t)$.

PROOF. First of all, we note that we need not worry about the propagation of the property of being an embedding. From basic differential topology we know that embeddings of compact manifolds into other (not necessarily compact) manifolds form stable classes (see, e.g. **[GP74]**). Therefore if $F(\cdot, 0)$ is an embedding, then each of $F(\cdot, t)$ will be too, for t sufficiently close to 0. Therefore, all we need to worry about is the curvature flow equation being true.

Consider the following following PDE system that we want to be satisfied by a smooth family $\tilde{F} : M^n \times I \to \mathbb{R}^{n+1}$:

$$\frac{\partial}{\partial t}\tilde{F} = \Delta_{\tilde{g}(t)}\tilde{F} + v^k \nabla_k \tilde{F} \tag{$\dagger_1$}$$

with the initial conditions $\tilde{F}(\cdot, 0) = F_0$. For each time t we endow M^n with a metric $\tilde{g}(t)$, the pullback metric $\tilde{F}(\cdot, t)^* g_{\mathbb{R}^{n+1}}$ where $g_{\mathbb{R}^{n+1}}$ is the standard metric on $\mathbb{R}^{n+1}$. This metric $\tilde{g}(t)$ in turn gives rise to the Levi-Civita connection on M^n, whose Christoffel symbols we label $\tilde{\Gamma}^k_{ij}(t)$. At time $t = 0$, we choose to label the Christoffel symbols $(\Gamma_0)^k_{ij}$. In conclusion, $\Delta_{\tilde{g}(t)}$ denotes the Laplace-Beltrami operator with respect to the metric $\tilde{g}(t)$.

The vector field v above is chosen so that:

$$v^k = \tilde{g}^{ij}\left(\tilde{\Gamma}^k_{ij} - (\Gamma_0)^k_{ij}\right)$$

Notice that this is well-defined because the difference between two pairs of Christoffel symbols transforms like a tensor. Given this definition of v, in local coordinates ($\dagger_1$) reads:

$$\begin{aligned}
\frac{\partial}{\partial t}\tilde{F} &= \Delta_{\tilde{g}(t)}\tilde{F} + v^k \nabla_k \tilde{F} \\
&= \tilde{g}^{ij}\nabla_i \nabla_j \tilde{F} + v^k \nabla_k \tilde{F} \\
&= \tilde{g}^{ij}\left(\frac{\partial^2}{\partial x^i \partial x^j}\tilde{F} - \tilde{\Gamma}^k_{ij}\frac{\partial}{\partial x^k}\tilde{F}\right) + \tilde{g}^{ij}\left(\tilde{\Gamma}^k_{ij} - (\Gamma_0)^k_{ij}\right)\frac{\partial}{\partial x^k}\tilde{F} \\
&= \tilde{g}^{ij}\frac{\partial^2}{\partial x^i \partial x^j}\tilde{F} - \tilde{g}^{ij}(\Gamma_0)^k_{ij}\frac{\partial}{\partial x^k}\tilde{F}
\end{aligned}$$

which is a quasilinear parabolic PDE system. This is because $\left(\tilde{g}^{ij}\right)$ is a positive definite matrix depending on just first order derivatives of $\tilde{F}$. In other words, the seemingly odd and possibly nonlinear PDE system ($\dagger_1$) is actually a simple quasilinear parabolic PDE system when expressed in local coordinates. From standard parabolic PDE theory and since M^n is compact, we know then that ($\dagger_1$) does admit a solution on some $[0, T)$ for some $T > 0$ (see, e.g. [**LSU68**], [**Eid69**], or the more up to date [**Tay96b**]). As discussed earlier, after possibly shrinking $T > 0$ we can ensure that $\tilde{F}(\cdot, t)$ continues to be an embedding.

Let us now transform this into a solution of our original problem. Indeed let $Q : M^n \times I \to M^n$ be a solution to the initial value problem:

$$\frac{\partial}{\partial t} Q^k = -v^k \tag{$\dagger_2$}$$

with initial data $Q(\cdot, 0) \equiv \text{id}$. After possibly shrinking $T > 0$ some more, we can assume that it is small enough for $Q(\cdot, t)$ to be a diffeomorphism (diffeomorphisms on compact manifolds also form a stable class) and also small enough for (†$_1$) and (†$_2$) to hold on $[0, T)$. Now define $F : M^n \times [0, T) \to \mathbb{R}^{n+1}$ by:

$$F(p, t) := \tilde{F}(Q(p, t), t)$$

The chain rule now gives:

$$\begin{aligned} \frac{\partial}{\partial t} F &= \frac{\partial}{\partial t} \tilde{F} + \frac{\partial}{\partial t} Q^k \cdot \nabla_k \tilde{F} \overset{(\dagger_1),(\dagger_2)}{=} \left(\Delta_{\tilde{g}(t)} \tilde{F} + v^k \nabla_k \tilde{F} \right) - v^k \nabla_k \tilde{F} \\ &= \Delta_{\tilde{g}(t)} \tilde{F} = \Delta_{g(t)} F \end{aligned}$$

where $g(t)$ is the pullback metric $Q(\cdot, t)^* \tilde{g}(t)$. Notice then that each $F(\cdot, t)$ is an embedding and:

$$g(t) = Q(\cdot, t)^* \tilde{g}(t) = Q(\cdot, t)^* \tilde{F}(\cdot, t)^* g_{\mathbb{R}^{n+1}} = F(\cdot, t)^* g_{\mathbb{R}^{n+1}}$$

so that $g(t)$ is the pullback metric of the embedded surface $F(M^n, t)$. Then from standard Riemannian geometry we know that $\Delta_{g(t)} F$ is simply the mean curvature vector, and hence we have constructed a solution F to the initial value mean curvature flow problem (MCF-IVP). $\square$

2.3. Geometric Quantities

We start this section off by urging the reader to look at Chapter A of the appendix in order to gain familiarity with the differentiation notation that is used throughout this work, and also understand the "intrinsic" vs "extrinsic" function nomenclature. In short, intrinsic functions are time dependent functions on the background manifold M^n (e.g. any geometric quantity like the metric) while extrinsic functions are time dependent functions on the ambient space $\mathbb{R}^{n+1}$ (e.g. the function $|x|^2 + 2nt$). One can convert extrinsic functions into intrinsic ones by pulling them back to M^n.

We begin with a proposition that is a corollary of (A.3.1) and (A.3.2) for mean curvature flow.

PROPOSITION 2.3.1. Let (M_t) move by mean curvature in $\mathcal{U} \times I$, and let $f : \mathcal{U} \times I \to \mathbb{R}^m$, $m \geq 1$. We have:

$$\frac{\partial}{\partial t} f(x,t) = D_t f(x,t) + H(x,t) \cdot Df(x,t) \tag{2.3.1.a}$$

$$\Delta_{M_t} f(x,t) = \mathrm{div}_{M_t} Df(x,t) + H(x,t) \cdot Df(x,t) \tag{2.3.1.b}$$

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f(x,t) = D_t f(x,t) - \mathrm{div}_{M_t} Df(x,t) \tag{2.3.1.c}$$

for all $t \in I$, $x \in M_t \cap \mathcal{U}$, under appropriate differentiability conditions. We need f to be (jointly) differentiable for (2.3.1.a), $f(\cdot, t)$ to be twice differentiable in $\mathcal{U}$ for each $t \in I$ for (2.3.1.b), and both hypotheses for (2.3.1.c).

PROOF. Let us begin with (2.3.1.a). Let $p \in M^n$ be the unique point of the background manifold such that $F(p,t) = x$. Then since (M_t) moves by mean curvature in $\mathcal{U} \times I$, we have:

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

and hence the claim follows from proposition (A.3.2). The second part, (2.3.1.b), is nothing but (A.3.1). Finally, (2.3.1.c) is obtained by subtracting (2.3.1.a) and (2.3.1.b). $\square$

At this point we introduce a list of geometric quantities that will be of interest to us in this work. Einstein summation notation will be used significantly in the remainder of this section for notational convenience.

For each t the background manifold M^n can be endowed with the pullback metric $g = g(t)$ of the Euclidean metric as induced on M_t. In local coordinates, the metric is given by:

$$g_{ij} = \frac{\partial}{\partial x^i} F \cdot \frac{\partial}{\partial x^j} F$$

Additionally, for each t the manifold M_t admits a second fundamental form $h = h(t)$ which in local coordinates is:

$$h_{ij} = \frac{\partial}{\partial x^i} \nu \cdot \frac{\partial}{\partial x^j} F$$

for some choice of a unit normal field ν. The sign of h depends on our choice of the unit normal field. For the remainder of this section only we will use H to denote the scalar mean curvature:

$$H = \text{tr}\, h = g^{ij} h_{ij} = \text{div}_{M_t} \nu$$

Notice that scalar mean curvature H depends on our choice of orientation. On the other hand, the mean curvature vector $-H\nu$ (which is denoted by H in all other sections) is independent of our choice of orientation.

Finally, one last geometric quantity that will be very important is the squared norm $|A|^2$ of the second fundamental form h, which essentially measures curvature and is given by:

$$|A|^2 = g^{ik} g^{j\ell} h_{ij} h_{k\ell}$$

Notice that A is just the second fundamental form h, and $|A|^2$ is the tensor's norm. We will tend to use h for the tensor itself and A when we wish to measure norms. The tensor's higher order derivatives $\nabla^m A$, $m \geq 1$, are defined simply by taking covariant derivatives of the second fundamental form tensor, and their norms' evolution will play a fundamental role in our regularity theory.

We will assume that the reader has a working knowledge of differential geometry in higher dimensions and is comfortable with the use of abstract manifolds, tensors, geodesics, normal coordinates, second fundamental forms, covariant derivatives, Riemann's curvature tensor symmetries, Gauss' equation, Codazzi's equation, and so forth. For a reference consult **[Lan95]** or **[Lee97]**.

The following propositions study all our geometric quantities' evolution under the flow. Results of this type were originally proven by Huisken **[Hui84]**, but can also be found in **[Zhu02]** and **[Eck04]**.

PROPOSITION 2.3.2. Let (M_t) move by mean curvature in $\mathcal{U} \times I$. For all $(p,t) \in M^n \times I$ such that $F(p,t) \in \mathcal{U}$:

$$\frac{\partial}{\partial t} g_{ij} = -2H h_{ij} \tag{2.3.2.a}$$

$$\frac{\partial}{\partial t} g^{ij} = 2H h^{ij} \tag{2.3.2.b}$$

The area element in turn evolves according to:

$$\frac{\partial}{\partial t} \sqrt{g} = -H^2 \sqrt{g} \tag{2.3.2.c}$$

PROOF. Let ν be a choice of a unit normal field to M_t. Since (M_t) moves by mean curvature we can commute the mixed partial derivatives below to get:

$$\frac{\partial}{\partial t}\frac{\partial}{\partial x^i}F \cdot \frac{\partial}{\partial x^j}F = \frac{\partial}{\partial x^i}\frac{\partial}{\partial t}F \cdot \frac{\partial}{\partial x^j}F = \frac{\partial}{\partial x^i}\left(-H\nu\right) \cdot \frac{\partial}{\partial x^j}F = -H\left(\frac{\partial}{\partial x^i}\nu \cdot \frac{\partial}{\partial x^j}F\right) = -H h_{ij}$$

where one term of the product rule in the second to last step has dropped out since ν is normal to the tangent space. Summarizing, and recalling that h is a symmetric tensor:

$$\frac{\partial}{\partial t} g_{ij} = \frac{\partial}{\partial t}\left(\frac{\partial}{\partial x^i}F \cdot \frac{\partial}{\partial x^j}F\right) = \frac{\partial}{\partial t}\frac{\partial}{\partial x^i}F \cdot \frac{\partial}{\partial x^j}F + \frac{\partial}{\partial x^i}F \cdot \frac{\partial}{\partial t}\frac{\partial}{\partial x^j}F = -2H h_{ij}$$

Establishing the second identity is but an exercise in Einstein summation notation:

$$\begin{aligned}
\frac{\partial}{\partial t} g^{ij} &= \left(\frac{\partial}{\partial t} g^{ik}\right)\delta^j_k = \left(\frac{\partial}{\partial t} g^{ik}\right) g_{k\ell} g^{\ell j} = \left(\frac{\partial}{\partial t}\left(g^{ik} g_{k\ell}\right) - g^{ik}\frac{\partial}{\partial t} g_{k\ell}\right) g^{\ell j} \\
&= \left(\frac{\partial}{\partial t}\delta^i_\ell - g^{ik}\frac{\partial}{\partial t} g_{k\ell}\right) g^{\ell j} = -g^{ik}\left(\frac{\partial}{\partial t} g_{k\ell}\right) g^{\ell j} = 2H g^{ik} h_{k\ell} g^{\ell j} = 2H h^{ij}
\end{aligned}$$

as claimed. In order to prove the final claim we use the known linear algebra fact that for a family $A = A(t)$ of square non-singular matrices (a_{ij}) with inverses $\left(a^{ij}\right)$, the determinant evolves according to:

$$\frac{d}{dt}(\det A) = \det A \ \text{tr}\left(A^{-1}\frac{d}{dt}A\right) = (\det A)\, a^{ij}\frac{d}{dt}a_{ij} \tag{$\dagger$}$$

In consequence, the area element $\sqrt{g} = \sqrt{\det(g_{ij})}$ satisfies:

$$\frac{\partial}{\partial t}\sqrt{g} = \frac{1}{2\sqrt{g}}\frac{\partial}{\partial t}g \overset{(\dagger)}{=} \frac{1}{2\sqrt{g}}\, g\, g^{ij}\, \frac{\partial}{\partial t}g_{ij} = \frac{1}{2}\,\sqrt{g}\, g^{ij}\,\left(-2Hh_{ij}\right) = -H\, g^{ij}h_{ij}\,\sqrt{g} = -H^2\,\sqrt{g}$$

and the result follows. □

PROPOSITION 2.3.3. Let (M_t) move by mean curvature in $\mathcal{U} \times I$. For all $(p,t) \in M^n \times I$ such that $F(p,t) \in \mathcal{U}$ and for a choice of a unit normal ν:

$$\frac{\partial}{\partial t}\nu = \nabla H \tag{2.3.3.a}$$

PROOF. Since the time derivative of ν is a tangent vector (differentiate $|\nu|^2 = 1$ in t), we can express it in terms of the basis vectors:

$$\begin{aligned}\frac{\partial}{\partial t}\nu &= g^{ij}\left(\frac{\partial}{\partial t}\nu \cdot \frac{\partial}{\partial x^j}F\right)\frac{\partial}{\partial x^i}F = g^{ij}\left(-\nu \cdot \frac{\partial}{\partial t}\frac{\partial}{\partial x^j}F\right)\frac{\partial}{\partial x^i}F = g^{ij}\left(-\nu \cdot \frac{\partial}{\partial x^j}\frac{\partial}{\partial t}F\right)\frac{\partial}{\partial x^i}F \\ &= g^{ij}\left(-\nu \cdot \frac{\partial}{\partial x^j}\left(-H\nu\right)\right)\frac{\partial}{\partial x^i}F = g^{ij}\left(\nu \cdot \frac{\partial}{\partial x^j}\left(H\nu\right)\right)\frac{\partial}{\partial x^i}F\end{aligned}$$

By the product rule again and recalling that spatial derivatives of ν are tangent vectors:

$$\frac{\partial}{\partial t}\nu = g^{ij}\left(\nu \cdot \left(\frac{\partial}{\partial x^i}H\right)\nu\right)\frac{\partial}{\partial x^i}F = g^{ij}\left(\frac{\partial}{\partial x^i}H\right)\frac{\partial}{\partial x^i}F = \nabla H$$

and the result follows. □

PROPOSITION 2.3.4. Let (M_t) move by mean curvature in $\mathcal{U} \times I$. For all $(p,t) \in M^n \times I$ such that $F(p,t) \in \mathcal{U}$:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)h_{ij} = -2Hh_{i\ell}g^{\ell m}h_{mj} + |A|^2 h_{ij} \tag{2.3.4.a}$$

PROOF. By the Codazzi equation ∇h is totally symmetric, and by the symmetries of the Riemann curvature tensor combined with the Gauss equations we see that:

$$\begin{aligned}
\Delta_{M_t} h_{ij} &= g^{mn}\nabla_m\nabla_n h_{ij} = g^{mn}\nabla_m\nabla_i h_{jn} \\
&= g^{mn}\nabla_i\nabla_m h_{jn} + g^{mn}\left(h_{mj}h_{ik} - h_{mk}h_{ij}\right)g^{k\ell}h_{\ell n} + g^{mn}\left(h_{mn}h_{ik} - h_{mk}h_{in}\right)g^{k\ell}h_{\ell j} \\
&= g^{mn}\nabla_i\nabla_j h_{mn} + g^{mn}\left(h_{mj}h_{ik} - h_{mk}h_{ij}\right)g^{k\ell}h_{\ell n} + g^{mn}\left(h_{mn}h_{ik} - h_{mk}h_{in}\right)g^{k\ell}h_{\ell j}
\end{aligned}$$

After expanding and canceling the second term with the fifth:

$$\begin{aligned}
\Delta_{M_t} h_{ij} &= g^{mn}\nabla_i\nabla_j h_{mn} - g^{mn}h_{mk}h_{ij}g^{k\ell}h_{\ell n} + g^{mn}h_{mn}h_{ik}g^{k\ell}h_{\ell j} \\
&= \nabla_i\nabla_j H - |A|^2 h_{ij} + H h_{ik}g^{k\ell}h_{\ell j} \qquad (\dagger)
\end{aligned}$$

Moving on to the time derivative, we see that the product rule followed by the usual trick of commuting F's derivatives to pull the time derivative in give us:

$$\frac{\partial}{\partial t}h_{ij} = -\frac{\partial}{\partial t}\left(\nu\cdot\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}F\right) = -\frac{\partial}{\partial t}\nu\cdot\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}F - \nu\cdot\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}\frac{\partial}{\partial t}F$$

so by (2.3.3.a) and the fact that our surfaces move by mean curvature:

$$\begin{aligned}
\frac{\partial}{\partial t}h_{ij} &= -\left(g^{\ell m}\left(\frac{\partial}{\partial x^\ell}H\right)\frac{\partial}{\partial x^m}F\right)\cdot\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}F + \nu\cdot\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}(H\nu) \\
&= \left(g^{\ell m}\left(\frac{\partial}{\partial x^\ell}H\right)\frac{\partial}{\partial x^m}F\right)\cdot\left(h_{ij}\nu - \Gamma^k_{ij}\frac{\partial}{\partial x^k}F\right) + \nu\cdot\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}(H\nu)
\end{aligned}$$

The $h_{ij}\nu$ term drops because we're dotting with tangent vectors, and two terms fall on the other side once we expand the second derivative of $H\nu$, because (first) derivatives of ν are themselves normal to ν. Summarizing:

$$\begin{aligned}
\frac{\partial}{\partial t}h_{ij} &= -\left(g^{\ell m}\left(\frac{\partial}{\partial x^\ell}H\right)\frac{\partial}{\partial x^m}F\right)\cdot\left(\Gamma^k_{ij}\frac{\partial}{\partial x^k}F\right) + \nu\cdot\left(\left(\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}H\right)\nu + H\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}\nu\right) \\
&= -\Gamma^k_{ij}g^{\ell m}g_{mk}\frac{\partial}{\partial x^k}H + \nu\cdot\left(\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}H\right)\nu + H\nu\cdot\left(\frac{\partial}{\partial x^i}\left(g^{\ell m}h_{j\ell}\frac{\partial}{\partial x^m}F\right)\right) \\
&= \frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}H - \Gamma^k_{ij}\frac{\partial}{\partial x^k}H + H\nu\cdot\left(g^{\ell m}h_{j\ell}\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^m}F\right)
\end{aligned}$$

because all terms except the double derivative of F have dropped in the product rule above, after dotting with ν. But the second derivative of F dotted with ν gives us components of the second fundamental form, so:

$$\frac{\partial}{\partial t}h_{ij} = \nabla_i\nabla_j H - Hh_{j\ell}g^{\ell m}h_{mi}$$

Subtracting (†) gives the required result. □

Before proceeding to computing curvature evolutions we need the following evolution lemmas; the first one is true in complete generality while the second one is originally due to Hamilton [**Ham82**] in the case of Ricci flow, and was later adapted to mean curvature flow by Huisken [**Hui84**]. One may safely skip the proof of these lemmas, since they are but a matter of lengthy tensor gymnastics.

LEMMA 2.3.5 (Evolution of Christoffel Symbols). For *any* smooth flow the Christoffel symbols of the Levi-Civita connection evolve according to:

$$\frac{\partial}{\partial t}\Gamma^k_{ij} = \left(\nabla_i\left(\frac{\partial}{\partial t}g\right)_{j\ell} + \nabla_j\left(\frac{\partial}{\partial t}g\right)_{i\ell} - \nabla_\ell\left(\frac{\partial}{\partial t}g\right)_{ij}\right)$$

PROOF. We differentiate the Levi-Civita connection/Christoffel symbol formula in time and use the product rule:

$$\begin{aligned}\frac{\partial}{\partial t}\Gamma^k_{ij} &= \frac{1}{2}\left(\frac{\partial}{\partial t}g^{k\ell}\right)\left(\frac{\partial}{\partial x^i}g_{j\ell} + \frac{\partial}{\partial x^j}g_{i\ell} - \frac{\partial}{\partial x^\ell}g_{ij}\right) + \frac{1}{2}g^{k\ell}\left(\frac{\partial}{\partial x^i}\frac{\partial}{\partial t}g_{j\ell} + \frac{\partial}{\partial x^j}\frac{\partial}{\partial t}g_{i\ell} - \frac{\partial}{\partial x^\ell}\frac{\partial}{\partial t}g_{ij}\right)\\ &= \left(\frac{\partial}{\partial t}g^{k\ell}\right)\Gamma^m_{ij}g_{m\ell} + \frac{1}{2}g^{k\ell}\left(\frac{\partial}{\partial x^i}\frac{\partial}{\partial t}g_{j\ell} + \frac{\partial}{\partial x^j}\frac{\partial}{\partial t}g_{i\ell} - \frac{\partial}{\partial x^\ell}\frac{\partial}{\partial t}g_{ij}\right)\end{aligned}$$

Now the terms in the latter parenthesis we can write in terms of the connection (i.e. what we want to have appear in the right hand side) and then subtract off the balancing Christoffel symbol terms. After cancellations, we get:

$$\begin{aligned}\frac{\partial}{\partial t}\Gamma^k_{ij} &= \left(\frac{\partial}{\partial t}g^{k\ell}\right)\Gamma^m_{ij}g_{m\ell} + \frac{1}{2}g^{k\ell}\left(\Gamma^m_{ij}\frac{\partial}{\partial t}g_{m\ell} + \Gamma^m_{ji}\frac{\partial}{\partial t}g_{m\ell}\right) + \frac{1}{2}g^{k\ell}\left(\nabla_i\left(\frac{\partial}{\partial t}g\right)_{j\ell} + \nabla_j\left(\frac{\partial}{\partial t}g\right)_{i\ell} - \nabla_\ell\left(\frac{\partial}{\partial t}g\right)_{ij}\right)\\ &= \left(\frac{\partial}{\partial t}g^{k\ell}\right)\Gamma^m_{ij}g_{m\ell} + g^{k\ell}\Gamma^m_{ij}\frac{\partial}{\partial t}g_{m\ell} + \frac{1}{2}g^{k\ell}\left(\nabla_i\left(\frac{\partial}{\partial t}g\right)_{j\ell} + \nabla_j\left(\frac{\partial}{\partial t}g\right)_{i\ell} - \nabla_\ell\left(\frac{\partial}{\partial t}g\right)_{ij}\right)\end{aligned}$$

and the result follows since the first two terms add up to zero, after the product rule. □

We proceed with Hamilton's lemma. For tensors σ, τ the expression $\sigma * \tau$ will denote any linear combination of σ, τ with products contracted using the metric.

LEMMA 2.3.6 (Hamilton's Tensor Evolution Lemma). Suppose that σ, τ are time dependent covariant tensors on M^n evolving per:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\tau = \sigma$$

at least in neighborhoods where we have flow by mean curvature. Then for those same points:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\nabla\tau = A * A * \nabla\tau + A * \nabla A * \tau + \nabla\sigma$$

PROOF. The previous lemma (2.3.5) and the metric evolution formula (2.3.2.a) give:

$$\frac{\partial}{\partial t}\Gamma_{ij}^k = -g^{k\ell}\left(\nabla_i (Hh)_{j\ell} + \nabla_j (Hh)_{i\ell} - \nabla_\ell (Hh)_{ij}\right) = A * \nabla A$$

In interchanging time derivatives with covariant derivatives on tensors all the correction terms are products of our original tensor with time derivatives of Christoffel symbols; having just computed the latter, we see that:

$$\frac{\partial}{\partial t}(\nabla\tau) = \nabla\left(\frac{\partial}{\partial t}\tau\right) + \left\{\frac{\partial}{\partial t}\Gamma_{ij}^k\right\} * \tau = \nabla\left(\frac{\partial}{\partial t}\tau\right) + A * \nabla A * \tau \qquad (\dagger_1)$$

The standard derivative interchange result and the Gauss equation (that gives $Rm = A * A$) show:

$$\Delta_{M_t}(\nabla\tau) = \nabla(\Delta_{M_t}\tau) + \nabla Rm * \tau + Rm * \nabla\tau = \Delta_{M_t}(\nabla\tau) + A * A * \nabla\tau + A * \nabla A * \tau \qquad (\dagger_2)$$

Subtracting $(\dagger_2)$ from $(\dagger_1)$:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\nabla\tau = \nabla\left(\frac{\partial}{\partial t}\tau\right) - \nabla(\Delta_{M_t}\tau) + A * \nabla A * \tau + A * A * \nabla\tau = \nabla\sigma + A * \nabla A * \tau + A * A * \nabla\tau$$

as claimed. □

PROPOSITION 2.3.7. Let (M_t) move by mean curvature in $\mathcal{U} \times I$. For all $(p,t) \in M^n \times I$ such that $F(p,t) \in \mathcal{U}$:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)|A|^2 = -2|\nabla A|^2 + 2|A|^4 \tag{2.3.7.a}$$

while for the covariant derivatives $\nabla^m A$, $m \geq 1$, we have the inequality:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)|\nabla^m A|^2 \leq -2|\nabla^{m+1}A|^2 + C_{n,m-1}\left(1 + |\nabla^m A|^2\right),\ m \geq 1 \tag{2.3.7.b}$$

where $C_{n,m}$ is a constant depending on n, m, and the values of $|A|, |\nabla A|, \ldots, |\nabla^m A|$ at our point.

PROOF. Let us first deal with the time derivative of $|A|^2$:

$$\frac{\partial}{\partial t}|A|^2 = \frac{\partial}{\partial t}\left(g^{ik}g^{j\ell}h_{ij}h_{k\ell}\right) = \left(\frac{\partial}{\partial t}g^{ik}\right)g^{j\ell}h_{ij}h_{k\ell} + g^{ik}\left(\frac{\partial}{\partial t}g^{j\ell}\right)h_{ij}h_{k\ell} + g^{ik}g^{j\ell}\left(\frac{\partial}{\partial t}h_{ij}\right)h_{k\ell} + g^{ik}g^{j\ell}h_{ij}\left(\frac{\partial}{\partial t}h_{k\ell}\right)$$

We know how to differentiate the parenthetical expressions by (2.3.2.b), (2.3.4.a). The first two terms represent the evolution of the metric, while the latter two represent that of A. After cancellations we are left with:

$$\frac{\partial}{\partial t}|A|^2 = 2|A|^4 + \langle h_{ij}, \Delta_{M_t} h_{ij}\rangle = 2|A|^4 + \langle A, \Delta_{M_t} A\rangle \tag{$\dagger_1$}$$

Let us now compute $\Delta_{M_t}|A|^2$. For simplicity let us perform this computation in normal coordinates. We adopt the convention that we sum over repeated indices. By the product rule:

$$\begin{aligned}\Delta_{M_t}|A|^2 &= \nabla_m\nabla_m\left(h_{ij}h_{ij}\right) = 2\nabla_m\left(h_{ij}\nabla_m h_{ij}\right)\\ &= 2\nabla_m h_{ij}\nabla_m h_{ij} + 2h_{ij}\nabla_m\nabla_m h_{ij} = 2|\nabla A|^2 + 2\left\langle A, \Delta_{M_t} A\right\rangle\end{aligned} \tag{$\dagger_2$}$$

Now (2.3.7.a) follows by subtracting ($\dagger_2$) from ($\dagger_1$).

To show (2.3.7.b) we first observe that for all $m \geq 0$:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \nabla^m A = \sum_{i+j+k=m} \nabla^i A * \nabla^j A * \nabla^k A * \nabla^m A \tag{$\dagger_3$}$$

With all the tools we have at hand now, proving ($\dagger_3$) is a simple induction; the base case $m = 0$ is just the second fundamental form evolution (2.3.4) and the inductive step is given precisely by Hamilton's lemma (2.3.6).

Let's proceed to compute the time derivative of $|\nabla^m A|^2$. We still use the product rule as in $m = 0$, after which half the terms will now come purely from the evolution of $\nabla^m A$ and the other half come from the evolution of the metric which we know varies according to $A * A$. In particular:

$$\frac{\partial}{\partial t}|\nabla^m A|^2 = 2\left\langle \nabla^m A, \frac{\partial}{\partial t}\nabla^m A\right\rangle + \sum_{i+j+k=m} \nabla^i A * \nabla^j A * \nabla^k A * \nabla^m A$$

Substituting ($\dagger_3$) and absorbing all products of tensors into the last sum:

$$\frac{\partial}{\partial t}|\nabla^m A|^2 = 2\left\langle \nabla^m A, \Delta_{M_t}\left(\nabla^m A\right)\right\rangle + \sum_{i+j+k=m} \nabla^i A * \nabla^j A * \nabla^k A * \nabla^m A \tag{$\dagger_4$}$$

We can also compute $\Delta_{M_t}|\nabla^m A|^2$ essentially as we did in the case $m = 0$. Using normal coordinates and multi-index notation observe that:

$$\begin{aligned}\Delta_{M_t}|\nabla^m A|^2 &= \nabla_\alpha \nabla_\alpha |\nabla^m A|^2 = 2\nabla_\alpha \left\langle \nabla^m A, \nabla_\alpha \nabla^m A\right\rangle \\ &= 2\left\langle \nabla_\alpha \nabla^m A, \nabla_\alpha \nabla^m A\right\rangle + 2\left\langle \nabla^m A, \nabla_\alpha \nabla_\alpha \nabla^m A\right\rangle = 2|\nabla^{m+1} A|^2 + 2\left\langle \nabla^m A, \Delta_{M_t}\left(\nabla^m A\right)\right\rangle\end{aligned} \tag{$\dagger_5$}$$

Subtracting ($\dagger_4$) and ($\dagger_5$) gives precisely:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) |\nabla^m A|^2 = -2|\nabla^{m+1} A|^2 + \sum_{i+j+k=m} \nabla^i A * \nabla^j A * \nabla^k A * \nabla^m A$$

To get the required inequality (2.3.7.b) for $m \geq 1$ we just capitalize on Hölder's inequality:

$$\begin{aligned}
\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)|\nabla^m A|^2 &= -2|\nabla^{m+1}A|^2 + \sum_{i+j+k=m} \nabla^i A * \nabla^j A * \nabla^k A * \nabla^m A \\
&= -2|\nabla^{m+1}A|^2 + \sum_{\substack{i+j+k=m \\ i,j,k<m}} \nabla^i A * \nabla^j A * \nabla^k A * \nabla^m A + A * A * \nabla^m A * \nabla^m A \\
&\leq -2|\nabla^{m+1}A|^2 + C_{n,m-1}|\nabla^m A| + C_{n,0}|\nabla^m A|^2 \\
&\leq -2|\nabla^{m+1}A|^2 + C_{n,m-1}\left(1 + |\nabla^m A|^2\right)
\end{aligned}$$

and the result follows. $\square$

2.4. Parabolic Rescaling

One tool that will prove to be remarkably useful throughout the remainder of this work is that of dilating solutions to our mean curvature flow problem. Dilations are used frequently in simple rescaling arguments and also in blow up arguments, both of which are encountered in this work. Unfortunately we cannot just dilate the space component of our manifold without breaking motion by mean curvature. Instead, we rescale parabolically by dilating both space and time appropriately. In other words, we rescale our (x,t) coordinate system according to:

$$\left\{\begin{array}{lcl} x & = & \lambda y + x_0 \\ t & = & \lambda^2 s + t_0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{lcl} y & = & \lambda^{-1}(x - x_0) \\ s & = & \lambda^{-2}(t - t_0) \end{array}\right\}$$

and turn it into a new (y,s) coordinate system which is centered at the point $x_0 \in \mathbb{R}^{n+1}$, and the time $t_0 \in \mathbb{R}$, with a dilation factor $\lambda > 0$. This gives rise to a new family of submanifolds $\left(M_s^{(x_0,t_0),\lambda}\right)$:

$$M_s^{(x_0,t_0),\lambda} := \frac{1}{\lambda}\left(M_{\lambda^2 s + t_0} - x_0\right)$$

which according to the following proposition also moves by mean curvature in the time variable s.

PROPOSITION 2.4.1. Let (M_t) move by mean curvature in $\mathcal{U} \times I$. For any $x_0 \in \mathbb{R}^{n+1}$, $t_0 \in \mathbb{R}$, and dilating factor $\lambda > 0$, the parabolically rescaled family $\left(M_s^{(x_0,t_0),\lambda}\right)$ moves by mean curvature in $\lambda^{-1}(\mathcal{U} - x_0) \times \lambda^{-2}(I - t_0)$.

PROOF. This is a straightforward differentiation result. Suppose that the (M_t) are given by $F : M^n \times I \to \mathbb{R}^{n+1}$, and consider $G : M^n \times \left(\lambda^{-2}(I - t_0)\right) \to \mathbb{R}^{n+1}$ given by:

$$G(p,s) := \lambda^{-1}\left(F\left(p, \lambda^2 s + t_0\right) - x_0\right)$$

Evidently the mean curvature vector $H_G(p,s)$ of $M_s^{(x_0,t_0),\lambda}$ at $G(p,t)$ is $\lambda \cdot H_F\left(p, \lambda^2 s + t_0\right)$, where H_F denotes the mean curvature vectors of the original submanifolds. Therefore, at each point $p \in M^n$ such that $F\left(p, \lambda^2 s + t_0\right) \in \mathcal{U}$, or equivalently $G(p,s) \in \lambda^{-1}(\mathcal{U} - x_0)$, we have:

$$\frac{\partial}{\partial s}G(p,s) = \lambda\,\frac{\partial}{\partial t}F\left(p, \lambda^2 s + t_0\right) = \lambda H_F\left(p, \lambda^2 s + t_0\right) = H_G(p,s)$$

and the result follows. $\square$

2.5. Weak Maximum Principles

In this section we establish a localized weak maximum principle that will later be used to establish curvature bounds to all orders. We first prove the following lemma from analysis.

LEMMA 2.5.1. For any non-negative $\varphi \in \mathcal{C}_c^2(\mathbb{R}^n)$ and any $x \in \mathbb{R}^n$ at which $\varphi(x) \neq 0$ we have:

$$\frac{|D\varphi(x)|^2}{\varphi(x)} \leq 2\left|D^2\varphi\right|$$

PROOF. Let $\varepsilon > 0$ be arbitrary. Observe that

$$\frac{|D\varphi|^2}{\varphi + \varepsilon} \in \mathcal{C}_c^1(\mathbb{R}^n)$$

and hence this function has a global maximum, x_0, where of course the gradient must vanish:

$$\frac{2}{\varphi(x_0)+\varepsilon} D^2\varphi_{x_0} D\varphi(x_0) - \frac{|D\varphi(x_0)|^2}{(\varphi(x_0)+\varepsilon)^2} D\varphi(x_0) = 0$$

Thus, for an arbitrary $x \in \mathbb{R}^n$ we have, in view of the fact that x_0 was the global maximum and $|D^2\varphi|$ is an operator norm:

$$\frac{|D\varphi(x)|^2}{\varphi(x)+\varepsilon} \leq \frac{|D\varphi(x_0)|^2}{\varphi(x_0)+\varepsilon} \leq 2\left|D^2\varphi_{x_0}\right| \leq 2\left|D^2\varphi\right|$$

Let us now fix a $\delta > 0$, and let $\varepsilon > 0$ be arbitrary as above. In the set $\left\{\varphi \geq \frac{\varepsilon}{\delta}\right\}$ we have:

$$\frac{|D\varphi|^2}{\varphi} = \frac{|D\varphi|^2}{\varphi+\varepsilon}\left(1 + \frac{\varepsilon}{\varphi}\right) \leq 2\left|D^2\varphi\right|(1+\delta)$$

Since $\varepsilon > 0$ was arbitrary, we conclude that the estimate above is true in $\{\varphi > 0\}$. The result now follows in view of the fact that $\delta > 0$ was also arbitrary. □

The following lemma we will refer to as a localized weak maximum principle (per Ecker **[Eck04]**). We will state it in complete generality although we will only need a special case of it. Notice that we state the theorem for intrinsic functions f, but since we're only using intrinsic derivatives (see Chapter A of the appendix) the same result does apply to extrinsic functions just as well, provided we have joint differentiability.

LEMMA 2.5.2 (Localized Weak Maximum Principle). Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, and let $\mathcal{O}$ be the subset of the background product manifold $M^n \times (t_1, t_0)$ in which mean curvature motion occurs. Let $f : \mathcal{O} \to \mathbb{R}$ be non-negative, once differentiable in time and twice differentiable in space, satisfying:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq -\delta f^2 + d(p,t) f + a(p,t) \cdot \nabla f + K$$

for $d(p,t) \geq 0$, $|a(p,t)| \leq a_0 \sqrt{1 + d(p,t)}$, and constants $a_0 \geq 0$, $\delta > 0$, and K. Let $\varphi \in \mathcal{C}^2\left(\mathcal{U} \times (t_1, t_0)\right)$ be non-negative and such that $M_t \cap \text{support } \varphi(\cdot, t)$ is a compact subset of $\mathcal{U}$ for each $t \in (t_1, t_0)$. If

$$|\varphi| + |D\varphi| + |D^2\varphi| + |D_t \varphi| \leq c_\varphi$$

in $\mathcal{U} \times (t_1, t_0)$, for some constant c_φ, then for all $t < T$ in (t_1, t_0) we have:

$$\max_{M_T} f\varphi \leq \max_{M_t} f\varphi + C(n) c_\varphi (1 + |K|)(1 + a_0^2)(1 + 1/\delta)$$

for some dimensionality constant $C(n) > 0$.

PROOF. All subsets that are mentioned, e.g. $\mathcal{U} \times (t_1, t_0)$ or $\{\varphi > 0\}$, will be thought of as being pulled back to subsets of $\mathcal{O}$ and φ will be thought of as being precomposed with $F : \mathcal{O} \to \mathcal{U}$. From (2.3.1) we get:

$$\left|\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \varphi\right| \leq C_1(n) c_\varphi$$

in $\mathcal{U} \times (t_1, t_0)$. By the product rule (A.3.3.a) and the inequality above we estimate:

$$\begin{aligned}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(f\varphi) &= \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \cdot \varphi + f \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \varphi - 2\nabla f \cdot \nabla \varphi \\ &\leq \left(-\delta f^2 - df + a \cdot \nabla f + K\right) \varphi + C_1(n) c_\varphi f - 2\nabla f \cdot \nabla\varphi \end{aligned} \qquad (\dagger_1)$$

Evidently in the set $\{\varphi > 0\}$ we have:

$$-2\nabla f \cdot \nabla \varphi = -\frac{2}{\varphi} \nabla\varphi \cdot \nabla(f\varphi) + 2\frac{|\nabla\varphi|^2}{\varphi} f \leq -\frac{2}{\varphi}\nabla\varphi \cdot \nabla(f\varphi) + 4c_\varphi f \qquad (\dagger_2)$$

where we've used the freshly established calculus estimate (2.5.1) on the right hand side. Similarly:

$$(a \cdot \nabla f)\,\varphi = a \cdot \nabla\,(f\varphi) - (a \cdot \nabla\varphi)\,f \tag{$\dagger_3$}$$

Finally, by Cauchy-Schwarz and the same calculus estimate we have in $\{\varphi > 0\}$ that:

$$\begin{aligned}|a \cdot \nabla\varphi| \leq |a||\nabla\varphi| = 2 \cdot \sqrt{1+d}\sqrt{\varphi} \cdot \frac{1}{2}\frac{|a|}{\sqrt{1+d}}\frac{|\nabla\varphi|}{\sqrt{\varphi}} &\leq 2 \cdot \sqrt{1+d}\sqrt{\varphi} \cdot \frac{1}{2}a_0\sqrt{2|D^2\varphi|} \\ &\leq 2 \cdot \sqrt{1+d}\sqrt{\varphi} \cdot \frac{1}{\sqrt{2}}\sqrt{c_\varphi}a_0 \leq (1+d)\varphi + \frac{1}{2}c_\varphi a_0^2\end{aligned} \tag{$\dagger_4$}$$

Substituting ($\dagger_3$), then ($\dagger_2$), then ($\dagger_4$) into ($\dagger_1$), we see that in $\{\varphi > 0\}$ we have:

$$\begin{aligned}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(f\varphi) &\leq \left(-\delta f^2 - df + a \cdot \nabla f + K\right)\varphi + C_1(n)c_\varphi f - 2\nabla f \cdot \nabla\varphi \\ &\leq -\delta f^2\varphi - df\varphi + a \cdot \nabla(f\varphi) - (a \cdot \nabla\varphi)\,f + K\varphi + C_1(n)c_\varphi f - 2\nabla f \cdot \nabla\varphi \\ &\leq -\delta f^2\varphi - df\varphi + a \cdot \nabla(f\varphi) - (a \cdot \nabla\varphi)\,f + K\varphi + C_1(n)c_\varphi f - \frac{2}{\varphi}\nabla\varphi \cdot \nabla(f\varphi) + 4c_\varphi f \\ &\leq -\delta f^2\varphi - df\varphi + a \cdot \nabla(f\varphi) + (1+d)f\varphi + \frac{1}{4}c_\varphi a_0^2 f + K\varphi + C_1(n)c_\varphi f - \frac{2}{\varphi}\nabla\varphi \cdot \nabla(f\varphi) + 4c_\varphi f \\ &= -\delta f^2\varphi + f\varphi + a \cdot \nabla(f\varphi) + \frac{1}{4}c_\varphi a_0^2 f + K\varphi + C_1(n)c_\varphi f - \frac{2}{\varphi}\nabla\varphi \cdot \nabla(f\varphi) + 4c_\varphi f\end{aligned}$$

After grouping like terms and using $|\varphi| \leq c_\varphi$ once more, we conclude that:

$$\begin{aligned}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(f\varphi) &\leq -\delta f^2\varphi + \left(a - \frac{2}{\varphi}\nabla\varphi\right) \cdot \nabla(f\varphi) + c_\varphi f + \frac{1}{4}c_\varphi a_0^2 f + |K|c_\varphi + C_1(n)c_\varphi f + 4c_\varphi f \\ &\leq -\delta f^2\varphi + \left(a - \frac{2}{\varphi}\nabla\varphi\right) \cdot \nabla(f\varphi) + C_2(n)c_\varphi(1+|K|)(1+a_0^2)(1+f)\end{aligned} \tag{$\ddagger$}$$

in the set $\{\varphi > 0\}$, for a possibly new constant $C_2(n)$.

Consider the function $m : (t_1, t_0) \to \mathbb{R}$ defined by:

$$m(t) := \max_{M_t} f\varphi$$

It is obviously well defined since $M_t \cap \text{support}\,\varphi(\cdot,t)$ is a compact subset of $\mathcal{U}$ for each $t \in (t_1,t_0)$, meaning the supremum is attained. Let $t < T$ be arbitrary in (t_1,t_0), and suppose for now that $m(t) < m(T)$. Let t^* be the first time in $(t,T]$ at which $m(t^*) = m(T)$, and thus $m < m(t^*)$ on (t,t^*). Let $x^* \in M_{t^*} \cap \mathcal{U}$ be the point at which the maximum is attained and let $p^* \in \mathcal{O}$ be its preimage under $F(\cdot,t^*)$. From calculus and the no-boundary assumption on our manifolds we know that:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(f\varphi)(p^*,t^*) \geq 0 \quad \text{and} \quad \nabla(f\varphi)(p^*,t^*) = 0$$

Plugging into (‡) and multiplying by $\varphi(p^*,t^*)$, we conclude that:

$$\begin{aligned}
\delta f(p^*,t^*)^2\varphi(p^*,t^*)^2 &\leq C_2(n)c_\varphi(1+|K|)(1+a_0^2)\left(1+f(p^*,t^*)\right)\varphi(p^*,t^*) \\
&= C_2(n)c_\varphi(1+|K|)(1+a_0^2)\varphi(p^*,t^*) + C_2(n)c_\varphi(1+|K|)(1+a_0^2)f(p^*,t^*)\varphi(p^*,t^*) \\
&= C_2(n)c_\varphi^2(1+|K|)(1+a_0^2) + C_2(n)c_\varphi(1+|K|)(1+a_0^2)f(p^*,t^*)\varphi(p^*,t^*)
\end{aligned}$$

where we have used $\varphi(p^*,t^*) \leq c_\varphi$. Applying $2ab \leq a^2+b^2$ on the second term to break off the $f\varphi$:

$$\begin{aligned}
\delta f(p^*,t^*)^2\varphi(p^*,t^*)^2 &\leq C_2(n)c_\varphi^2(1+|K|)(1+a_0^2) + C_2(n)^2c_\varphi^2(1+|K|)^2(1+a_0^2)^2\frac{1}{\delta} + \frac{\delta}{2}f(p^*,t^*)^2\varphi(p^*,t^*)^2 \\
&\leq C_3(n)c_\varphi^2(1+|K|)^2(1+a_0^2)^2(1+1/\delta) + \frac{\delta}{2}f(p^*,t^*)^2\varphi(p^*,t^*)^2
\end{aligned}$$

for a possibly different constant $C_3(n)$. Rearranging and multiplying by $\frac{2}{\delta}$, and recalling that (p^*,t^*) was the point at which $m(t^*)$ attained its maximum, we see that:

$$m(t^*)^2 = f(p^*,t^*)^2\varphi(p^*,t^*)^2 \leq C_3(n)c_\varphi^2(1+|K|)^2(1+a_0^2)^2(1+1/\delta)\frac{2}{\delta} \leq C_4(n)c_\varphi^2(1+|K|)^2(1+a_0^2)^2(1+1/\delta)^2$$

so that $m(t^*) \leq C(n)c_\varphi(1+|K|)(1+a_0^2)(1+1/\delta)$ for a constant $C(n)$.

That is, for $t < T$ in (t_1,t_0) such that $m(t) < m(T)$ we have $m(t^*) \leq C(n)c_\varphi(1+|K|)(1+a_0^2)(1+1/\delta)$, where t^* is the first time in $(t,T]$ such that $m(t^*) = m(T)$. In other words, when $m(t) < m(T)$ for some $t < T$ we conclude that $m(T) \leq C(n)c_\varphi(1+|K|)(1+a_0^2)(1+1/\delta)$. On the other hand, when $m(t) < m(T)$ fails for a pair $t < T$, we just have

$m(T) \leq m(t)$ instead. In any case, then, $m(T)$ is bounded from above by the sum of these two non-negative separate upper bounds, and the result follows. $\square$

We conclude this section with a brief mention of other maximum principles that hold for compact sets. They will not be needed in the sequel, so we only only provide sketches of their proofs.

LEMMA ((Very) Weak Maximum Principle). Let N be a compact manifold, and $f : N \times [t_1, t_0] \to \mathbb{R}$ be continuous. Suppose that for every $t \in (t_1, t_0]$ we have:

$$\frac{\partial}{\partial t} f(p, t) \geq 0 \text{ and the derivative exists}$$

for every $p \in N$ that attains the minimum of $f(\cdot, t)$. Then the minimum of f on N is non-decreasing in time.

PROOF. (Sketch.) It evidently suffices to fix $t_* < t^*$ in $(t_1, t_0]$ and show that $\min_{N \times [t_*, t^*]} f = \min_N f(\cdot, t_*)$.

Let us first assume that the time derivative in the statement is strictly positive. We show that the global minimum in the left hand side cannot be attained at a $(p, t) \in N \times (t_*, t^*]$. Indeed if that were the case then the time derivative would be positive at that p (by assumption) which contradicts that the time derivative would have to be non-negative since we're working at a minimum. In the general case where the inequality is not strict, add a term $\varepsilon(t - t_1)$ to f to make it be strict, and then let $\varepsilon \searrow 0$. $\square$

A corollary of this is the following interesting grow-apart lemma.

LEMMA. The distance between two smooth solutions $(M_t^1), (M_t^2)$ to mean curvature flow is non-decreasing in time.

PROOF. (Sketch.) Let the corresponding immersions be $F_i : M^i \times [t_1, t_0] \to \mathbb{R}^{n+1}$ for $i = 1, 2$. Set $N := M^1 \times M^2$ and $f(p_1, p_2, t) := |F_1(p_1, t) - F_2(p_2, t)|^2$ and observe that at a pair of points (p_1, p_2) minimizing distance at time t we have $\frac{\partial}{\partial t} f(p_1, p_2, t) = \Delta_{M^1} h_1 + \Delta_{M^2} h_2 \geq 0$ where $h_1(p) := |F_1(p, t) - F_2(p_2, t)|^2$ and $h_2(p) := |F_1(p_1, t) - F_2(p, t)|^2$. The result follows from our (very) weak maximum principle. $\square$

The very weak maximum principle of course gives rise to the classical weak maximum principle:

LEMMA. Let (M_t) move by mean curvature flow in $[t_1, t_0]$ and suppose that $f : M^n \times [t_1, t_0]$ is differentiable once in time, twice in space, and satisfies

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq v^i \nabla_i f$$

for some (not necessarily bounded) vector field v. Then $t \mapsto \max_{M_t} f$ is non-increasing. This also holds for extrinsically defined f that are additionally jointly differentiable. □

Finally, we have the additional result that initially (compact) embedded solutions remain embedded for as long as they are smooth:

LEMMA (Preserving Embeddedness). Let $F : M^n \times [t_1, t_0) \to \mathbb{R}^{n+1}$ be a smooth family of immersions moving by mean curvature flow. If $F(\cdot, t_1) : M^n \to \mathbb{R}^{n+1}$ is an embedding, then so is $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$ for all $t \in [t_1, t_0)$.

PROOF. (Sketch.) By compactness it suffices to just show injectivity. Suppose that there was an intersection point eventually and let t_* be the first such time. Then after possibly rotating, for times near t_* and around the intersection point, our two branches of the immersed manifold $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$ can be written as graphs of u_1, u_2 satisfying:

$$D_t u_i = \sqrt{1 + |Du_i|^2}\, \mathrm{div}_{\mathbb{R}^n} \frac{Du_i}{\sqrt{1 + |Du_i|^2}}$$

for $i = 1, 2$. Then their difference function $w := u_1 - u_2$ satisfies a uniformly parabolic PDE. By the strong maximum principle for parabolic PDE **[Lie05]**, any vanishing of w would imply vanishing for all previous times which contradicts that t_* is the first time of intersection. □

CHAPTER 3

Integral Estimates

So far we have only looked at mean curvature flow from a strictly differential geometric point of view by studying the evolution of several geometric quantities associated with our manifolds such as metrics, second fundamental forms, etc. In this chapter we will change gears and study the evolution of our surfaces from a measure theoretic point of view. What lets us proceed in this direction is the area element evolution formula (2.3.2.c):

$$\frac{\partial}{\partial t}\sqrt{g} = -|H|^2\sqrt{g}$$

With this formula at hand we can study the evolution in time of integrals over our surfaces. See Chapter B of the appendix for more details on integration.

Formally we're only treating the case of smooth, compact, embedded submanifolds moving by mean curvature. However, extra care has been taken in composing the proofs of this chapter to have them continue to hold under the milder assumption that our surfaces are locally $\mathscr{H}^n$-finite and satisfy only the first variation formula (3.1.3)–even with an inequality sign "$\leq$" in place of equality. Under this hypothesis we could rid ourselves of compactness and embeddedness assumptions. This gives our whole presentation a Brakke flow twist, where the first variation formula serves the purpose of a definition. Despite all this, we chose to formally treat (3.1.3) as a proposition rather than as a definition (and subsequently have to assume compactness and embeddedness to check it) because the rest of this work treats mean curvature flow from a smooth surface standpoint.

This chapter is structured as follows. In section 3.1 we introduce the notion of test functions. In section 3.2 we introduce a particular type of test function that will play a central role in the chapter, and in section 3.3 we use it to establish forward area growth estimates and estimates for the spacetime integral of $|H|^2$. In section 3.4 we introduce a heat-kernel-like function that we then use in 3.5 to prove Huisken's monotonicity formula. Then in section 3.6 we establish upper and lower bounds on a parabolic-type Hausdorff density, and in particular conclude Brakke's clearing out lemma which asserts that a manifold will necessarily leave an open ball in which it does not have sufficiently large presence. We conclude with section 3.7, where Gauss density is introduced.

3.1. Test Functions

Before proceeding any further we need to introduce the notion of a test function, which will refer to a function that is nice enough to be integrated against our family of manifolds within our open set $\mathcal{U}$ of interest.

DEFINITION 3.1.1. Let $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ be open. We call $\varphi : \mathcal{U} \to \mathbb{R}$ a time independent test function on $\mathcal{U}$ if $\varphi \in \mathcal{C}^2_c(\mathcal{U})$. $\square$

Most of the time, of course, we will want to study time dependent test functions.

DEFINITION 3.1.2. Let $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ be open and I be an interval of times. We say that $\varphi : \mathcal{U} \times I \to \mathbb{R}$ is a (time dependent) test function on $\mathcal{U} \times I$ if $\varphi \in \mathcal{C}^1(\mathcal{U} \times I)$, for every $t \in I$ we have $\varphi(\cdot, t) \in \mathcal{C}^2(\mathcal{U})$, and for every $t \in I$ there exists a $\delta > 0$ and a compact set $K \subset \mathcal{U}$ such that support $\varphi(\cdot, t')$, support $D_t\varphi(\cdot, t') \subseteq K$ for every $t' \in I \cap B_\delta(t)$.

We will generally assume that all our test functions are time dependent. What is useful about test functions is that they have finite integrals over the M_t, since the latter are locally $\mathscr{H}^n$-finite.

The evolution equation for the area element gives us the following basic but powerful result regarding differentiation under the integral sign in the mean curvature flow case. Consult chapter B of the appendix for the proof.

THEOREM 3.1.3 (First Variation Formula). Let (M_t) move by mean curvature flow in $\mathcal{U} \times I$. If φ is a test function on $\mathcal{U}$, then:

$$\frac{d}{dt}\int_{M_t} \varphi = \int_{M_t} H \cdot D\varphi - |H|^2\varphi$$

$\square$

There are many ways in fact to rewrite the result of differentiating the integral of a test function, the most important of which are listed in this corollary. Despite its triviality, it is a really important result that will be used throughout the text. Notice that our previous result was proven to be true for time independent φ. It is crucial that we generalize this to φ that might also depend on time.

COROLLARY 3.1.4. Let (M_t) move by mean curvature in $\mathcal{U} \times I$. If φ is a test function on $\mathcal{U} \times I$, then:

$$\begin{aligned}
\frac{d}{dt}\int_{M_t}\varphi &= \int_{M_t} D_t\varphi + H\cdot D\varphi - |H|^2\varphi \\
&= \int_{M_t} D_t\varphi - \operatorname{div}_{M_t} D\varphi - |H|^2\varphi \\
&= \int_{M_t} \frac{\partial}{\partial t}\varphi - |H|^2\varphi \\
&= \int_{M_t} \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi - |H|^2\varphi \\
&= \int_{M_t} \left(\frac{\partial}{\partial t} + \Delta_{M_t}\right)\varphi - |H|^2\varphi
\end{aligned}$$

PROOF. Consider the function $f : I \times I \to \mathbb{R}$ defined by:

$$f(t,s) := \int_{M_t} \varphi(\cdot, s)$$

Notice that the first variation formula (3.1.3) tells us that:

$$D_t f(t,s) = \int_{M_t} H\cdot D\varphi(\cdot,s) - |H|^2\varphi(\cdot,s)$$

and standard differentiation under the integral sign tells us that:

$$D_s f(t,s) = \int_{M_t} D_t\varphi(\cdot,s)$$

Combining these two results, by the chain rule we have:

$$\frac{d}{dt}\int_{M_t}\varphi = \frac{d}{dt}\left(f(t,t)\right) = D_t f(t,t) + D_s f(t,t) = \int_{M_t} D_t\varphi(\cdot,t) + H\cdot D\varphi(\cdot,t) - |H|^2\varphi(\cdot,t)$$

which gives the first identity. Now, recall from (2.3.1) that:

$$\Delta_{M_t}\varphi = \operatorname{div}_{M_t} D\varphi + H\cdot D\varphi \Rightarrow H\cdot D\varphi = \Delta_{M_t}\varphi - \operatorname{div}_{M_t} D\varphi$$

The $\Delta_{M_t}\varphi$ term integrates out to zero by the divergence theorem, since M_t has no boundary in $\mathcal{U}$, and hence we get the second equality from the first since $H\cdot D\varphi$ and $\operatorname{div}_{M_t} D\varphi$ integrate out to the same thing. The third equality follows simply from (2.3.1.a), while the fourth and fifth equalities follow by subtracting or adding a Laplacian (since we know it integrates to zero). $\square$

3.2. Standard Test Function

In this section we introduce an extrinsic function that plays a dominant role in our study of mean curvature flow. Let us begin by considering the smooth function $\psi : \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R}$ given by:

$$\psi(x,t) := |x|^2 + 2nt$$

Evidently $D\psi = 2x$, and hence $\operatorname{div}_{M_t} D\psi = 2n$. At the same time, $D_t\psi = 2n$, and hence by (2.3.1) we see that:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\psi = D_t\psi - \operatorname{div}_{M_t}\psi = 0$$

Now consider the non-negative function $\varphi : \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R}$ given by:

$$\varphi_\varrho(x,t) := \left(1 - \frac{\psi(x,t)}{\varrho^2}\right)_+^3$$

where $(\cdot)_+$ denotes the non-negative part of a function. Denote its translates $\varphi_\varrho(x - x_0, t - t_0)$ by $\varphi_{(x_0,t_0),\varrho}(x,t)$.

PROPOSITION 3.2.1. Let $x_0 \in \mathbb{R}^{n+1}$, $t_0 \in \mathbb{R}$, and $\varrho > 0$. The standard parabolic test function $\varphi_{(x_0,t_0),\varrho}$ is $\mathcal{C}^2(\mathbb{R}^{n+1} \times \mathbb{R})$, and satisfies the differential inequality:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi_\varrho \le 0 \tag{3.2.1.a}$$

and the estimates:

$$0 \le \varphi_{(x_0,t_0),\varrho}(\cdot,t) \le \left(\frac{\varrho^2 - 2n(t-t_0)}{\varrho^2}\right)_+^3 \quad \text{for } t \in \mathbb{R} \tag{3.2.1.b}$$

$$\left\{\varphi_{(x_0,t_0),\varrho}(\cdot,t) > 0\right\} = \left\{D_t\varphi_{(x_0,t_0),\varrho}(\cdot,t) \neq 0\right\} = B_{(\varrho^2 - 2n(t-t_0))_+^{1/2}}(x_0) \quad \text{for } t \in \mathbb{R} \tag{3.2.1.c}$$

When $\sigma > 0$ we additionally have

$$\inf_{B_\sigma(x_0)} \varphi_{(x_0,t_0),\varrho}(\cdot,t) = \left(\frac{\varrho^2 - \sigma^2 - 2n(t-t_0)}{\varrho^2}\right)_+^3 \quad \text{for } t \in \mathbb{R} \tag{3.2.1.d}$$

PROOF. All of (3.2.1.b), (3.2.1.c), (3.2.1.d) can be shown with simple algebraic manipulations, which we omit. To establish (3.2.1.a), we see that with:

$$\eta(f) := \left(1 - \frac{f}{\varrho^2}\right)_+^3$$

the chain rule (A.3.3.b) gives:

$$\begin{aligned}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi_{(x_0,t_0),\varrho} &= \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\eta\left(\varphi_{(x_0,t_0),\varrho}\right)\\ &= \eta'\left(\varphi_{(x_0,t_0),\varrho}\right)\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi_{(x_0,t_0),\varrho} - \eta''\left(\varphi_{(x_0,t_0),\varrho}\right)\left|\nabla\varphi_{(x_0,t_0),\varrho}\right|^2 \leq 0\end{aligned}$$

because the first term vanishes (the second factor is identically zero) and the second term is non-positive. □

This support estimate enables us to treat $\varphi_{(x_0,t_0),\varrho}$ as a test function on $\mathcal{U} \times I$, as long as $\varrho^2 - 2n(t - t_0)$ is sufficiently small (e.g. ϱ is sufficiently small and $t - t_0$ is not very negative). In particular:

PROPOSITION 3.2.2. The function $\varphi_{(x_0,t_0),\varrho}$ is a test function on $B_R(x_0) \times (t_1, \infty)$ when $x_0 \in \mathbb{R}^{n+1}$, $t_0 \in \mathbb{R}$, and $\varrho^2 - 2n(t_1 - t_0) \leq R^2$.

PROOF. By definition φ is $\mathcal{C}^2$, so we need only check the part of a test function's definition pertaining to the function's support. Recall that support $\varphi(\cdot, t') = B_{(\varrho^2 - 2n(t'-t_0))_+^{1/2}}(x_0)$. Therefore for any $t > t_1$ we can choose $0 < \delta < t - t_1$ so that for $|t' - t| < \delta$ we have:

$$\text{support } \varphi(\cdot, t') = \text{support } D_t\varphi(\cdot, t') = \overline{B}_{(\varrho^2 - 2n(t'-t_0))_+^{1/2}}(x_0) = \overline{B}_{(\varrho^2 - 2n(-\delta + t - t_1 + t_1 - t_0))_+^{1/2}}(x_0)$$

which is indeed compact and contained in $B_{(\varrho^2 - 2n(t_1 - t_0))_+^{1/2}}(x_0) \subseteq B_R(x_0)$, and hence the result follows. □

One of the first powerful applications of this test function is in the proof of a lower bound on the speed of convergence of manifolds to points that they reach in the ambient space, which will later be essential to our proof of Brakke's clearing out lemma. Before proceeding, we first need to establish what we mean by "points that are reached."

DEFINITION 3.2.3. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, and let $x_0 \in \mathcal{U}$. We say that the our manifolds (M_t) reach x_0 at time t_0, or that x_0 is reached by (M_t) at time t_0, if there exist sequences $t_j \nearrow t_0$ and $x_j \in M_{t_j} \cap \mathcal{U}$ such that $x_j \to x_0$. □

We establish that our surfaces converge to the points that they reach with at least a "square root" speed.

PROPOSITION 3.2.4. Let (M_t) move by mean curvature in $B_\varrho(x_0) \times \left(t_0 - \frac{1}{2n+1}\varrho^2, t_0\right)$, and suppose that x_0 is reached by (M_t) at t_0. Then:

$$\mathscr{H}^n\left(M_t \cap B_{\sqrt{(2n+1)(t_0-t)}}(x_0)\right) > 0$$

for every $t \in \left(t_0 - \frac{1}{2n+1}\varrho^2, t_0\right)$.

PROOF. Suppose, for the sake of contradiction, that there did exist some $T_0 \in \left(t_0 - \frac{1}{2n+1}\varrho^2, t_0\right)$ for which:

$$\mathscr{H}^n\left(M_{T_0} \cap B_{\sqrt{(2n+1)(t_0-T_0)}}(x_0)\right) = 0 \tag{‡}$$

Define $\varrho_0 := \sqrt{(2n+1)(t_0 - T_0)}$, and observe that $\varrho_0 \leq \varrho$. The standard test function $\varphi_{(x_0,T_0),\varrho_0}$ is a test function on $B_\varrho(x_0) \times (T_0, \infty)$ by (3.2.2). By the differentiation formula (3.1.4) and (3.2.1.a) on $\varphi_{(x_0,t_0),\varrho_0}$ we obtain:

$$\frac{d}{dt}\int_{M_t} \varphi_{(x_0,T_0),\varrho_0} \leq 0$$

for every $t \in (T_0, t_0)$. At the same time by (3.2.1.c) we have $\{\varphi_{(x_0,T_0),\varrho_0}(\cdot, T_0) \neq \emptyset\} = B_{\varrho_0}(x_0)$, so:

$$\int_{M_{T_0}} \varphi_{(x_0,T_0),\varrho_0} = \int_{M_{T_0}\cap B_{\varrho_0}(x_0)} \varphi_{(x_0,T_0),\varrho_0} = 0$$

in view of (‡). Consequently, the established monotonicity yields that

$$\int_{M_t} \varphi_{(x_0,T_0),\varrho_0} = 0 \quad \text{and hence } \mathscr{H}^n\left(M_t \cap B_{\sqrt{\varrho_0^2 - 2n(t-T_0)}}(x_0)\right) = 0$$

for every $t \in [T_0, t_0)$, since $\{\varphi_{(x_0,T_0),\varrho_0}(\cdot, t) \neq 0\} = B_{\sqrt{\varrho_0^2-2n(t-T_0)}}(x_0)$ by (3.2.1.c). What's more, by definition of ϱ_0 we have $0 < t_0 - T_0 = \varrho_0^2 - 2n(t_0 - T_0)$, so:

$$\mathscr{H}^n\left(M_t \cap B_{\sqrt{t_0-T_0}}(x_0)\right) = \mathscr{H}^n\left(M_t \cap B_{\sqrt{\varrho_0^2-2n(t_0-T_0)}}(x_0)\right) \leq \mathscr{H}^n\left(M_t \cap B_{\sqrt{\varrho_0^2-2n(t-T_0)}}(x_0)\right) = 0$$

for every $t \in [T_0, t_0)$. Since the M_t are smooth surfaces, this in turn implies that $M_t \cap B_r(x_0) = \emptyset$ for all $t \in [T_0, t_0)$, with $r := \sqrt{t_0 - T_0} > 0$. This contradicts that x_0 is reached at time t_0, and the result follows. $\square$

3.3. Area and $|H|^2$ Estimates

The standard parabolic test function allows us to prove our first fundamental localization result, which lets us bound forward-time areas and spacetime $|H|^2$ integral estimates by older time areas. In particular, we can prove that:

THEOREM 3.3.1. Let (M_t) move by mean curvature in $B_{\varrho_0}(x_0) \times \left(t_0 - \frac{1}{2n}\varrho_0^2, t_0\right)$. Fix $\theta_1, \theta_2 > 0$ such that $\theta_1^2 + 2n\theta_2^2 < 1$. Then there exists a constant $C = C(\theta_1, \theta_2)$ such that for all $0 < \varrho \le \varrho_0$:

$$\mathscr{H}^n(M_t \cap B_{\theta_1 \varrho}(x_0)) + \int_{t_0 - \theta_2^2 \varrho^2}^{t} \int_{M_\tau \cap B_{\theta_1 \varrho}(x_0)} |H|^2 \le C(n, \theta_1, \theta_2) \mathscr{H}^n(M_{t_0 - \theta_2^2 \varrho^2} \cap B_\varrho(x_0))$$

for every $t \in [t_0 - \theta_2^2 \varrho^2, t_0)$. The constant is, in fact, $C(n, \theta_1, \theta_2) := (1 - \theta_1^2 - 2n\theta_2^2)^{-3}$.

PROOF. Let us define $T_0 := t_0 - \theta_2^2 \varrho^2$, so that $t_0 = T_0 + \theta_2^2 \varrho^2$, and let us write φ for $\varphi_{(x_0, T_0), \varrho}$.

From (3.2.2) we know that φ is a test function on $B_\varrho(x_0) \times (T_0, T_0 + \theta_2^2 \varrho^2)$, and hence from the general differentiation identity (3.1.4) and the standard test function's identity (3.2.1.a), we know that:

$$\frac{d}{dt} \int_{M_t} \varphi = \int_{M_t} \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) \varphi - |H|^2 \varphi \le - \int_{M_t} |H|^2 \varphi$$

for every $t \in (T_0, T_0 + \theta_2^2 \varrho^2)$. Therefore, by the fundamental theorem of calculus on $[T_0, t]$ we have:

$$\int_{M_t} \varphi - \int_{M_{T_0}} \varphi \le - \int_{T_0}^{t} \int_{M_\tau} |H|^2 \varphi$$

and rearranging gives:

$$\int_{M_t} \varphi + \int_{T_0}^{t} \int_{M_\tau} |H|^2 \varphi \le \int_{M_{T_0}} \varphi \tag{†}$$

From (3.2.1.b) and (3.2.1.c) we can estimate the right hand side as follows:

$$\int_{M_{T_0}} \varphi \le \int_{M_{T_0} \cap \{\varphi(x, T_0) \neq 0\}} \varphi = \int_{M_{T_0} \cap B_\varrho(x_0)} \varphi \le \int_{M_{T_0} \cap B_\varrho(x_0)} \left(\frac{\varrho^2 - 2n(T_0 - T_0)}{\varrho^2} \right)_+^3 = \mathscr{H}^n(M_{T_0} \cap B_\varrho(x_0))$$

To estimate the left hand side from below we, let t be as above and $\tau \in [T_0, t]$. Then from (3.2.1.d) we have:

$$\inf_{B_{\theta_1\varrho}(x_0)} \varphi(\cdot,\tau) = \left(\frac{\varrho^2 - \theta_1^2\varrho^2 - 2n(\tau - T_0)}{\varrho^2}\right)_+^3 \geq \left(\frac{(1-\theta_1^2)\varrho^2 + 2n\theta_2^2\varrho^2}{\varrho^2}\right)_+^3 = (1-\theta_1^2+2n\theta_2^2)_+^3 = (1-\theta_1^2-2n\theta_2^2)^3$$

Therefore, we may estimate the left hand side of (†) from below by:

$$\begin{aligned}
\int_{M_t}\varphi + \int_{T_0}^t\int_{M_\tau}|H|^2\varphi &\geq \int_{M_t\cap B_{\theta_1\varrho}(x_0)}\varphi + \int_{T_0}^t\int_{M_\tau\cap B_{\theta_1\varrho}(x_0)}|H|^2\varphi \\
&\geq \int_{M_t\cap B_{\theta_1\varrho}(x_0)}(1-\theta_1^2-2n\theta_2^2)^3 + \int_{T_0}^t\int_{M_\tau\cap B_{\theta_1\varrho}(x_0)}|H|^2(1-\theta_1^2-2n\theta_2^2)^3 \\
&\geq (1-\theta_1^2-2n\theta_2^2)^3\left(\mathscr{H}^n(M_t\cap B_{\theta_1\varrho}(x_0)) + \int_{T_0}^t\int_{M_\tau\cap B_{\theta_1\varrho}(x_0)}|H|^2\right)
\end{aligned}$$

Combining these two estimates into (†) we see that:

$$\mathscr{H}^n(M_t\cap B_{\theta_1\varrho}(x_0)) + \int_{T_0}^t\int_{M_\tau\cap B_{\theta_1\varrho}(x_0)}|H|^2 \leq (1-\theta_1^2-2n\theta_2^2)^{-3}\mathscr{H}^n(M_{T_0}\cap B_\varrho(x_0))$$

which is the required result upon recalling that $T_0 = t_0 - \theta_2^2\varrho^2$. □

3.4. Backwards Heat Kernel

There is a second auxiliary function that will play a fundamental role in our monotonicity formulas, which resembles a time-reversed heat kernel for n dimensions. It is the non-negative smooth function $\Phi : \mathbb{R}^{n+1} \times (-\infty, 0) \to \mathbb{R}$:

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

As with the standard parabolic test function, we will also use the translated functions $\Phi_{(x_0,t_0)}(x,t) := \Phi(x-x_0, t-t_0)$, for $t < t_0$. Both Φ and its translates satisfy the following important equation:

PROPOSITION 3.4.1. Let Φ be the backwards heat kernel, or any of its translates. Then

$$D_t\Phi + \mathrm{div}_M D\Phi + \frac{|\nabla^\perp \Phi|^2}{\Phi} = 0$$

where the (tangential) divergence and the (normal) gradient $\nabla^\perp$ are both relative to an n-dimensional manifold M^n.

PROOF. Without loss of generality we may work with the un-translated Φ, since the result is translation invariant. We begin by expanding $D_t\Phi$:

$$\begin{aligned}
D_t\Phi &= D_t\left(\frac{1}{(-4\pi t)^{n/2}} \exp\left(\frac{|x|^2}{4t}\right)\right) \\
&= -\frac{n}{2}\frac{1}{-4\pi t}\frac{1}{(-4\pi t)^{-n/2}}(-4\pi)\exp\left(\frac{|x|^2}{4t}\right) + \left(-\frac{|x|^2}{4t^2}\right)\frac{1}{(-4\pi t)^{n/2}}\exp\left(\frac{|x|^2}{4t}\right) \\
&= -\frac{n}{2t}\Phi - \frac{|x|^2}{4t^2}\Phi
\end{aligned}$$

We now compute the (ambient) spatial derivative. For $i = 1, \ldots, n+1$:

$$D_i\Phi = \frac{1}{2t}\frac{1}{(-4\pi t)^{n/2}} \exp\left(\frac{|x|^2}{4t}\right) x_i$$

Combining this for all i, we get the following four equations for free:

$$D\Phi = \frac{1}{2t}\Phi x \quad \text{and} \quad \nabla\Phi = \frac{1}{2t}\Phi x^t \quad \text{and} \quad \nabla^\perp\Phi = \frac{1}{2t}\Phi x^\perp \quad \text{and} \quad \frac{|\nabla^\perp\Phi|^2}{\Phi} = \frac{|x^\perp|^2}{4t^2}\Phi$$

Let us finally compute the tangential divergence:

$$\begin{aligned}
\mathrm{div}_M D\Phi = \mathrm{div}_M\left(\frac{1}{2t}\Phi x\right) &= \frac{1}{2t}\nabla\Phi \cdot x + \frac{n}{2t}\Phi \\
&= \frac{1}{4t^2}\Phi x^t \cdot x + \frac{n}{2t}\Phi = \frac{|x^t|^2}{4t^2}\Phi + \frac{n}{2t}\Phi
\end{aligned}$$

Adding everything up:

$$
\begin{aligned}
D_t\Phi + \operatorname{div}_M D\Phi + \frac{|\nabla^\perp \Phi|^2}{\Phi} &= \left(-\frac{n}{2t}\Phi - \frac{|x|^2}{4t^2}\Phi\right) + \left(\frac{|x^t|^2}{4t^2}\Phi + \frac{n}{2t}\Phi\right) + \frac{|x^\perp|^2}{4t^2}\Phi \\
&= \left(\frac{|x^t|^2}{4t^2} + \frac{|x^\perp|^2}{4t^2} - \frac{|x|^2}{4t^2}\right)\Phi = 0
\end{aligned}
$$

and the result follows. □

Notice that since we have made the effort to compute the derivatives of Φ, we get the following for free:

PROPOSITION 3.4.2. Let $x_0 \in \mathbb{R}^{n+1}$ and $t_0 \in \mathbb{R}$. Then for $t < t_0$ we have:

$$\sup_{\mathbb{R}^{n+1}} \Phi_{(x_0,t_0)}(\cdot, t) = \frac{1}{(4\pi(t_0 - t))^{n/2}} \tag{3.4.2.a}$$

If additionally $\sigma > 0$, then:

$$\inf_{B_\sigma(x_0)} \Phi_{(x_0,t_0)}(\cdot, t) = \frac{1}{(4\pi(t_0 - t))^{n/2}} \exp\left(-\frac{\sigma^2}{4(t_0 - t)}\right) \tag{3.4.2.b}$$

$$\sup_{\mathbb{R}^{n+1}\setminus B_\sigma(x_0)} \Phi_{(x_0,t_0)}(\cdot, t) = \frac{1}{(4\pi(t_0 - t))^{n/2}} \exp\left(-\frac{\sigma^2}{4(t_0 - t)}\right) \tag{3.4.2.c}$$

and hence $\Phi_{(x_0,t_0)} \to 0$ uniformly away from x_0, as $t \nearrow t_0$. Additionally, for $x \in \mathbb{R}^{n+1}$, $t < t_0$:

$$\Phi_{(x_0,t_0)}(x, t) \leq C(n)|x - x_0|^{-n} \tag{3.4.2.d}$$

where $C(n) = (2n^{-1}\pi)^{-n/2} \exp\left(-\frac{n}{2}\right)$.

PROOF. Without loss of generality, we may work at $(x_0, t_0) = (0, 0)$. The first two inequalities are trivial since Φ is decreasing in $|x|$. Let us prove the last part then.

Fix $x \in \mathbb{R}^{n+1}$ and view $\Phi(x, t)$ as a function of $t < t_0$. It is evidently smooth, and we know that its derivative is:

$$D_t\Phi = -\frac{n}{2t}\Phi - \frac{|x|^2}{4t^2}\Phi$$

which is positive for $t < \frac{|x|^2}{2n}$, zero at that point, and negative afterwards, meaning $t \mapsto \Phi(x, t)$ attains its global maximum at $t = \frac{|x|^2}{2n}$. In other words:

$$\sup_{t<0} \Phi(x, t) = \frac{1}{(2n^{-1}\pi)^{n/2}|x|^n} \exp\left(-\frac{n}{2}\right)$$

which concludes the proof. □

3.5. Huisken's Monotonicity Formula

The most useful property of the backwards heat kernel is that it is used in proving monotonicity results. The monotonicity formula below was originally discovered by Huisken [**Hui90**].

THEOREM 3.5.1 (Huisken's Monotonicity Formula). Let (M_t) move by mean curvature in $\mathcal{U} \times I$. Fix $x_0 \in \mathbb{R}^{n+1}$ and $t_0 \in \mathbb{R}$. For any test function φ on $\mathcal{U} \times I$ we have:

$$\frac{d}{dt}\int_{M_t} \Phi_{(x_0,t_0)}\varphi = \int_{M_t} \Phi\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi - \left|H - \frac{\nabla^{\perp}\Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2 \Phi_{(x_0,t_0)}\varphi$$

for all $t \in I$ with $t < t_0$.

PROOF. Let us write Φ for $\Phi_{(x_0,t_0)}$ for brevity. Evidently $\Phi\varphi$ is also a test function, only on the shrunk domain $\mathcal{U} \times (I \cap (-\infty, t_0))$ instead of $\mathcal{U} \times I$. But this imposes no problems, since we are working at times $t < t_0$ anyway.

By the differentiation identity (3.1.4) and the (standard) product rule we see that:

$$\begin{aligned} \frac{d}{dt}\int_{M_t} \Phi\varphi &= \int_{M_t} \frac{\partial}{\partial t}(\Phi\varphi) - |H|^2\Phi\varphi \\ &= \int_{M_t} \frac{\partial}{\partial t}\Phi \cdot \varphi + \Phi\frac{\partial}{\partial t}\varphi - |H|^2\Phi\varphi \end{aligned} \tag{†}$$

Of course the divergence theorem tells us that:

$$\int_{M_t} \Delta_{M_t}\Phi \cdot \varphi = \int_{M_t} \Phi\Delta_{M_t}\varphi$$

In other words, we may add the term $\Delta_{M_t}\Phi \cdot \varphi - \Phi\Delta_{M_t}\varphi$ in the integral in (†) without changing the value of the integral. Consequently:

$$\begin{aligned} \frac{d}{dt}\int_{M_t} \Phi\varphi &= \int_{M_t} \frac{\partial}{\partial t}\Phi \cdot \varphi + \Delta_{M_t}\Phi \cdot \varphi + \Phi\frac{\partial}{\partial t}\varphi - \Phi\Delta_{M_t}\varphi - |H|^2\Phi\varphi \\ &= \int_{M_t} \left(\left(\frac{\partial}{\partial t} + \Delta_{M_t}\right)\Phi - |H|^2\Phi\right)\varphi + \Phi\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi \end{aligned} \tag{‡}$$

At this point observe that:

$$\left(\frac{\partial}{\partial t} + \Delta_{M_t}\right)\Phi - |H|^2\Phi = D_t\Phi + \operatorname{div}_{M_t} D\Phi + 2H\cdot D\Phi - |H|^2\Phi$$

because of the intrinsic/extrinsic function relationship results, (2.3.1.a) and (2.3.1.b). But by (3.4.1) the first two terms can be collapsed into one term as follows:

$$= -\frac{|\nabla^\perp\Phi|^2}{\Phi} + 2H\cdot D\Phi - |H|^2\Phi = -\frac{|\nabla^\perp\Phi|^2}{\Phi} + 2H\cdot\nabla^\perp\Phi - |H|^2\Phi$$

where in the last step we replaced $D\Phi$ by $\nabla^\perp\Phi$ since we're taking the dot product with H, which is normal to the tangent space. Completing the square gives:

$$= -\frac{|\nabla^\perp\Phi|^2}{\Phi} + 2\sqrt{\Phi}H\cdot\frac{1}{\sqrt{\Phi}}\nabla^\perp\Phi - |H|^2\Phi = -\left|H - \frac{\nabla^\perp\Phi}{\Phi}\right|^2\Phi$$

The result follows by plugging this back into (‡). □

An immediate corollary is:

COROLLARY 3.5.2. Let (M_t) move by mean curvature in $\mathcal{U}\times I$. Fix $x_0\in\mathbb{R}^{n+1}$ and $t_0\in\mathbb{R}$. If φ is a test function on $\mathcal{U}\times I$ such that:

$$\varphi\geq 0 \quad\text{and}\quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi\leq 0$$

then

$$\frac{d}{dt}\int_{M_t}\Phi_{(x_0,t_0)}\varphi\leq 0$$

for $t\in I$, $t<t_0$.

3.6. Density Estimates; Brakke Clearing Out Lemma

In this section we establish bounds on our manifolds' spacetime density. Our first result is an important result originally due to Brakke which shows that points that are reached (in spacetime) by our manifolds have sufficiently high density around them. In other words, it is not possible for a point $x_0 \in \mathbb{R}^{n+1}$ to be reached at some time $t_0 \in \mathbb{R}$ without there being sufficient overlap between M_t, $t < t_0$, and balls around x_0.

We start with a particularly stated version of Brakke's clearing out lemma which argues that mass disappears locally after finite time. We essentially follow Evans and Spruck **[ES92b]**, whose proof requires that $n \geq 2$. One way to prove the one dimensional case is to shrink the ball ever so slightly to exclude all non-closed curves around the edges, and then argue that all remaining curves are closed and will die sufficiently quickly (by Grayson's theorem). But in any case we are only interested in $n \geq 2$ in this work:

LEMMA 3.6.1. Let (M_t) move by mean curvature in $B_1(0) \times \left(-\frac{1}{2n+1}, 0\right)$. There exists a constant $C(n)$ such that if:

$$\mathscr{H}^n \left(M_{-\beta} \cap B_1(0)\right) \leq \eta$$

for some $\beta \in \left(0, \frac{1}{2n+1}\right)$ and $\eta \geq 0$, then:

$$\mathscr{H}^n \left(M_t \cap B_{\sqrt{1-(2n+1)\cdot(t+\beta)}}(0)\right) = 0$$

for all $t \in \left(-\frac{1}{2n+1}, 0\right)$ such that $t \geq -\beta + C(n)\eta^{\frac{1}{n+3}}$ we have:

PROOF. We will use a function similar to the standard parabolic test function, only with a minor modification. Namely, let us take:

$$\psi(x,t) := \left(1 - |x|^2 - (2n+1)(t+\beta)\right)_+ \quad \text{and} \quad \varphi := \psi^3$$

The function φ is still smooth as well as a test function on $B_1(0) \times (-\beta, \infty)$, and it satisfies:

$$\nabla\varphi = -6\psi^2 x^t \tag{$\dagger_1$}$$

and from the chain rule, (A.3.3.b):

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi = -3\psi^2 - 24\psi\left|x^t\right|^2 \tag{$\dagger_2$}$$

For $t > -\beta$ let us now define the auxiliary function:

$$G(t) := \int_{M_t} \varphi$$

By the differentiation under the integral sign identity (3.1.4) we see that:

$$G'(t) = \int_{M_t} \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi - |H|^2\varphi \overset{(\dagger_1)}{=} \int_{M_t} -3\psi^2 - 24\psi\left|x^t\right|^2 - |H|^2\varphi$$

or after rearranging:

$$\int_{M_t} 3\psi^2 + 24\psi\left|x^t\right|^2 + |H|^2\varphi = -G'(t) \tag{$\dagger_3$}$$

Let $t \geq -\beta$ be arbitrary. By Hölder's inequality applied with the conjugate exponents $p = \frac{n+2}{n-1}$, $q = \frac{n+2}{3}$, we have:

$$\begin{aligned} G(t) &= \int_{M_t} \varphi^{\frac{n}{n+2}}\varphi^{\frac{2}{n+2}} \leq \left(\int_{M_t} \varphi^{\frac{n}{n+2}\cdot\frac{n+2}{n-1}}\right)^{\frac{n-1}{n+2}} \left(\int_{M_t} \varphi^{\frac{2}{n+2}\cdot\frac{n+2}{3}}\right)^{\frac{3}{n+2}} \\ &= \left(\int_{M_t} \varphi^{\frac{n}{n-1}}\right)^{\frac{n-1}{n+2}} \left(\int_{M_t} \varphi^{\frac{2}{3}}\right)^{\frac{3}{n+2}} =: J_1(t)^{\frac{n-1}{n+2}} J_2(t)^{\frac{3}{n+2}} \end{aligned} \tag{$\dagger_4$}$$

We will estimate $J_1(t), J_2(t)$ independently of each other and then combine our results back into ($\dagger_4$). To study the left integral we employ the Michael-Simon inequality **[Sim83]**:

$$\left(\int_M f^{\frac{n}{n-1}}\right)^{\frac{n-1}{n}} \leq C(n)\int_M |\nabla f| + |H|f$$

where $f \geq 0$ and $\dim M = n$ (assumed to be $n \geq 2$). We then have that:

$$J_1(t) := \int_{M_t} \varphi^{\frac{n}{n-1}} \leq C(n)\left(\int_{M_t} |\nabla\varphi| + |H|\varphi\right)^{\frac{n}{n-1}} \overset{(\dagger_1)}{=} C(n)\left(\int_{M_t} 6\psi^2\left|x^t\right| + |H|\varphi\right)^{\frac{n}{n-1}}$$

By the Cauchy-Schwarz inequality $2ab \leq a^2 + b^2$ with $a = 3\psi^{3/2}$ and $b = \psi^{1/2}\left|x^t\right|$, and also that $\psi \leq 1$ for $t \geq -\beta$ (so that $\psi^3 \leq \psi^2$) we have:

$$J_1(t) \leq C(n)\left(\int_{M_t} 9\psi^3 + \psi\left|x^t\right|^2 + |H|\varphi\right)^{\frac{n}{n-1}} \leq C(n)\left(\int_{M_t} 9\psi^2 + \psi\left|x^t\right|^2 + |H|\varphi\right)^{\frac{n}{n-1}}$$

By the same Cauchy-Schwarz inequality on $|H|\varphi$, with $a = |H|\varphi^{1/2}$ and $b = \frac{1}{2}\varphi^{1/2} = \frac{1}{2}\psi^{3/2}$, followed by $\psi^3 \leq \psi^2$, we conclude that:

$$J_1(t) \leq C(n)\left(\int_{M_t} 9\psi^2 + \psi\left|x^t\right|^2 + |H|\varphi^2 + \frac{1}{4}\psi^2\right)^{\frac{n}{n-1}} \overset{(\dagger_3)}{\leq} C(n)(-G'(t))^{\frac{n}{n-1}}$$

for a new constant $C(n)$. As for $J_2(t)$, we immediately have:

$$J_2(t) = \int_{M_t} \varphi^{2/3} = \int_{M_t} \psi^2 \leq \int_{M_t} 3\psi^2 + 24\psi\left|x^t\right|^2 + |H|^2\varphi = -G'(t)$$

Combining these two estimates into $(\dagger_4)$ gives:

$$G(t) \leq J_1(t)^{\frac{n-1}{n+2}} J_2(t)^{\frac{3}{n+2}} \leq C(n)\left(-G'(t)\right)^{\frac{n}{n-1}\cdot\frac{n-1}{n+2}}\left(-G'(t)\right)^{\frac{3}{n+2}} = C(n)\left(-G'(t)\right)^{\frac{n+3}{n+2}}$$

for a new constant $C(n)$. Rearranging yields $G'(t)G(t)^{-\frac{n+2}{n+3}} \leq -C(n)$, for some new constant $C(n)$. By the fundamental theorem of calculus, we conclude that:

$$\begin{aligned}
G(t)^{\frac{1}{n+3}} &\leq G(-\beta)^{\frac{1}{n+3}} - C(n)\cdot(t+\beta) = \left(\int_{M_{-\beta}} \varphi\right)^{\frac{1}{n+3}} - C(n)\cdot(t+\beta) \\
&\leq \mathscr{H}^n\left(M_{-\beta} \cap B_1(0)\right)^{\frac{1}{n+3}} - C(n)\cdot(t+\beta) \\
&\leq \eta^{\frac{1}{n+3}} - C(n)\cdot(t+\beta)
\end{aligned}$$

And hence $G(t) = 0$ for all $t \geq -\beta + C(n)\eta^{\frac{1}{n+3}}$ for a new constant $C(n)$ that we get after rearranging, and the result follows by noting that $\{\varphi(\cdot, t) > 0\} = B_{\sqrt{1-(2n+1)\cdot(t+\beta)}}(0)$. □

In combination with the proposition on the speed of convergence to limit points (3.2.4), we obtain Brakke's clearing out lemma.

THEOREM 3.6.2 (Brakke Clearing Out Lemma). Let (M_t) move by mean curvature in $B_{\varrho_0}(x_0) \times \left(t_0 - \frac{1}{2n+1}\varrho_0^2, t_0\right)$, and suppose that it reaches x_0 at time t_0. Then for any $0 < \varrho \leq \varrho_0$ and $\beta \in \left(0, \frac{1}{2n+1}\right)$ we have the lower bound:

$$\frac{\mathscr{H}^n\left(M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0)\right)}{\varrho^n} \geq \eta(n,\beta)$$

for a fixed $\eta(n,\beta) > 0$. In fact, we may take $\eta(n,\beta) := \frac{\beta^{n+3}}{2^{n+3}C(n)^{n+3}}$ where $C(n)$ is as in (3.6.1).

PROOF. Indeed let us choose:

$$\eta(n,\beta) := \frac{\beta^{n+3}}{2^{n+3}C(n)^{n+3}}$$

and show that it works. We will proceed by contradiction, and assume that:

$$\frac{\mathscr{H}^n\left(M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0)\right)}{\varrho^n} < \eta(n,\beta)$$

for some $0 < \varrho \leq \varrho_0$, and $\beta \in \left(0, \frac{1}{2n+1}\right)$. Rescale parabolically around (x_0, t_0) with scale ϱ, and observe that:

$$\left(M_s^{(x_0,t_0),\varrho}\right) \text{ moves by mean curvature in } B_1(0) \times \left(-\frac{1}{2n+1}, 0\right)$$

and additionally by assumption

$$\mathscr{H}^n\left(M_{-\beta}^{(x_0,t_0),\varrho} \cap B_1(0)\right) < \frac{\beta^{n+3}}{C(n)^{n+3}}$$

Then by our preliminary version of Brakke's clearing out lemma (3.6.1) we have:

$$\mathscr{H}^n\left(M_s^{(x_0,t_0),\varrho} \cap B_{\sqrt{1-(2n+1)(s+\beta)}}(0)\right) = 0 \tag{$\dagger_1$}$$

for every $s \geq -\beta + C(n)\left(\frac{\beta^{n+3}}{2^{n+3}C(n)^{n+3}}\right)^{\frac{1}{n+3}} = -\frac{\beta}{2}$ in $\left(-\frac{1}{2n+1}, 0\right)$. At the same time, since $\beta \in \left(0, \frac{1}{2n+1}\right)$ we have for all s that:

$$-(2n+1)s < 1 - (2n+1)(s+\beta) \tag{$\dagger_2$}$$

Consequently for an arbitrary choice of $s \in \left(-\frac{\beta}{2}, 0\right)$ we have:

$$0 \overset{(\dagger_1)}{=} \mathscr{H}^n \left(M_s^{(x_0,t_0),\varrho} \cap B_{\sqrt{1-(2n+1)(s+\beta)}}(0) \right)$$
$$\overset{(\dagger_2)}{\geq} \mathscr{H}^n \left(M_s^{(x_0,t_0),\varrho} \cap B_{\sqrt{-(2n+1)s}}(0) \right) = \varrho^{-n} \mathscr{H}^n \left(M_{t_0+\varrho^2 s} \cap B_{\sqrt{-(2n+1)\varrho^2 s}}(x_0) \right)$$

This, however, contradicts that the point x_0 is reached at time t_0 which guarantees by (3.2.4) that the latter set has positive $\mathscr{H}^n$-measure, and the result follows, i.e.

$$\frac{\mathscr{H}^n \left(M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0) \right)}{\varrho^n} \geq \eta(n,\beta) = \frac{\beta^{n+3}}{2^{n+3} C(n)^{n+3}}$$

for all $\beta \in \left(0, \frac{1}{2n+1}\right)$, $0 < \varrho \leq \varrho_0$, and where $C(n)$ is the same constant as in the special case of Brakke's clearing out lemma, (3.6.1). □

Let us now establish an upper density bound. Recall that in (3.3.1) we established forward-time local bounds on the area ($\mathscr{H}^n$-measure) of our surfaces in terms of the area of previous surfaces, using rudimentary estimates and the standard test functions $\varphi_{(x_0,t_0),\varrho}$. The backwards heat kernel $\Phi_{(x_0,t_0)}$ and Huisken's monotonicity formula can be used to help us estimate, additionally, the n-dimensional Hausdorff density of our surfaces from above. Notice that this bound is similar in spirit to the forward area estimate, (3.3.1).

THEOREM 3.6.3. Let (M_t) move by mean curvature in $B_{\varrho_0}(x_0) \times \left(t_0 - \frac{1}{2n}\varrho_0^2, t_0\right)$. Let $\theta > 0$ be such that $(1+2n)\theta^2 < 1$. Then there exists a constant $C = C(n,\theta)$ such that for all $0 < \varrho \leq \varrho_0$:

$$\sup_{0<\sigma\leq\theta\varrho} \quad \sup_{t\in(t_0-\sigma^2,t_0)} \frac{\mathscr{H}^n \left(M_t \cap B_\sigma(x_0)\right)}{\sigma^n} \leq C(n,\theta) \quad \frac{\mathscr{H}^n(M_{t_0-\theta^2\varrho^2} \cap B_\varrho(x_0))}{\varrho^n}$$

In fact, $C(n,\theta) := 2^{n/2} e^{1/4} (1-(1+2n)\theta^2)^{-3} \theta^{-n}$

PROOF. Write $T_0 := t_0 - \theta^2 \varrho^2$. Consider the function:

$$f := \varphi_{(x_0,T_0),\varrho}$$

which we know to be a test function on $B_\varrho(x_0) \times (T_0, \infty)$ and hence also on $B_\varrho(x_0) \times (T_0, t_0)$. Furthermore:

$$f \geq 0 \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \leq 0$$

and hence from the monotonicity formula (3.5.2) we know that:

$$\int_{M_t} \Phi_{(x_0,t_0+\sigma^2)} f \leq \int_{M_{T_0}} \Phi_{(x_0,t_0+\sigma^2)} f \tag{$\dagger_1$}$$

for every $t \in (T_0, t_0)$. Now observe that:

$$\begin{aligned}
\int_{M_{T_0}} \Phi_{(x_0,t_0+\sigma^2)} f &= \int_{M_{T_0} \cap \{f(\cdot,T_0)>0\}} \Phi_{(x_0,t_0+\sigma^2)} f \\
&= \int_{M_{T_0} \cap B_\varrho(x_0)} \Phi_{(x_0,t_0+\sigma^2)} f \\
&\leq \sup_{\mathbb{R}^{n+1}} \Phi_{(x_0,t_0+\sigma^2)}(\cdot, T_0) \quad \sup_{\mathbb{R}^{n+1}} f(\cdot, T_0) \quad \mathscr{H}^n\left(M_{T_0} \cap B_\varrho(x_0)\right)
\end{aligned}$$

This we can estimate from above by (3.4.2.a) and (3.2.1.b):

$$\begin{aligned}
\int_{M_{T_0}} \Phi_{(x_0,t_0+\sigma^2)} f &\leq \frac{1}{(4\pi(t_0+\sigma^2-T_0))^{n/2}} \quad \mathscr{H}^n\left(M_{T_0} \cap B_\varrho(x_0)\right) \\
&= \frac{1}{(4\pi(\sigma^2+\theta^2\varrho^2))^{n/2}} \quad \mathscr{H}^n\left(M_{T_0} \cap B_\varrho(x_0)\right) \\
&\leq \frac{1}{(4\pi\theta^2\varrho^2)^{n/2}} \quad \mathscr{H}^n\left(M_{T_0} \cap B_\varrho(x_0)\right)
\end{aligned} \tag{$\dagger_2$}$$

On the other hand, for $t \in (t_0 - \sigma^2, t_0) \subseteq (T_0, t_0)$ we also have:

$$\begin{aligned}
\int_{M_t} \Phi_{(x_0,t_0+\sigma^2)} f &\geq \int_{M_t \cap B_\sigma(x_0)} \Phi_{(x_0,t_0+\sigma^2)} f \\
&\geq \inf_{B_\sigma(x_0)} \Phi_{(x_0,t_0+\sigma^2)}(\cdot, t) \quad \inf_{B_\sigma(x_0)} f(\cdot, t) \quad \mathscr{H}^n\left(M_t \cap B_\sigma(x_0)\right) \\
&= \frac{1}{(4\pi(t_0+\sigma^2-t))^{n/2}} \quad \exp\left(-\frac{\sigma^2}{4(t_0+\sigma^2-t)}\right) \quad \left(\frac{\varrho^2-\sigma^2-2n(t-T_0)}{\varrho^2}\right)_+^3 \quad \mathscr{H}^n\left(M_t \cap B_\sigma(x_0)\right)
\end{aligned}$$

where the last step follows from (3.4.2.b) and (3.2.1.d). Recall that $t \in (t_0 - \sigma^2, t_0)$. Then $\sigma^2 < t_0 + \sigma^2 - t < 2\sigma^2$, and $t - T_0 = t - t_0 + \theta^2\varrho^2 < \theta^2\varrho^2$. Thus:

$$\begin{aligned} \int_{M_t} \Phi_{(x_0, t_0+\sigma^2)} f &\geq \frac{1}{(8\pi\sigma^2)^{n/2}} \quad e^{-1/4} \quad \left(\frac{\varrho^2 - \theta^2\varrho^2 - 2n\theta^2\varrho^2}{\varrho^2}\right)_+^3 \quad \mathscr{H}^n\left(M_t \cap B_\sigma(x_0)\right) \\ &= \frac{1}{(8\pi\sigma^2)^{n/2}} \quad e^{-1/4} \quad (1-(1+2n)\theta^2)^3 \quad \mathscr{H}^n\left(M_t \cap B_\sigma(x_0)\right) \end{aligned} \tag{$\dagger_3$}$$

Combining ($\dagger_1$), ($\dagger_2$), ($\dagger_3$) we conclude that for all $t \in (t_0 - \sigma^2, t_0)$:

$$\frac{1}{(8\pi\sigma^2)^{n/2}} \quad e^{-1/4} \quad (1-(1+2n)\theta^2)^3 \quad \mathscr{H}^n\left(M_t \cap B_\sigma(x_0)\right) \leq \frac{1}{(4\pi\theta^2\varrho^2)^{n/2}} \quad \mathscr{H}^n\left(M_{T_0} \cap B_\varrho(x_0)\right)$$

Rearranging:

$$\frac{\mathscr{H}^n\left(M_t \cap B_\sigma(x_0)\right)}{\sigma^n} \leq 2^{n/2} e^{1/4} (1-(1+2n)\theta^2)^{-3} \theta^{-n} \quad \frac{\mathscr{H}^n\left(M_{T_0} \cap B_\varrho(x_0)\right)}{\varrho^n}$$

which is the required result. $\square$

3.7. Gauss Density

The backwards heat kernel also lets us define a new kind of (parabolic) density for (local) solutions of mean curvature flow, which is essentially a condensed form of a spacetime density estimator. This density will let us prove sharp estimates later on that we will need to have available to study what happens to mean curvature flow at the first singular time in the next chapter. We will need to establish a sequence of lemmas before being able to give a rigorous definition of Gauss density.

LEMMA 3.7.1. Let (M_t) move by mean curvature in $B_{\varrho_0}(x_0) \times (t_0 - \frac{1}{2n}\varrho_0^2, t_0)$. Then for every $\varrho < \varrho_0$ and $\mathcal{C}^2$ time independent χ_ϱ with $\mathbf{1}_{B_{\varrho/2}(x_0)} \leq \chi_\varrho \leq \mathbf{1}_{B_\varrho(x_0)}$ the limit:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \chi_\varrho$$

exists and is finite.

PROOF. A direct computation using (2.3.1.c) shows that:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\chi_\varrho = D_t\chi_\varrho - \mathrm{div}_{M_t}\chi_\varrho = -\mathrm{div}_{M_t}\chi_\varrho = -\Delta_{\mathbb{R}^{n+1}}\chi_\varrho + D^2\chi_\varrho(\nu,\nu) \leq C_1(n)|D^2\chi_\varrho|$$

for some constant C_1 depending only on n. Notice that inside $B_{\varrho/2}(x_0)$ and outside $\overline{B}_\varrho(x_0)$ the left hand side vanishes identically, since χ_ϱ is constant in these sets. Taking this into account, we may refine our estimate above as:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\chi_\varrho \leq C_1(n)|D^2\chi_\varrho|\mathbf{1}_{\overline{B}_\varrho(x_0)\setminus B_{\varrho/2}(x_0)} \tag{$\dagger_1$}$$

Additionally, we know from (3.4.2.d) that we can estimate the backwards heat kernel away from x_0, so that:

$$\Phi_{(x_0,t_0)}\mathbf{1}_{\overline{B}_\varrho(x_0)\setminus B_{\varrho/2}(x_0)} \leq C_2(n)\varrho^{-n}\mathbf{1}_{\overline{B}_\varrho(x_0)} \tag{$\dagger_2$}$$

for some constant C_2 depending only on n.

There is one final auxiliary step. Recall that $\chi_\varrho \leq \mathbf{1}_{B_\varrho(x_0)}$. Since $\varrho < \varrho_0$, there exists a $\sigma(\varrho)$ such that $\varrho < \sigma(\varrho) < \varrho_0$, and a $\theta_1(\varrho) \in (0,1)$ such that $\varrho < \theta_1(\varrho)\sigma(\varrho)$. Let $\theta_2(\varrho)$ be such that $\theta_1(\varrho)^2 + 2n\theta_2(\varrho)^2 < 1$. From the forward area estimate (3.3.1) we know that:

$$\sup_{t\in(t_0-\theta_2(\varrho)^2\sigma(\varrho)^2,t_0)} \mathscr{H}^n(M_t\cap\overline{B}_\varrho(x_0)) \le C_3(n,\varrho)\mathscr{H}^n(M_{t_0-\theta_2(\varrho)^2\sigma(\varrho)^2}\cap B_{\sigma(\varrho)}(x_0)) \tag{$\dagger_3$}$$

where $C_3(n,\varrho)$ was actually a constant depending on n,θ_1,θ_2 in (3.3.1), but θ_1,θ_2 both depend on n and ϱ here, so we may relabel the constant as such. Notice of course that $(t_0-\theta_2(\varrho)^2\sigma(\varrho)^2,t_0)\subset\left(t_0-\frac{1}{2n}\varrho_0^2,t_0\right)$, so everything above is well-defined.

Using Huisken's monotonicity (3.5.1) we can estimate that for $t\in(t_0-\theta_2(\varrho)^2\sigma(\varrho)^2,t_0)$:

$$\begin{aligned}
\frac{d}{dt}\int_{M_t}\Phi_{(x_0,t_0)}\chi_\varrho &= \int_{M_t}\Phi_{(x_0,t_0)}\left(\frac{\partial}{\partial t}-\Delta_{M_t}\right)\chi_\varrho - \left|H-\frac{\nabla^\perp\Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2\Phi_{(x_0,t_0)}\chi_\varrho \\
&\le \int_{M_t}\Phi_{(x_0,t_0)}\left(\frac{\partial}{\partial t}-\Delta_{M_t}\right)\chi_\varrho \\
&\le C_1(n)|D^2\chi_\varrho|\int_{M_t}\Phi_{(x_0,t_0)}\mathbf{1}_{\overline{B}_\varrho(x_0)\setminus B_{\varrho/2}(x_0)} \\
&\le C_1(n)C_2(n)|D^2\chi_\varrho|\varrho^{-n}\int_{M_t}\mathbf{1}_{\overline{B}_\varrho(x_0)} \\
&\le C_1(n)C_2(n)C_3(n,\varrho)|D^2\chi_\varrho|\varrho^{-n}\mathscr{H}^n(M_{t_0-\theta_2(\varrho)^2\sigma(\varrho)^2}\cap B_{\sigma(\varrho)}(x_0))
\end{aligned}$$

where we've used ($\dagger_1$), ($\dagger_2$), and ($\dagger_3$) in the last three steps. Label the right hand side as C, knowing that it is independent of t as long as $t\in(t_0-\theta_2(\varrho)^2\sigma(\varrho)^2,t_0)$. Then function:

$$t\mapsto -Ct+\int_{M_t}\Phi_{(x_0,t_0)}\chi_\varrho$$

is non-increasing on $(t_0-\theta_2(\varrho)^2\sigma(\varrho)^2,t_0)$. It is also bounded from below (by $-Ct_0$), so:

$$\lim_{t\nearrow t_0}\left(-Ct+\int_{M_t}\Phi_{(x_0,t_0)}\chi_\varrho\right)\quad\text{exists}$$

and hence:

$$\lim_{t\nearrow t_0}\int_{M_t}\Phi_{(x_0,t_0)}\chi_\varrho\quad\text{exists}$$

also, and the result follows. $\square$

Notice that *a priori* the limit above may depend on χ_ϱ, or at least our choice of ϱ. In fact it turns out that there is no such dependence, as shown by the following lemma.

LEMMA 3.7.2. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, and let $x_0 \in \mathcal{U}$. Then for any function $h \in \mathcal{C}_c^0(\mathcal{U})$ such that $h \equiv 1$ in a neighborhood of x_0 the limit:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} h$$

exists, is finite, and independent of h.

PROOF. Let us first show that for $g \in \mathcal{C}_c^0(\mathcal{U})$ such that $g \equiv 0$ in a neighborhood of x_0, we have:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} g = 0 \tag{‡}$$

Let the neighborhood where g vanishes be $\mathcal{V}$, and let K denote the support of g. Of course $\mathcal{V}, K \subseteq \mathcal{U}$.

Recall that from (3.4.2.c) we know that $\Phi_{(x_0,t_0)} \to 0$ uniformly away from x_0, meaning:

$$\lim_{t \nearrow t_0} \int_{M_t \setminus \mathcal{U}} \left|\Phi_{(x_0,t_0)} g\right| \leq |g| \limsup_{t \nearrow t_0} \int_{(M_t \cap K) \setminus \mathcal{V}} |\Phi_{(x_0,t_0)}| = 0$$

where $|g|$ denotes the supremum bound of g. The last equality follows from a non-classical dominated convergence theorem, since $\Phi_{(x_0,t_0)} \to 0$ uniformly on the domain of integration; see (B.2.2). This shows that:

$$\lim_{t \nearrow t_0} \int_{M_t \setminus \mathcal{V}} \Phi_{(x_0,t_0)} g = 0 \tag{$\dagger_1$}$$

On the other hand, in the interior of $\mathcal{V}$ we know that $g = 0$, and hence:

$$\int_{M_t \cap \mathcal{V}} \Phi_{(x_0,t_0)} g = 0 \tag{$\dagger_2$}$$

which clearly implies that the limit as $t \nearrow t_0$ also exists and equals 0. Adding up ($\dagger_1$) and ($\dagger_2$) we get (‡).

Let us now return to the original problem. Choose $\varrho_0 > 0$ such that $B_{\varrho_0}(x_0) \subseteq \mathcal{U}$ and $\varrho_0^2 < 2n(t_0 - t_1)$. Then evidently (M_t) also moves by mean curvature in $B_{\varrho_0}(x_0) \times \left(t_0 - \frac{1}{2n}\varrho_0^2, t_0\right)$. Choose any $\varrho < \varrho_0$ and a χ_ϱ as in (3.7.1). Then:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \chi_\varrho =: \ell \tag{$\star_1$}$$

exists. Now let h be as in the problem statement, and observe that $g := h - \chi_\varrho \in \mathcal{C}^0_c(\mathcal{U})$ and $g = 0$ in a neighborhood of x_0, so:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} (h - \chi_\varrho) = 0 \tag{$\star_2$}$$

Therefore, by combining ($\star_1$) and ($\star_2$) we get that for $t < t_0$:

$$\int_{M_t} \Phi_{(x_0,t_0)} h = \int_{M_t} \Phi_{(x_0,t_0)} \chi_\varrho + \int_{M_t} \Phi_{(x_0,t_0)} (h - \chi_\varrho) \to \ell + 0 = \ell \quad \text{as} \quad t \nearrow t_0$$

In other words the limit exists and equals ℓ, and so it is also independent of h. □

It is also important to realize that the domain $\mathcal{U}$ does not play an important role in this computation; indeed, as long as x_0 is contained in our open set, the limit above exists and is independent of the domain.

COROLLARY 3.7.3. Let (M_t) move by mean curvature in both $\mathcal{U} \times (t_1, t_0)$ and $\mathcal{U}' \times (t_1, t_0)$, with $x_0 \in \mathcal{U} \cap \mathcal{U}'$. Then for any two functions $h \in \mathcal{C}^0_c(\mathcal{U})$, $h' \in \mathcal{C}^0_c(\mathcal{U}')$ such that $h, h' \equiv 1$ in a neighborhood of x_0 we have:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} h = \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} h'$$

PROOF. Choose $\varrho > 0$ such that $\overline{B}_\varrho(x_0) \subseteq \mathcal{U} \cap \mathcal{U}'$ and observe that by (3.7.2) we have:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} h = \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \chi_\varrho = \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} h'$$

which gives the required result. □

We are finally at a position to define (localized) Gauss densities by combining the previous three results, which guarantee that our quantity is well defined.

DEFINITION 3.7.4 (Gauss Density). Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, and let $x_0 \in \mathcal{U}$. We define Gauss density at (x_0, t_0) to be:

$$\Theta(M_t, x_0, t_0) := \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} h$$

for some (and hence any) $h \in \mathcal{C}_c^0(\mathcal{U})$ such that $h \equiv 1$ in a neighborhood of x_0. □

We are free to drop $\mathcal{U}$ from our notation because we have shown that it is irrelevant, according to (3.7.3). Unfortunately, the requirements that h be time independent and identically 1 in an entire neighborhood of x_0 are somewhat restrictive. Fortunately, we may remove that restriction:

THEOREM 3.7.5. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, and $x_0 \in \mathcal{U}$. For any $f \in \mathcal{C}^0 (\mathcal{U} \times [t_1, t_0])$ such that support $f(\cdot, t) \subseteq K$ for all $t \in [t_1, t_0]$ and a fixed compact set $K \subseteq \mathcal{U}$, we have:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} f = f(x_0, t_0)\Theta(M_t, x_0, t_0)$$

PROOF. Let us first show that for $g \in \mathcal{C}^0 (\mathcal{U} \times [t_1, t_0])$ such that $g(x_0, t_0) = 0$ and $g(\cdot, t) \subseteq K$ for all $t \in [t_1, t_0]$ and a fixed $K \subseteq \mathcal{U}$ we have:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} g = 0 \tag{‡}$$

Indeed, let $\varepsilon > 0$ be arbitrary. The continuity of g lets us choose $\varrho > 0$, $\delta > 0$ such that:

$$|g(x, t)| \leq C\varepsilon \quad \text{for} \quad x \in B_\varrho(x_0), t > t_0 - \delta$$

where $C := \frac{1}{1+\Theta(M_t,x_0,t_0)}$ is some constant. Without loss of generality, $\varrho > 0$ is small enough that $\overline{B}_{2\varrho}(x_0) \subseteq \mathcal{U}$.

Additionally, by compactness of $K \times [t_1, t_0]$ and the fact that support $g(\cdot, t) \subseteq K$ for all $t \in [t_1, t_0]$, we know that g is bounded from above in $\mathcal{U} \times [t_1, t_0]$. This crude estimate gives us that for $t \in (t_1, t_0)$:

$$\int_{M_t \setminus B_\varrho(x_0)} \left|\Phi_{(x_0,t_0)} g\right| \leq |g| \int_{M_t \setminus B_\varrho(x_0)} \Phi_{(x_0,t_0)} \mathbf{1}_K \to 0 \quad \text{as} \quad t \nearrow t_0$$

where the last step follows from a non-classical dominated convergence theorem since $\Phi_{(x_0,t_0)} \to 0$ uniformly away from x_0; see (B.2.2). That is, we have established that:

$$\limsup_{t \nearrow t_0} \int_{M_t \setminus B_\varrho(x_0)} \left|\Phi_{(x_0,t_0)} g\right| = 0 \tag{$\dagger_1$}$$

Now recall that $\varrho > 0$ was chosen such that $|g| < C\varepsilon$ in $B_\varrho(x_0) \times (t_0 - \delta, t_0]$, and also small enough that $\overline{B}_{2\varrho}(x_0) \subseteq \mathcal{U}$. If $\chi_{2\varrho}$ is as in (3.7.1), we know that support $\chi_{2\varrho} \subseteq \overline{B}_{2\varrho}(x_0) \subseteq \mathcal{U}$, and that by construction $\chi_{2\varrho} \geq 1$ in $B_\varrho(x_0)$. Then for $t \in (t_0 - \delta, t_0)$:

$$\begin{aligned}
\int_{M_t \cap B_\varrho(x_0)} \left|\Phi_{(x_0,t_0)} g\right| &\leq C\varepsilon \int_{M_t \cap B_\varrho(x_0)} \Phi_{(x_0,t_0)} \\
&\leq C\varepsilon \int_{M_t \cap B_\varrho(x_0)} \Phi_{(x_0,t_0)} \chi_{2\varrho} \\
\Rightarrow \limsup_{t \nearrow t_0} \int_{M_t \cap B_\varrho(x_0)} \left|\Phi_{(x_0,t_0)} g\right| &\leq C\varepsilon \Theta(M_t, x_0, t_0) \\
&= \frac{1}{1 + \Theta(M_t, x_0, t_0)} \varepsilon \Theta(M_t, x_0, t_0) < \varepsilon
\end{aligned} \tag{$\dagger_2$}$$

Combining ($\dagger_1$) and ($\dagger_2$) we see that:

$$\limsup_{t \nearrow t_0} \int_{M_t} \left|\Phi_{(x_0,t_0)} g\right| \leq \limsup_{t \nearrow t_0} \int_{M_t \setminus B_\varrho(x_0)} \left|\Phi_{(x_0,t_0)} g\right| + \limsup_{t \nearrow t_0} \int_{M_t \cap B_\varrho(x_0)} \left|\Phi_{(x_0,t_0)} g\right| < \varepsilon$$

But $\varepsilon > 0$ was arbitrary, so ($\ddagger$) follows.

Let us now return to our problem. By definition of Gauss densities we know that:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} f(x_0, t_0) \chi_{2\varrho} = f(x_0, t_0) \Theta(M_t, x_0, t_0) \tag{$\star_1$}$$

Also observe that $g := f - f(x_0, t_0)\chi_{2\varrho} \in \mathcal{C}^0(\mathcal{U} \times [t_1, t_0])$, $g(x_0, t_0) = 0$, and support $g(\cdot, t) \subseteq K$ for all $t \in [t_1, t_0]$ and a fixed $K \subseteq \mathcal{U}$. Therefore by ($\ddagger$) we have:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \left(f - f(x_0,t_0)\chi_{2\varrho}\right) = 0 \tag{$\star_2$}$$

Combining ($\star_1$) and ($\star_2$) gives:

$$\begin{aligned}\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} f &= \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} f(x_0,t_0)\chi_{2\varrho} + \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \left(f - f(x_0,t_0)\chi_{2\varrho}\right) \\ &= f(x_0,t_0)\Theta(M_t, x_0, t_0)\end{aligned}$$

which is the required result. $\square$

Of course, we may apply this result to the standard test function and get the following trivial result which is important in and of itself, so we label it as a theorem.

THEOREM 3.7.6. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, $x_0 \in \mathcal{U}$. For any $\varrho > 0$ such that $\varphi_{(x_0,t_0),\varrho}$ is a test function on $\mathcal{U} \times (t_2, \infty)$ for some $t_2 < t_0$, we must have that

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho} = \Theta(M_t, x_0, t_0) \tag{3.7.6.a}$$

and

$$\Theta(M_t, x_0, t_0) \leq \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho} \tag{3.7.6.b}$$

for every $t \in (t_1, t_0)$ such that $t > t_2$.

In particular, this is true for any $\varrho \leq \frac{1}{\sqrt{1+2n}} \mathrm{dist}(x_0; \mathbb{R}^{n+1} \setminus \mathcal{U})$ and for $t_2 := t_0 - \varrho^2$.

PROOF. The last remark is trivial. For ϱ as above:

$$B_{\sqrt{1+2n}\varrho}(x_0) \subseteq \mathcal{U} \quad \text{and} \quad \varrho^2 - 2n\left((t_0 - \varrho^2) - t_0\right) = (1+2n)\varrho^2$$

so that by (3.2.2) we know that $\varphi_{(x_0,t_0),\varrho}$ is a test function on $B_{\sqrt{1+2n}\varrho}(x_0) \times (t_0 - \varrho^2, \infty)$, and hence also on $\mathcal{U} \times (t_0 - \varrho^2, \infty)$, so we can apply the first part of the theorem, to which we now return.

Let $\delta > 0$ be small enough that support $\varphi_{(x_0,t_0),\varrho}(\cdot, t)$ is a fixed compact subset of $\mathcal{U}$ for all $t \in [t_0 - \delta, t_0]$. We can do this because $\varphi_{(x_0,t_0),\varrho}$ is, by assumption, a test function on $\mathcal{U} \times (t_2, \infty)$ for $t_2 < t_0$.

Of course $\varphi_{(x_0,t_0),\varrho} \in \mathcal{C}^0(\mathcal{U} \times [t_0 - \delta, t_0])$. Hence by applying (3.7.5) on $\mathcal{U} \times (t_0 - \delta, t_0)$ we have:

$$\lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho} = \varphi_{(x_0,t_0),\varrho}(x_0, t_0)\Theta(M_t, x_0, t_0) = \Theta(M_t, x_0, t_0)$$

as claimed. Finally, recall that $\varphi_{(x_0,t_0),\varrho}$ also satisfies:

$$\varphi_{(x_0,t_0),\varrho} \geq 0 \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \varphi_{(x_0,t_0),\varrho} \leq 0$$

so that by Huisken's monotonicity (3.5.2) the map:

$$t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho}$$

is decreasing for $t \in (t_2, t_0)$, and hence for any one of those times t we have:

$$\int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho} \geq \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho} = \Theta(M_t, x_0, t_0)$$

which completes our proof. $\square$

A direct corollary shows that Gauss density is actually an upper semicontinuous function of its arguments.

THEOREM 3.7.7. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$. For every $x_0 \in \mathcal{U}$, and every pair of sequences $t_j \in (t_1, t_0)$, $x_j \in \mathcal{U}$, $j \geq 1$, such that $t_j \nearrow t_0$ and $x_j \to x_0$ we have:

$$\Theta(M_t, x_0, t_0) \geq \limsup_{j \to \infty} \Theta(M_t, x_j, t_j)$$

PROOF. Since $x_j \to x_0$, it follows that $\inf_{j \geq 0} \operatorname{dist}(x_j; \mathbb{R}^{n+1} \setminus \mathcal{U}) > 0$. Choosing $\varrho > 0$ that does not exceed this infimum, we know that (3.7.6) applies for all $\varphi_{(x_j,t_j),\varrho}$, $j \geq 0$.

Fix $t \in (t_1, t_0)$. Since $t_j \nearrow t_0$, we can find a j_0 such that $t_j > t$ for all $j \geq j_0$. Then for all $j \geq j_0$ we have by (3.7.6) that:

$$\Theta(M_t, x_j, t_j) \leq \int_{M_t} \Phi_{(x_j,t_j)} \varphi_{(x_j,t_j),\varrho}$$

Let $j \to \infty$:

$$\limsup_{j \to \infty} \Theta(M_t, x_j, t_j) \leq \limsup_{j \to \infty} \int_{M_t} \Phi_{(x_j,t_j)} \varphi_{(x_j,t_j),\varrho}$$

For any fixed t all the functions $\varphi_{(x_j,t_j),\varrho}(\cdot, t)$, $j \geq 0$, have their support contained within a fixed compact set K of $\mathcal{U}$. Therefore, the classical dominated convergence theorem (B.2.1) applies on the right hand side and tells us that the $\limsup$ is actually a $\lim$ and it can be switched around with the integral sign:

$$\begin{aligned}
\limsup_{j \to \infty} \Theta(M_t, x_j, t_j) &\leq \int_{M_t} \lim_j \Phi_{(x_j,t_j)} \varphi_{(x_j,t_j),\varrho} \\
&= \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho}
\end{aligned}$$

This is true for all $t \in (t_1, t_0)$. Notice that the left hand side is independent of t, and hence taking $t \nearrow t_0$ we conclude by using (3.7.6) that:

$$\limsup_{j \to \infty} \Theta(M_t, x_j, t_j) \leq \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho} = \Theta(M_t, x_0, t_0)$$

which is the required result. $\square$

Another corollary of the previous theorem in combination with Brakke's clearing out lemma allows us to bound Gauss density from below for points x_0 that are reached at time t_0. This allows us to separate points that are reached from those that are not reached, since the latter will trivially have a vanishing Gauss density. Refer to Chapter C of the appendix for an alternate proof of this, specifically tailored to smooth manifolds.

PROPOSITION 3.7.8. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, and suppose that the point $x_0 \in \mathcal{U}$ is reached by (M_t). Then:

$$\Theta(M_t, x_0, t_0) \geq C(n)$$

for a contant $C(n) > 0$.

PROOF. Let $\varrho_0 > 0$ be any radius such that:

$$\Theta(M_t, x_0, t_0) = \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho_0}$$

according to (3.7.6). Let $\beta \in \left(0, \frac{1}{2n+1}\right)$ be arbitrary for now. Observe that:

$$\Theta(M_t, x_0, t_0) = \lim_{\varrho \searrow 0} \int_{M_{t_0 - \beta\varrho^2}} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho_0} \geq \liminf_{\varrho \searrow 0} \int_{M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0)} \Phi_{(x_0,t_0)} \varphi_{(x_0,t_0),\varrho_0}$$

which by the backwards heat kernel and standard test function lower bound inequalities (3.4.2.b), (3.2.1.d) we estimate:

$$\begin{aligned}
&\geq \liminf_{\varrho \searrow 0} \inf_{B_\varrho(x_0)} \Phi_{(x_0,t_0)}(\cdot, t_0 - \beta\varrho^2) \quad \inf_{B_\varrho(x_0)} \varphi_{(x_0,t_0),\varrho_0}(\cdot, t_0 - \beta\varrho^2) \quad \mathscr{H}^n \left(M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0)\right) \\
&\geq \liminf_{\varrho \searrow 0} \frac{1}{(4\pi\beta\varrho^2)^{n/2}} \exp\left(-\frac{\varrho^2}{4\beta\varrho^2}\right) \cdot \left(\frac{\varrho_0^2 - \varrho^2 - 2n\beta\varrho^2}{\varrho_0^2}\right)_+^3 \cdot \mathscr{H}^n \left(M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0)\right)
\end{aligned}$$

The last term can be estimated from below by Brakke's clearing out lemma (3.6.2):

$$
\begin{aligned}
&\geq \liminf_{\varrho \searrow 0} \frac{1}{(4\pi\beta\varrho^2)^{n/2}} \exp\left(-\frac{\varrho^2}{4\beta\varrho^2}\right) \cdot \left(\frac{\varrho_0^2 - \varrho^2 - 2n\beta\varrho^2}{\varrho_0^2}\right)_+^3 \cdot \eta(n,\beta) \cdot \varrho^n \\
&= \liminf_{\varrho \searrow 0} \frac{1}{(4\pi\beta)^{n/2}} \exp\left(-\frac{1}{4\beta}\right) \cdot \left(\frac{\varrho_0^2 - \varrho^2 - 2n\beta\varrho^2}{\varrho_0^2}\right)_+^3 \cdot \eta(n,\beta) \\
&= \frac{1}{(4\pi\beta)^{n/2}} \exp\left(-\frac{1}{4\beta}\right) \cdot \eta(n,\beta)
\end{aligned}
$$

Choosing, for example, $\beta = \frac{1}{2(2n+1)} \in \left(0, \frac{1}{2n+1}\right)$ gives us a lower bound $\Theta(M_t, x_0, t_0) \geq C(n)$, as claimed. $\square$

CHAPTER 4

Assessing Curvature

In Chapter 2 we studied the evolution of the geometry of our submanifolds, while in Chapter 3 we studied the evolution of their measure theoretic structure. This chapter combines geometric and measure theoretic techniques and lays the groundwork for beginning to understand regularity as we approach the first singular time. By the end of this chapter, we will have shown how estimates on the measure theoretic "flatness" of our surfaces are strong enough to control all derivatives of the curvature tensor up to the first singular time.

There are four types of flatness we will be interested in:

Geometric			Measure Theoretic		
$\mathcal{C}^2$	$\rightsquigarrow$	$\sup\limits_{\substack{t\in(t_0-\varrho^2,t_0)\\ x\in M_t\cap B_\varrho(x_0)}} \lvert A(x,t)\rvert^2$	$\mathscr{L}^2$	$\rightsquigarrow$	$\sup\limits_{t\in(t_0-\varrho^2,t_0)} \dfrac{1}{\varrho^{n+2}}\displaystyle\int_{M_t\cap B_\varrho(x_0)} (x-x_0)^2_{n+1}$
$\mathcal{C}^\infty$	$\rightsquigarrow$	$\sup\limits_{\substack{t\in(t_0-\varrho^2,t_0)\\ x\in M_t\cap B_\varrho(x_0)}} \sup\limits_{m\geq 0} \lvert\nabla^m A(x,t)\rvert^2$	$\mathscr{L}^\infty$	$\rightsquigarrow$	$\sup\limits_{\substack{t\in(t_0-\varrho^2,t_0)\\ x\in M_t\cap B_\varrho(x_0)}} \dfrac{1}{\varrho^2}(x-x_0)^2_{n+1}$

With this notation at hand, the plan of action is easy to describe in this chapter:

$$\mathscr{L}^2 \Rightarrow \mathscr{L}^\infty \Rightarrow \mathcal{C}^2 \Rightarrow \mathcal{C}^\infty \quad \text{in terms of flatness}$$

It is most convenient, however, to work nonlinearly. We begin by showing in section 4.1 that due to the parabolic nature of our geometric evolution, uniform curvature estimates ($\mathcal{C}^2$) propagate to estimates for derivatives of the curvature tensor of all orders ($\mathcal{C}^\infty$). In section 4.2 we take a step into measure theoretic flatness and show that mean-square flatness ($\mathscr{L}^2$) effectively controls uniform flatness ($\mathscr{L}^\infty$) in a neighborhood of the center point. We then conclude with section 4.3 where we link the notions of geometric and measure theoretic flatness and prove that sufficiently uniformly flat surfaces have bounded curvature. As a corollary, we see that mean-square estimates suffice for uniform curvature bounds.

4.1. Curvature Controls all its Derivatives

We can use the localized weak maximum principle to inductively establish bounds on curvature derivatives of every order given only a bound on $|A|^2$. Most of the work towards this goal is done by the lemma below, which will in turn give the general case from after a parabolic rescaling and an inductive argument. We do essentially follow **[Eck04]**, but we tackle all higher dimensional cases simultaneously.

LEMMA 4.1.1. Let (M_t) move by mean curvature in $B_1(0) \times (-1, 0)$. Let $m \geq 0$ be given, and suppose that:

$$|A|^2, |\nabla A|^2, \ldots, |\nabla^m A|^2 \leq K_m$$

in $B_1(0) \times (-1, 0)$. For a fixed $\theta \in (0, 1)$ we then have:

$$|\nabla^{m+1} A|^2 \leq \frac{C(n, m, K_m)}{(1-\theta)^4}$$

in $B_\theta(0) \times (-\theta^2, 0)$, and a constant $C(n, m, K_m)$.

PROOF. The curvature evolution equations from (2.3.7) will play a key role in this lemma. For convenience $C(n, m, K_m)$ will denote any constant depending on n, m, K_m, and will change throughout the proof. For the evolution of $|\nabla^{m+1} A|^2$ (2.3.7.b) tells us that:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) |\nabla^{m+1} A|^2 \leq -2|\nabla^{m+2} A|^2 + C(n, m, K_m)(1 + |\nabla^{m+1} A|^2)$$

The evolution of $|\nabla^m A|^2$, however, can fall under either (2.3.7.b) or (when $m = 0$) (2.3.7.a). On the other hand, in view of the fact that $m \geq 0$ and thus $|A|^2 \leq K_m$, we can conclude that $|\nabla A^m|$ evolves similarly to $|\nabla A^{m+1}|$ independently of whether or not $m = 0$ as follows:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) |\nabla^m A|^2 \leq -2|\nabla^{m+1} A|^2 + C(n, m, K_m)(1 + |\nabla^m A|^2)$$

for some constant $C(n, m, K_m) \geq 1$, where the catch is that K_m depends on $|\nabla^m A|^2$ too. Let us, then, consider the non-negative function:

$$f := |\nabla^{m+1} A|^2 \left(\Lambda + |\nabla^m A|^2\right), \quad \Lambda := 15K_m$$

defined in the preimage $\mathcal{O} \subseteq M^n \times (-1, 0)$ of $B_1(0)$. By the parabolic product rule (A.3.3.a) and the evolution equations for $|\nabla^{m+1} A|^2$, $|\nabla^m A|^2$ described above, we see that:

$$\begin{aligned}
\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f &= \left(\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) |\nabla^{m+1} A|^2\right) \left(\Lambda + |\nabla^m A|^2\right) \\
&\quad + |\nabla^{m+1} A|^2 \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \left(\Lambda + |\nabla^m A|^2\right) \\
&\quad - 2\nabla \left(|\nabla^{m+1} A|^2\right) \cdot \nabla \left(\Lambda + |\nabla^m A|^2\right) \\
&\leq \left(-2|\nabla^{m+2} A|^2 + C(n, m, K_m)\left(1 + |\nabla^{m+1} A|^2\right)\right) \left(\Lambda + |\nabla^m A|^2\right) \\
&\quad + |\nabla^{m+1} A|^2 \left(-2|\nabla^{m+1} A|^2 + C(n, m, K_m)\left(1 + |\nabla^m A|^2\right)\right) \\
&\quad - 2\nabla \left(|\nabla^{m+1} A|^2\right) \cdot \nabla \left(|\nabla^m A|^2\right)
\end{aligned} \tag{‡}$$

Let us estimate the very last term by Cauchy-Schwarz:

$$\begin{aligned}
-2\nabla \left(|\nabla^{m+1} A|^2\right) \cdot \nabla \left(|\nabla^m A|^2\right) &\leq 2 \left|\nabla|\nabla^{m+1} A|^2\right| \left|\nabla|\nabla^m A|^2\right| \\
&= 2 \left(2|\nabla^{m+1} A| \; \left|\nabla|\nabla^{m+1} A|\right|\right) \left(2|\nabla^m A| \; |\nabla|\nabla^m A||\right) \\
&= 8|\nabla^m A| \; |\nabla^{m+1} A| \; |\nabla|\nabla^m A|| \; \left|\nabla|\nabla^{m+1} A|\right|
\end{aligned}$$

By Kato's tensor inequality we have $|\nabla|\nabla^m A|| \leq |\nabla^{m+1} A|$ and $\left|\nabla|\nabla^{m+1} A|\right| \leq |\nabla^{m+2} A|$. Substituting into the equation above, and then using $2ab \leq a^2 + b^2$ to break the product up:

$$\begin{aligned}
-2\nabla \left(|\nabla^{m+1} A|^2\right) \cdot \nabla \left(|\nabla^m A|^2\right) &\leq 8|\nabla^m A| \; |\nabla^{m+1} A|^2 \; |\nabla^{m+2} A| \\
&= 2 \cdot \sqrt{2}|\nabla^{m+2} A|\sqrt{\Lambda + |\nabla^m A|^2} \cdot \frac{4}{\sqrt{2}\sqrt{\Lambda + |\nabla^m A|^2}} |\nabla^m A| \; |\nabla^{m+1} A|^2 \\
&\leq 2|\nabla^{m+2} A|^2 \left(\Lambda + |\nabla^m A|^2\right) + 8\frac{|\nabla^m A|^2}{\Lambda + |\nabla^m A|^2} |\nabla^{m+1} A|^4
\end{aligned}$$

Substituting this into (‡), we get a cancellation on the $|\nabla^{m+2} A|^2$ term:

$$
\begin{aligned}
\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f &\leq C(n,m,K_m)\left(1 + |\nabla^{m+1}A|^2\right)\left(\Lambda + |\nabla^m A|^2\right) \\
&\quad + |\nabla^{m+1}A|^2\left(-2|\nabla^{m+1}A|^2 + C(n,m,K_m)\left(1+|\nabla^m A|^2\right)\right) \\
&\quad + 8\frac{|\nabla^m A|^2}{\Lambda + |\nabla^m A|^2}|\nabla^{m+1}A|^4 \\
&= \left(-2 + 8\frac{|\nabla^m A|^2}{\Lambda + |\nabla^m A|^2}\right)|\nabla^{m+1}A|^4 + C(n,m,K_m)\left(1+|\nabla^m A|^2\right)|\nabla^{m+1}A|^2 \\
&\quad + C(n,m,K_m)\left(1 + |\nabla^{m+1}A|^2\right)\left(\Lambda + |\nabla^m A|^2\right)
\end{aligned}
$$

At this point we make the observation that

$$
t \mapsto \frac{t}{\Lambda + t}
$$

is increasing, so in view of that $|\nabla^m A|^2 \leq K_m$ and $\Lambda = 15K_m$, we have:

$$
\begin{aligned}
\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f &\leq \left(-2 + 8\frac{K_m}{\Lambda + K_m}\right)|\nabla^{m+1}A|^4 + C(n,m,K_m)\left(1+|\nabla^m A|^2\right)|\nabla^{m+1}A|^2 \\
&\quad + C(n,m,K_m)\left(1 + |\nabla^{m+1}A|^2\right)\left(\Lambda + |\nabla^m A|^2\right) \\
&= -\frac{3}{2}|\nabla^{m+1}A|^4 + C(n,m,K_m)\left(1+|\nabla^m A|^2\right)|\nabla^{m+1}A|^2 \\
&\quad + C(n,m,K_m)\left(1 + |\nabla^{m+1}A|^2\right)\left(15K_m + |\nabla^m A|^2\right) \\
&= -\frac{3}{2}|\nabla^{m+1}A|^4 + C(n,m,K_m)|\nabla^m A|^2 + C(n,m,K_m)
\end{aligned}
$$

where parenthetical expression of the second term and entire last term were absorbed into a new constant $C(n,m,K_m)$, since $|\nabla^m A|^2 \leq K_m$. We now use $2ab \leq a^2 + b^2$ to break off $|\nabla^{m+1}A|^2$ from the middle term:

$$
\begin{aligned}
\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f &= -\frac{3}{2}|\nabla^{m+1}A|^4 + 2\cdot\sqrt{2}C(n,m,K_m)\cdot\frac{1}{\sqrt{2}}|\nabla^{m+1}A|^2 + C(n,m,K_m) \\
&\leq -|\nabla^{m+1}A|^4 + 2C(n,m,K_m)^2 + C(n,m,K_m) \leq -|\nabla^{m+1}A|^4 + C(n,m,K_m)
\end{aligned}
$$

for a newer yet constant $C(n,m,K_m)$. But $f = |\nabla^{m+1}A|^2(15K_m + |\nabla^m A|^2) \leq 16K_m|\nabla^{m+1}A|^2$, and hence:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq -\frac{1}{(16K_m)^2} f^2 + C(n, m, K_m)$$

in the subset $\mathcal{O} \subset M^n \times (-1, 0)$ at which motion by mean curvature occurs. We are in a position to apply the localized weak maximum principle, (2.5.2). Let $\sigma \in (0, 1)$, $t_0 \in (-1, 0)$ be arbitrary, and choose:

$$\varphi_{\sigma,t_0}(x, t) := (t - t_0) \left(\sigma^2 - |x|^2\right)_+^3$$

which has $c_{\varphi_{t_0}} \leq C_0$, independently of σ, t_0. Observe that support $\varphi_{\sigma,t_0}(\cdot, t) \subseteq \overline{B_\sigma}(0)$, so that $M_t \cap$ support $\varphi_{\sigma,t_0}(\cdot, t)$ is a compact subset of $B_1(0)$ for all $t \in (t_0, 0)$. The localized weak maximum principle (2.5.2) gives:

$$\max_{M_t} f \varphi_{\sigma,t_0} \leq \max_{M_{t_0}} f \varphi_{\sigma,t_0} + C(n) C_0 \left(1 + C(n, m, K_m)\right) \left(1 + (16K_m)^2\right) = C(n, m, K_m)$$

for every $t \in (t_0, 0)$, where we've dropped the maximum over M_{t_0} since φ_{σ,t_0} vanishes identically at at time t_0, and we've absorbed all constants into a new $C(n, m, K_m)$. Rearranging and recalling the definitions of f, φ_{σ,t_0}:

$$\begin{aligned}
C(n, m, K_m) &\geq \max_{M_t} |\nabla^{m+1} A|^2 \left(\Lambda + |\nabla^m A|^2\right) (t - t_0) \left(\sigma^2 - |x|^2\right)_+^3 \\
&\geq \Lambda \max_{M_t} |\nabla^{m+1} A|^2 (t - t_0) \left(\sigma^2 - |x|^2\right)_+^3 \\
&\geq \Lambda \sup_{M_t \cap B_\theta(0)} |\nabla^{m+1} A|^2 (t - t_0) \left(\sigma^2 - \theta\right)^3
\end{aligned}$$

Rearranging once again and absorbing $\Lambda = 15K_m$ into $C(n, m, K_m)$:

$$\sup_{M_t \cap B_\theta} |\nabla^{m+1} A|^2 \leq \frac{C(n, m, K_m)}{(t - t_0)(\sigma^2 - \theta^2)^3}$$

for all $t_0 \in (-1, 0)$, $t \in (t_0, 0)$, $\sigma \in (0, 1)$, $0 < \theta < \sigma$. Letting $t_0 \searrow -1$ and $\sigma \nearrow 1$, we get:

$$\sup_{M_t \cap B_\theta} |\nabla^{m+1} A|^2 \leq \frac{C(n, m, K_m)}{(t + 1)(1 - \theta^2)^3}$$

and result now follows by restricting our attention to $t \in (-\theta^2, 0)$, and recalling that $\theta^2 < \theta$. $\square$

THEOREM 4.1.2. Let (M_t) move by mean curvature in $B_\varrho(x_0) \times (t_0 - \varrho^2, t_0)$, and suppose that:

$$|A|^2 \leq \frac{c_0}{\varrho^2}$$

in $B_\varrho(x_0) \times (t_0 - \varrho^2, t_0)$. For a fixed $\theta \in (0,1)$ we can then estimate all higher curvature derivatives according to:

$$|\nabla^m A|^2 \leq \frac{C(n, m, \theta, c_0)}{\varrho^{2(m+1)}}$$

in $B_{\theta\varrho}(x_0) \times (t_0 - \theta^2\varrho^2, t_0)$.

PROOF. Most of the work has been done in the previous lemma. Let us rescale our family (M_t) parabolically around (x_0, t_0) with a factor of ϱ, to obtain a new family $\left(M_s^{(x_0,t_0),\varrho}\right)$ moving by mean curvature in $B_1(0) \times (-1, 0)$, which satisfies $|\tilde{A}|^2 \leq c_0$ in $B_1(0) \times (-1, 0)$. Here $\tilde{A}$ is the curvature of the rescaled family of submanifolds.

Fix some $\theta \in (0, 1)$ and $m \geq 1$. By inductively using the previous lemma (4.1.1) in the domains:

$$B_1(0) \times (-1, 0) \quad \rightsquigarrow \quad B_{\theta^{1/m}} \times \left(-\left(\theta^{1/m}\right)^2, 0\right) \quad \rightsquigarrow \quad B_{\theta^{2/m}} \times \left(-\left(\theta^{2/m}\right)^2, 0\right) \quad \rightsquigarrow \quad \cdots \quad \rightsquigarrow \quad B_\theta(0) \times (-\theta^2, 0)$$

we obtain the corresponding estimates:

$$|\tilde{A}|^2 \leq c_0 \quad \rightsquigarrow \quad |\nabla\tilde{A}|^2 \leq \frac{c_1(n, \theta, c_0)}{\left(1 - \theta^{1/m}\right)^4} \quad \rightsquigarrow \quad |\nabla^2\tilde{A}|^2 \leq \frac{c_2(n, \theta, c_0)}{\left(1 - \theta^{1/m}\right)^{4\cdot 2}} \quad \rightsquigarrow \quad \cdots \quad \rightsquigarrow \quad |\nabla^m\tilde{A}|^2 \leq \frac{c_m(n, \theta, c_0)}{\left(1 - \theta^{1/m}\right)^{4m}}$$

Let us expand on the constant dependence mechanism. Per lemma (4.1.1), each upper bound c_k depends on n, k, the shrinking factor $\theta^{1/m}$, and the maximum of previous derivatives' upper bounds $c_0, \ldots, c_{k-1}$ effective in a domain slightly larger (by the inverse shrinking factor) than the one we wish to establish bounds in. Inductively, then, each constant c_k depends merely on $n, k, \theta^{1/m}$ and the base case bound c_0, which explains our notation above.

After absorbing all θ, m dependence into a new constant, the last estimate reads:

$$|\nabla^m\tilde{A}|^2 \leq \frac{c_m(n, \theta, c_0)}{\left(1 - \theta^{1/m}\right)^{4m}} = c_m(n, \theta, c_0)$$

This yields the required estimate upon undoing the parabolic rescaling and returning to (M_t). $\square$

4.2. Mean-Square Flatness Implies Uniform Flatness

In this section we establish a relationship between two kinds of measure theoretic flatness of a manifold. Namely, we guarantee that a manifold that is mean-squared ($\mathscr{L}^2$) close to being flat will actually also be uniformly close to being flat (i.e. almost $\mathscr{L}^\infty$ flat on an entire neighborhood). In the next section we see how that translates into a curvature bound.

We need the following mean value inequality lemma before proceeding any further:

LEMMA 4.2.1. Let (M_t) be move by mean curvature in $B_{\varrho_0}(x_0) \times (t_0 - \varrho_0^2, t_0)$, and let $f \in \mathcal{C}^2\left(B_{\varrho_0}(x_0) \times (t_1, t_0)\right) \cap \mathcal{C}^0\left(B_{\varrho_0}(x_0) \cap (t_1, t_0]\right)$ be such that:

$$f \geq 0 \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq 0 \quad \text{or alternatively just} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f = 0$$

in $B_{\varrho_0}(x_0) \times (t_0 - \varrho_0^2, t_0)$. If M_t reaches x_0 at time t_0 then for any $0 < \varrho < \varrho_0$ we have:

$$f(x_0, t_0)^2 \Theta(M_t, x_0, t_0) \leq \frac{C(n)}{\varrho^{n+2}} \int_{t_0 - \varrho^2}^{t_0} \int_{M_t \cap B_\varrho(x_0)} f^2$$

for some constant $C(n)$.

PROOF. By the product rule (A.3.3.a) we have:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f^2 = 2f\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f - 2|\nabla f|^2 \leq -2|\nabla f|^2$$

where the last inequality follows by our assumptions on f. By the product rule once again, and for any test function φ we have:

$$\begin{aligned}
\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(f^2\varphi^2) &= \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f^2 \cdot \varphi^2 + f^2\left(\frac{\partial}{\partial t}\right)\varphi^2 - 2\nabla f^2 \cdot \nabla\varphi^2 \\
&\leq -2|\nabla f|^2\varphi^2 + f^2\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi^2 - 2\nabla f^2 \cdot \nabla\varphi^2 \\
&= -2|\nabla f|^2\varphi^2 + f^2\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi^2 - 8\varphi\nabla f \cdot f\nabla\varphi
\end{aligned}$$

Estimating the last term by $2x \cdot y \leq \frac{1}{2}|x|^2 + 2|y|^2$ we get:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(f^2\varphi^2) \leq -2|\nabla f|^2\varphi^2 + f^2\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi^2 + 2|\nabla f|^2\varphi^2 + 8f^2|\nabla\varphi|^2$$
$$= f^2\left(\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi^2 + 8|\nabla\varphi|^2\right)$$

Denote the term inside the parentheses by g. Then by Huisken's monotonicity (3.5.1) we have for $t \in (t_0 - \varrho_0^2, t_0)$:

$$\frac{d}{dt}\int_{M_t}\Phi_{(x_0,t_0)}f^2\varphi^2 = \int_{M_t}\Phi_{(x_0,t_0)}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(f^2\varphi^2) - \left|H - \frac{\nabla^\perp\Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2\Phi_{(x_0,t_0)}f^2\varphi^2$$
$$\leq \int_{M_t}\Phi_{(x_0,t_0)}f^2 g \tag{$\dagger_1$}$$

At this point let us define $C_\varrho := B_\varrho(x_0) \times (t_0 - \varrho^2, t_0)$. Choose φ to be a test function such that $\mathbf{1}_{C_{\varrho/2}} \leq \varphi \leq \mathbf{1}_{C_\varrho}$ and

$$|\varphi| + \varrho|D\varphi| + \varrho^2\left(|D^2\varphi| + |D_t\varphi|\right) \leq C_0(n) \tag{$\dagger_2$}$$

From our differential identities (2.3.1) and (A.3.1) we have:

$$g := \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi^2 + |\nabla\varphi|^2 = D_t\varphi^2 - \Delta_{\mathbb{R}^{n+1}}\varphi^2 + D^2\varphi^2(\nu,\nu) \leq \frac{C_1(n)}{\varrho^2}\mathbf{1}_{\overline{C_\varrho}\setminus C_{\varrho/2}}$$

where the last inequality follows from ($\dagger_2$). Finally, from backwards heat kernel estimate (3.4.2.d) we know that:

$$\Phi_{(x_0,t_0)}\mathbf{1}_{\overline{C_\varrho}\setminus C_{\varrho/2}} \leq C_2(n)\varrho^{-n} \tag{$\dagger_3$}$$

Summarizing, from ($\dagger_1$) we get:

$$\frac{d}{dt}\int_{M_t}\Phi_{(x_0,t_0)}f^2\varphi^2 \leq \int_{M_t}\Phi_{(x_0,t_0)}f^2 g = \int_{M_t\cap\{\varphi\neq 0\}}\Phi_{(x_0,t_0)}f^2 g = \int_{M_t\cap B_\varrho(x_0)}\Phi_{(x_0,t_0)}f^2 g$$
$$= \int_{M_t\cap B_\varrho(x_0)}\Phi_{(x_0,t_0)}f^2\left(\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\varphi^2 + |\nabla\varphi|^2\right)$$

Combining with ($\dagger_2$), ($\dagger_3$), and setting $C(n) := C_1(n)C_2(n)$:

$$\frac{d}{dt}\int_{M_t}\Phi_{(x_0,t_0)}f^2\varphi^2 \le \frac{C_1(n)}{\varrho^2}\int_{M_t\cap B_\varrho(x_0)}\Phi_{(x_0,t_0)}f^2\mathbf{1}_{\overline{C_\varrho}\setminus C_{\varrho/2}} \le \frac{C(n)}{\varrho^{n+2}}\int_{M_t\cap B_\varrho(x_0)} f^2$$

Integrating over $t \in [t_0-\varrho^2, t_0-\delta]$, $\delta > 0$ small, and keeping in mind that φ vanishes at $t = t_0 - \varrho^2$, we have:

$$\int_{M_{t_0-\delta}}\Phi_{(x_0,t_0)}f^2\varphi^2 = \frac{C(n)}{\varrho^{n+2}}\int_{t_0-\varrho^2}^{t_0-\delta}\int_{M_t\cap B_\varrho(x_0)} f^2 \le \frac{C(n)}{\varrho^{n+2}}\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)} f^2$$

Letting $\delta \searrow 0$, we conclude by the extraction theorem (3.7.5) that:

$$f(x_0,t_0)^2\varphi(x_0,t_0)^2\Theta(M_t,x_0,t_0) \le \frac{C(n)}{\varrho^{n+2}}\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)} f^2$$

which is the required result, in view of our lower bound result on densities at points that are reached, (3.7.8). □

The following lemma shows how $\mathscr{L}^2$ estimates translate into $\mathscr{L}^\infty$ flatness at our center point.

LEMMA 4.2.2. Let (M_t) move by mean curvature in $B_{\varrho_0}(x_0)\times(t_0-\varrho_0^2, t_0)$. For any $0<\varrho<\varrho_0$ and $\theta\in(0,1)$:

$$\sup_{t\in(t_0-\theta^2\varrho^2,t_0)}\ \sup_{x\in M_t\cap M_{\theta\varrho}(x_0)}(x-x_0)_{n+1}^2 \le \frac{C(n)}{(1-\theta)^{n+2}\varrho^{n+2}}\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}(x-x_0)_{n+1}^2$$

where $C(n)$ is the same constant as in (4.2.1), and $(\cdot)_{n+1}$ denotes the $(n+1)$-st component of an element of $\mathbb{R}^{n+1}$.

PROOF. Consider the function $f(x,t) := (x-x_0)_{n+1}$ which evidently satisfies:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f = 0$$

Fix $t_1\in(t_0-\theta^2\varrho^2, t_0)$ and $x_1\in M_{t_1}\cap B_{\theta\varrho}(x_0)$. Notice that $B_{(1-\theta)\varrho_0}(x_1)\subseteq B_{\varrho_0}(x_0)$, and $(t_1-(1-\theta)^2\varrho^2, t_1)\subseteq(t_0-\varrho^2,t_0)$, because $\theta^2+(1-\theta)^2<1$. Since (M_t) reaches x_1 at time t_1 and f satisfies the conditions of the theorem, applying the mean value inequality (4.2.1) to $B_{(1-\theta)\varrho_0}(x_1)\times(t_1-(1-\theta)^2\varrho_0^2, t_1)$ we get:

$$(x_1-x_0)_{n+1}^2 \le \frac{C(n)}{(1-\theta)^{n+2}\varrho^{n+2}}\int_{t_1-(1-\theta)^2\varrho^2}^{t_1}\int_{M_t\cap B_{(1-\theta)\varrho}(x_1)}(x-x_0)_{n+1}^2 \le \frac{C(n)}{(1-\theta)^{n+2}\varrho^{n+2}}\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}(x-x_0)_{n+1}^2$$

But our choice of $t_1\in(t_0-\theta^2\varrho^2,t_0)$ and $x_1\in M_{t_1}\cap B_{\theta\varrho}(x_0)$ was arbitrary, so the result follows. □

We actually don't even need to center both estimates at x_0! If the manifolds are $\mathscr{L}^2$-flat near (x_0, t_0), then they will actually also be $\mathscr{L}^\infty$-flat around (y, τ) for (y, τ) sufficiently close to (x_0, t_0). The result is proved below.

LEMMA 4.2.3. Let (M_t) move by mean curvature in $B_{\varrho_0}(x_0) \times (t_0 - \varrho_0^2, t_0)$, and suppose additionally that $0 < \varrho < \varrho_0$ is such that:

$$\sup_{t\in(t_0-\varrho^2,t_0)} \frac{1}{\varrho^{n+2}} \int_{M_t\cap B_\varrho(x_0)} (x - x_0)_{n+1}^2 \leq K$$

for some $K < \infty$. Let $\eta, \theta \in (0, 1)$ be such that $\eta + \theta < 1$. Then for all $\tau \in \big(t_0 - \theta^2\varrho^2, t_0\big)$ and all $y \in M_\tau \cap B_{\theta\varrho}(x_0)$ we have:

$$\sup_{t\in(\tau-\eta^2\varrho^2,\tau)} \quad \sup_{x\in M_\tau\cap B_{\eta\varrho}(y)} \quad (x - y)_{n+1}^2 \leq \frac{2C(n)K}{(1 - \eta - \theta)^{n+2}} \quad \varrho^2$$

where $C(n)$ is the same constant as in (4.2.1).

PROOF. Notice that since $\tau \in \big(t_0 - \theta^2\varrho^2, t_0\big)$ and $y \in M_\tau \cap B_{\theta\varrho}(x_0)$, (4.2.2) gives:

$$\begin{aligned}
(y - x_0)_{n+1}^2 &\leq \frac{C(n)}{(1-\theta)^{n+2}\varrho^{n+2}} \int_{t_0-\varrho^2}^{t_0} \int_{M_t\cap B_\varrho(x_0)} (z - x_0)_{n+1}^2 \, d\mathscr{H}^n(z) \\
&\leq \frac{C(n)}{(1-\theta)^{n+2}\varrho^{n}} \sup_{t\in(t_0-\varrho^2,t_0)} \int_{M_t\cap B_\varrho(x_0)} (z - x_0)_{n+1}^2 \, d\mathscr{H}^n(z) \\
&\leq \frac{C(n)K\varrho^2}{(1-\theta)^{n+2}}
\end{aligned}$$

Now choose $t \in (\tau - \eta^2\varrho^2, \tau)$ and $x \in M_t \cap B_{\eta\varrho}(y)$. Notice that since $\theta^2 + \eta^2 < (\eta + \theta)^2$, we have $(\tau - \eta^2\varrho^2, \tau) \subseteq (t_0 - \theta^2\varrho^2 - \eta^2\varrho^2, t_0) \subseteq \big(t_0 - (\eta + \theta)^2\varrho^2, t_0\big)$. Additionally, we have $x \in M_t \cap B_{\eta\varrho}(y) \subseteq B_{(\eta+\theta)\varrho}(x_0)$. Therefore, by applying (4.2.2) with $\eta + \theta$ in place of θ, we get:

$$\begin{aligned}
(x - x_0)_{n+1}^2 &\leq \frac{C(n)}{(1-\eta-\theta)^{n+2}\varrho^{n+2}} \int_{t_0-\varrho^2}^{t_0} \int_{M_t\cap B_\varrho(x_0)} (z - x_0)_{n+1}^2 \, d\mathscr{H}^n(z) \\
&\leq \frac{C(n)}{(1-\eta-\theta)^{n+2}\varrho^{n}} \sup_{t\in(t_0-\varrho^2,t_0)} \int_{M_t\cap B_\varrho(x_0)} (z - x_0)_{n+1}^2 \, d\mathscr{H}^n(z) \\
&\leq \frac{C(n)K\varrho^2}{(1-\eta-\theta)^{n+2}}
\end{aligned}$$

for all $t \in (\tau - \eta^2\varrho^2, \tau)$ and $x \in M_t \cap B_{\eta\varrho}(y)$.

Summarizing, we know from basic algebra that $(x-y)^2_{n+1} \leq 2(x-x_0)^2_{n+1} + 2(y-x_0)^2_{n+1}$, and hence combining our two previous estimates gives:

$$\begin{aligned}
\sup_{t\in(\tau-\eta^2\varrho^2,\tau)} \quad \sup_{x\in M_\tau\cap B_{\eta\varrho}(y)} \quad (x-y)^2_{n+1} &\leq 2(y-x_0)^2_{n+1} + 2 \sup_{t\in(\tau-\eta^2\varrho^2,\tau)} \quad \sup_{x\in M_\tau\cap B_{\eta\varrho}(y)} \quad (x-x_0)^2_{n+1} \\
&\leq \frac{C(n)K\varrho^2}{(1-\theta)^{n+2}} + \frac{C(n)K\varrho^2}{(1-\eta-\theta)^{n+2}} \\
&\leq \frac{2C(n)K}{(1-\eta-\theta)^{n+2}} \quad \varrho^2
\end{aligned}$$

which is the required estimate. □

4.3. Measure Theoretic Flatness Controls Curvature

In this final section we will establish a relationship between geometric flatness (section 4.1) and measure theoretic flatness (section 4.2). In particular, we will guarantee that we can control the curvature norm ($\mathcal{C}^2$) inside neighborhoods within which our manifold is close to being uniformly flat ($\mathscr{L}^\infty$). This requires a blow up argument similar in nature to Allard's regularity theorem, only this is substantially weaker because it requires a closeness assumption for arbitrarily small radii, rather than for a fixed small radius. The section will be concluded by noticing that we have enough tools to interrelate all types of flatness and get uniform curvature bounds to all orders of differentiation from just mean-squared flatness estimates.

We begin with a special case lemma requiring that we work away from the first singular time.

LEMMA 4.3.1. There exist constants $\varepsilon_0, c_0 > 0$ such that if (M_t) moves by mean curvature in $B_3(0) \times \left(-3^2, \delta\right)$, for some $\delta > 0$, and satisfies:

$$\sup_{t \in (\tau - \sigma^2, \tau)} \sup_{x \in M_t \cap B_\sigma(y)} (x-y)_{n+1}^2 \leq \varepsilon_0 \sigma^2$$

for every $\sigma \in (0,1)$, $\tau \in (-1,0)$, and $y \in M_\tau \cap B_1(0)$, then:

$$(1-\sigma)^2 \sup_{t \in (\sigma^2, 0)} \sup_{x \in M_t \cap B_\sigma} |A(x,t)|^2 \leq c_0$$

for all $\sigma \in (0,1)$, where $|A|^2$ denotes the squared norm of the second fundamental form.

PROOF. We argue by contradiction. If the claim were false then for every $j \geq 1$ we would be able to find a (M_t^j) moving by mean curvature in $B_3(0) \times \left(-3^2, \delta_j\right)$ such that:

$$\sup_{t \in (\tau - \sigma^2, \tau)} \sup_{x \in M_t^j \cap B_\sigma(y)} (x-y)_{n+1}^2 \leq \frac{\sigma^2}{j}$$

for every $\sigma \in [0,1]$, $\tau \in (-1,0)$, and $y \in M_\tau^j \cap B_1(0)$, but at the same time:

$$\sup_{\sigma \in [0,1]} \left((1-\sigma)^2 \sup_{t \in (-\sigma^2, 0)} \sup_{x \in M_t^j \cap B_\sigma(0)} |A^j(x,t)|^2 \right) \to \infty \quad \text{as } j \to \infty$$

Notice that since we have assumed that our solutions exist (and are smooth) past time $t = 0$ the operator $|A^j(x,t)|$ is continuous up to $t = 0$, and hence the expressions above can be equivalently rewritten as:

$$\sup_{t\in[\tau-\sigma^2,\tau]} \sup_{x\in M_t^j\cap\overline{B_\sigma(y)}} (x-y)_{n+1}^2 \leq \frac{\sigma^2}{j} \tag{$\dagger_1$}$$

and

$$\gamma_j^2 := \sup_{\sigma\in[0,1]} \left((1-\sigma)^2 \sup_{t\in[-\sigma^2,0]} \sup_{x\in M_t^j\cap\overline{B_\sigma(0)}} |A^j(x,t)|^2 \right) \to \infty \quad \text{as } j\to\infty$$

Notice that by replacing σ with $1-\sigma$, we may further rewrite

$$\gamma_j^2 = \sup_{\sigma\in[0,1]} \left(\sigma^2 \sup_{t\in[-(1-\sigma)^2,0]} \sup_{x\in M_t^j\cap\overline{B_{1-\sigma}(0)}} |A^j(x,t)|^2 \right) \tag{$\dagger_2$}$$

By continuity of $|A^j(x,t)|^2$ and since $\gamma_j^2 < \infty$ in view of smoothness up to time $t = 0$, we can choose $\sigma_j \in [0,1]$, $\tau_j \in [-(1-\sigma)^2, 0]$, and $y_j \in M_{\tau_j} \cap \overline{B_{1-\sigma_j}(0)}$ at which maximization occurs in ($\dagger_2$), i.e.:

$$\sigma_j^2 |A^j(y_j,\tau_j)|^2 = \gamma_j^2$$

Let us also define $\lambda_j := |A^j(y_j,\tau_j)|^{-1}$. Notice that by the maximality of γ_j^2 over $\sigma\in[0,1]$ in ($\dagger_2$), setting $\sigma = \frac{\sigma_j}{2}$ gives:

$$\frac{\sigma_j^2}{4} \sup_{t\in\left[-\left(1-\frac{\sigma_j^2}{4}\right),0\right]} \sup_{x\in M_t^j\cap\overline{B_{1-\frac{\sigma_j}{2}}(0)}} |A^j(x,t)|^2 \leq \gamma_j^2 = \sigma_j^2 |A(y_j,\tau_j)|^2$$

or equivalently:

$$\sup_{t\in\left[-\left(1-\frac{\sigma_j^2}{4}\right),0\right]} \sup_{x\in M_t^j\cap\overline{B_{1-\frac{\sigma_j}{2}}(0)}} |A^j(x,t)|^2 \leq 4|A(y_j,\tau_j)|^2$$

Elementary computations show that $\left[\tau_j - \frac{\sigma_j^2}{4}, \tau_j\right] \subseteq \left[-\left(1-\frac{\sigma_j^2}{4}\right), 0\right]$ and $\overline{B_{\frac{\sigma_j}{2}}(y_j)} \subseteq \overline{B_{1-\frac{\sigma_2}{2}}(0)}$, and hence for every $j \geq 1$:

$$\sup_{t\in\left[\tau_j-\frac{\sigma_j^2}{4},\tau_j\right]} \sup_{x\in M_t^j\cap\overline{B_{\frac{\sigma_j}{2}}(y_j)}} |A^j(x,t)|^2 \le 4|A^j(y_j,\tau_j)|^2$$

Let us make a parabolic change of coordinates. For every $j \ge 1$ the family:

$$\tilde{M}_s^j := \lambda_j^{-1}\left(M^j_{\lambda_j^2 s+\tau_j} - y_j\right)$$

moves by mean curvature in $B_{\lambda_j^{-1}\frac{\sigma_j}{2}}(0) \times \left(\lambda_j^{-2}\frac{\sigma_j^2}{4} - \delta_j', \delta_j'\right)$ for some $\delta_j' > 0$, and satisfies:

$$|\tilde{A}^j(0,0)|^2 = 1 \text{ and } 0 \in \tilde{M}_s^j \tag{$\dagger_3$}$$

for every $s \in \left[-\lambda_j^{-2}\frac{\sigma_j^2}{2}, 0\right]$, and it also satisfies the estimate:

$$\sup_{s\in\left[-\lambda_j^2\frac{\sigma_j^2}{4},0\right]} \sup_{y\in\tilde{M}_s^j\cap\overline{B_{\lambda_j^{-1}\frac{\sigma_j}{2}}(0)}} |\tilde{A}^j(y,s)|^2 \le 4$$

Recall that $\gamma_j \to \infty$ and $\lambda_j = \frac{\sigma_j}{\gamma_j} \le \frac{1}{\gamma_j} \to 0$. Therefore, for every $R > 0$ there is a $j_0(R) \ge 1$ such that:

$$\sup_{s\in[-R^2,0]} \sup_{y\in\tilde{M}_s^j\cap\overline{B_R(0)}} |\tilde{A}^j(y,s)|^2 \le 4$$

for every $j \ge j_0(R)$. As a matter of fact, we get uniform bounds on all derivatives $|\nabla^m \tilde{A}^j|^2$ on slightly smaller balls by the curvature estimate propagation theorem (4.1.2) from an earlier section. Consequently, for fixed $s \le 0$ we know by Arzelà-Ascoli (E.3.1) that after possibly passing to a subsequence $\tilde{M}_s^j$ converges in $\mathcal{C}^\infty$ on compact sets to some smooth limiting surface $\tilde{M}_s^\infty$. Notice that by taking $j \to \infty$ and $s = 0$ in ($\dagger_3$), $\mathcal{C}^\infty$ convergence gives:

$$|\tilde{A}^\infty(0,0)|^2 = 1 \text{ and } 0 \in M_0^\infty \tag{$\dagger_4$}$$

Now fix $R > 0$, let $j \ge j_0(R)$, and return to ($\dagger_1$). For $\tau = \tau_j$, $y = y_j$, and $\frac{\sigma}{2}$ in place of σ, the inequality reads:

$$\sup_{t\in\left[\tau_j-\frac{\sigma^2}{4},\tau_j\right]} \sup_{x\in M_t^j\cap\overline{B_{\frac{\sigma}{2}}(y_j)}} (x-y_j)_{n+1}^2 \le \frac{\sigma^2}{4j}$$

In parabolic coordinates this reads:

$$\sup_{s\in\left[-\lambda_j^{-2}\frac{\sigma^2}{4},0\right]} \quad \sup_{y\in\tilde{M}_s^j\cap\overline{B_{\lambda_j^{-1}\frac{\sigma}{2}}(0)}} \quad y_{n+1}^2 \leq \lambda_j^{-2}\frac{\sigma_j^2}{4j}$$

Letting $\sigma = \lambda_j^{-1}\frac{\sigma_j}{2}$, we see that for $j \geq j_0(R)$:

$$\sup_{s\in[-R^2,0]} \quad \sup_{y\in\tilde{M}_s^j\cap\overline{B_R(0)}} \quad y_{n+1}^2 \leq \frac{R}{j}$$

Letting $j \to \infty$ we conclude that $\tilde{M}_s^\infty \cap \overline{B_R(0)} \subset \mathbb{R}^n \times \{0\}$ for every $s \in [-R^2, 0]$, but this contradicts the curvature statement in ($\dagger_4$). Consequently, the required result follows. □

In combination with the previously established result relating $\mathscr{L}^2$ to $\mathscr{L}^\infty$ flatness, we are now in a position to use $\mathscr{L}^2$ estimates to get curvature bounds.

PROPOSITION 4.3.2. There exist constants $\varepsilon_0, c_0 > 0$ such that if (M_t) moves by mean curvature in $B_{\varrho_0}(x_0) \times \left(t_0 - \varrho_0^2, t_0\right)$, and for some $0 < \varrho < \varrho_0$ satisfies:

$$\sup_{t\in(t_0-\sigma^2,t_0)} \quad \frac{1}{\sigma^{n+2}}\int_{M_t\cap B_\sigma(x_0)} (x-y)_{n+1}^2 \, d\mathscr{H}^n(x) \leq \frac{\varepsilon_0}{2\cdot 3^n C(n)}$$

for all $0 < \sigma < \varrho$, where $C(n)$ is as in the 2nd mean value inequality (4.2.1), then:

$$|A(x,t)|^2 \leq \frac{c_0}{(1-\theta)^2\varrho^2}$$

for all $\theta \in \left(0, \frac{1}{3}\right)$, $t \in (t_0 - \theta^2\varrho^2, t_0)$, and $x \in M_t \cap B_{\theta\varrho}(x_0)$, where $|A|^2$ denotes the squared norm of the second fundamental form.

PROOF. Choose any $\tau \in (t_0 - \varrho_0^2 + \varrho^2, t_0)$, which is possible since $\varrho < \varrho_0$. Then (M_t) moves by mean curvature in $B_\varrho(x_0) \times (\tau - \varrho^2, \tau + \delta)$, for some $\delta > 0$. If we rescale according to:

$$M'_{t'} := \frac{3}{\varrho}\left(M_{\frac{\varrho^2}{3^2}t'+\tau} - x_0\right)$$

then $(M_{t'})$ moves by mean curvature in $B_3(0) \times (-3^2, \delta')$ for some $\delta' > 0$. Rescaling the given inequality (which is actually scale-invariant), and writing σ' in place of $\frac{\sigma}{\varrho}$ to be notationally compatible with (4.3.1), gives:

$$\sup_{t' \in (-3^2(\sigma')^2, 0)} \quad \frac{1}{3^{n+2}(\sigma')^{n+2}} \int_{M'_{t'} \cap B_{3\sigma'}(0)} x_{n+1}^2 \, d\mathscr{H}^n(x) \leq \frac{\varepsilon_0}{2 \cdot 3^n C(n)}$$

for all $\sigma' \in (0,1)$. Employ (4.2.3) on $B_{3\sigma'}(0) \times (-3^2(\sigma')^2, 0)$ with $\theta = \eta = \frac{1}{3}$ to turn the $\mathscr{L}^2$ estimate into an $\mathscr{L}^\infty$ estimate:

$$\sup_{t' \in (\tau' - 3^2(\sigma')^2, \tau')} \quad \sup_{x \in M'_{t'} \cap B_{3\sigma'}(y')} (x - y')_{n+1}^2 \leq \frac{2C(n)}{\left(1 - \frac{1}{3} - \frac{1}{3}\right)^n} \cdot \frac{\varepsilon_0}{2 \cdot 3^n C(n)} \cdot (\sigma')^2 = \varepsilon_0 (\sigma')^2$$

for every $\sigma' \in (0,1)$, $\tau' \in (-1, 0)$, and $y' \in M_{\tau'} \cap B_{\sigma'}(0)$. In view of the previous lemma, and since $M'_{t'}$ exists smoothly past 0 by our choice of τ, we conclude that:

$$|A'(x,t)|^2 \leq \frac{c_0}{(1 - \sigma')^2}$$

for all $\sigma' \in (0,1)$, $t \in (-(\sigma')^2, 0)$, and $x \in M_t \cap B_{\sigma'}(0)$. Rescaling to our original size, and writing θ rather than σ', we conclude that:

$$|A(x,t)|^2 \leq \frac{c_0}{(1 - \theta)^2 \varrho^2} \tag{‡}$$

for all $\theta \in (0,1)$, $t \in \left(\tau - \frac{\theta^2 \varrho^2}{3^2}, \tau\right)$, $x \in B_{\frac{\theta\varrho}{3}}(x_0)$.

Notice that the right hand side is independent of $\tau \in (t_0 - \varrho_0^2 + \varrho^2, t_0)$, so we may let $\tau \nearrow t_0$ and write θ in place of $\frac{\theta}{3}$, to conclude that (‡) is thus true when $\theta \in \left(0, \frac{1}{3}\right)$, $t \in \left(t_0 - \theta^2 \varrho^2, t_0\right)$, and $x \in B_{\theta\varrho}(0)$. $\square$

The following result is now a combination of all results in this Chapter. Its proof is now a trivial matter, but we label it a theorem nonetheless because it is of paramount importance.

THEOREM 4.3.3. There exist constants $\varepsilon_0, c_0 > 0$ such that if (M_t) moves by mean curvature in $B_{\varrho_0}(x_0) \times \big(t_0 - \varrho_0^2, t_0\big)$, and for some $0 < \varrho < \varrho_0$ satisfies:

$$\sup_{t \in (t_0 - \sigma^2, t_0)} \quad \frac{1}{\sigma^{n+2}} \int_{M_t \cap B_\sigma(x_0)} (x-y)_{n+1}^2 \, d\mathscr{H}^n(x) \leq \frac{\varepsilon_0}{2 \cdot 3^n C(n)}$$

for all $0 < \sigma < \varrho$, where $C(n)$ is as in the 2nd mean value inequality (4.2.1), then:

$$|\nabla^m A(x,t)|^2 \leq \frac{C(n,m,\theta)}{\varrho^2}$$

for all $\theta \in \big(0, \frac{1}{3}\big)$, $m \geq 0$, $t \in (t_0 - \theta^2 \varrho^2, t_0)$, and $x \in M_t \cap B_{\theta\varrho}(x_0)$.

PROOF. This is a straightforward result if we glue our pieces together carefully. The assumptions of the theorem allow us to use (4.3.2) and obtain a curvature estimate, which in turn translates into an estimate to derivatives of all orders by (4.1.2). $\square$

CHAPTER 5

The First Singular Time

This chapter is devoted to the proof of the main regularity theorem for mean curvature flow, which asserts that under certain conditions on the family (M_t) of submanifolds moving by mean curvature in $\mathcal{U} \times (t_1, t_0)$, the set of singularities that forms at t_0 cannot be "too big," in the sense that it has $\mathscr{H}^n$-measure zero. We will require that the reader has some knowledge of geometric measure theory and is comfortable with concepts such as Hausdorff densities and approximate tangent spaces; if not, refer to [**Sim83**] as needed.

Our assumptions throughout the chapter are:

(1) *(rectifiability at time t_0)* there exists a set M_{t_0} which is countably n-rectifiable, and that
(2) *(unit density and area continuity at time t_0)* the surfaces M_t converge to M_{t_0} in the sense of measures, i.e.

$$\lim_{t \nearrow t_0} \int_{M_t} \varphi = \int_{M_{t_0}} \varphi$$

for every time independent $\varphi \in \mathcal{C}_c^0(\mathcal{U})$.

Under these hypotheses we will show that $\mathscr{H}^n$-a.e. point in $\mathcal{U}$ constitutes a regular point of the flow at the first singular time. Our adopted definition is:

A point $x_0 \in \mathcal{U}$ will be said to be **regular** at time t_0 provided there exists a sufficiently small neighborhood of x_0 in which connected components of the flow approaching x_0 at time t_0 converge to a smooth n-dimensional submanifold around x_0. These points are often also referred to as being **non-singular**.

Our approach to Brakke's regularity theorem in this chapter is laid out as follows. We commence with section 5.1 which proves crude but consequential Hausdorff density and $|H|^2$ bounds that we capitalize on in section 5.2 where we prove a result that allows us to approximate integrals on the limiting surface by integrals on nearby smooth surfaces, at least around a large class of points. Finally, we conclude with section 5.3 that proves near $\mathscr{L}^2$-flatness of our manifolds around $\mathscr{H}^n$-a.e. point of the limiting surface, which in view of Chapter 4 and an Arzelà-Ascoli type compactness theorem from Chapter E of the appendix yields Brakke's regularity theorem.

5.1. Density Bounds and $|H|^2$ Finiteness

It is possible to obtain time independent bounds on the area of our manifolds within a domain slightly smaller than our original one without assuming any sort of regularity at time t_0. This is a simple application of the forward area estimate (3.3.1), and it's a rather weak result but it is needed in what follows.

LEMMA 5.1.1. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$. For any point in $\mathcal{U}$ we can find a d_0 such that $(t_0 - d_0^2, t_0) \subseteq (t_1, t_0)$, the ball B_{d_0} centered at the given point is compactly contained in $\mathcal{U}$, and

$$\sup_{t \in (t_0 - d_0^2, t_0)} \frac{\mathscr{H}^n \left(M_t \cap B_{d_0}\right)}{d_0^n} < \infty$$

In particular, for any center point in $\mathcal{U}$ we may take $d_0 := \frac{1-\eta}{2\sqrt{1+2n}} \delta$ where δ is the smaller of $\sqrt{t_0 - t_1}$ and the distance from the center of the ball to $\mathbb{R}^{n+1} \setminus \mathcal{U}$, and $\eta \in (0,1)$ is arbitrary.

PROOF. Evidently, it is sufficient to show that that $\mathscr{H}^n(M_t \cap B_{d_0})$ is bounded from above independently of t, rather than deal with the ratio. Choose some ball $B_{(1-\eta)\delta} \subseteq \mathcal{U}$ with δ as above. Let $\theta_0 := \frac{1}{2\sqrt{1+2n}}$ and observe that $(1+2n)\theta_0^2 < 1$. Then by the forward area estimate (3.3.1) on $B_{(1-\eta)\delta}$ and $d_0 := \theta_0(1-\eta)\delta$ we have:

$$\mathscr{H}^n(M_t \cap B_{d_0}) \leq C(n, \theta_0) \mathscr{H}^n(M_{t_0 - d_0^2} \cap B_{(1-\eta)\delta})$$

for all $t \in [t_0 - d_0^2, t_0)$. The right hand side is independent of t, and the result follows. □

Let us introduce some notation in order to facilitate brevity in the future. We say that (M_t) has single center/single radius density ratios bounded by A_0 in $B_{d_0} \times (t_0 - d_0^2, t_0)$ provided:

$$\sup_{t \in (t_0 - d_0^2, t_0)} \frac{\mathscr{H}^n \left(M_t \cap B_{d_0}\right)}{d_0^n} \leq A_0$$

What we just showed, then, is that any family of submanifolds moving by mean curvature in a domain must have bounded single center/single radius density ratios in a domain slightly smaller than the original one.

LEMMA 5.1.2. Let (M_t) move by mean curvature in $B_{d_0} \times (t_0 - d_0^2, t_0)$, where its single center/single radius density ratios are bounded by A_0. Then there exists a constant $c_0(n) \geq 2$ such that for $d := \frac{d_0}{c_0(n)}$ we have:

$$\int_{t_0-d^2}^{t_0} \int_{M_t \cap B_d} |H|^2 < \infty \tag{5.1.2.a}$$

and also:

$$\frac{\mathscr{H}^n\left(M_t \cap B_\varrho(x_0)\right)}{\varrho^n} \leq A \tag{5.1.2.b}$$

for some constant $A = C(n)A_0$ and all $x_0 \in B_d$, $0 < \varrho \leq d$, and $t \in (t_0 - \varrho^2, t_0)$. In particular, we may choose $c_0 := 4\sqrt{1+2n}$ in which case $C(n) = 2^{5n/2}e^{1/4}(3/4)^{-3}(1+2n)^{n/2}$.

If (M_t) additionally satisfies the unit density and area continuity hypotheses in $B_{d_0} \times (t_0 - d_0^2, t_0)$, then the density bound above extends to t_0 as well.

PROOF. Define $\theta_0 := \frac{1}{4\sqrt{1+2n}}$. Since $(1+2n)\theta_0^2 < 1$, the $|H|^2$ spacetime estimate (3.3.1) on $B_{d_0} \times (t_0 - d_0^2, t_0)$ tells us that:

$$\int_{t_0-\theta_0^2 d_0^2}^{t} \int_{M_\tau \cap B_{\theta_0 d_0}} |H|^2 \leq C(n,\theta_0)\mathscr{H}^n\left(M_{t_0-\theta_0^2 d_0^2} \cap B_{d_0}\right)$$

for every $t \in [t_0 - \theta_0^2 d_0^2, t_0)$. Since the right hand side is independent of t, we may let $t \nearrow t_0$ and conclude, by the monotone convergence theorem, that for $d := \theta_0 d_0$:

$$\int_{t_0-d^2}^{t_0} \int_{M_t \cap B_d} |H|^2 \leq C(n,\theta_0)\mathscr{H}^n\left(M_{t_0-d^2} \cap B_{d_0}\right) < \infty \tag{$\dagger_1$}$$

Now choose $x_0 \in B_d$ and observe that since $\theta_0 \leq \frac{1}{2}$, we have $B_d(x_0) \subseteq B_{d_0/2}(x_0) \subseteq B_{d_0}$. Then by the density upper bound theorem (3.6.3) on the domain $B_{d_0/2}(x_0)$ we have, for $\theta_1 := 2\theta_0$ which still satisfies $(1+2n)\theta_1^2 < 1$:

$$\frac{\mathscr{H}^n(M_t \cap B_\varrho(x_0))}{\varrho^n} \leq C(n,\theta_1)\frac{\mathscr{H}^n\left(M_{t_0-\theta_1^2(d_0/2)^2} \cap B_{d_0/2}(x_0)\right)}{(d_0/2)^n}$$

for every $0 < \varrho \leq \theta_1 \cdot d_0/2 = (2\theta_0)(d_0/2) = \theta_0 d_0 = d$ and every $t \in (t_0 - \varrho^2, t_0)$. We may estimate the right hand side from above using given our A_0 bound on single center/single radius density ratios, and conclude that:

$$\frac{\mathscr{H}^n(M_t \cap B_\varrho(x_0))}{\varrho^n} \le C(n,\theta_1)\frac{\mathscr{H}^n\left(M_{t_0-d^2} \cap B_{d_0/2}(x_0)\right)}{(d_0/2)^n} \le C(n,\theta_1)\frac{\mathscr{H}^n\left(M_{t_0-d^2} \cap B_{d_0}\right)}{(d_0/2)^n} \le 2^n A_0 C(n,\theta_1)$$

The density upper bound theorem had $C(n,\theta) = 2^{n/2}e^{1/4}(1-(1+2n)\theta^2)^{-3}\theta^{-n}$, so with $\theta_1 = 2\theta_0 = \frac{1}{2\sqrt{1+2n}}$:

$$\frac{\mathscr{H}^n(M_t \cap B_\varrho(x_0))}{\varrho^n} \le 2^{5n/2}e^{1/4}(3/4)^{-3}(1+2n)^{n/2}A_0 = A$$

for $t \in (t_0 - d^2, t_0)$ and $0 < \varrho \le d$ and $x \in B_d$.

To extend to $t = t_0$ we need a different argument, because our manifolds may decompose into something as bad as a countably $\mathscr{H}^n$-rectifiable set. We need to use the unit density and area continuity hypothesis at t_0. Fix $t = t_0$, and ϱ. Let $\varepsilon > 0$ be small, and let χ be a continuous function such that $\mathbf{1}_{B_{\varrho-\varepsilon}(x_0)} \le \chi \le \mathbf{1}_{B_\varrho(x_0)}$. Then:

$$\frac{\mathscr{H}^n(M_{t_0} \cap B_{\varrho-\varepsilon}(x_0))}{\varrho^n} \le \varrho^{-n}\int_{M_{t_0}} \chi = \varrho^{-n}\lim_{t\nearrow t_0}\int_{M_t}\chi \le \limsup_{t\nearrow t_0}\frac{\mathscr{H}^n(M_t \cap B_\varrho(x_0))}{\varrho^n} \le A$$

Now $\varepsilon > 0$ was arbitrary, so let us choose $\varepsilon_n \searrow 0$ and observe that $\cup_n \left(M_{t_0} \cap B_{\varrho-\varepsilon_n}(x_0)\right) = M_{t_0} \cap B_\varrho(x_0)$ so taking the limit $n \to \infty$ in the estimate above we get the required extension to t_0. $\square$

Let us introduce some more notation for future brevity. We will say that the family of submanifolds (M_t) satisfies $|H|^2$ finiteness in $B_d \times (t_0 - d^2, t_0)$ provided (5.1.2.a) holds, i.e.

$$\int_{t_0-d^2}^{t_0}\int_{M_t \cap B_d} |H|^2 < \infty$$

We will also say density ratios are uniformly bounded by A in $B_d \times (t_0 - d^2, t_0)$ for radii up to d provided (5.1.2.b):

$$\frac{\mathscr{H}^n\left(M_t \cap B_\varrho(x_0)\right)}{\varrho^n} \le A$$

for all $x_0 \in B_d$, $0 < \varrho \le d$, $t \in (t_0 - d^2, t_0)$. In summary, lemma (5.1.2) says that single center/single radius density bounds guarantee $|H|^2$ finiteness and uniform density bounds in a slightly smaller domain.

Notice that in talking about bounded density ratios for radii up to d in the domain $B_d \times (t_0 - d^2, t_0)$, we will be looking at balls $B_d(x_0)$, $x_0 \in B_d$, which may very well stretch outside of the domain $B_d \times (t_0 - d^2, t_0)$ we started with. This will be a non-issue for us, because in our applications we will always arrange for there to be movement by mean curvature in a large enough surrounding domain, e.g. a $B_{d_0} \times (t_0 - d_0^2, t_0)$ as in the lemma.

5.2. α-Good Points and Integral Approximations

In this section we aim to show that local integrals over the limiting surface M_{t_0} can be approximated well by integrals over nearby surfaces M_t. Clearly we cannot expect this to be true everywhere, but rather only around sufficiently "good" points of the limiting surface. The following set of points will prove to suffice:

DEFINITION 5.2.1. For any $\alpha > 0$ and family of submanifolds (M_t) we define the set $G^{\alpha}_{t_0}$:

$$G^{\alpha}_{t_0} := \left\{ x_0 \in \mathbb{R}^{n+1} \ : \ \limsup_{\sigma \searrow 0} \sigma^{-n} \int_{t_0-\sigma^2}^{t_0} \int_{M_t \cap B_\sigma(x_0)} |H|^2 \leq \alpha^2 \right\}$$

and we refer to it as the set of "α-good" points at time t_0. □

In other words, near "good" points x_0 the spacetime integral of $|H|^2$ over a domain $B_\sigma(x_0) \times (t_0 - \sigma^2, t_0)$ decays at least as fast as σ^n. Most remarkably, $|H|^2$ finiteness over $B_d \times (t_0 - d^2, t_0)$ guarantees that almost every point in B_d is α-good. This does not even require motion by mean curvature!

LEMMA 5.2.2. Let (M_t) be a family of submanifolds satisfying $|H|^2$ finiteness in $B_d \times (t_0 - d^2, t_0)$. Then for any $\alpha > 0$, $\mathscr{H}^n$-a.e. $x_0 \in B_d$ is an α-good point. In other words, $\mathscr{H}^n(B_d \setminus G^{\alpha}_{t_0}) = 0$.

PROOF. We proceed by a covering argument. Let $0 < \delta < d$ be arbitrary. For $x \in B_d \setminus G^{\alpha}_{t_0}$, by definition of (non-)good points we can choose $\varrho(x) > 0$ such that $\varrho(x) < \frac{\delta}{10}$, $B_{\varrho(x)}(x) \subseteq B_d$, and:

$$\varrho(x)^{-n} \int_{t_0-\varrho(x)^2}^{t_0} \int_{M_t \cap B_\varrho(x)} |H|^2 \geq \alpha^2 \tag{$\dagger_1$}$$

Evidently, the collection $\left\{ B_{\varrho(x)}(x) \ : \ x \in B_d \setminus G^{\alpha}_{t_0} \right\}$ covers the domain $B_d \setminus G^{\alpha}_{t_0}$ so by the 5-lemma (see, e.g. **[Sim83]**) we can choose a sequence $\{x_j\}$ such that the balls $B_{\varrho(x_j)}(x_j)$ are pairwise disjoint, and:

$$B_d \setminus G^{\alpha}_{t_0} \subseteq \cup_j B_{5\varrho(x_j)}(x_j)$$

But $5\varrho(x_j) < \delta/2$, and hence $\left\{ B_{5\varrho(x_j)}(x_j) \right\}$ is an admissible covering of $B_d \setminus G^{\alpha}_{t_0}$ with diameters below δ, and hence we can bound the δ-diameter Hausdorff outer measure as follows:

$$\begin{aligned}\mathscr{H}^n_\delta(B_d \setminus G^\alpha_{t_0}) &\leq \sum_j 5^n \varrho(x_j)^n \omega_n \leq 5^n \omega_n \sum_j \varrho(x_j)^n \\ &\overset{(\dagger_1)}{\leq} 5^n \omega^n \alpha^{-2} \sum_j \int_{t_0-\varrho(x_j)^2}^{t_0} \int_{M_t \cap B_{\varrho(x_j)}(x_j)} |H|^2\end{aligned}$$

But by definition $\varrho(x_j) < \delta$, so:

$$\begin{aligned}\mathscr{H}^n_\delta(B_d \setminus G^\alpha_{t_0}) &\leq 5^n \omega^n \alpha^{-2} \sum_j \int_{t_0-\delta^2}^{t_0} \int_{M_t \cap B_{\varrho(x_j)}(x_j)} |H|^2 \\ &\leq 5^n \omega^n \alpha^{-2} \int_{t_0-\delta^2}^{t_0} \int_{M_t \cap B_d} |H|^2\end{aligned}$$

By assumption the latter integral is finite (since $\delta < d$), so letting $\delta \searrow 0$ makes the right hand side vanish, and the result follows. □

Recall that we did not define α-good points just to study the growth of $|H|^2$ integrals. The following result motivates our interest in α-good points: near them we can approximate integrals over the limiting surface M_{t_0} by integrals over nearby surfaces M_t.

PROPOSITION 5.2.3. Let (M_t) move by mean curvature in $B_d \times (t_0 - d^2, t_0)$, where it satisfies the unit density and area continuity hypothesis at t_0, $|H|^2$ finiteness, and has density ratios for radii up to d bounded by A. Then for any α-good point x_0 at time t_0 there exists a $0 < \varrho_0 \leq d$ such that:

$$\sup_{t \in (t_0 - \varrho^2, t_0)} \left| \int_{M_t} \varphi - \int_{M_{t_0}} \varphi \right| \leq \left(4\alpha^2 |\varphi| + 2\alpha\sqrt{A} \cdot \varrho |D\varphi| \right) \varrho^n$$

for every $0 < \varrho \leq \varrho_0$ and $\varphi \in \mathcal{C}^1_c(B_\varrho(x_0))$

PROOF. Choose an α-good point x_0 at time t_0, and let $\varrho_0 > 0$ be small enough that $B_{\varrho_0}(x_0) \subseteq B_d$ and:

$$\int_{t_0-\varrho^2}^{t_0} \int_{M_t \cap B_\varrho(x_0)} |H|^2 \leq 4\alpha^2 \varrho^n \tag{$\dagger_1$}$$

for every $0 < \varrho \leq \varrho_0$. This is certainly possible, due to our definition of "α-good points."

Let $\varphi \in \mathcal{C}^1_c(B_\varrho(x_0))$. Now let $\tau \in (t_0 - \varrho^2, t_0)$ and $\delta > 0$ be small. By the fundamental theorem of calculus we have:

$$\begin{aligned}\left|\int_{M_{t_0-\delta}}\varphi-\int_{M_\tau}\varphi\right| &\leq \left|\int_\tau^{t_0-\delta}\int_{M_t} H\cdot D\varphi-|H|^2\varphi\right| \\ &\leq \int_{t_0-\varrho^2}^{t_0}\int_{M_t}|H||D\varphi|+|H|^2|\varphi|\end{aligned}$$

We let $\tau \searrow t_0-\varrho^2$ as well as $\delta \searrow 0$ and employ area continuity at t_0 to get:

$$\sup_{t\in(t_0-\varrho^2,t_0)}\left|\int_{M_{t_0}}\varphi-\int_{M_t}\varphi\right| \leq \int_{t_0-\varrho^2}^{t_0}\int_{M_t}|H||D\varphi|+|H|^2|\varphi| \tag{$\dagger_2$}$$

It remains to bound the right hand side as appropriate. Since x_0 is an α-good point and $\varrho \leq \varrho_0$ we evidently have:

$$\begin{aligned}\int_{t_0-\varrho^2}^{t_0}\int_{M_t}|H|^2|\varphi| &= \int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap\{\varphi>0\}}|H|^2|\varphi| = \int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}|H|^2|\varphi| \\ &\leq |\varphi|\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}|H|^2 \overset{(\dagger_1)}{\leq} 4\alpha^2|\varphi|\varrho^n\end{aligned} \tag{$\dagger_3$}$$

We estimate the remaining part of the right hand side as follows:

$$\begin{aligned}\int_{t_0-\varrho^2}^{t_0}\int_{M_t}|H||D\varphi| &= \int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}|H||D\varphi| \leq |D\varphi|\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}|H|\cdot 1 \\ &\leq |D\varphi|\left(\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}|H|^2\right)^{1/2}\left(\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}1\right)^{1/2}\end{aligned}$$

where the last step follows by Cauchy-Schwarz. By the spacetime integral bound ($\dagger_1$) on $|H|^2$:

$$\leq |D\varphi|\cdot 2\alpha\varrho^{n/2}\cdot\left(\int_{t_0-\varrho^2}^{t_0}\int_{M_t\cap B_\varrho(x_0)}1\right)^{1/2}$$

Our density ratio bounds assumption gives $\mathscr{H}^n(M_t\cap B_\varrho(x_0)) \leq A\varrho^n$, so:

$$\leq |D\varphi|\cdot 2\alpha\varrho^{n/2}\cdot\left(\varrho^2\cdot A\varrho^n\right)^{1/2} = 2\alpha\sqrt{A}\varrho|D\varphi|\cdot\varrho^n \tag{$\dagger_4$}$$

The result follows directly by combining ($\dagger_2$), ($\dagger_3$), and ($\dagger_4$). □

We can trivially extend this estimate to good points in a parabolically rescaled context.

COROLLARY 5.2.4. Let (M_t) move by mean curvature in $B_d \times (t_0 - d^2, t_0)$, where it satisfies the unit density and area continuity hypothesis at t_0, $|H|^2$ finiteness, and has density ratios for radii up to d bounded by A. Then for any α-good point x_0 at time t_0 there exists a $0 < \varrho_0 \leq d$ such that:

$$\sup_{s \in (-R^2, 0)} \left| \int_{M_s^{(x_0,t_0),\lambda}} \psi - \int_{M_0^{(x_0,t_0),\lambda}} \psi \right| \leq \left(4\alpha^2 |\psi| + 2\alpha\sqrt{A} \cdot R|D\psi| \right) R^n$$

for every $0 < R \leq \lambda^{-1}\varrho_0$ and $\psi \in \mathcal{C}_c^1(B_R(0))$

PROOF. We choose the same radius cutoff ϱ_0 as in (5.2.3), and let $\varrho := \lambda R$, which evidently satisfies $\varrho \leq \varrho_0$. We define $\varphi(x) := \psi\left(\lambda^{-1}(x - x_0)\right)$, observe that $\varphi \in \mathcal{C}_c^1(B_\varrho(x_0))$, $|\varphi| = |\psi|$, and $|D\varphi| = \lambda^{-1}|D\psi|$. Therefore we may apply the previously established integral comparison theorem (5.2.3) and change into parabolic coordinates, at which point the result follows. $\square$

5.3. Brakke Regularity Theorem

In this section we will complete the proof of the main regularity theorem, which has been the main focus of this work. We will essentially prove that $\mathscr{H}^n$-a.e. point is either not reached by our manifolds, or if it is then the manifolds are sufficiently $\mathscr{L}^2$-flat around it to apply our uniform curvature estimates from Chapter 4.

We begin by proving it in the special case of domains of the form $B_d \times (t_0 - d^2, t_0)$ under the additional assumption of density ratio bounds and $|H|^2$ finiteness, and then we conclude with some clean-up work and extend to general domains.

For notational simplification, let us introduce the following definition. We say that $x_0 \in \mathcal{U}$ is ε-regular at t_0 in the $\mathscr{L}^2$ sense if x_0 has an approximate tangent space $T_{x_0}M_{t_0}$ and there exists a $\varrho > 0$ such that $\overline{B_\varrho(x_0)} \subseteq \mathcal{U}$, $(t_0 - \varrho^2, t_0) \subset (t_1, t_0)$, and:

$$\sup_{t \in (t_0 - \sigma^2, t_0)} \quad \frac{1}{\sigma^{n+2}} \int_{M_t \cap B_\sigma(x_0)} (x - x_0)_\perp^2 \; d\mathscr{H}^n(x) \leq \varepsilon$$

for all $0 < \sigma < \varrho$. Here $(\cdot)_\perp$ denotes the (scalar) orthogonal component of a vector in $\mathbb{R}^{n+1}$ with respect to $T_{x_0}M_{t_0}$.

Notice that this is not a new definition; although we did not give them a name, in Chapter 4 we considered ε-regular points in the case $T_{x_0}M_{t_0} = \mathbb{R}^{n+1} \times \{0\}$ and showed that for $\varepsilon > 0$ sufficiently small, we can obtain curvature bounds (for all orders of differentiation) around ε-regular points.

PROPOSITION 5.3.1. Let (M_t) move by mean curvature in $B_d \times (t_0 - d^2, t_0)$ under the rectifiability, unit density, and area continuity hypotheses at t_0, satisfy $|H|^2$ finiteness and have density ratios bounded by A for radii up to d. Then there exists a constant $\alpha_1(n, A) > 0$ such that:

$$x_0 \text{ is not reached by } (M_t) \text{ at time } t_0 \text{ for } \mathscr{H}^n\text{-a.e. } x \in \left(B_d \cap G^\alpha_{t_0}\right) \setminus M_{t_0}$$

for all $0 < \alpha < \alpha_1$. Furthermore, for any $\varepsilon > 0$ there exists a constant $\alpha_2(n, A, \varepsilon) > 0$ such that:

$$x_0 \text{ is } \varepsilon\text{-regular at time } t_0 \text{ for } \mathscr{H}^n\text{-a.e. } x \in M_{t_0} \cap \left(B_d \cap G^\alpha_{t_0}\right)$$

for all $0 < \alpha < \alpha_2$.

PROOF. The proofs for both cases employ versions of the integral comparison theorems (5.2.3), (5.2.4). In the first case we wish to show that a certain point cannot be reached by contradicting Brakke's clearing out lemma (3.6.2). In the second case we wish to show directly that the $\mathscr{L}^2$-flatness integral can be bounded from above by $\varepsilon > 0$.

Case 1: Points that cannot be reached.

In this part we will work on $B_d \setminus M_{t_0}$.

We know that there exists a constant $C(n)$ such that for any $\varrho > 0$ we can construct a $\mathcal{C}^1_c$ cutoff $\mathbf{1}_{B_{\frac{\varrho}{2}}} \leq \varphi \leq \mathbf{1}_{B_\varrho}$ with $\varrho|D\varphi| \leq C(n)$. Let $M(n, A) := 2^n \left(4 + 2\sqrt{A}C(n)\right)$ and $\beta_0 := \frac{1}{2(2n+1)} \in \left(0, \frac{1}{2n+1}\right)$. For $\eta(n, \beta_0)$ as in Brakke's clearing out lemma (3.6.2) define:

$$\alpha_1(n, A) := \min\left\{1, \frac{\eta(n, \beta_0)}{2M(n, A)}\right\}$$

Choose any $\alpha > 0$ such that $\alpha < \alpha_1(n, A)$.

Then for each $\varrho \leq d$ we can construct a $\mathcal{C}^1_c$ cutoff function $\mathbf{1}_{B_{\frac{\varrho}{2}}} \leq \varphi \leq \mathbf{1}_{B_\varrho}$ such that:

$$2^n \left(4\alpha^2|\varphi| + 2\alpha\sqrt{A}\varrho|D\varphi|\right) \leq 2^n \left(4\alpha \cdot 1 + 2\alpha\sqrt{A} \cdot C(n)\right) < \frac{\eta(n, \beta_0)}{2} \tag{$\dagger_3$}$$

We know from geometric measure theory (see, e.g. **[Sim83]**) that $\mathscr{H}^n$-a.e. point in $B_d \setminus M_{t_0}$ has n-dimensional Hausdorff density zero, and from (5.2.2) that $\mathscr{H}^n$-a.e. point in B_d is in $G^\alpha_{t_0}$. Then $\mathscr{H}^n$-a.e. $x_0 \in B_d \setminus M_{t_0}$ is such that $\Theta^n(M_{t_0}, x_0) = 0$ and $x_0 \in G^\alpha_{t_0}$.

We claim that all such x_0 cannot be reached by (M_t) at t_0.

Indeed, fix such an x_0. Let $\varrho_0 \leq d$ be as in the integral comparison theorem (5.2.3). Since $\Theta^n(M_{t_0}, x_0) = 0$, we may choose a $\varrho > 0$ small enough that $\varrho < \varrho_0$ and:

$$\frac{\mathscr{H}^n(M_{t_0} \cap B_\varrho(x_0))}{\varrho^n} < \frac{\eta(n, \beta_0)}{2^{n+1}} \tag{$\dagger_4$}$$

For our cut-off function $\mathbf{1}_{B_{\frac{\varrho}{2}}(x_0)} \leq \varphi \leq \mathbf{1}_{B_\varrho(x_0)}$ we see that:

$$\begin{aligned}\frac{\mathscr{H}^n\left(M_{t_0-\beta_0\frac{\varrho^2}{4}}\cap B_{\frac{\varrho}{2}}(x_0)\right)}{\frac{\varrho^n}{2^n}} &= \frac{2^n}{\varrho^n}\mathscr{H}^n\left(M_{t_0-\frac{1}{4n}\frac{\varrho^2}{4}}\cap B_{\frac{\varrho}{2}}(x_0)\right)\\ &\leq \frac{2^n}{\varrho^n}\int_{M_{t_0-\frac{1}{16n}\varrho^2}\cap B_\varrho(x_0)}\varphi\\ &\leq \frac{2^n}{\varrho^n}\sup_{t\in(t_0-\varrho^2,t_0)}\int_{M_t\cap B_\varrho(x_0)}\varphi\end{aligned}$$

which by the integral approximation theorem (5.2.3) around α-good points gets bounded by:

$$\begin{aligned}&\leq \frac{2^n}{\varrho^n}\left((4\alpha^2|\varphi|+2\alpha\sqrt{A}\varrho|D\varphi|)\varrho^n+\int_{M_{t_0}\cap B_\varrho(x_0)}\varphi\right)\\ &= 2^n(4\alpha^2|\varphi|+2\alpha\sqrt{A}\varrho|D\varphi|)+\frac{2^n}{\varrho^n}\int_{M_{t_0}\cap B_\varrho(x_0)}\varphi\\ &\overset{(\dagger_3),(\dagger_4)}{<}\frac{\eta(n,\beta_0)}{2}+2^n\frac{\eta(n,\beta_0)}{2^{n+1}}=\eta(n,\beta_0)\end{aligned}$$

Consequently x_0 is not reached by (M_t), because Brakke's clearing out lemma (3.6.2) states that if x_0 were reached then the area ratio would actually be bounded from below by $\eta(n,\beta_0)$.

Case 2: Points that can be reached.

In this part we will work on $M_{t_0}\cap B_d$.

Fix $\chi\in\mathcal{C}_c^1(\mathbb{R}^{n+1})$ such that $\mathbf{1}_{B_1(0)}\leq\chi\leq\mathbf{1}_{B_2(0)}$ and define $\psi:=y_{n+1}^2\chi$. Let $M(n,A):=2^n\left(4|\psi|+4\sqrt{A}|D\psi|\right)$. Define:

$$\alpha_2(n,A,\varepsilon):=\min\left\{1,\frac{\varepsilon}{2M(n,A)}\right\}$$

Choose any $\alpha>0$ such that $\alpha<\alpha_2(n,A,\varepsilon)$. Then:

$$\left(4\alpha^2|\psi|+2\alpha\sqrt{A}\cdot 2\cdot|D\psi|\right)2^n<\frac{\varepsilon}{2}\qquad(\dagger_1)$$

From basic geometric measure theory (see, e.g. [**Sim83**]) we know that $\mathscr{H}^n$-a.e. point in $M_{t_0} \cap B_d$ has an approximate tangent space to M_{t_0}, and from (5.2.2) that $\mathscr{H}^n$-a.e. point in $M_{t_0} \cap B_d$ is in $G^{\alpha}_{t_0}$. Then $\mathscr{H}^n$-a.e. $x_0 \in M_{t_0} \cap B_d$ has an approximate tangent space $T_{x_0} M_{t_0}$ and is such that $x_0 \in G^{\alpha}_{t_0}$.

We claim that all such x_0 are ε-regular.

Fix such an x_0. Without loss of generality, $T_{x_0} M_{t_0} = \mathbb{R}^n \times \{0\}$. Notice that $M_0^{(x_0,t_0),\lambda}$ converges locally uniformly to $T_{x_0} M_{t_0} = \mathbb{R}^n \times \{0\}$, as $\lambda \searrow 0$, so by the classical dominated convergence theorem (B.2.1) there exists a $\varrho > 0$ such that:

$$\int_{M_0^{(x_0,t_0),\lambda}} \psi < \frac{\varepsilon}{2} + \int_{T_{x_0} M_{t_0}} \psi = \frac{\varepsilon}{2} + \int_{\mathbb{R}^n \times \{0\}} \chi y_{n+1}^2 = \frac{\varepsilon}{2} \tag{$\dagger_2$}$$

for every $\lambda < \varrho$. Without loss of generality, we may take $\varrho < \frac{\varrho_0}{2}$ where ϱ_0 is as in the parabolically rescaled integral comparison theorem (5.2.4). Notice then for all $\lambda < \varrho$ we have $2 \leq \lambda^{-1} \varrho_0$, and hence by the same theorem for a choice of radius $R = 2$ we have:

$$\sup_{s \in (-1,0)} \left| \int_{M_s^{(x_0,t_0),\lambda}} \psi - \int_{M_0^{(x_0,t_0),\lambda}} \psi \right| \leq \left(4\alpha^2 |\psi| + 2\alpha \cdot 2 \cdot |D\psi|\right) 2^n \overset{(\dagger_1)}{<} \frac{\varepsilon}{2}$$

where we have restricted the supremum to the left to $(-1, 0)$ instead of $\left(-2^2, 0\right)$, without affecting the direction of the inequality of course. Rearranging, we see that:

$$\sup_{s \in (-1,0)} \int_{M_s^{(x_0,t_0),\lambda}} \psi \leq \int_{M_0^{(x_0,t_0),\lambda}} \psi + \frac{\varepsilon}{2} \overset{(\dagger_2)}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for every $\lambda < \varrho$. Recalling that $\psi = \chi y_{n+1}^2$, with $\mathbf{1}_{B_1(0)} \leq \psi \leq \mathbf{1}_{B_2(0)}$, we conclude that:

$$\sup_{s \in (-1,0)} \int_{M_s^{(x_0,t_0),\lambda} \cap B_1(0)} y_{n+1}^2 \, d\mathscr{H}^n(y) \leq \sup_{s \in (-1,0)} \int_{M_s^{(x_0,t_0),\lambda}} \psi < \varepsilon$$

for every $\lambda < \varrho$. Scaling back to normal coordinates, writing $t = \lambda^2 s + t_0$, and using σ in place of λ:

$$\sup_{t \in (t_0 - \sigma^2, t_0)} \quad \frac{1}{\sigma^{n+2}} \int_{M_t \cap B_\sigma(x_0)} (x - x_0)_{n+1}^2 \, d\mathscr{H}^n(x) < \varepsilon$$

for every $0 < \sigma < \varrho$, which concludes our check that x_0 is ε-regular at t_0. $\square$

We can now prove the final theorem.

THEOREM 5.3.2. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$ under the rectifiability, unit density, and area continuity hypotheses at t_0. If $\text{reg}_{t_0}(M_t)$ denotes the set of regular points of the flow at time t_0, then $\mathscr{H}^n\left(\mathcal{U} \setminus \text{reg}_{t_0}(M_t)\right) = 0$.

PROOF. Let $\mathcal{V}$ be an arbitrary open set that is compactly contained in $\mathcal{U}$. Let δ be the smaller among $\sqrt{t_0 - t_1}$ and dist$(\mathcal{V}; \mathbb{R}^{n+1} \setminus \mathcal{U})$, and define $d_0 := \frac{1}{4\sqrt{1+2n}}\delta$ as in (5.1.1) with $\eta = \frac{1}{2}$. Then for any $x \in \mathcal{V}$ we may apply that lemma and conclude bounded single center/single radius density ratios:

$$\sup_{t \in (t_0 - d_0^2, t_0)} \frac{\mathscr{H}^n\left(M_t \cap B_{d_0}(x)\right)}{d_0^n} < \infty$$

Cover the open set $\mathcal{V}$ with a (finite) collection of balls $B_d(x_j)$, $j \geq 1$, where $d = \frac{d_0}{c_0(n)}$ is as in (5.1.2). Then by (5.1.2) we see that $|H|^2$ finiteness and bounded density ratios for radii up to d, i.e. (5.1.2.a) and (5.1.2.b), are both true on $B_d(x_j) \times (t_0 - d^2, t_0)$, for all $j \geq 1$.

Let $\varepsilon > 0$ be arbitrary. Denote the set of points that are either not reached by (M_t) at time t_0, or that are ε-regular at time t_0 by $\text{reg}_{\varepsilon, t_0}(M_t)$. By the special case of Brakke's regularity theorem (5.3.1) just established, we see that $\mathscr{H}^n\left(B_d(x_j) \setminus \text{reg}_{\varepsilon, t_0}(M_t)\right) = 0$, for all $j \geq 1$. Adding up:

$$\mathscr{H}^n\left(\mathcal{V} \setminus \text{reg}_{\varepsilon, t_0}(M_t)\right) \leq \sum_j \mathscr{H}^n\left(B_d(x_j) \setminus \text{reg}_{\varepsilon, t_0}(M_t)\right) = 0$$

Since we can exhaust $\mathcal{U}$ by a countable family of $\mathcal{V}$, we have $\mathscr{H}^n\left(\mathcal{U} \setminus \text{reg}_{\varepsilon, t_0}(M_t)\right) = 0$ for all $\varepsilon > 0$.

Let us now proceed to relate the sets $\text{reg}_{\varepsilon, t_0}(M_t)$ and $\text{reg}_{t_0}(M_t)$ for sufficiently small choices of ε. In particular, we will show that for ε_0 as in the curvature estimation theorem (4.3.3) and $\varepsilon > 0$ small enough that

$$\varepsilon < \frac{\varepsilon_0}{2 \cdot 3^n C(n)}$$

it will necessarily follow that $\text{reg}_{\varepsilon, t_0}(M_t) \subseteq \text{reg}_{t_0}(M_t)$. By definition there are two reasons for which a x_0 can belong to $\text{reg}_{\varepsilon, t_0}(M_t)$, so it suffices to check that in both cases the point x_0 is regular.

If x_0 is not reached by (M_t) then there exists an open neighborhood around x_0 that is not breached by the (M_t) near time t_0. Therefore there is no flow at all around x_0, so x_0 is vacuously a regular point.

The more interesting case is that of x_0 that are reached by (M_t) and are ε-regular at time t_0. Since $\varepsilon < \frac{\varepsilon_0}{2\cdot 3^n C(n)}$, and after rotating so that $T_{x_0}M_{t_0} = \mathbb{R}^n \times \{0\}$, (4.3.3) guarantees (e.g. with $\theta = \frac{1}{2}$) that we can control the norms of all curvature derivatives $|\nabla^m A|^2$ in some domain of the form $B_\varrho(x_0) \times (t_0 - \varrho^2, t_0)$ by:

$$\sup_{t\in(t_0-\varrho^2,t_0)} \sup_{x\in M_t\cap B_\varrho(x_0)} |\nabla^m A|^2 \leq \frac{C_m}{\varrho^{2(m+1)}} \quad \text{for all } m$$

By our version of Arzelà-Ascoli discussed in Chapter E of the appendix, our manifolds converge in $\mathcal{C}^\infty$ to a smooth limiting surface consisting of points that are reached by the flow, so x_0 is regular.

In any case, for sufficiently small $\varepsilon > 0$ we have established that $\operatorname{reg}_{\varepsilon,t_0}(M_t) \subseteq \operatorname{reg}_{t_0}(M_t)$, so:

$$\mathscr{H}^n\left(\mathcal{U} \cap \operatorname{sing}_{t_0}(M_t)\right) \leq \mathscr{H}^n\left(\mathcal{U} \setminus \operatorname{reg}_{\varepsilon,t_0}(M_t)\right) = 0$$

and Brakke's regularity theorem follows. □

Remark: Let us identify M_{t_0} with the set of limit points of the flow, since we have concluded that they are identical up to sets of $\mathscr{H}^n$-measure zero. With a slightly more delicate argument it is possible to show that for $\mathscr{H}^n$-a.e. point $x_0 \in M_{t_0}$ there exists an open set $\mathcal{U}$ containing x_0 such that $M_{t_0} \cap \mathcal{U}$ is a smooth n-dimensional manifold: a stronger regularity theorem than in **[Eck04]**, that is more in line with **[Bra78]** and **[Whi05]**. One may work towards this direction as follows. We begin by showing that for sufficiently small neighborhoods of x_0 and t sufficiently close to t_0, all disjoint connected components of M_t must be nearly parallel; if they were not, then they would intersect near x_0 and hence not be disjoint. Then for a sufficiently small neighborhood of x_0 and t sufficiently close to t_0, there would only be one component of M_t near x_0; otherwise, we could obtain two distinct sequences of nearly parallel components approaching x_0, contradicting $\Theta^n(M_{t_0}, x) = 1$ and area continuity. Since we are working close to x_0, there would exist a neighborhood of x_0 inside which M_t has a unique component for t near t_0. By the same Arzelà-Ascoli argument as above, these M_t would converge to a smooth manifold around x_0 but we now know that there are no other limiting points near x_0.

Afterword

We are now finished with our proof of Brakke's regularity theorem. What's more, with the exception of the appendices which complement the earlier discussion, we are also essentially finished with this thesis. That is not to say, however, that we are finished with our study of regularity. This afterword serves as an introduction to possible next steps one may wish to take to carry Brakke's results one step further and possibly generalize them. We will briefly discuss the following:

(1) Is Brakke's dimension of the singular set optimal?
(2) How restrictive are Brakke's hypotheses?
(3) How do we even pin our limiting surface down?

We will also try to motivate possible techniques (as conceived by the author) that may prove to be fruitful in getting a better understanding of the questions above. There has not been significant progress on the author's behalf, but one wishes that future researchers will find these motivations useful.

Hausdorff Dimesion of Singular Set

The first question that comes up is that of the dimensionality of the singular set. Can we not do better than say that the singular set has $\mathscr{H}^n$-measure zero, and hence Hausdorff dimension at most n?

In the special case of mean convex initial data White [**Whi00**] was able to improve Brakke's regularity theorem both by discarding the unit density and area continuity hypotheses and by establishing a better Hausdorff dimension bound. He showed that at each time t the singular set has $\mathscr{H}^{n-1}$-measure zero, that at almost every time it actually has $\mathscr{H}^{n-3}$-measure zero, and that in the two dimensional case the surfaces are smooth manifolds for almost all times. Observe that in view of the shrinking torus $\mathbb{T}^2$, the $n-1$ Hausdorff dimension bound is optimal.

In fact, all examples that have been constructed to date end up with a singular set of dimension at most $n-1$, so it is conjectured (but not known) that the optimal dimensionality bound in the general case is also $n-1$. Nevertheless, Brakke's dimension n bound is still the best that can be obtained in the general case.

Unit Density and Area Continuity Hypotheses

It is essential for the proof of Brakke's theorem to work that the limiting surface be countably n-rectifiable, have unit density, and for there to be no jump in area as we get to the first singular time; these were our rectifiability, unit density, and area continuity hypotheses. The regularity result will *not* hold if it turns out that we can somehow get certain very pathological limiting surfaces where the density exceeds one on sets of positive measure.

A strong maximum principle argument such as the one given in the preservation of embeddedness result of Chapter 2 guarantees that we cannot have sheets merging to form double density on entire open sets. However, it is very well conceivable that sets of positive $\mathscr{H}^n$-measure (but no interior points) may form where density exceeds one. The following example was even discussed by Brakke himself in [**Bra78**].

Fix a radius $R > 0$ and bound $C > 0$. We aim to construct a surface whose mean curvature is bounded by C that is simply a double density plane away from a ball of radius R and a catenoid within $R/2$ of the center. This can easily be done by shrinking the central neck of the catenoid enough (depending on R, C) to get the catenoid's sheets to almost merge together within $R/2$ of the center, and then interpolate a smooth two valued function between the double density plane and the catenoid's sheets. (See figure below.) Let us call this a catenoid patch.

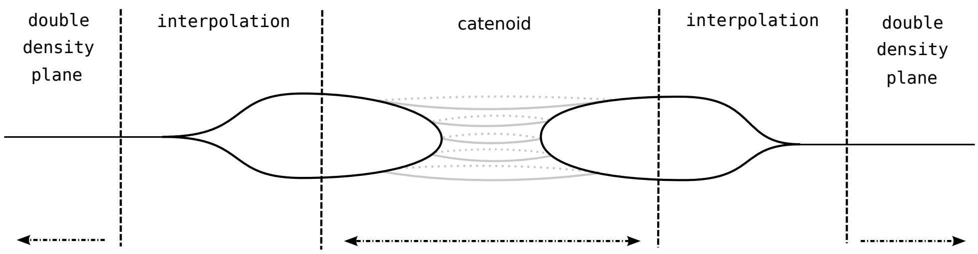


Figure 1: A catenoid patch such as the one described above.

On to our pathological construction. Start with the plane $\mathbb{R}^2 \times \{0\}$ and assign to it a density of two. Inductively remove small disks from the double density plane and replace them with catenoid patches so that no two catenoid patches overlap, the disks have a union that is dense in $\mathbb{R}^2 \times \{0\}$, but a set of positive $\mathscr{H}^2$-measure has been left untouched and continues to have double density. This can be done by appropriately choosing center points and radii within the field of rational numbers. Now this surface has no graph components because it has a dense set of holes. In particular then, Brakke's conclusion will fail if we somehow end up with a limiting surface like this.

Nonetheless, no examples of surfaces moving by mean curvature flow have been constructed to this date where the unit density and area continuity requirements are not satisfied automatically at the first singular time! One is left wondering, then, if the requirements are simply there for technical reasons or perhaps are necessary because of possible pathologies (such as the one described) might occur.

What is most off-putting about Brakke's hypotheses is that, in a sense, they are the best case scenario for a limiting surface and let us disregard possible bad singularities right away. The fact that we cannot (and have not yet been able to) construct such bad singularities does not mean they might not be out there.

The reader must observe that the unit density and area continuity hypotheses are very much interrelated. If we were to discard the unit density hypothesis, for example, then we would have to study the area continuity variant

$$\int_{M_t} f \to \int_{M_{t_0}} f\theta_0$$

for a density function θ_0 that is not necessarily equal to one at $\mathscr{H}^n$-a.e. point of M_{t_0}. This raises the question of what a likely candidate for the density function θ_0 is.

Our conjecture is that (under no unit density assumptions) the limit above is true with $\theta_0 = \Theta(M_t, \cdot, t_0)$, the Gauss density function. The intuition behind this is that Gauss density is a parabolic density function and takes averages of densities from sufficiently nearby points and times. Unfortunately, checking

$$\int_{M_t} f \to \int_{M_{t_0}} f\Theta(M_t, \cdot, t_0)$$

in the general case appears to be rather tricky, despite its apparent likelihood from specific examples.

Notice that if this conjecture were to be true, then Brakke's unit density hypothesis would be exactly that Gauss density is $\mathscr{H}^n$-a.e. equal to one on the limiting surface, in which case Brakke's theorem follows directly from (and is even generalized by) White's local regularity theorem [**Whi05**].

Limiting Surface Candidates

As the reader might recall, we began Chapter 5 by assuming that there is *some* set $M_{t_0} \subset \mathbb{R}^{n+1}$ to which our surfaces converge in a particular way, but we never actually explicitly discussed how such a set may be obtained in the first place; it has to be given to us on silver platter. Only as a side-consequence of our proof of Brakke's theorem do we get enough intuition about this elusive M_{t_0} to say that $\mathscr{H}^n$-a.e. point in it is reached by (M_t) and vice versa. The catch is that we merely established that *if* there exists a particularly nice (rectifiable, with unit density, and area continuity) limiting set candidate, *then* this limiting set is going to consist of precisely points that we would like it to contain (i.e. points reached by our surfaces).

But who is to guarantee the existence of such a candidate, and how might one go about constructing it in the first place? This question is of fundamental importance; if approached carefully it might allow one to get a good enough understanding of limiting surfaces to do away with Brakke's unit density and/or area continuity assumptions.

Let us leave the question of density to the side for a minute. Since we know *a posteriori* that up to $\mathscr{H}^n$-null sets the limiting surface is going to consist of points that are reached by our flow, an obvious set candidate is:

$$M_{t_0} := \left\{x_0 \in \mathbb{R}^{n+1} \text{ that is reached by } (M_t) \text{ as } t \nearrow t_0\right\}$$

Recall that there is a lower bound on the Gauss density of all points that are reached by our flow; in the case of smooth surfaces we even know from the appendix that this lower bound is 1. Combined with the fact that points that are not reached by the flow have Gauss density zero, an equivalent formulation of our candidate surface is:

$$M_{t_0} := \left\{x_0 \in \mathbb{R}^{n+1} \;:\; \Theta(M_t, x_0, t_0) \geq 1\right\}$$

The advantage of this definition is that our candidate M_{t_0} is immediately seen to be a closed set and hence $\mathscr{H}^n$-measurable, since Gauss density is upper semicontinuous.

We can even bound its $\mathscr{H}^n$-measure from above in the compact case. Let us restrict our attention to global solutions for simplicity. Denote by $N(\varrho)$ to be the largest number of non-overlapping balls of radius ϱ that one can center on M_{t_0}, and let these balls be $\{B_\varrho(x_j)\}$ for $1 \leq j \leq N$. For $\beta = \frac{1}{4n}$, and with repeated use of our standard test function and backwards heat kernel upper and lower bound estimates, we have:

$$\begin{aligned}\mathscr{H}^n\left(M_{t_0-\beta\varrho^2}\right) &\geq \sum_{j=1}^{N}\mathscr{H}^n\left(M_{t_0-\beta\varrho^2}\cap B_\varrho(x_j)\right) \geq (4\pi\beta)^{n/2}\varrho^n\sum_{j=1}^{N}\int_{M_{t_0-\beta\varrho^2}\cap B_\varrho(x_j)}\Phi_{(x_j,t_0)} \\ &\geq \frac{1}{2}(4\pi\beta)^{n/2}\varrho^n\sum_{j=1}^{N}\int_{M_{t_0-\beta\varrho^2}}\Phi_{(x_j,t_0)}\varphi_{(x_j,t_0),\frac{\varrho}{\sqrt{2}}} \geq \frac{1}{2}(4\pi\beta)^{n/2}\varrho^n\sum_{j=1}^{N}\Theta(M_t,x_j,t_0) \geq \frac{1}{2}\left(\frac{\pi}{n}\right)^{n/2}\varrho^n N(\varrho)\end{aligned}$$

And since the left hand side is decreasing in ϱ in view of the fact that surface areas decrease by mean curvature flow, we realize that we have bounded $N(\varrho)\varrho^n$ from above as $\varrho \searrow 0$. From basic measure theory, we know this implies that the $\mathscr{H}^n$-measure is also finite. Ostensibly we could perhaps work harder and deduce a stronger result: that $\mathscr{H}^n\left(M_{t_0}\right)$ is bounded from above by the $\mathscr{H}^n\left(M_t\right)$.

What is more impending, however, is checking that this candidate M_{t_0} is indeed a limiting set and deciding what density to assign to it. This is where one starts to question the leniency of the unit density assumption; under such an assumption we would have to have:

$$\int_{M_t} f \to \int_{\{\Theta(M_t,\,\cdot\,,t_0)\geq 1\}} f$$

which may seem absurd given that Gauss density assesses exactly the concentration of mass around a point in spacetime. It is far more likely that we will have to endow M_{t_0} with its own not-identically-one density function θ_0, which we conjectured in the previous section would have to be $\Theta(M_t,\,\cdot\,,t_0)$.

Bear in mind that we are not questioning the degree to which Brakke's assumptions may hold true in general, but we are simply trying to provide tools to check for their validity. In particular, we postulate that checking for unit density possibly amounts to having Gauss density equal to one $\mathscr{H}^n$-a.e., and checking for area continuity possibly amounts to having M_t converge as a Radon measure to the (closed) set of points that are reached by the flow.

APPENDIX A

Function Representation

In our study of surface evolution in this work we start with a fixed n-dimensional (abstract) manifold M^n which we embed in $\mathbb{R}^{n+1}$ with a family of embeddings varying smoothly on some time interval $I \subseteq \mathbb{R}$. In other words, we will have a smooth map

$$F : M^n \times I \to \mathbb{R}^{n+1}$$

such that $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$ is an embedding for each $t \in I$. We denote the image submanifolds by:

$$M_t := F(M^n, t) \subset \mathbb{R}^{n+1}$$

Naturally, we are going to want to study functions on the M_t. There are two ways to do this: we can either have functions defined strictly on (open subsets of) our manifolds, or we can have functions defined on (an open subset of) the ambient space $\mathbb{R}^{n+1}$ and restrict them to our manifolds when necessary. We devote this chapter to introducing notation and studying basic properties of these two different but interrelated approaches to defining functions. This chapter does not require that the M_t move by mean curvature.

A.1. Intrinsic Functions

These are functions that pertain only to the manifolds we are studying and not their neighboring points within the ambient space. In other words, they are functions whose extensions to the ambient space not only don't interest us but are also likely not to even exist, especially when our manifolds are not properly embedded in $\mathbb{R}^{n+1}$.

There is no simple way to go about having functions defined on just M_t for each $t \in I$, so we instead think of intrinsic functions as being defined on the background manifold M^n and also having time dependence. In other words, intrinsic functions are of the form:

$$f : \mathcal{O} \to \mathbb{R}^m \quad \text{for some } m \geq 1$$

where $\mathcal{O} \subseteq M^n \times I$ is open. We will use these types of functions to study geometric quantities (and their evolution) related to the submanifolds M_t such as the metric itself, the second fundamental form, the normal vector field, and so on.

We can consider tangential derivatives of the form

$$Xf : \mathcal{O} \to \mathbb{R}^m \quad \text{for a choice of a vector field } X \text{ on } M^n$$

or even

$$\frac{\partial}{\partial x^i} f : \mathcal{O} \to \mathbb{R}^m \quad \text{for a choice of local coordinates on } M^n$$

where differentiation is performed component-wise when f is vector-valued ($m \geq 2$).

Time derivatives are just derivatives in the t-direction in the product manifold $M^n \times I$, and are essentially the same as the coordinate derivatives above. They will be denoted by:

$$\frac{\partial}{\partial t} f : \mathcal{O} \to \mathbb{R}^m$$

Now for each $t \in I$, we can endow M^n by pulling back the metric induced on M_t from $\mathbb{R}^{n+1}$. This turns M^n into a Riemannian manifold (with a time dependent metric).

When f is scalar valued we may want to study the (spatial) tangential gradient

$$\nabla f : \mathcal{O} \to \left(\bigcup_{t \in I} TM_t \right) \subseteq \mathbb{R}^{n+1}$$

where we are thinking of all the tangent spaces as subspaces of $\mathbb{R}^{n+1}$ for the last inclusion to be true. Of course the time component is especially crucial here, because the induced metric on M^n changes in time. If t is unclear from the context we may label the tangential gradient $\nabla^{M_t} f$ for extra emphasis.

We can also study the Laplace-Beltrami operator

$$\Delta_{M_t} f : \mathcal{O} \to \mathbb{R}^m$$

where differentiation is performed component-wise when f is a vector-valued function.

Finally, when X is an $\mathbb{R}^{n+1}$-valued function or some vector field on $\mathcal{O}$, we define the tangential divergence operator in the standard way, and denote it:

$$\mathrm{div}_{M_t} X : \mathcal{O} \to \mathbb{R}$$

Summarizing, intrinsically defined functions $f : \mathcal{O} \to \mathbb{R}^m$ (or vector fields), for some open $\mathcal{O} \subseteq M^n \times I$, are equipped with the following differential symbols:

Table 1: Manifold differential symbols

$\frac{\partial}{\partial x^i}$	local coordinate differential in i-th direction
$\frac{\partial}{\partial t}$	time differential
∇, ∇^{M_t}	tangential gradient (M_t metric)
Δ_{M_t}	Laplace-Beltrami operator (M_t metric)
div_{M_t}	divergence operator (M_t metric)

A.2. Extrinsic Functions

Another class of functions we will tend to work with are functions that pertain only to the ambient space and are virtually unrelated to our manifolds M_t. We will not always be interested in behavior in all of $\mathbb{R}^{n+1}$, but rather in some open subset thereof. In other words these will be functions of the form:

$$f : \mathcal{U} \times I \to \mathbb{R}^m \quad \text{for some } m \geq 1$$

where $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ is open. We will use these functions to study exogenous quantities (and their evolution) such as the area and the density of the M_t within some small neighborhood of the ambient space. These functions will be used primarily in chapters having to do with integral estimates.

Notice that as functions on the flat space $\mathcal{U} \times I$, we may differentiate f in the classical sense of multivariable calculus. Spatial partial derivatives will be denoted by:

$$D_i f : \mathcal{U} \times I \to \mathbb{R}^m \quad \text{for } 1 \leq i \leq n+1$$

and the time derivative will be denoted by:

$$D_t f : \mathcal{U} \times I \to \mathbb{R}^m$$

The spatial Jacobian will be denoted by Df. If f is scalar valued, Df will also be used to denote the gradient:

$$Df : \mathcal{U} \times I \to \mathbb{R}^{n+1}$$

Higher order derivatives will be denoted accordingly; for instance, the classical (spatial) Hessian will be denoted by $D^2 f$. The Laplace operator will be denoted by:

$$\Delta_{\mathbb{R}^{n+1}} f : \mathcal{U} \times I \to \mathbb{R}^m$$

where differentiation is performed component-wise on vector-valued functions.

We will also use the divergence operator for vector fields $X : \mathcal{U} \times I \to \mathbb{R}^{n+1}$, which we will denote by:

$$\mathrm{div}_{\mathbb{R}^{n+1}} X : \mathcal{U} \times I \to \mathbb{R}$$

Notice that we have taken care to introduce separate differentiation symbols for extrinsic functions. The D-symbols (D, D_i, D_t, etc.) will be referred to as the classical calculus differential symbols and are to be separated from the symbols used in the previous section on intrinsic functions such as ∇, $\frac{\partial}{\partial x^i}$, $\frac{\partial}{\partial t}$ which will be referred to as the manifold differential symbols. We will explain the connection shortly, in section A.3.

Summarizing, extrinsically defined functions $f : \mathcal{U} \times I \to \mathbb{R}^m$ (or vector fields) are equipped with the following differential symbols:

Table 2: Classical calculus differential symbols

D_i	partial (spatial) derivative in the i-th direction
D	spatial Jacobian matrix, or gradient vector if $m = 1$
D_t	partial time derivative
D^2	Hessian operator in ambient space
$\Delta_{\mathbb{R}^{n+1}}$	Laplace operator in the ambient space
$\mathrm{div}_{\mathbb{R}^{n+1}}$	divergence operator in the ambient space

As a final side note, it was mentioned earlier that extrinsic functions will be used primarily in integral estimates. Most of the time we will be integrating against the $\mathscr{H}^n$-measure, so we will abbreviate

$$\int_{M_t} f \quad \text{instead of the more cumbersome} \int_{M_t} f(x,t)\, d\mathscr{H}^n(x)$$

unless otherwise stated.

A.3. Differential Identities

Both intrinsic and extrinsic functions are in a sense flawed. Intrinsic functions depend on the background manifold *too much.* With the exception of well known geometric quantities (e.g. metric, second fundamental form), intrinsic functions are difficult to construct since it is hard to explicitly define a function on a background manifold M^n whose structure is oftentimes unknown. Extrinsic functions on the other hand do not depend on the background manifold M^n *at all.* Despite being easy to construct, they hold no regard whatsoever about our family of submanifolds (M_t) moving in $\mathbb{R}^{n+1}$, because their calculus differential symbols miss the (M_t) entirely.

What does strike a perfect balance, however, is defining an extrinsic function, restricting it to our submanifolds M_t, and finally pulling back to M^n to obtain an intrinsic function. This way we gain the option of having an explicit formula in $\mathbb{R}^{n+1}$ at hand, being able to work with the simple calculus differentiation symbols (D, D_i, D_t, etc.), but also being able study differential properties with respect to the manifolds through the pullback function. So given an extrinsic

$$f : \mathcal{U} \times I \to \mathbb{R}^m$$

for an open $\mathcal{U} \subseteq \mathbb{R}^{n+1}$, we can define the intrinsic

$$h : \mathcal{O} \to \mathbb{R}^m$$

via $h(p,t) := f(F(p,t),t)$, where $\mathcal{O} = \{(p,t) \in M^n \times I \ : \ F(p,t) \in \mathcal{U}\}$. We may now write things like

$$\frac{\partial}{\partial x^i} f, \ \frac{\partial}{\partial t} f, \ \nabla f, \ \mathrm{div}_{M_t} f, \ \Delta_{M_t} f, \ \ldots$$

on the subset $M_t \cap \mathcal{U}$ of the ambient space to refer exactly to the intrinsic quantities

$$\frac{\partial}{\partial x^i} h, \ \frac{\partial}{\partial t} h, \ \nabla h, \ \mathrm{div}_{M_t} h, \ \Delta_{M_t} h, \ \ldots$$

pushed forward to $M_t \subset \mathbb{R}^{n+1}$ after precomposing with the inverse map $F(\cdot,t)^{-1} : M_t \to M^n$. These intrinsic differentiation operators will of course differ from the classical D-differentiation operators (as they should). For one, $\nabla^{M_t} f$ is tangential to M_t while Df is not (and has no reason to be). Most importantly, $\frac{\partial}{\partial t} f$ takes into account the evolution of the M_t, through the precomposition with $F(\cdot,t)$, while $D_t f$ most certainly does not.

Fortunately, we can relate the intrinsic differentiation symbols to the extrinsic ones as the following results show.

PROPOSITION A.3.1. Let $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ be open, and $f : \mathcal{U} \times I \to \mathbb{R}$ be differentiable in the spatial variables. Then:

$$\nabla^{M_t} f(x,t) = \pi_{x,t}\left(Df(x,t)\right) \tag{A.3.1.a}$$

$$\mathrm{div}_{M_t} Df(x,t) = \Delta_{\mathbb{R}^{n+1}} f(x,t) - D^2 f\big|_{(x,t)}(\nu,\nu) \tag{A.3.1.b}$$

$$\Delta_{M_t} f(x,t) = \mathrm{div}_{M_t} Df(x,t) + H(x,t) \cdot Df(x,t) \tag{A.3.1.c}$$

for every $t \in I$, $x \in M_t \cap \mathcal{U}$. Here $\pi_{x,t}$ is the projection operator from $\mathbb{R}^{n+1}$ to the tangent space $T_x M_t$, ν is a unit vector normal to this space, and $H(x,t)$ is the mean curvature vector at $x \in M_t$.

PROOF. We will drop the (x,t) in our computations below for notational brevity. The first identity is very well known (see, e.g. **[Sim83]**). To prove the second identity, let $\{\tau_i\}_{i=1}^n$ be a (local) basis of TM_t that is orthonormal at x, and let ν be a unit normal to the tangent space. Then by the definition of $\Delta_{\mathbb{R}^{n+1}}$ and tangential divergence we have:

$$\begin{aligned}\Delta_{\mathbb{R}^{n+1}} f &= \sum_{i=1}^n \tau_i \cdot D_{\tau_i} Df + \nu \cdot D_\nu Df \\ &= \mathrm{div}_{M_t} Df + D^2 f(\nu,\nu)\end{aligned}$$

which is the required result, upon rearranging. Finally, to obtain the third identity we perform the following computation:

$$\begin{aligned}\Delta_{M_t} f &= \mathrm{div}_{M_t} \nabla f = \mathrm{div}_{M_t}\left(Df - (Df \cdot \nu)\,\nu\right) \\ &= \mathrm{div}_{M_t} Df - \mathrm{div}_{M_t}\left((Df \cdot \nu)\,\nu\right) \\ &= \mathrm{div}_{M_t} Df - (Df \cdot \nu)\mathrm{div}_{M_t} \nu \\ &= \mathrm{div}_{M_t} Df + H \cdot Df\end{aligned}$$

where we've used the well known fact that $H = -\left(\mathrm{div}_{M_t} \nu\right)\nu$ (see, e.g. **[Sim83]**), as well as the identity $\mathrm{div}_{M_t}(fX) = \nabla f \cdot X + f \mathrm{div}_{M_t} X$, where in our case the dot product term vanishes as $X = \nu$ is normal to the tangent space. $\square$

Notice that the previous proposition applies to time independent functions as much as it does to time dependent ones, since we are not interested in the evolution of our manifolds. The following proposition deals with the case in which we *are* interested in the evolution of our surfaces. It is not specific to mean curvature flow.

PROPOSITION A.3.2. Let $\mathcal{U} \subseteq \mathbb{R}^{n+1}$ be open, and $f : \mathcal{U} \times I \to \mathbb{R}$ be (jointly) differentiable in the product space. Assume, further, that our surfaces (M_t) evolve according to $F : M^n \times I \to \mathbb{R}^{n+1}$, i.e. $M_t = F(M^n, t)$. Then:

$$\frac{\partial}{\partial t} f(x,t) = D_t f(x,t) + \frac{\partial}{\partial t} F(p,t) \cdot Df(x,t)$$

for every $t \in I$, $x \in M_t \cap \mathcal{U}$. Here $p \in M^n$ is the unique point such that $F(p,t) = x$.

PROOF. Recall that to get an intrinsic derivative we define $h : \mathcal{O} \to \mathbb{R}$ according to $h(p,t) := f(F(p,t),t)$, where $\mathcal{O} := \{(p,t) \in M^n \times I \ : \ F(p,t) \in \mathcal{U}\}$. By the chain rule on the right hand side and for $(p,t) \in \mathcal{O}$ we have:

$$\frac{\partial}{\partial t} h(p,t) = D_t f(x,t) + \frac{\partial}{\partial t} F(p,t) \cdot Df(F(p,t),t)$$

Writing this back in terms of f and x, the result follows. □

The following result introduces versions of the product rule and the chain rule that we will come across several times in the course of this work. It works for both intrinsic and extrinsic functions.

PROPOSITION A.3.3. Let (M_t) be a family of surfaces evolving according to $F : M^n \times I \to \mathbb{R}^{n+1}$. Let $\mathcal{O} \subseteq M^n \times I$ be open. Let $f, g : \mathcal{O} \to \mathbb{R}$ be differentiable spatially and in time (not necessarily jointly). Also let η be a twice differentiable function on $\mathbb{R}$ (or an open superset of the range of f). Then:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)(fg) = \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \cdot g + f \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) g - 2\nabla f \cdot \nabla g \tag{A.3.3.a}$$

and

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \eta(f) = \eta'(f) \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f - \eta''(f) |\nabla f|^2 \tag{A.3.3.b}$$

In particular, this result also holds if $f, g : \mathcal{U} \times I \to \mathbb{R}$ are extrinsic functions which are jointly differentiable and twice spatially differentiable.

PROOF. The proof is trivial, simply the result of straightforward differentiations. □

APPENDIX B

Integration Tools

We have occasionally had the need to use various theorems on exchanging limits with integrals throughout the course of the text. The purpose of this appendix chapter is to justify all the theorems that have been used. In particular, we prove the first variation formula on which all of Chapter 3 built, and finally go over variants of dominated convergence theorems that have been used.

B.1. First Variation Formula

Our first variation formula depends strongly on our assumption that our manifolds are compact. We do not actually know whether or not it breaks for non-compact manifolds–refer to Chapter D of the appendix for a discussion of this issue.

THEOREM B.1.1 (First Variation Formula). Let (M_t) move by mean curvature in $\mathcal{U} \times I$, and let $\varphi \in \mathcal{C}_c^1(\mathcal{U})$. For any $t \in I$ we have the first variation formula:

$$\frac{d}{dt}\int_{M_t} \varphi = \int_{M_t} H \cdot D\varphi - |H|^2\varphi$$

Part of the conclusion here is that the derivative above exists.

PROOF. Fix a time instant $t_0 \in I$ at which we want to show the identity above. Since this is a local result, we may restrict our attention to compact subset of I containing t_0, which we may assume to be I itself.

For $(p, t) \in M^n \times I$:

$$\begin{aligned}\frac{\partial}{\partial t}\left(\varphi(F(p,t))\sqrt{g}(p,t)\right) &= \left(D\varphi(F(p,t)) \cdot \frac{\partial}{\partial t}F(p,t)\right)\sqrt{g}(p,t) + \varphi(F(p,t))\frac{\partial}{\partial t}\sqrt{g}(p,t)\\ &= \left(H(p,t)\cdot\varphi(F(p,t))\right)\sqrt{g}(p,t) - |H(p,t)|^2\sqrt{g}(p,t)\end{aligned}$$

This is obviously true when $F(p,t) \in \mathcal{U}$ in view of mean curvature motion and the area element evolution function (2.3.2.c). For all other (p,t) both φ and $D\varphi$ happen to vanish, which means the formula is still true.

The set of points $(p,t) \in M^n \times I$ mapping to support $\varphi \subset \mathcal{U}$ is compact, as a closed subset of the compact space $M^n \times I$. Since all our functions are continuous, there exists a $C_0 < \infty$ bounding the right hand side at all points (p,t) in this compact subset. Of course C_0 is also an effective upper bound at all other $(p,t) \in M^n \times I$, since both φ and $D\varphi$ vanish at those points. Consequently, for every $(p,t) \in M^n \times I$ and a positive lower bound of $\sqrt{g}(\cdot, t_0)$ on M^n we have:

$$\left| \frac{\partial}{\partial t} \left(\varphi(F(p,t)) \sqrt{g}(p,t) \right) \right| \leq C_0 C_1^{-1} \sqrt{g}(p,t_0)$$

This is a time independent majorizing function of our time derivative. Note that it is $\mathscr{L}^1$:

$$\int_{M^n} C_0 C_1^{-1} \sqrt{g}(p,t_0) = C_0 C_1^{-1} \mathscr{H}^n \left(M_{t_0} \right) < \infty$$

The standard differentiation under the integral sign theorem tells us that:

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=t_0} \int_{M_t} \varphi &= \left.\frac{d}{dt}\right|_{t=t_0} \int_{M^n} \varphi(F(\cdot,t)) \sqrt{g}(\cdot,t) \\
&= \int_{M^n} \left.\frac{\partial}{\partial t}\right|_{t=t_0} \left(\varphi(F(\cdot,t)) \sqrt{g}(\cdot,t) \right) \\
&= \int_{M^n} \left(H(p,t_0) \cdot \varphi(F(p,t_0)) \right) \sqrt{g}(p,t_0) - |H(p,t_0)|^2 \sqrt{g}(p,t_0) \\
&= \int_{M_{t_0}} H \cdot D\varphi - |H|^2 \varphi
\end{aligned}$$

as claimed. ◻

B.2. Dominated Convergence Theorems

We have used various versions of the dominated convergence theorem in our text, which we set to prove in this section. We do not need compactness from our manifolds for our results here, so we will not assume it. The reader is assumed to be familiar with the standard dominated convergence theorem (see, e.g. **[Coh94]**).

LEMMA B.2.1. Let (X, μ) be a measure space, and let $f_n : X \to \mathbb{R}$ be a sequence of measurable functions converging pointwise to a function $f : X \to \mathbb{R}$. If $|f_n| \leq g$ for all n and a fixed $g \in \mathscr{L}^1$, then:

$$\lim_n \int_X f_n \, d\mu = \int_X f \, d\mu$$

In particular, the conclusion also holds on locally finite spaces (X, μ), when our measurable f_n converge uniformly to some f, and all have supports within a fixed compact set.

PROOF. This is verbatim the standard version of dominated convergence, and can be found in any textbook on Lebesgue Integration (see, e.g. **[Coh94]**). $\square$

The more interesting variant of dominated convergence is when we let our domains vary in time. We do not require that they vary by mean curvature per se, but we do require that they vary in such a way that the first variation formula (3.1.3) holds. As a consequence, it does apply for compact manifolds moving by mean curvature since we have established the first variation formula in this case.

LEMMA B.2.2. Let (M_t) be possibly non-compact, but locally $\mathscr{H}^n$-finite, manifolds moving in $\mathcal{U} \times (t_1, t_0)$ in such a way that the first variation formula (3.1.3) holds. Let $f : \mathcal{U} \times (t_1, t_0) \to \mathbb{R}$ be continuous, with support $f(\cdot, t) \subseteq K$ for all $t \in (t_1, t_0)$, where $K \subset \mathcal{U}$ is a fixed compact set. If $f(\cdot, t) \to 0$ uniformly in $\mathcal{U}$ as $t \nearrow t_0$, then:

$$\lim_{t \nearrow t_0} \int_{M_t} f = 0$$

PROOF. If $\delta > 0$ denotes the smaller between $\sqrt{t_0 - t_1}$ and the distance between K and the boundary of $\mathcal{U}$ and $\theta_0 := \frac{1}{2\sqrt{1+2n}}$, let us cover the compact set K with finitely many balls $B_{\theta_0 \delta/2}(x_j)$, $1 \leq j \leq N$. Then for all $t \in (t_1, t_0)$:

$$\begin{aligned}\int_{M_t} |f| &= \int_{M_t \cap \{f(\cdot,t)\neq 0\}} |f| \le \int_{M_t \cap K} |f| \\ &\le \sum_{j=1}^{N} \int_{M_t \cap B_{\theta_0 \delta/2}(x_j)} |f| \\ &\le \sup_K |f(\cdot,t)| \quad \sum_{j=1}^{N} \mathscr{H}^n \left(M_t \cap B_{\theta_0\delta/2}(x_j)\right) \end{aligned} \tag{†}$$

We now employ the forward area estimate (3.3.1) which holds even in the noncompact case provided we have the first variation formula (3.1.3), as with every other result in Chapter 3. Since $(1+2n)\theta_0^2 < 1$ and $B_{\delta/2}(x_j) \subseteq \mathcal{U}$, the forward area estimate guarantees that:

$$\mathscr{H}^n \left(M_t \cap B_{\theta_0\delta/2}(x_j)\right) \le C(n)\mathscr{H}^n \left(M_{t_0-\theta_0^2\delta^2/4} \cap B_{\delta/2}(x_j)\right)$$

for all $t \in [t_0 - \theta_0^2\delta^2/4, t_0)$ and all $j = 1, \dots, N$. The constant above should have been $C(n, \theta_0)$, but θ_0 depends on n so we have absorbed this into a new constant $C(n)$. In any case, restricting ourselves to $t \in [t_0 - \theta_0^2\delta^2/4, t_0)$ in (†), we have:

$$\int_{M_t} |f| \le \sup_K |f(\cdot,t)| C(n) \sum_{j=1}^{N} \mathscr{H}^n \left(M_{t_0-\theta_0^2\delta^2/4} \cap B_{\delta/2}(x_j)\right)$$

The last sum is finite and independent of t, so we the result follows by letting $t \nearrow t_0$ in view of which the supremum vanishes, and the result follows. $\square$

APPENDIX C

Computing Gauss Density

Although we have defined the notion of Gauss density and we have proved a number of results that are directly linked to it, we don't have a concrete understanding of it yet. We have seen how to bound $\Theta(M_t, x_0, t_0)$ from below for points reached by (M_t) in (3.7.8), but we don't have a closed form expression for the lower bound. Our goal in this chapter is to show that for smooth solutions of mean curvature, Gauss density equals 1 at all points prior to the first singular time. In view of the density function's established upper semicontinuity, it is also at least one at the first singular time. As a direct corollary, we prove another variant of a mean value inequality satisfied by mean curvature flows.

Unlike in the main text, this bound is now established independently of Brakke's clearing out lemma, primarily because we wish to have a dimension-independent bound. With this bound at hand, we can actually get a different proof of Brakke's clearing out lemma and obtain a closed form expression for the density lower bound postulated by it. In other words, if we know that our solutions are smooth then we can first compute Gauss densities and bound them explicitly from below, and then come up with a cleaner version of Brakke's clearing out lemma.

C.1. Gauss density equals one

We will need the two following auxiliary lemmas from measure theory:

LEMMA C.1.1. Let μ be a Radon measure on a metric space X such that:

$$\mu(B_R(x_0)) \leq A_0 R^p$$

for every $R > R_0$, where $x_0 \in X$, $R_0, p > 0$ are fixed. Then for every $\gamma > 0$ we have:

$$\int_{X \setminus B_R(x_0)} e^{-\gamma |x-x_0|^2} \, d\mu(x) \leq C(A_0, p, \gamma, R) \, (< \infty)$$

where $C(A_0, p, \gamma, R) \to 0$ as $R \to \infty$.

PROOF. Without loss of generality we may take $r > \max\{1, R_0\}$. Then:

$$\int_{\mathbb{R}^N} e^{-\gamma|x-x_0|^2}\, d\mu(x) = \int_{B_r(x_0)} e^{-\gamma|x-x_0|^2}\, d\mu(x) + \sum_{j=1}^{\infty} \int_{B_{r^{j+1}}(x_0)\setminus B_{r^j}(x_0)} e^{-\gamma|x-x_0|^2}\, d\mu(x)$$

Now $t \mapsto e^{-\gamma t^2}$ is a decreasing function of $t > 0$, and hence:

$$\leq \mu(B_r(x_0)) + \sum_{j=1}^{\infty} e^{-\gamma r^{2j}} \mu(B_{r^{j+1}}(x_0) \setminus B_{r^j}(x_0)) \leq \mu(B_r(x_0)) + \sum_{j=1}^{\infty} e^{-\gamma r^{2j}} \mu(B_{r^{j+1}}(x_0))$$

Employing the polynomial measure growth assumption, we conclude:

$$\leq A_0 \left(r^p + \sum_{j=1}^{\infty} e^{-\gamma r^{2j}} r^{p(j+1)} \right) < \infty$$

and hence $x \mapsto e^{-\gamma|x-x_0|^2}$ is an $\mathscr{L}^1(\mu)$ function. Therefore, its integral over $X \setminus B_R(x_0)$ decays to 0 as $R \to \infty$, i.e.:

$$\int_{X\setminus B_R(x_0)} e^{-\gamma|x-x_0|^2}\, d\mu(x) \leq C(A_0, p, \gamma, R) \to 0 \quad \text{as } R \to \infty$$

□

LEMMA C.1.2. If P is any n-dimensional subspace of $\mathbb{R}^{n+1}$ then:

$$\int_P \Phi(\cdot, s)\, d\mathscr{H}^n = 1$$

for any $s < 0$, where Φ is the backwards heat kernel.

PROOF. This is a straightforward integration result. □

At this point we can proceed to compute the Gauss density at points where our flow is still smooth.

PROPOSITION C.1.3. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0 + \delta)$. Then for any $x_0 \in M_{t_0} \cap \mathcal{U}$ we have:

$$\Theta(M_t, x_0, t_0) = 1$$

PROOF. The computation is remarkably tricky and employs the density upper bound estimate (3.6.3), so we will split the proof into a sequence of steps.

Step 0: Implications of Smoothness.

A significant part of this proof actually does not need smoothness at time t_0. In fact, only one step in the proof (step 3.v) requires that we know something structural about time t_0 and in fact we don't actually even need smoothness per se but rather convergence, in some sense, of M_s^λ to $T_{x_0}M_{t_0}$ as $\lambda \searrow 0$. But smoothness will do, since the purpose of this result is to compute densities prior to the first singular time anyway. Since our result is rotation invariant, we may suppose without loss of generality that $T_{x_0}M_{t_0} = \mathbb{R}^n \times \{0\}$.

Step 1: Shrinking the Domain.

Let $\varrho_0 > 0$ be such that $\overline{B_{\varrho_0}(x_0)} \subseteq \mathcal{U}$ and $(t_0 - \varrho_0^2, t_0) \subseteq (t_1, t_0)$ and (M_t) can be written as a graph of a function $u : B_{\varrho_0}(\pi(x_0)) \times (t_0 - \varrho_0^2, t_0) \to \mathbb{R}$ within a cylinder $C_{\varrho_0,h}(x_0) \subset \mathcal{U}$ centered at x_0 with radius ϱ_0 and height $2h > 2\varrho_0$. The latter can be arranged by (E.1.3). Here π denotes projection onto $\mathbb{R}^n$.

Now choose $\varrho > 0$ such that:

$$\varrho \leq \frac{1}{2(1+2n)}\varrho_0 \tag{$\dagger_1$}$$

In particular, $\varrho \leq \frac{1}{\sqrt{1+2n}}\varrho_0$ so (3.7.6) is applicable with the test function $\varphi_{(x_0,t_0),\varrho}$ on the time interval $(t_0 - \varrho^2, t_0)$. In other words:

$$\begin{aligned}\Theta(M_t, x_0, t_0) &= \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)}(x,t)\varphi_{(x_0,t_0),\varrho}(x,t)\, d\mathscr{H}^n(x) \\ &= \lim_{t \nearrow t_0} \int_{M_t} \frac{1}{(4\pi(t_0-t))^{n/2}} \exp\left(-\frac{|x-x_0|^2}{4(t_0-t)}\right) \left(1 - \frac{|x-x_0|^2 + 2n(t-t_0)}{\varrho^2}\right)_+^3 d\mathscr{H}^n(x)\end{aligned}$$

Step 2: Rewriting in Parabolic Coordinates.

We now aim to make a parabolic change of coordinates, only we wish to let the scaling λ vary rather than the parabolic time s. Fix $s = -1/2$ and write $t = \lambda(t)^2 s + t_0$. Change our integration variable to y given by $x = \lambda y + x_0$, so that:

$$\begin{aligned}
\Theta(M_t, x_0, t_0) \quad &= \lim_{t \nearrow t_0} \int_{M^{\lambda(t)}_{-1/2}} \frac{1}{\lambda(t)^n (2\pi)^{n/2}} \exp\left(-\frac{\lambda(t)^2|y|^2}{2\lambda(t)^2}\right) \quad \left(1 - \lambda(t)^2 \frac{|y|^2 - n}{\varrho^2}\right)^3_+ \quad \lambda(t)^n \, d\mathscr{H}^n(y) \\
&= \lim_{\lambda \searrow 0} \int_{M^{\lambda}_{-1/2}} \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{|y|^2}{2}\right) \quad \left(1 - \lambda^2 \frac{|y|^2 - n}{\varrho^2}\right)^3_+ \quad d\mathscr{H}^n(y) \\
&= \lim_{\lambda \searrow 0} \int_{M^{\lambda}_{-1/2}} \Phi(y, -1/2) \quad \varphi_{\lambda^{-1}\varrho}(y, -1/2) \quad d\mathscr{H}^n(y)
\end{aligned} \tag{$\ddagger_0$}$$

where the second identity follows from the first by a change of variables in the limit, and the third identity follows by observing that the expressions we got are the standard heat kernel and the standard parabolic test function, respectively. Notice that in studying the limit we have to restrict our attention to

$$\lambda < \varrho\sqrt{2} \tag{$\dagger_2$}$$

so that $t = \lambda^2 s + t_0 \in (t_0 - \varrho^2, t_0)$.

Step 3: Approximation Argument Setup.

Let $\varepsilon > 0$ be given. We will employ an $\varepsilon - \delta$ approximation argument to show that the quantity below is arbitrarily close to zero. We will explain everything, including our choice of λ, R, χ_R, in the sequel. For now we just write this down as a guideline.

$$\begin{aligned}
\Theta(M_t, x_0, t_0) - 1 \quad &= \left(\Theta(M_t, x_0, t_0) \quad - \quad \int_{M^{\lambda}_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho}\right) \\
&+ \left(\int_{M^{\lambda}_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho} \quad - \quad \int_{M^{\lambda}_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho}\chi_R\right) \\
&+ \left(\int_{M^{\lambda}_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho}\chi_R \quad - \quad \int_{T_{x_0}M_{t_0}} \Phi\chi_R\right) \\
&+ \left(\int_{T_{x_0}M_{t_0}} \Phi\chi_R \quad - \quad 1\right)
\end{aligned}$$

We wish to show that for a choice of R, χ_R, all the four expressions in the parentheses above (say $E_1, \ldots, E_4$) can be made arbitrarily small in terms of just λ.

Step 3.i: Estimating E_4

We know that $y \mapsto \Phi(y, -1/2)$ is an $\mathscr{L}^1$ function with respect to the measure $\mathscr{H}^n|_{T_{x_0} M_{t_0}}$, so there exists an $R_1 > 0$ such that:

$$\left| \int_{T_{x_0} M_{t_0} \setminus B_R(0)} \Phi(y, -1/2) \, d\mathscr{H}^n(y) \right| < \frac{\varepsilon}{4} \tag{$\ddagger_1$}$$

for every $R \geq R_1$.

Step 3.ii: Estimating E_2

Let us apply the density upper bound result (3.6.3) to get:

$$\sup_{0<\sigma\leq \frac{1}{2\sqrt{1+2n}}\varrho_0} \quad \sup_{t \in (t_0-\sigma^2, t_0)} \quad \frac{\mathscr{H}^n(M_t \cap B_\sigma(x_0))}{\sigma^n} \leq C(n) \quad \frac{\mathscr{H}^n\left(M_{t_0 - \frac{1}{4(1+2n)}\varrho_0^2} \cap B_{\varrho_0}(x_0)\right)}{\varrho_0^n} =: C \tag{$\dagger_3$}$$

Consider the Radon measure μ on $\mathbb{R}^{n+1}$ given by:

$$d\mu := \varphi_{\lambda^{-1}\varrho}(\cdot, -1/2) \quad \mathbf{1}_{M^\lambda_{-1/2}} \quad d\mathscr{H}^n$$

We wish to establish a polynomial growth bound on μ that is independent of λ. Without loss of generality we need only show this for $R > \frac{1}{\sqrt{2}}$, because small values of R are irrelevant. In the case of $R > \frac{1}{\sqrt{2}}$ we have $s = -1/2$ belonging to $(-R^2, 0)$, in which case the bound $\lambda^2 s > -\lambda^2 R^2$ comes in handy. We will consider two different cases for $R > \frac{1}{\sqrt{2}}$, that of $\lambda R \leq \frac{1}{2\sqrt{1+2n}}\varrho_0$ and that of $\lambda R > \frac{1}{2\sqrt{1+2n}}\varrho_0$.

If $\lambda R \leq \frac{1}{2\sqrt{1+2n}}\varrho_0$, then we want to employ ($\dagger_3$) with $\sigma := \lambda R$ and $t = \lambda^2 s + t_0 \in (t_0 - \sigma^2, t_0)$. In fact:

$$\begin{aligned}
\mu(B_R(0)) &= \int_{M^\lambda_{-1/2}\cap B_R(0)} \varphi_{\lambda^{-1}\varrho}(y,-1/2) \quad d\mathscr{H}^n(y) \le \sup \varphi_{\lambda^{-1}\varrho}(\cdot,-1/2) \quad \mathscr{H}^n\left(M^\lambda_{-1/2}\cap B_R(0)\right) \\
&\overset{(3.2.1.b)}{\le} \left(1+\frac{n\lambda^2}{\varrho^2}\right)^3 \mathscr{H}^n\left(M^\lambda_{-1/2}\cap B_R(0)\right) \overset{(\dagger_2)}{\le} (1+2n)^3\,\mathscr{H}^n\left(M^\lambda_{-1/2}\cap B_R(0)\right) \\
&= (1+2n)^3\lambda^{-n}\mathscr{H}^n\left(M_{-\lambda^2/2+t_0}\cap B_{\lambda R}(x_0)\right) \overset{(\dagger_3)}{\le} (1+2n)^3\lambda^{-n}C\lambda^n R^n = (1+2n)^3 C R^n
\end{aligned}$$

If $\lambda R > \frac{1}{2\sqrt{1+2n}}\varrho_0$ then the radius is too big for the previous argument to work verbatim. In this case we make use of the fact that the support of $\varphi_{\lambda^{-1}\varrho}$ is small enough to compensate for λR being large. Indeed from (3.2.1.c) we know that $\varphi_{\lambda^{-1}\varrho}(\cdot,-1/2)$ is identically zero outside:

$$\begin{aligned}
B_{\sqrt{\lambda^{-2}\varrho^2+n}}(0) &= \lambda^{-1}\left(B_{\sqrt{\varrho^2+n\lambda^2}}(x_0) \quad - x_0\right) \\
&\overset{(\dagger_2)}{\subseteq} \lambda^{-1}\left(B_{\sqrt{1+2n}\varrho}(x_0) \quad - x_0\right) \\
&\overset{(\dagger_1)}{\subseteq} \lambda^{-1}\left(B_{\frac{1}{2\sqrt{1+2n}}\varrho_0}(x_0) \quad - x_0\right)
\end{aligned}$$

Consequently for $\lambda R > \frac{1}{2\sqrt{1+2n}}\varrho_0$ we estimate:

$$\begin{aligned}
\mu(B_R(0)) &= \int_{M^\lambda_{-1/2}\cap B_R(0)} \varphi_{\lambda^{-1}\varrho}(y,-1/2) \quad d\mathscr{H}^n(y) \\
&= \int_{M^\lambda_{-1/2}\cap B_R(0)\cap B_{\sqrt{\lambda^{-2}\varrho^2+n}}(0)} \varphi_{\lambda^{-1}\varrho}(y,-1/2) \quad d\mathscr{H}^n(y) \\
&\overset{(3.2.1.b)}{\le} (1+2n)^3\mathscr{H}^n\left(M^\lambda_{-1/2}\cap B_R(0)\cap B_{\sqrt{\lambda^{-2}\varrho^2+n}}(0)\right) \\
&= (1+2n)^3\lambda^{-n}\mathscr{H}^n\left(M_{-\lambda^2/2+t_0}\cap B_{\lambda R}(x_0)\cap B_{\frac{1}{2\sqrt{1+2n}}\varrho_0}(x_0)\right) \\
&= (1+2n)^3\lambda^{-n}\mathscr{H}^n\left(M_{-\lambda^2/2+t_0}\cap B_{\frac{1}{2\sqrt{1+2n}}\varrho_0}(x_0)\right)
\end{aligned}$$

where the last equality follows simply because $\lambda R > \frac{1}{2\sqrt{1+2n}}\varrho_0$. We can now employ ($\dagger_3$) as before with $\sigma := \frac{1}{2\sqrt{1+2n}}\varrho_0$ and $t = \lambda^2 s + t_0$. Notice that $t = -\lambda^2/2 + t_0 > -\varrho^2 + t_0 = t_0 - \sigma^2$, so we can indeed employ this step to obtain:

$$\mu(B_R(0)) \quad \leq (1+2n)^3\lambda^{-n}C\left(\frac{1}{2\sqrt{1+2n}}\varrho_0\right)^n \leq (1+2n)^3\lambda^{-n}C\lambda^n R^n = (1+2n)^3 C R^n$$

In any case, then, for all $R > \frac{1}{\sqrt{2}}$ we have the polynomial growth bound:

$$\mu(B_R(0)) \leq (1+2n)^3 C R^n$$

which is independent of λ. By employing the measure growth lemma (C.1.1) for $\gamma = 1/2$ we conclude that there exists a $R_2 > 0$ such that for every $R \geq R_2$, and all λ:

$$\left| \int_{M^{\lambda}_{-1/2} \setminus B_R(0)} \Phi(y, -1/2) \quad \varphi_{\lambda^{-1}\varrho}(y, -1/2) \quad d\mathscr{H}^n(y) \right| < \frac{\varepsilon}{4} \tag{$\ddagger_2$}$$

Step 3.iii: Choosing R, χ_R.

Define $R := \max\{R_1, R_2\}$ from steps ($\ddagger_1$) and ($\ddagger_2$), and let χ_R be a continuous function such that $\mathbf{1}_{B_R} \leq \chi_R \leq \mathbf{1}_{B_{2R}}$.

Step 3.iv: Estimating E_1

Of course we know that there exists a $\lambda_1 > 0$ such that:

$$\left| \Theta(M_t, x_0, t_0) - \int_{M^{\lambda}_{-1/2}} \Phi(y, -1/2) \quad \varphi_{\lambda^{-1}\varrho}(y, -1/2) \quad d\mathscr{H}^n(y) \right| < \frac{\varepsilon}{4} \tag{$\ddagger_3$}$$

for every $\lambda \leq \lambda_1$, directly by the definition of limits and ($\ddagger_0$).

Step 3.v: Estimating E_3

We finally make use of our smoothness assumption by writing the (M_t) as graphs locally near (x_0, t_0). Recall by step 1 that within a cylinder $C_{\varrho_0,h}(x_0)$, $h > \varrho_0$, we can write each M_t, $t \in (t_0 - \varrho_0^2, t_0)$, as the graph of a function $u(\cdot, t) : B_{\varrho_0}(\pi(x_0)) \to \mathbb{R}$.

We will be working with the integral:

$$\int_{M^\lambda_{-1/2}} \Phi(y,-1/2) \quad \varphi_{\lambda^{-1}\varrho}(y,-1/2) \quad \chi_R(y) \quad d\mathscr{H}^n(y) \tag{$\dagger_4$}$$

The support of $\varphi_{\lambda^{-1}\varrho}(\cdot,-1/2)$ we recall from step 3.ii falls within $\lambda^{-1}\left(B_{\frac{1}{2\sqrt{1+2n}}\varrho_0}(x_0) \quad - x_0\right)$ because of our assumptions on λ, ϱ at ($\dagger_1$), ($\dagger_2$). Recall that

$$M_{-\lambda^2/2+t_0} \cap C_{\varrho_0,h}(x_0)$$

can be viewed as the graph of $u(\cdot,t)$ at $t=-\lambda^2/2+t_0$, and hence

$$M^\lambda_{-1/2} \cap \lambda^{-1}\left(C_{\frac{1}{2\sqrt{1+2n}}\varrho_0,h}(x_0) \quad - x_0\right)$$

can be viewed as the graph of $\hat{u}^\lambda(\cdot,s)$ at $s=-1/2$, where:

$$\hat{u}^\lambda(y,s) := \frac{u(\lambda y+\pi(x_0),\lambda^2 s+t_0)-u(\pi(x_0),t_0)}{\lambda}$$

and π is the projection onto $\mathbb{R}^n$. Notice $|D\hat{u}^\lambda| = |Du|$, and that by Taylor's theorem $\hat{u}^\lambda(y,s) \to Du(\pi(x_0),t_0)$ uniformly on compact sets as $\lambda \searrow 0$.

Unfortunately the domain of $\hat{u}^\lambda(\cdot,s)$ is expanding as $\lambda \searrow 0$. But the presence of χ_R in the integrand dictates that we need only pay attention to the pre-compact set $B_{2R}(0)$ in the domain, since $\chi_R = 0$ identically outside of $B_{2R}(0)$. If $B^n_\varrho(\pi(x_0)) \subseteq \mathbb{R}^n$ is the open set on which the graph function $u(\cdot,t)$ is defined, then we proceed as follows:

$$\begin{aligned}
&\int_{M^\lambda_{-1/2}} \Phi(y,-1/2) \quad \varphi_{\lambda^{-1}\varrho}(y,-1/2) \quad \chi_R(y) \quad d\mathscr{H}^n(y)\\
= &\int_{\lambda^{-1}(B^n_\varrho(\pi(x_0))-\pi(x_0))\cap B_{2R}(0)} \Phi \quad \varphi_{\lambda^{-1}\varrho} \quad \chi_R(y,\hat{u}^\lambda(y,-1/2)) \quad \sqrt{1+|D\hat{u}^\lambda(y,-1/2)|^2} \quad d\mathscr{H}^n(y)
\end{aligned}$$

with $\Phi, \varphi_{\lambda^{-1}\varrho}$ both evaluated at $((y,\hat{u}^\lambda(y,-1/2)),-1/2)$. Taking $\lambda \searrow 0$ we know that $\hat{u}^\lambda(y,-1/2)$ converges uniformly on compact subsets to $Du(\pi(x_0),t_0)\cdot y$. In other words, the coordinates $(y,\hat{u}^\lambda(y,-1/2))$ converge uniformly on compact subsets to $(y, Du(\pi(x_0),t_0)\cdot y)$, which represent $T_{x_0}M_{t_0} \cap B_{2R}(0)$. Since $B_{2R}(0)$ is precompact, convergence is uniform. Also:

$$\varphi_{\lambda^{-1}\varrho}(y,-1/2) = \left(1 - \lambda^2 \frac{|y|^2 - n}{\varrho^2}\right)_+^3$$

converges uniformly to 1 for $y \in B_{2R}(0)$, as $\lambda \searrow 0$. Therefore, since $B_{2R}(0)$ is pre-compact we may employ classical dominated convergence–see (B.2.1)–and conclude that the integral in ($\dagger_4$) converges to:

$$\int_{T_{x_0}M_{t_0} \cap B_{2R}} \Phi\chi_R = \int_{T_{x_0}M_{t_0}} \Phi\chi_R$$

as $\lambda \searrow 0$. Therefore, there exists a $\lambda_2 > 0$ such that:

$$\left| \int_{M^\lambda_{-1/2}} \Phi(y,-1/2) \quad \varphi_{\lambda^{-1}\varrho}(y,-1/2) \quad \chi_R(y)\, d\mathscr{H}^n(y) - \int_{T_{x_0}M_{t_0}} \Phi(y,-1/2) \quad \chi_R(y) \quad d\mathscr{H}^n(y) \right| < \frac{\varepsilon}{4} \qquad (\ddagger_4)$$

for every $\lambda \leq \lambda_2$.

Step 4: Conclusion.

Choose $\lambda := \min\{\lambda_1, \lambda_2\}$, where λ_1, λ_2 are as in ($\ddagger_3$) and ($\ddagger_4$). Then, from ($\ddagger_1$), ($\ddagger_2$), ($\ddagger_3$), and ($\ddagger_4$) we have:

$$\begin{aligned}
|\Theta(M_t, x_0, t_0) - 1| = & \left| \Theta(M_t, x_0, t_0) \quad - \quad \int_{M^\lambda_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho} \right| \\
& + \left| \int_{M^\lambda_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho} \quad - \quad \int_{M^\lambda_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho}\chi_R \right| \\
& + \left| \int_{M^\lambda_{-1/2}} \Phi\varphi_{\lambda^{-1}\varrho}\chi_R \quad - \quad \int_{T_{x_0}M_{t_0}} \Phi\chi_R \right| \\
& + \left| \int_{T_{x_0}M_{t_0}} \Phi\chi_R \quad - \quad 1 \right| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Notice that it is required that the solution be smooth up to the time t_0 at which we are evaluating. Can we extend our result to points on the limiting surface without imposing any kind of regularity?

C.2. Another Mean Value Inequality

Thankfully, it turns out that the upper semicontinuity we had previously established is just what we need. It requires us to assume no extra regularity on our manifolds, and lets us conclude that even in the limiting case Gauss density is at least 1, and prove an interesting mean value inequality for mean curvature flow.

COROLLARY C.2.1. Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$. If x_0 is reached by (M_t) at time t_0, then $\Theta(M_t, x_0, t_0) \geq 1$.

LEMMA C.2.2 (Mean Value Inequality). Let (M_t) move by mean curvature in $\mathcal{U} \times (t_1, t_0)$, and let f be a test function on $\mathcal{U} \times (t_1, t_0]$ such that:

$$f \geq 0 \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq 0$$

If M_t reaches x_0 at time t_0, then:

$$f(x_0, t_0) \leq \int_{M_t} \Phi_{(x_0,t_0)} f$$

for every $t \in (t_1, t_0)$.

PROOF. Since M_t reaches x_0 at time t_0, from (C.2.1) we know that $\Theta(M_t, x_0, t_0) \geq 1$. Combining this with the extraction theorem (3.7.5) we obtain that:

$$f(x_0, t_0) \leq f(x_0, t_0)\Theta(M_t, x_0, t_0) = \lim_{t \nearrow t_0} \int_{M_t} \Phi_{(x_0,t_0)} f$$

and by our assumption on f and Huisken's monotonicity (3.5.2) we know:

$$t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} f$$

that f is a decreasing function on f, and the result follows. $\square$

C.3. Alternative proof of Brakke Clearing Out

We can actually use the fact that Gauss density is at least one up to the first singular time to establish a sharper version of Brakke's clearing out lemma, where all constants used are explicit and our time domain stretches up to a factor of $\frac{1}{2n}$ rather than the smaller $\frac{1}{2n+1}$ postulated in (3.6.2). The second version of Brakke's clearing out lemma is as follows.

THEOREM C.3.1. Let (M_t) move by mean curvature in $B_{\varrho_0}(x_0) \times \left(t_0 - \frac{1}{2n}\varrho_0^2, t_0\right)$. Let $\beta \in \left(0, \frac{1}{2n}\right)$, and suppose that x_0 is reached by (M_t) at time t_0. Then there exists a positive constant $\eta(n,\beta)$ such that for all $0 < \varrho < \varrho_0$:

$$\frac{\mathscr{H}^n(M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0))}{\varrho^n} \geq \eta(n,\beta)$$

The constant is, in fact, $\eta(n,\beta) := (1-2n\beta)^3(4\pi\beta)^{n/2}$.

PROOF. Let $0 < \varrho < \varrho_0$, as above. Observe that for any $0 < \sigma \leq \varrho$ we are guaranteed by (3.2.2) that $\varphi_{(x_0,t_0),\sigma}$ is a test function on $B_{\varrho_0}(x_0) \times \left(t_0 - \frac{1}{2n}\varrho_0^2 + \frac{1}{2n}\sigma^2, t_0\right)$, and so by (3.7.6) we know that for every $t \in \left(t_0 - \frac{1}{2n}\varrho_0^2 + \frac{1}{2n}\sigma^2, t_0\right)$ we have:

$$\Theta(M_t, x_0, t_0) \leq \int_{M_t} \Phi_{(x_0,t_0)}\varphi_{(x_0,t_0),\sigma}$$

In particular, we may specialize to $t = t_0 - \alpha\sigma^2$ for some choice of $\alpha > 0$ such that $\alpha\sigma^2 < \frac{1}{2n}\varrho_0^2 - \frac{1}{2n}\sigma^2$. Then:

$$\begin{aligned}
\Theta(M_t, x_0, t_0) &\leq \int_{M_{t_0-\alpha\sigma^2}} \Phi_{(x_0,t_0)}\varphi_{(x_0,t_0),\sigma} \\
&= \int_{M_{t_0-\alpha\sigma^2}\cap\{\varphi_{(x_0,t_0),\sigma}(\cdot,t_0-\alpha\sigma^2)>0\}} \Phi_{(x_0,t_0)}\varphi_{(x_0,t_0),\sigma} \\
&= \int_{M_{t_0-\alpha\sigma^2}\cap B_{\sqrt{1+2n\alpha}\sigma}(x_0)} \Phi_{(x_0,t_0)}\varphi_{(x_0,t_0),\sigma}
\end{aligned}$$

We can estimate the integral from above by (3.4.2.a) and (3.2.1.b) and obtain:

$$\begin{aligned}\Theta(M_t, x_0, t_0) &\leq \sup_{\mathbb{R}^{n+1}} \Phi_{(x_0,t_0)}(\cdot, t_0 - \alpha\sigma^2) \quad \sup_{\mathbb{R}^{n+1}} \varphi_{(x_0,t_0),\sigma}(\cdot, t_0 - \alpha\sigma^2) \quad \mathscr{H}^n(M_{t_0-\alpha\sigma^2} \cap B_{\sqrt{1+2n\alpha}\sigma}(x_0)) \\ &= \frac{1}{(4\pi\alpha\sigma^2)^{n/2}}(1+2n\alpha)^3 \mathscr{H}^n\left(M_{t_0-\alpha\sigma^2} \cap B_{\sqrt{1+2n\alpha}\sigma}(x_0)\right)\end{aligned}$$

Rearranging:

$$\frac{\mathscr{H}^n\left(M_{t_0-\alpha\sigma^2} \cap B_{\sqrt{1+2n\alpha}\sigma}(x_0)\right)}{\sigma^n} \geq \frac{(4\pi\alpha)^{n/2}}{(1+2n\alpha)^3} \quad \Theta(M_t, x_0, t_0)$$

Now choose $\alpha := \frac{\beta}{1-2n\beta} > 0$ and $\sigma := \sqrt{1-2n\beta}\varrho$, observe that this choice of parameters does satisfy:

$$\alpha\sigma^2 < \frac{1}{2n}\varrho_0^2 - \frac{1}{2n}\sigma^2$$

and conclude that:

$$\frac{\mathscr{H}^n\left(M_{t_0-\beta\varrho^2} \cap B_\varrho(x_0)\right)}{\varrho^n} \geq (1-2n\beta)^3(4\pi\beta)^{n/2} \quad \Theta(M_t, x_0, t_0) \geq (1-2n\beta)^3(4\pi\beta)^{n/2}$$

as claimed. The last inequality follows since $\Theta(M_t, x_0, t_0) \geq 1$, according to (C.2.1) since the point x_0 is reached at time t_0. $\square$

APPENDIX D

Non-Compact Manifolds

Throughout the thesis we have assumed that our background manifold M^n, whose embeddings into $\mathbb{R}^{n+1}$ we flow by mean curvature, is compact. This was a decision made near the completion of the thesis, with the purpose of simplifying our arguments and working around complications that can occur *a priori* in the non-compact case. The purpose of this chapter is to give the reader some insight as to what sorts of issues can occur, and motivate one's choice of adopting a compactness hypothesis. We also extend Huisken's monotonicity formula to the non-compact case and prove a non-compact maximum principle.

D.1. Geometric Issues: Non-existence, Improperness

One issue that is almost always encountered by PDE in non-compact settings is that of non-existence of global solutions. In chapter 2 we showed that for any initial embedding F_0 there exists an $\varepsilon > 0$ such that F_0 can be extended to a smooth family of embeddings on $[0, \varepsilon)$.

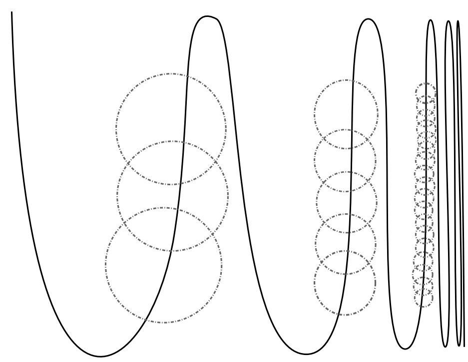


Figure 1: Local nature of ε in non-compact case. The circles' radius represents the magnitude of ε around each point.

This result, however, depends very strongly on our surface's compactness. It is not an issue of the argument we are using, but rather an inherent issue of our inability to control things from decaying in non-compact scenarios.

The same short-time existence argument does show that around any given point x we can extend the flow locally, but the catch is that ε would depend on the center point x. In the compact case this is a non-issue because we can find a uniform lower bound for our $\varepsilon(x)$, but in the non-compact case this is not at all true, as the previous diagram suggests.

Another issue that usually pops up in the non-compact case is the annoyance of having to distinguish between injective immersions and embeddings, as well as having to ask oneself questions regarding properness. This is not at all a big issue and is usually dealt with by stating explicitly which results hold for what types of morphisms, but this clutters the presentation and stains readability. In the compact case all these terms essentially coincide, which makes one's presentation an easier yet task.

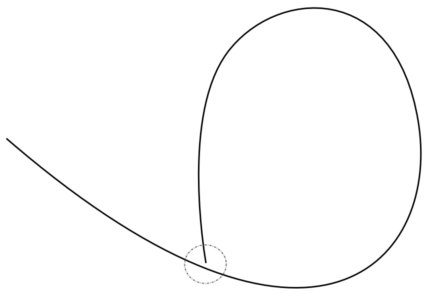


Figure 2: Prototypical example of a non-compact immersed submanifold that is not embedded.

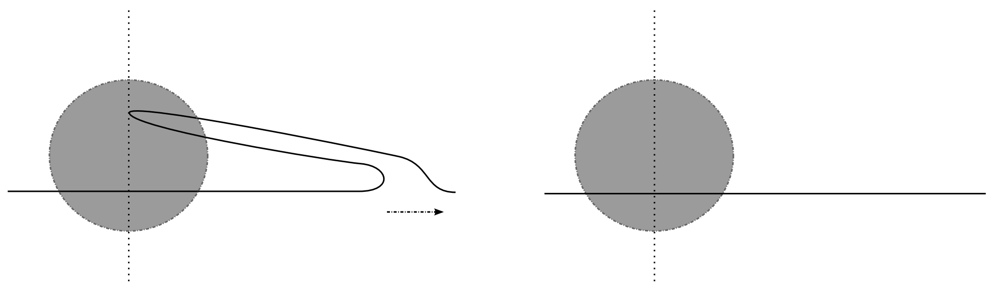
D.2. Measure Theoretic Issues: First Variation Discontinuity

Non-existence and improperness are nuances one can easily put up with, and indeed the original goal of the thesis was to be generic enough to work around these issues and treat the non-compact case. That is indeed what had happened, until it suddenly became evident that the first variation formula, on which most of this work is built, may break in the non-compact case!

It is not a matter of it being an inequality rather than an equality: one can check we never actually needed equality for our results. The problem is that even with compactly supported smooth cutoff functions, the first variation integral may conceivably fail to be differentiable, or even continuous, when our background manifold M^n is not compact! What is most unfortunate is that even though we have examples of smooth families of properly embedded manifolds moving in time where the integral is discontinuous, we have not actually been able to come up with a counterexample that moves by mean curvature flow.

Having been unable to either prove or disprove this concern, we decided to restrict the entire thesis to the compact case where the first variation formula works flawlessly. Another possibility would have been to adopt the first variation formula as our definition of mean curvature motion, but that would not have boded well with the non-measure-theoretic results, especially the geometric ones in Chapter 2.

Consider a smooth family of properly embedded curves $\gamma(\cdot, t) : \mathbb{R} \to \mathbb{R}^2$ moving as suggested by the diagram below:



(a) The base of the fold shifts out as $t \searrow 0$, but the tip of the fold is fixed.

(b) The fold disappears at $t = 0$.

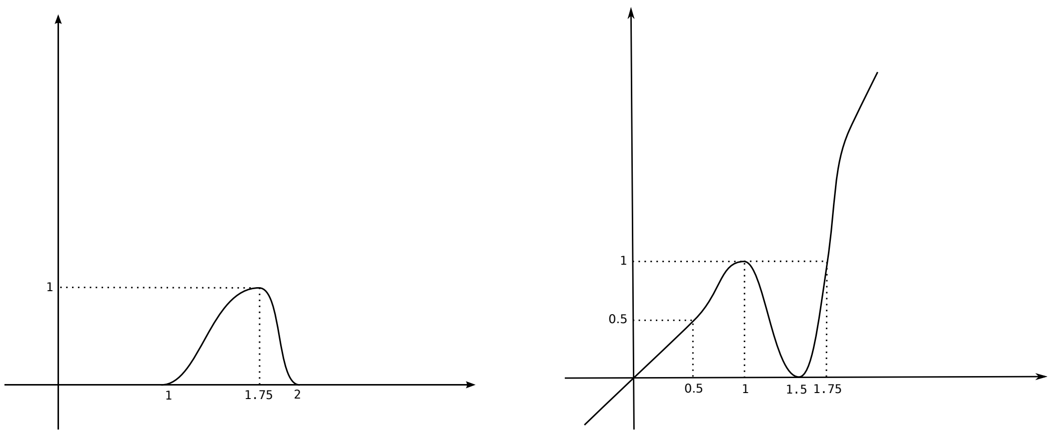
Figure 3: A smooth family of curves with a fold that slides out as $t \searrow 0$ and disappears at $t = 0$. The vertical dotted line represents the y-axis, $x = 0$, which the tip of the folds always touches at the same intercept.

Evidently, then, for a cutoff function χ around the gray-shaded disk the function:

$$t \mapsto \int_{\gamma(\cdot,t)} \chi \, d\mathscr{H}^1$$

is discontinuous at $t = 0$ because an entire chunk of the curve vanishes suddenly.

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a smooth function that vanishes identically outside of $[1, 2]$, increases on $[1, 1.75]$, reaches $\varphi(1.75) = 1$, and then decreases on $[1.75, 2]$. Let $\psi : \mathbb{R} \to \mathbb{R}$ be a smooth function that is just the identity on $(-\infty, 0.5]$, increases on $[0.5, 1]$, reaches $\psi(1) = 1$, decreases on $[1, 1.5]$, reaches $\psi(1.5) = 0$, and increases on $[1.5, \infty)$ by also passing through $\psi(1.75) = 1$. One goes about constructing such smooth functions with simple tricks from analysis.



(a) $\varphi : \mathbb{R} \to \mathbb{R}$ (b) $\psi : \mathbb{R} \to \mathbb{R}$

Figure 4: Two auxiliary functions we need to construct the sliding fold.

At this point it is a matter of calculations to check that:

$$\gamma(x,t) := \begin{pmatrix} \frac{\psi(tx)}{t} \\ \sin \pi h(tx) \end{pmatrix} \quad \text{for } t > 0 \quad \text{and} \quad \gamma(x,0) := \begin{pmatrix} x \\ 0 \end{pmatrix}$$

is a smooth family of properly embedded curves $\gamma(\cdot, t)$ that look essentially like our sliding fold diagram above. The fact that γ is differentiable may seem surprising but it's really a consequence of $h(tx)$ being precisely tx for t small (depending on x).

D.3. Non-Compact Huisken Monotonicity Formula

Ecker **[Eck04]** often makes strong differentiability claims which only turn out to hold weakly. For instance, he claimed that the distance function between two non-intersecting compact solutions to mean curvature has non-negative derivative, while it is only sensible to expect for this distance to be merely a non-decreasing function–something we have shown in Chapter 2.

The original goal behind this section was to substantiate remarks made in **[Eck04]** about Huisken's monotonicity formula (3.5.1) and its extension to non-compact cases. We wish to understand the extent to which his claims on Huisken monotonicity are true.

There is an obvious objection that is voiced by just the previous section: it is conceivable that the first variation formula will not hold for our non-compact surfaces, in which case there is nothing to study. We could just restrict our attention to non-compact families of surfaces moving by mean curvature for which the first variation formula (3.1.3) works. In our attempt to mimic Brakke flow theory, we will simply adopt the first variation formula as our definition of mean curvature flow. Notice then that all results of Chapter 3 are available for use, because they do not require compactness but merely the first variation formula.

We prove that Ecker's differentiability claims are indeed true weakly. For an introduction to weak differentiability and Sobolev spaces we refer the reader to any standard text in PDE (see, e.g. **[Fol95]**, **[Tay96a]**). Ostensibly, strong differentiability could break for isolated time instances when there is significant curvature contribution in (3.5.1). This is discussed later in this section.

PROPOSITION D.3.1. Let (M_t) move by the first variation formula (3.1.3) for $t \in [t_1, t_0)$, and let $x_0 \in \mathbb{R}^{n+1}$ be arbitrary. Then the weighted mass function:

$$t \mapsto \int_{M_t} \Phi_{(x_0,t_0)}$$

is a non-increasing (but not necessarily finite) function of $t \in [t_1, t_0)$.

PROOF. Fix $T_1 \in [t_1, t_0)$, and let $R > 0$, $\theta \in (0, 1)$ be arbitrary. Consider the standard test function $\varphi_{(x_0,T_1),R}$. Then by (3.5.2) we know that:

$$t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,T_1),R}$$

is a decreasing function of t, and consequently for every $t \in (T_1, t_0)$:

$$\int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,T_1),R} \leq \int_{M_{T_1}} \Phi_{(x_0,t_0)} \varphi_{(x_0,T_1),R} \tag{$\dagger_1$}$$

Let us estimate both sides of the inequality. Starting with the right hand side of ($\dagger_1$), we get:

$$\int_{M_{T_1}} \Phi_{(x_0,t_0)} \varphi_{(x_0,T_1),R} \leq \sup_{\mathbb{R}^{n+1}} \varphi_{(x_0,T_1),R}(\cdot, T_1) \quad \int_{M_{T_1}} \Phi_{(x_0,t_0)} \leq \int_{M_{T_1}} \Phi_{(x_0,t_0)} \tag{$\dagger_2$}$$

Let us look at the left hand side of ($\dagger_1$):

$$\begin{aligned} \int_{M_t} \Phi_{(x_0,t_0)} \varphi_{(x_0,T_1),R} &\geq \int_{M_t \cap B_{\theta R}(x_0)} \Phi_{(x_0,t_0)} \varphi_{(x_0,T_1),R} \\ &\geq \inf_{B_{\theta R}(x_0)} \varphi_{(x_0,T_1),R}(\cdot, t) \quad \int_{M_t \cap B_{\theta R}(x_0)} \Phi_{(x_0,t_0)} \\ &= \left(1 - \theta^2 - \frac{2n(t - T_1)}{R^2}\right)_+^3 \quad \int_{M_t \cap B_{\theta R}(x_0)} \Phi_{(x_0,t_0)} \end{aligned} \tag{$\dagger_3$}$$

Combining ($\dagger_1$), ($\dagger_2$), and ($\dagger_3$), we see that for every $T_1 \in [t_1, t_0)$, $t \in (T_1, t_0)$, $R > 0$, and $\theta \in (0, 1)$ we have:

$$\left(1 - \theta^2 - \frac{2n(t - T_1)}{R^2}\right)_+^3 \quad \int_{M_t \cap B_{\theta R}(x_0)} \Phi_{(x_0,t_0)} \leq \int_{M_{T_1}} \Phi_{(x_0,t_0)}$$

Letting $R \to \infty$ and applying the monotone convergence theorem on the left (while keeping t fixed):

$$(1 - \theta^2)^3 \quad \int_{M_t} \Phi_{(x_0,t_0)} \leq \int_{M_{T_1}} \Phi_{(x_0,t_0)}$$

for every $T_1 \in [t_1, t_0)$, $t \in (T_1, t_0)$, $\theta \in (0, 1)$. Since $\theta \in (0, 1)$ was arbitrary, we conclude by letting $\theta \searrow 0$ that:

$$\int_{M_t} \Phi_{(x_0,t_0)} \leq \int_{M_{T_1}} \Phi_{(x_0,t_0)}$$

for every $T_1 \in [t_1, t_0)$, $t \in (T_1, t_0)$, and the result follows. $\square$

If we now additionally impose a finiteness constraint on the weighted mass of the first manifold, then we can establish uniform finiteness and weak differentiability of the weighted mass function.

PROPOSITION D.3.2. Let (M_t) move by the first variation formula (3.1.3) for $t \in [t_1, t_0)$, and let $x_0 \in \mathbb{R}^{n+1}$ be arbitrary. If:

$$\int_{M_{t_1}} \Phi_{(x_0,t_0)} < \infty$$

then the weighted mass function is real valued, non-increasing, in $\mathscr{W}^{1,1}_{loc}(t_1, t_0)$, and satisfies

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

weakly in $\mathscr{L}^1_{loc}(t_1, t_0)$.

PROOF. By the function's monotonicity established (D.3.1), the finiteness condition carries over to all later times and hence our weighted mass function is finite at all times and non-increasing.

Once again, let $R > 1$ be arbitrary, and let χ_R be a smooth function in $\mathbb{R}^{n+1}$ such that $\mathbf{1}_{B_R(0)} \leq \chi_R \leq \mathbf{1}_{B_{2R}(0)}$ and $R|D\chi_R| + R^2|D^2\chi_R| \leq C(n)$. Observe that this implies:

$$\left|\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\chi_R\right| \leq \frac{C(n)}{R} \tag{$\ddagger_1$}$$

Then χ_R is a test function, and hence by (3.5.1) we have for every $t \in (t_1, t_0)$:

$$\frac{d}{dt}\int_{M_t} \Phi_{(x_0,t_0)}\chi_R = \int_{M_t} -\left|H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2 \Phi_{(x_0,t_0)}\chi_R + \Phi_{(x_0,t_0)}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\chi_R$$

Let $\psi : \mathbb{R} \to \mathbb{R}$ be smooth and with compact support within (t_1, t_0). Multiplying both sides of the equation above by ψ, integrating on (t_1, t_0), and integrating the left hand side by parts, we get:

$$-\int_{t_1}^{t_0} \psi'(t)\int_{M_t} \Phi_{(x_0,t_0)}\chi_R = \int_{t_1}^{t_0} \psi(t)\int_{M_t}\left(-\left|H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2 \Phi_{(x_0,t_0)}\chi_R + \Phi_{(x_0,t_0)}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right)\chi_R\right) \tag{$\ddagger_2$}$$

Rearranging:

$$
\begin{aligned}
&\int_{t_1}^{t_0} \psi(t) \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} \chi_R \\
\stackrel{|\cdot|}{\leq} \quad & \int_{t_1}^{t_0} |\psi'(t)| \int_{M_t} \Phi_{(x_0,t_0)} |\chi_R| \quad + \quad \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} \left| \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) \chi_R \right| \\
\stackrel{(\ddagger_1)}{\leq} \quad & \int_{t_1}^{t_0} |\psi'(t)| \int_{M_t} \Phi_{(x_0,t_0)} \quad + \quad \frac{C(n)}{R} \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} \\
\stackrel{\text{(D.3.1)}}{\leq} \quad & \int_{t_1}^{t_0} |\psi'(t)| \int_{M_{t_1}} \Phi_{(x_0,t_0)} \quad + \quad \frac{C(n)}{R} \int_{t_1}^{t_0} |\psi(t)| \int_{M_{t_1}} \Phi_{(x_0,t_0)} \\
\stackrel{|\cdot|}{\leq} \quad & (t_0 - t_1)|\psi'| \int_{M_{t_1}} \Phi_{(x_0,t_0)} \quad + \quad \frac{C(n)}{R} (t_0 - t_1)|\psi| \int_{M_{t_1}} \Phi_{(x_0,t_0)}
\end{aligned}
$$

We now let $R \to \infty$ above and apply Fatou's lemma:

$$
\int_{t_1}^{t_0} \psi(t) \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} \leq (t_0 - t_1)|\psi'| \int_{M_{t_1}} \Phi_{(x_0,t_0)}
$$

In other words:

$$
\left(t \mapsto \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} \right) \in \mathscr{L}^1_{loc}(t_1, t_0) \tag{$\ddagger_3$}
$$

Let us return to ($\ddagger_2$). We let $R \to \infty$. The left hand side and the right-most term are handled by the classical dominated convergence theorem, (B.2.1). But ($\ddagger_3$) tells us that the middle term can also be handled by classical dominated convergence, and thus:

$$
\int_{t_1}^{t_0} \psi'(t) \int_{M_t} \Phi_{(x_0,t_0)} = \int_{t_1}^{t_0} \psi(t) \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} \tag{$\ddagger_4$}
$$

Since ψ was an arbitrary smooth function with compact support within (t_1, t_0), we conclude:

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

$\mathscr{L}^1_{loc}$-weakly in (t_1, t_0), so the result follows. $\square$

Let us now begin to consider incorporating other integrands in our weighted mass function. The following result generalizes Huisken's monotonicity formula (3.5.1) in the non-compact case.

PROPOSITION D.3.3. Let (M_t) move by the first variation formula (3.1.3) in (t_1, t_0), and let $x_0 \in \mathbb{R}^{n+1}$ be arbitrary. Suppose, additionally, that $f : \mathbb{R}^{n+1} \times (t_1, t_0) \to \mathbb{R}$ is $\mathcal{C}^1$ in the product space, and each $f(\cdot, t)$ is $\mathcal{C}^2$. Let us additionally impose the growth constraint:

$$\left(t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} \left(|f| + |Df| + |D^2 f| + |D_t f|\right)\right) \in \mathscr{L}^1_{loc}(t_1, t_0)$$

Then the weighted mass function integrated against f:

$$t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} f$$

is $\mathscr{W}^{1,1}_{loc}(t_1, t_0)$ and satisfies:

$$\frac{d}{dt} \int_{M_t} \Phi_{(x_0,t_0)} f = \int_{M_t} - \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} f + \Phi_{(x_0,t_0)} \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) f$$

weakly in $\mathscr{L}^1_{loc}(t_1, t_0)$.

PROOF. Once again, let $R > 1$ be arbitrary, and let χ_R be a smooth function on $\mathbb{R}^{n+1}$ which satisfies the inequalities $\mathbf{1}_{B_R(0)} \leq \chi_R \leq \mathbf{1}_{B_{2R}(0)}$ and $R|D\chi_R| + R^2|D^2\chi_R| \leq C(n)$. Observe that this implies:

$$\left| \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) \chi_R \right| \leq \frac{C(n)}{R} \tag{$\ddagger_1'$}$$

Then χ_R is a test function, and hence by (3.5.1) we have for every $t \in (t_1, t_0)$:

$$\frac{d}{dt} \int_{M_t} \Phi_{(x_0,t_0)} f\chi_R = \int_{M_t} - \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} f\chi_R + \Phi_{(x_0,t_0)} \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) (f\chi_R)$$

Let ψ be smooth and with compact support $[p, q] \subset (t_1, t_0)$. Using the product rule (A.3.3.a), then multiplying both sides of the equation above by ψ, then integrating on (t_1, t_0), then integrating the left hand side by parts, finally rearranging we get:

$$\begin{aligned}
&\int_{t_1}^{t_0} \psi(t) \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} f \chi_R \qquad (\ddagger_2') \\
= \; & \int_{t_1}^{t_0} \psi'(t) \int_{M_t} \Phi_{(x_0,t_0)} f \chi_R \quad + \quad \int_{t_1}^{t_0} \psi(t) \int_{M_t} \Phi_{(x_0,t_0)} f \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) \chi_R \\
& + \quad \int_{t_1}^{t_0} \psi(t) \int_{M_t} \Phi_{(x_0,t_0)} \chi_R \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) f \quad - \quad 2 \int_{t_1}^{t_0} \psi(t) \int_{M_t} \Phi_{(x_0,t_0)} \nabla f \cdot \nabla \chi_R
\end{aligned}$$

We may estimate this from above as follows:

$$\begin{aligned}
\leq \; & \int_{t_1}^{t_0} |\psi'(t)| \int_{M_t} \Phi_{(x_0,t_0)} |f| |\chi_R| \quad + \quad \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} |f| \left| \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) \chi_R \right| \\
& + \quad \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} |\chi_R| \left| \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) f \right| \quad + \quad 2 \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} |\nabla f| |\nabla \chi_R| \\
\overset{(\ddagger_1')}{\leq} \; & \int_{t_1}^{t_0} |\psi'(t)| \int_{M_t} \Phi_{(x_0,t_0)} |f| \quad + \quad \frac{C(n)}{R} \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} |f| \\
& + \quad \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} \left(|D_t f| + |D^2 f| \right) \quad + \quad \frac{2C(n)}{R} \int_{t_1}^{t_0} |\psi(t)| \int_{M_t} \Phi_{(x_0,t_0)} |Df| \\
\leq \; & \left(|\psi'| + \left(1 + \frac{2C(n)}{R} \right) |\psi| \right) \int_{t_1}^{t_0} \mathbf{1}_{[p,q]} \int_{M_t} \Phi_{(x_0,t_0)} \left(|f| + |Df| + |D^2 f| + |D_t f| \right)
\end{aligned}$$

Therefore by letting $R \to \infty$ and employing Fatou's lemma we get:

$$\int_{t_1}^{t_0} \psi(t) \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} f \leq (|\psi'| + |\psi|) \int_{t_1}^{t_0} \mathbf{1}_{[p,q]} \int_{M_t} \Phi_{(x_0,t_0)} \left(|f| + |Df| + |D^2 f| + |D_t f| \right)$$

or in other words we get that:

$$\left(t \mapsto \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} f \right) \in \mathscr{L}^1_{loc}(t_1, t_0) \qquad (\ddagger_3')$$

Now we can finally let $R \to \infty$ in $(\ddagger_2')$, because $(\ddagger_3')$ guarantees that we can handle the left hand side by dominated convergence, (B.2.1):

$$\int_{t_1}^{t_0} \psi(t) \int_{M_t} \left| H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}} \right|^2 \Phi_{(x_0,t_0)} f = \int_{t_1}^{t_0} \psi'(t) \int_{M_t} \Phi_{(x_0,t_0)} f \quad + \quad \int_{t_1}^{t_0} \psi(t) \int_{M_t} \Phi_{(x_0,t_0)} \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) f$$

Since ψ was an arbitrary smooth function with compact support within (t_1, t_0), we conclude that indeed

$$\frac{d}{dt}\int_{M_t} \Phi_{(x_0,t_0)} f = \int_{M_t} -\left|H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2 \Phi_{(x_0,t_0)} f + \Phi_{(x_0,t_0)}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f$$

weakly in $\mathscr{L}^1_{loc}(t_1, t_0)$, as claimed. □

Notice that we can't say a lot about strong differentiability on (t_1, t_0). Our weak differentiability together with basic Sobolev space theory (see, e.g. **[Fol95]**, **[Tay96a]**) guarantees strong differentiability a.e. on (t_1, t_0), but there is no working around the fact that there is an elusive set of measure zero on which strong differentiability might break. Strong differentiability will break if either:

$$\text{curvature:} \quad \left|H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2 \Phi_{(x_0,t_0)} f \quad \text{or parabolic evolution:} \quad \Phi_{(x_0,t_0)}\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f$$

are too large. Despite efforts by the author, everywhere-existence of the strong derivative remains unsolved. The main obstacle in coming up with a counterexample, at least where the curvature term blows up, is that non-compact manifolds with high curvature contribution are hard to visualize statically, and even harder to flow by mean curvature. Fortunately, we have no need for accurate pointwise strong differentiability claims so we can leave the matter at rest.

Consider, however, the case of

$$f \geq 0 \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq 0$$

where the weak differentiation formula tells us that, up to redefinition on a set of measure zero, the function:

$$t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} f$$

is (weakly) decreasing. Once again, we have an elusive set of times of measure zero, when ideally we would like to understand classical monotonicity for all time instants. Luckily, with a slight modification of our previous arguments we can do exactly that, with even weaker hypotheses!

PROPOSITION D.3.4. Let (M_t) move by the first variation formula (3.1.3) in (t_1, t_0), and let $x_0 \in \mathbb{R}^{n+1}$ be arbitrary. Suppose, additionally, that $f : \mathbb{R}^{n+1} \times (t_1, t_0) \to \mathbb{R}$ is $\mathcal{C}^1$ in the product space, that each $f(\cdot, t)$ is $\mathcal{C}^2$, and that:

$$f \geq 0 \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq 0$$

Let us additionally impose the growth constraint:

$$\left(t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} \left(|f| + |Df|\right)\right) \in \mathscr{L}^1_{loc}[t_1, t_0)$$

Then the weighted mass function integrated against f is a real-valued (except possibly at t_1, were it might be infinite), non-increasing function of $t \in [t_1, t_0)$.

PROOF. Fix $T_1 \in [t_1, t_0)$. Let $R > 1$ be arbitrary, and let χ_R be a smooth function in $\mathbb{R}^{n+1}$ such that $\mathbf{1}_{B_R(0)} \leq \chi_R \leq \mathbf{1}_{B_{2R}(0)}$ and $R|D\chi_R| + R^2|D^2\chi_R| \leq C(n)$. Once again, this implies:

$$\left|\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \chi_R\right| \leq \frac{C(n)}{R}$$

Then working as before and employing (A.3.3.a) and (3.5.1) for the function of compact support $f\chi_R$ we get, at $t \in (T_1, t_0)$:

$$\begin{aligned}
&\frac{d}{dt}\int_{M_t} \Phi_{(x_0,t_0)} f\chi_R \\
= \;& \int_{M_t} -\left|H - \frac{\nabla^\perp \Phi_{(x_0,t_0)}}{\Phi_{(x_0,t_0)}}\right|^2 \Phi_{(x_0,t_0)} f\chi_R \quad + \quad \int_{M_t} \Phi_{(x_0,t_0)} \chi_R \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \\
& + \quad \int_{M_t} \Phi_{(x_0,t_0)} f \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \chi_R \quad - \quad 2\int_{M_t} \Phi_{(x_0,t_0)} \nabla f \cdot \nabla \chi_R \\
\leq \;& \int_{M_t} \Phi_{(x_0,t_0)} f \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \chi_R \quad - \quad 2\int_{M_t} \Phi_{(x_0,t_0)} \nabla f \cdot \nabla \chi_R
\end{aligned}$$

Therefore by the fundamental theorem of calculus on $[T_1, t]$, $t \in (T_1, t_0)$, we have:

$$\int_{M_t} \Phi_{(x_0,t_0)} f\chi_R \quad \leq \quad \int_{M_{T_1}} \Phi_{(x_0,t_0)} f\chi R + \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} f \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) \chi_R - 2\int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} \nabla f \cdot \nabla \chi_R \qquad (\ddagger_1'')$$

The left hand side we may estimate from below as follows:

$$\int_{M_t} \Phi_{(x_0,t_0)} f \chi_R \geq \int_{M_t \cap B_R(0)} \Phi_{(x_0,t_0)} f \tag{$\ddagger_2''$}$$

because $f, \chi_R \geq 0$ and $\chi_R = 1$ on $B_R(0)$. The right hand side of ($\ddagger_2''$) we estimate from above by:

$$\begin{aligned}
& \int_{M_{T_1}} \Phi_{(x_0,t_0)} f \chi R \quad + \quad \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} f \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) \chi_R \quad - \quad 2 \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} \nabla f \cdot \nabla \chi_R \\
\leq \; & \int_{M_{T_1}} \Phi_{(x_0,t_0)} f |\chi R| \quad + \quad \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} f \left| \left(\frac{\partial}{\partial t} - \Delta_{M_t} \right) \chi_R \right| \quad + \quad 2 \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} |\nabla f| |\nabla \chi_R| \\
\leq \; & \int_{M_{T_1} \cap B_{2R}(0)} \Phi_{(x_0,t_0)} f \quad + \quad \frac{C(n)}{R} \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} f \quad + \quad \frac{2C(n)}{R} \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} |\nabla f|
\end{aligned} \tag{$\ddagger_3''$}$$

Combining ($\ddagger_1''$), ($\ddagger_2''$), and ($\ddagger_3''$) yields:

$$\int_{M_t \cap B_R(0)} \Phi_{(x_0,t_0)} f \leq \int_{M_{T_1} \cap B_{2R}(0)} \Phi_{(x_0,t_0)} f \quad + \quad \frac{C(n)}{R} \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} f \quad + \quad \frac{2C(n)}{R} \int_{T_1}^t \int_{M_\tau} \Phi_{(x_0,t_0)} |\nabla f|$$

Letting $R \to \infty$ and recalling our $\mathscr{L}^1_{loc}$ growth conditions gives:

$$\int_{M_t} \Phi_{(x_0,t_0)} f \leq \int_{M_{T_1}} \Phi_{(x_0,t_0)} f$$

for every $T_1 \in [t_1, t_0)$ and $t \in (T_1, t_0)$. In other words:

$$\left(t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} f \right) \quad \text{is non-increasing on} \quad [t_1, t_0)$$

Notice that *a priori* the values of the function above are not necessarily finite. But since we know that this function is also $\mathscr{W}^{1,1}_{loc}(t_1, t_0)$, then we know that it is finite a.e. on (t_1, t_0) and hence by monotonicity it is constrained to be finite everywhere on $[t_1, t_0)$, except possibly at t_1. $\square$

D.4. Non-Compact Weak Maximum Principle

This last monotonicity result has a very interesting application from the PDE point of view. We can use (D.3.4) to prove a weak maximum principle satisfied by non-compact solutions of our mean curvature flow problem which obey mild growth assumptions.

PROPOSITION D.4.1. Let (M_t) move by the first variation formula (3.1.3) in $[t_1, t_0)$. Let $f : \mathbb{R}^{n+1} \times [t_1, t_0) \to \mathbb{R}$ be $\mathcal{C}^1$ in the product space and $f(\cdot, t) \in \mathcal{C}^2$ for every $t \in [t_1, t_0)$. Suppose additionally that:

$$f \geq 0 \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) f \leq 0 \quad \text{and} \quad \left(t \mapsto \int_{M_t} \Phi_{(x_0,t_0)} \left(|f| + |Df|\right)\right) \in \mathscr{L}^1_{loc}[t_1, t_0)$$

Then:

$$t \mapsto \sup_{M_t} f(\cdot, t)$$

is a non-increasing (possibly infinite) function of $t \in [t_1, t_0)$.

PROOF. Let us write $\sup_{M_t} f$ for $\sup_{M_t} f(\cdot, t)$. Let $T_1 \in [t_1, t_0)$ be such that $\sup_{M_{T_1}} f < \infty$. Let $\Sigma := \sup_{M_{T_1}} f < \infty$. Define a non-negative, non-decreasing, convex $\mathcal{C}^2$ function $\gamma : \mathbb{R} \to \mathbb{R}$ that vanishes identically on $(-\infty, \Sigma]$, is strictly positive on (Σ, ∞), and $\gamma(x) = x - M$ for all $x \geq \Sigma + 1$. In other words γ is a non-negative convex function that vanishes up to Σ, and grows (essentially) linearly thereafter.

With such a γ and $g(x) := \gamma(f(x))$ we evidently have $g \geq 0$, $|g| + |Dg| \leq C(|f| + |Df|)$ for some constant $C > 0$, and since γ is non-decreasing and convex we may employ the chain rule (A.3.3.b) and get:

$$\left(\frac{\partial}{\partial t} - \Delta_{M_t}\right) g \leq 0$$

Therefore g satisfies the hypotheses of (D.3.4), so that for every $t \in (T_1, t_0)$ we have:

$$\int_{M_t} \Phi_{(x_0,t_0)} g \leq \int_{M_{T_1}} \Phi_{(x_0,t_0)} g$$

But $f \leq \Sigma$ on M_{T_1}, so $g = 0$ on M_{T_1} so the right hand side is zero. But the left hand side is non-negative, so $g = 0$ on M_t. Since $\gamma(x) > 0$ when $x > \Sigma$, we conclude that $\sup_{M_t} f \leq \Sigma = \sup_{M_{T_1}} f$ for all $t \in (T_1, t_0)$ when the latter quantity is finite. But the same is also obviously true when that quantity is infinite, so the result follows. $\square$

APPENDIX E

Representation by Graphs

In this chapter of the appendix we tackle the issue of representing n-dimensional manifolds (or smoothly moving families thereof) as graphs over a given affine space $T^n \subset \mathbb{R}^{n+1}$. In the last section we prove an Arzelà-Ascoli type compactness theorem for submanifolds which is used by our regularity theory.

The geometric results here have nothing to do with mean curvature flow, so we will not assume any such specific motion. We will also not require compactness or that our manifolds be embedded, unless otherwise stated. In particular, we will assume that we are working with possibly non-compact immersed n-dimensional submanifolds in $\mathbb{R}^{n+1}$.

E.1. Very Local Representations

We begin with a result which lets us write families $F : M^n \times I \to \mathbb{R}^{n+1}$ of immersed submanifolds locally in the intrinsic topology as graphs over a fixed affine space $T^n \subset \mathbb{R}^{n+1}$.

LEMMA E.1.1. Let M^n be a smooth n-dimensional manifold, $I \subseteq \mathbb{R}$ be an open set of times, and $F : M^n \times I \to \mathbb{R}^{n+1}$ be a smooth family of immersions $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$. Fix a point $p_0 \in M^n$, a time $t_0 \in I$, and an n-dimensional affine space $T^n \subset \mathbb{R}^{n+1}$ whose normal space $N = \text{span}\,\{\eta\}$ is such that:

$$N \cap T_{F(p_0,t_0)} F(M^n, t_0) = \{0\}$$

Denote by ξ_0 the projection of $F(p_0, t_0)$ onto T^n. Then there exists a $\varrho > 0$, a connected open set $J \subseteq I$ containing t_0, an open set $\mathcal{O} \subset M^n \times I$ containing (p_0, t_0), and a smooth function $u : B^n_\varrho(\xi_0) \times J \to \mathbb{R}$ such that:

$$\{F(p,t) \ : \ p \in M^n \cap \{(\cdot, t) \in \mathcal{O}\}\} = \text{graph}_{\,T^n} u(\cdot, t) = \left\{\xi + u(\xi, t)\eta \ : \ \xi \in B^n_\varrho(\xi_0)\right\}$$

for all fixed $t \in J$.

PROOF. Let $\pi : \mathbb{R}^{n+1} \to T^n$ denote the projection operation onto T^n. Consider the smooth map $G : M^n \times I \to T^n \times I$ acting by:

$$G(p,t) := (\pi(F(p,t)), t)$$

Observe that its differential satisfies:

$$dG = \left(\begin{array}{ccc|c} \ddots & & & \vdots \\ & \pi(d_pF) & & d_tF \\ & & \ddots & \vdots \\ \hline & 0 & & 1 \end{array}\right)$$

Recall that $N \cap T_{F(p_0,t_0)}F(M^n, t_0) = \{0\}$, which implies that at (p_0, t_0) the submatrix $\pi(d_pF)$ has full rank, and hence dG is of full rank at (p_0, t_0). Consequently, by the inverse function theorem there exist open sets $\mathcal{O}_0 \subset M^n \times I$, $\mathcal{W}_0 \subset T^n \times I$, such that $(p_0, t_0) \in \mathcal{O}_0$, $G : \mathcal{O}_0 \to \mathcal{W}_0$ is a smooth bijection, and its inverse $H : \mathcal{W}_0 \to \mathcal{O}_0$ is also smooth. By the openness of $\mathcal{W}_0$, there exists a $\varrho > 0$ and a connected open $J \subset I$ containing t_0 such that $B^n_\varrho(\xi_0) \times J \subset \mathcal{W}_0$. Let $\mathcal{O}$ be the open subset of $\mathcal{O}_0$ that is the preimage of $B^n_\varrho(\xi_0) \times J$.

Define our potential graph function $u : B^n_\varrho(\xi_0) \times J \to \mathbb{R}$ by:

$$u(\xi, t) := (F(H(\xi,t)) - \xi) \cdot \eta$$

It is obvious that u is smooth. Since G, H are inverses, we have $(\xi, t) = G(H(\xi,t)) = (\pi(F(H(\xi,t))), t)$ for all $(\xi, t) \in B^n_\varrho(\xi_0) \times J$. Reading off the first component gives:

$$\pi(F(H(\xi,t))) = \xi \tag{$\dagger_1$}$$

for all $(\xi, t) \in B^n_\varrho(\xi_0)$. If $(\cdot)^\perp$ is the (vector) orthogonal component with respect to T^n, ($\dagger_1$) gives:

$$\begin{aligned} F(H(\xi,t))^\perp &= \left(F(H(\xi,t))^\perp \cdot \eta\right)\eta = ((F(H(\xi,t)) - \pi(F(H(\xi,t)))) \cdot \eta)\,\eta \\ &= ((F(H(\xi,t)) - \xi) \cdot \eta)\,\eta = u(\xi,t)\eta \end{aligned} \tag{$\dagger_2$}$$

We now have the tools to check that u is the required graph representation. We will do this by showing that $\operatorname{graph}_{T^n} u(\cdot,t) \subseteq \{F(p,t) \;:\; p \in M^n \cap \{(\cdot,t) \in \mathcal{O}\}\}$ and also $\{F(p,t) \;:\; p \in M^n \cap \{(\cdot,t) \in \mathcal{O}\}\} \subseteq \operatorname{graph}_{T^n} u(\cdot,t)$ for all $t \in J$. So let $t \in J$ be arbitrary.

For all $\xi \in B^n_\varrho(\xi_0)$ we have:

$$\begin{aligned}\xi + u(\xi,t)\eta &\overset{(\dagger_2)}{=} \xi + F(H(\xi,t))^\perp \\ &\overset{(\dagger_1)}{=} \pi\left(F(H(\xi,t))\right) + F(H(\xi,t))^\perp \\ &= F(H(\xi,t)) \in \{F(p,t) \;:\; p \in M^n \cap \{(\cdot,t) \in \mathcal{O}\}\}\end{aligned}$$

and consequently, since ξ was arbitrary, we have $\operatorname{graph}_{T^n} u(\cdot,t) \subseteq \{F(p,t) \;:\; p \in M^n \cap \{(\cdot,t) \in \mathcal{O}\}\}$.

For the reverse inclusion let $p \in \{(\cdot,t) \in \mathcal{O}\}$ be arbitrary. Define ξ from the relation $(\xi,t) := G(p,t) \in B^n_\varrho(\xi_0) \times J$. Working essentially backwards, we observe that:

$$\begin{aligned}F(p,t) = F(H(\xi,t)) &= \pi\left(F(H(\xi,t))\right) + F(H(\xi,t))^\perp \\ &\overset{(\dagger_1)}{=} \xi + F(H(\xi,t))^\perp \\ &\overset{(\dagger_2)}{=} \xi + u(\xi,t)\eta \in \operatorname{graph}_{T^n} u(\cdot,t)\end{aligned}$$

and the result follows, in view of the arbitrariness of p. □

Notice that in the absence of motion, this result reads:

COROLLARY E.1.2. Let M^n be a smooth n-dimensional manifold immersed into $\mathbb{R}^{n+1}$ via $F : M^n \to \mathbb{R}^{n+1}$. Fix a point $p_0 \in M^n$ and an n-dimensional affine space $T^n \subset \mathbb{R}^{n+1}$ whose normal space $N = \operatorname{span}\{\eta\}$ is such that:

$$N \cap T_{F(p_0)}F(M^n) = \{0\}$$

Denote by ξ_0 the projection of $F(p_0)$ onto T^n. Then there exists a $\varrho > 0$, an open set $\mathcal{U} \subset M^n$ containing p_0, and a smooth function $u : B^n_\varrho(\xi_0) \to \mathbb{R}$ such that:

$$F(\mathcal{U}) = \operatorname{graph}_{T^n} u = \left\{\xi + u(\xi)\eta \;:\; \xi \in B^n_\varrho(\xi_0)\right\}$$

The drawback of the results above is that they depend on the intrinsic geometry of M^n. In certain cases, however, we want to be able to check that we get a graph representation within a small neighborhood in the ambient space. Fortunately, with compactness and embeddedness this can be arranged to happen in small cylindrical neighborhoods. Denote:

$$C_{\varrho,h,T^n}(x_0) := \text{open cylinder normal to } T^n \text{ centered at } x_0 \text{ with radius } \varrho \text{ and height } 2h$$

THEOREM E.1.3. Let M^n be a smooth n-dimensional compact manifold, $I \subseteq \mathbb{R}$ be an open set of times, and $F : M^n \times I \to \mathbb{R}^{n+1}$ be a smooth family of embeddings $F(\cdot, t) : M^n \to \mathbb{R}^{n+1}$. Fix a point $p_0 \in M^n$, a time $t_0 \in I$, and an n-dimensional affine space $T^n \subset \mathbb{R}^{n+1}$ whose normal space $N = \text{span}\,\{\eta\}$ is such that:

$$N \cap T_{F(p_0,t_0)} F(M^n, t_0) = \{0\}$$

Let x_0 be $F(p_0, t_0)$ and ξ_0 be its projection onto T^n. Then there exist $0 < \varrho < h$, a connected open set $J \subseteq I$ containing t_0, and a smooth function $u : B^n_\varrho(\xi_0) \times J \to \mathbb{R}$ such that for every $t \in J$:

$$M_t \cap C_{\varrho,h,T^n}(x_0) = \text{graph}_{\,T^n} u(\cdot, t) = \left\{\xi + u(\xi, t)\eta \;:\; \xi \in B^n_\varrho(\xi_0)\right\}$$

PROOF. By applying the previous lemma, (E.1.1), we know there exists a connected open set J_0 containing t_0, a radius ϱ_0, a smooth function $u : B^n_{\varrho_0}(\xi_0) \times J_0 \to \mathbb{R}$, and an open $\mathcal{O} \subseteq M^n \times J_0$ such that:

$$\{F(p, t) \;:\; p \in M^n \cap \{(\cdot, t) \in \mathcal{O}\}\} = \text{graph}_{\,T^n} u(\cdot, t) = \left\{\xi + u(\xi, t)\eta \;:\; \xi \in B^n_\varrho(\xi_0)\right\} \qquad (\dagger_1)$$

for all $t \in J_0$. We will need the following topological property: there exists an $\varepsilon > 0$ such that for every $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ we have:

$$\{F(\cdot, t) \in C_{\varepsilon,\varepsilon,T^n}(x_0)\} \subseteq \{(\cdot, t) \in \mathcal{O}\} \qquad (\dagger_2)$$

Intuitively, this says that only points within $\mathcal{O}$ map to a sufficiently small cylindrical neighborhood of $F(p_0, t_0)$ in the ambient space. Suppose, for the sake of contradiction, that this were false. Take a sequence $\varepsilon_n \searrow 0$ for which this fails. That is, for every n there exists a $p_n \in M^n$ and a $t_n \in (t_0 - \varepsilon_n, t_0 + \varepsilon_n)$ such that $F(p_n, t_n) \in C_{\varepsilon_n,\varepsilon_n,T^n}$ and $(p_n, t_n) \notin \mathcal{O}$.

Since M^n is compact, we may assume that $p_n \to p_\infty \in M^n$ after possibly passing to a subsequence. Since $\varepsilon_n \searrow 0$, we also have $t_n \to t_0$ and $F(p_\infty, t_0) = x_0$.

Since $F(\cdot, t_0) : M^n \to \mathbb{R}^{n+1}$ is injective and $F(p_\infty, t_0) = x_0 = F(p_0, t_0)$, we have $p_\infty = p_0$. On the other hand $(p_n, t_n) \notin \mathcal{O}$ for all n, and hence the limit point (p_∞, t_0) is also outside $\mathcal{O}$, and in particular it cannot be that $p_\infty = p_0$. This is a contradiction, and our ε-claim follows.

Now let $\delta > 0$ be small enough that $\delta < \varepsilon$, $\delta < \varrho_0$, $(t_0 - \delta, t_0 + \delta) \subseteq J_0$, and $\operatorname{graph}_{T^n} u\,(B^n_\delta(\xi_0) \times (t_0 - \delta, t_0 + \delta)) \subseteq C_{\varepsilon,\varepsilon}(x_0)$. The latter can be arranged because u is continuous and $\operatorname{graph}_{T^n} u(\xi_0, t_0) = x_0$. In fact, since the graph of $u(\cdot, t)$ on $B^n_\delta(\xi_0)$ cannot map outside the cylinder of radius δ, the last inclusion can be revised to:

$$\operatorname{graph}_{T^n} u\,(B^n_\delta(\xi_0) \times (t_0 - \delta, t_0 + \delta)) \subseteq C_{\delta,\varepsilon}(x_0) \tag{$\dagger_3$}$$

Notice that we're abusing notation here and using $\operatorname{graph}_{T^n} u$ to denote the function mapping into $\mathbb{R}^{n+1}$. We now claim that the result follows with $J := (t_0 - \delta, t_0 + \delta)$, $\varrho := \delta$, and $h := \varepsilon$. Indeed, let $t \in (t_0 - \delta, t_0 + \delta)$ be arbitrary.

By ($\dagger_3$) we know that u maps within the cylinder $C_{\delta,\varepsilon,T^n}$. By construction, though, it also maps within M_t. Thus:

$$\begin{aligned}
\operatorname{graph}_{T^n} u(\cdot, t)\big|_{B^n_\delta(\xi_0)} &\overset{(\dagger_3)}{\subseteq} M_t \cap C_{\delta,\varepsilon,T^n}(x_0) \\
&= \{F(\cdot, t) \ : \ M^n\} \cap C_{\delta,\varepsilon,T^n}(x_0) \\
&= \{F(\cdot, t) \ : \ M^n \cap \{F(\cdot, t) \in C_{\delta,\varepsilon,T^n}(x_0)\}\}
\end{aligned}$$

Now $C_{\delta,\varepsilon,T^n}(x_0) \subseteq C_{\varepsilon,\varepsilon,T^n}(x_0)$. But we also intersect again from the outside the brackets by $C_{\delta,\varepsilon,T^n}(x_0)$ to keep the cylinder radius at δ:

$$\begin{aligned}
\operatorname{graph}_{T^n} u(\cdot, t)\big|_{B^n_\delta(\xi_0)} &\subseteq \{F(\cdot, t) \ : \ M^n \cap \{F(\cdot, t) \in C_{\varepsilon,\varepsilon,T^n}(x_0)\}\} \cap C_{\delta,\varepsilon,T^n}(x_0) \\
&\overset{(\dagger_2)}{\subseteq} \{F(\cdot, t) \ : \ M^n \cap \{(\cdot, t) \in \mathcal{O}\}\} \cap C_{\delta,\varepsilon,T^n}(x_0) \\
&\overset{(\dagger_1)}{=} \operatorname{graph}_{T^n} u(\cdot, t)\big|_{B^n_{\varrho_0}(\xi_0)} \cap C_{\delta,\varepsilon,T^n}(x_0) \overset{(\dagger_3)}{=} \operatorname{graph}_{T^n} u(\cdot, t)\big|_{B^n_\delta(\xi_0)}
\end{aligned}$$

and the result follows. □

E.2. Controlling the Representation Radii

In this section we will not concern ourselves with any motion on behalf of our manifolds. Our no-movement result (E.1.2) from the previous section established that around any center point we can write a fixed immersed manifold locally as as the graph of a function. The catch is that the radius of representation may get arbitrarily small. In this section we show that we can always bound the radius of representation from below in terms of the curvature tensor.

We begin with a uniqueness lemma that will let us patch graph representations together.

LEMMA E.2.1. Let M^n be a smooth n-dimensional manifold injectively immersed into $\mathbb{R}^{n+1}$ by $F : M^n \to \mathbb{R}^{n+1}$. Let $T^n \subset \mathbb{R}^{n+1}$ be any affine space. Suppose that for $i = 1, 2$ the sets $\mathcal{U}_i \subseteq M^n$, $V_i \subseteq T^n$ are open, and there exist differentiable functions $u_i : V_i \to \mathbb{R}$ such that:

$$F(\mathcal{U}_i) = \text{graph}_{\,T^n} u_i$$

If $u_1(\xi_0) = u_2(\xi_0)$ for some $\xi_0 \in V_1 \cap V_2$, then $u_1 = u_2$ on the entire connected component of ξ_0 within $V_1 \cap V_2$.

PROOF. Consider the set $\Sigma := \{\xi \in V_1 \cap V_2 \ : \ u_1(\xi) = u_2(\xi)\}$. By assumption $\xi_0 \in \Sigma$ and by continuity the set Σ is relatively closed in $V_1 \cap V_2$. It remains, then, to check that it is relatively open.

Let $\xi \in \Sigma$ be arbitrary, i.e. such that $u_1(\xi) = u_2(\xi)$. Then $\text{graph}_{\,T^n} u_1(\xi) = \text{graph}_{\,T^n} u_2(\xi)$, and thus both are equal to $F(p)$ for some $p \in \mathcal{U}_1 \cap \mathcal{U}_2$ since F is injective. Since F is locally the graph of the differentiable u_1 and/or u_2 around p, we know that its tangent space $T_{F(p)}F(M^n)$ cannot intersect T^n orthogonally. Consequently, we may now apply the previous local graph representation result (E.1.2) to conclude that there exists an open set $\mathcal{U} \subseteq M^n$ containing p, a radius $\varrho > 0$, and a smooth function $u : B^n_\varrho(\xi) \to \mathbb{R}$ such that:

$$F(\mathcal{U}) = \text{graph}_{\,T^n} u$$

Shrink $\varrho > 0$ to arrange for $\mathcal{U} \subseteq \mathcal{U}_1 \cap \mathcal{U}_2$ and $B^n_\varrho(\xi) \subseteq V_1 \cap V_2$ to occur. For any $\xi' \in B^n_\varrho(\xi)$ and $i = 1, 2$:

$$\text{graph}_{\,T^n} u(\xi') \in F(\mathcal{U}) \subseteq F(\mathcal{U}_1 \cap \mathcal{U}_2) \subseteq F(\mathcal{U}_i) = \text{graph}_{\,T^n} u_i(V_i)$$

which implies that $u(\xi') = u_i(\xi')$ for $i = 1, 2$, and thus by transitivity $u_1(\xi') = u_2(\xi')$, so $\xi' \in \Sigma$. Since $\xi' \in B_\varrho(\xi)$ was arbitrary we get that Σ is relatively open, and the result follows. $\square$

From this point onward let us associate M^n with its image $F(M^n)$ that is immersed into $\mathbb{R}^{n+1}$, to avoid making any references to $F : M^n \to \mathbb{R}^{n+1}$. Points on (what is now) the submanifold $M^n \subset \mathbb{R}^{n+1}$ will be denoted by x variables.

DEFINITION E.2.2. Let M^n be an immersed submanifold of $\mathbb{R}^{n+1}$. We say that a given $\varrho > 0$ is a representation radius around $x_0 \in M^n$ relative to an affine space $T^n \subset \mathbb{R}^{n+1}$ if there exists a smooth function $u : B^n_\varrho(\xi_0) \to \mathbb{R}$ and an open set $\mathcal{U} \subseteq M^n$ containing x_0 such that:

$$\mathcal{U} = \text{graph}_{\,T^n} u\big|_{B^n_\varrho(\xi_0)}$$

Here ξ_0 denotes the projection of x_0 onto T^n. At this point we define:

$$\mathscr{R}_{x_0,T^n} := \{0\} \cup \{\varrho > 0 \text{ is a representation radius around } x_0 \text{ relative to } T^n\}$$

to be the set of representation radii around our center point. We have included 0 for convenience to ensure that $\mathscr{R}_{x_0,T^n}$ is always non-empty. □

LEMMA E.2.3. Let M^n be an immersed submanifold of $\mathbb{R}^{n+1}$, and let $x_0 \in M^n$ and $T^n \subset \mathbb{R}^{n+1}$ be given. The set $\mathscr{R}_{x_0,T^n}$ of representation radii is of the form:

$$\mathscr{R}_{x_0,T^n} = [0, \varrho^*_{x_0,T^n}]$$

for a (possibly infinite) $\varrho^*_{x_0,T^n} \geq 0$. If the tangent space at x_0 intersects the normal space to T^n trivially, then $\varrho^*_{x_0,T_n} > 0$.

PROOF. The set of radii is of course connected, since any representation radius ϱ with associated graph function u can be shrunk to give any smaller representation radius $0 < \varrho' < \varrho$. In other words, $\mathscr{R}_{x_0,T^n}$ is of the form $[0, \varrho^*)$ for $\varrho^* > 0$, or $[0, \varrho^*]$ for $\varrho^* \geq 0$. We claim that the former form is impossible.

Indeed, suppose for the sake of contradiction that $\mathscr{R}_{x_0,T^n} = [0, \varrho^*)$, $\varrho^* > 0$. Choose a sequence $0 < \varrho_n \nearrow \varrho^*$. By the uniqueness lemma (E.2.1), all the associated graph representation functions u_n must agree on their common domains. Therefore, we may define a joint graph representation function $u : B^n_{\varrho^*} \to \mathbb{R}$ by patching the u_n together. Thus $\varrho^* \in \mathscr{R}_{p_0,T^n}$, a contradiction. It follows that $\mathscr{R}_{x_0,T^n}$ is of the form $[0, \varrho^*]$.

The last claim follows straight from the local existence theorem (E.1.2) which guarantees that there exists a $\varrho > 0$ in $\mathscr{R}_{x_0,T^n}$, so by maximality $\varrho^* \geq \varrho > 0$. □

The following theorem now lets us estimate representation radii from below in cylindrical neighborhoods where M^n has no boundary and where we have uniform curvature, i.e. $|A|^2$, bounds.

THEOREM E.2.4. Let M^n be an immersed submanifold of $\mathbb{R}^{n+1}$, $x_0 \in M^n$, and $T^n \subset \mathbb{R}^{n+1}$ an affine space whose orthogonal complement is spanned by the unit vector η. Suppose that $C_{R,h,T^n}(x_0)$ is a cylinder around x_0 in which M^n has no boundary, i.e.

$$\left(\overline{M^n} \setminus M^n\right) \cap C_{R,h,T^n} = \emptyset$$

Suppose, further, that for some constant $C > 0$ we have the uniform estimate:

$$|A(x)|^2 \leq C^2 \quad \text{for } x \in M^n \cap C_{R,h,T^n}(x_0)$$

and also $|\nu(x_0) \cdot \eta| \geq \frac{1}{K}$ at the center point, for a constant $K < \frac{h}{R}$. As always, ν is a local choice of unit normals around x_0 (which choice we make is immaterial). Then

$$\varrho := \min\left\{R, \frac{h - KR}{2C(R^2 + h^2)^{1/2}}\right\} \in \mathscr{R}_{x_0,T^n}$$

and there exists a smooth $u : B^n_\varrho(\xi_0) \to \mathbb{R}$ such that $|Du| \leq \frac{h}{R}$ and:

$$\mathscr{C}_{x_0}\left(M^n \cap C_{\varrho,h,T^n}(x_0)\right) = \operatorname{graph}_{T^n} u$$

where $\mathscr{C}_{x_0}(\cdot)$ is short for x_0's connected component within the set in question.

PROOF. The fact that $\nu(x_0) \cdot \eta$ is non-zero guarantees that $T_{x_0}M^n$ intersects the orthogonal complement of T^n trivially, since the latter is spanned by η. In other words, the local graph representation theorem applies so (E.2.3) guarantees that $\varrho^* := \varrho^*_{x_0,T^n} > 0$. Denote the maximal representation function around x_0 by $u : B^n_{\varrho^*}(\xi_0) \to \mathbb{R}$, with ξ_0 the projection of x_0 onto T^n as always.

Let us first show that if $\varrho^* < R$ then we have the escape condition:

$$\sup_{B^n_{\varrho^*}(\xi_0)} |Du| \geq \frac{h}{R} \tag{$\dagger_1$}$$

We proceed by contradiction. Suppose that the inequality above were false, i.e. that there exists a $\theta \in (0,1)$ such that $|Du| \le \theta \cdot \frac{h}{R}$ on $B^n_{\varrho^*}(\xi_0)$. Then $u : B^n_{\varrho^*}(\xi_0) \to \mathbb{R}$ has a bounded derivative on its domain and can therefore be extended to a continuous function $v : \overline{B^n_{\varrho^*}}(\xi_0)$ that agrees with u automatically at all interior points.

For convenience, let us denote the boundary points $\overline{B^n_{\varrho^*}}(\xi_0) \setminus B^n_{\varrho^*}(\xi_0)$ of the n-dimensional ball by $\partial B^n_{\varrho^*}(\xi_0)$. Then for all boundary points $\xi \in \partial B^n_{\varrho^*}(\xi_0)$ we have:

$$|v(\xi) - u(\xi_0)| \le \sup_{B^n_{\varrho^*}(\xi_0)} |Du| \quad |\xi - \xi_0| \le \theta \cdot \frac{h}{R} \cdot \varrho^* \le \theta h$$

so that $\xi + v(\xi)\eta \in C_{R,h,T^n}(x_0)$. Since by assumption M^n has no boundary within this cylinder, we conclude:

$$\xi + v(\xi)\eta \in M^n \cap C_{R,h,T^n}(x_0)$$

Since $|Dv|$ is bounded in the interior points $B^n_{\varrho^*}(\xi_0)$, by taking limits we see that the tangent space $T_{\xi+v(\xi)\eta}M^n$ intersects the normal space to T^n (spanned by η) trivially, and hence the local graph representation lemma (E.1.2) applies in combination with the representation radius lemma (E.2.3) and tell us that $\mathscr{R}_{\xi+v(\xi)\eta,T^n}$ is nontrivial.

This process can be repeated for all $\xi \in \partial B^n_{\varrho^*}(\xi_0)$. Since the latter set is compact, there exists a $\delta > 0$ and a finite number of points $\xi_j \in \partial B^n_{\varrho^*}(\xi_0)$ such that the n-balls of radius δ around the ξ_j cover the boundary, and hence in particular:

$$\overline{B^n_{\varrho^*}}(\xi_0) \subseteq B^n_{\varrho^*}(\xi_0) \cup (\cup_j B^n_\delta(\xi_j)) \subseteq B^n_R(\xi_0)$$

By the uniqueness lemma the various graph representatives around the new ξ_j all agree on their common domains, and furthermore they all agree with u in each of their common domains with $B^n_{\varrho^*}(\xi_0)$. We may thus patch them all together to get a graph representative on a ball of radius slightly larger than ϱ^*, which contradicts the latter quantity's maximality. Therefore, the escape condition ($\dagger_1$) holds.

Let us return to the main problem. Our gradient assumption gives us:

$$\frac{1}{\sqrt{1+|Du(\xi_0)|^2}} = |\nu(x_0)\cdot\eta| > K \text{ and thus } |Du(\xi_0)| \le K$$

Recall that $K < \frac{h}{R}$. Make an arbitrary choice of $\theta \in (0,1)$ such that $K < \theta\frac{h}{R}$, and call the latter quantity M. At this point construct:

$$\sigma := \sup\left\{0 < \varrho \leq \varrho^* \ : \ |Du| \leq M \text{ on } B^n_\varrho(\xi_0)\right\}$$

This quantity is well defined since the set in question is bounded and non-empty, as $|Du(\xi_0)| \leq K < M$ by our choice of θ. We now claim that σ must satisfy this interesting *a priori* estimate:

$$\sigma \geq \min\left\{R, \frac{M-K}{2C\left(1+M^2\right)^{1/2}}\right\} \qquad (\dagger_2)$$

Once again we proceed by contradiction and suppose that σ is less than the quantity above. It is now that the curvature $|A|^2$ estimates come into play. Recalling the geometric fact that the second fundamental form can be recovered from the graph function according to:

$$h_{ij} = \frac{D_{ij}u}{\sqrt{1+|Du|^2}}$$

we see that on $B^n_\sigma(\xi_0)$, in combination with our gradient assumption $|Du| \leq M$:

$$|D^2u| \leq |A|\sqrt{1+|Du|^2} \leq C\left(1+M^2\right)^{1/2}$$

By the fundamental theorem of calculus we thus see that on $B^n_\sigma(\xi_0)$:

$$\begin{aligned}
|Du| &\leq |Du(\xi_0)| + \sigma \sup_{B^n_\sigma(\xi_0)} |D^2u| \\
&\leq K + \sigma C\left(1+M^2\right)^{1/2} \\
&\leq K + \frac{M-K}{2C\left(1+M^2\right)^{1/2}} \cdot C\left(1+M^2\right)^{1/2} \\
&\leq K + \frac{M-K}{2} = M \cdot \frac{1+\frac{K}{M}}{2}
\end{aligned}$$

The rightmost fraction is smaller than 1, which implies that $|Du| < M$ on $B^n_\sigma(\xi_0)$. The only way in which this does not contradict the maximality of σ is when we cannot increase σ any further, i.e. when $\sigma = \varrho^*$. But then this would imply that:

$$\sup_{B^n_{\varrho^*}(\xi_0)} |Du| \leq M \cdot \frac{1+\frac{K}{M}}{2} < M = \theta\frac{h}{R} \leq \frac{h}{R}$$

which violates our escape condition, ($\dagger_1$). In any case, then, we reach a contradiction and thus the *a priori* estimate ($\dagger_2$) must hold. Rewriting it in terms of just h, R, K, C, θ, we proved that:

$$\sigma \geq \min\left\{R, \frac{\theta\frac{h}{R} - K}{2C\left(1+\theta^2\frac{h^2}{R^2}\right)^{1/2}}\right\}$$

Recall that $\theta \in (0,1)$ was an arbitrary number such that $K < \theta\frac{h}{R}$. Taking $\theta \nearrow 1$, we may revise our estimate to:

$$\sigma \geq \min\left\{R, \frac{\frac{h}{R} - K}{2C\left(1+\frac{h^2}{R^2}\right)^{1/2}}\right\} = \min\left\{R, \frac{h-KR}{2C\left(R^2+h^2\right)^{1/2}}\right\}$$

In particular, if we let the expression to the right above be denoted by $\varrho > 0$ then we observe that $\varrho^* \geq \sigma \geq \varrho$, so $\varrho \in \mathscr{R}_{x_0,T^n}$ which completes the first part of the theorem. For the second part, recall that by definition of the set of representation radii there must exist an open set $\mathcal{U} \subseteq M^n$ containing x_0 such that:

$$\mathcal{U} = \text{graph}_{\,T^n} u\big|_{B^n_\varrho(\xi_0)}$$

Of course by our choice of σ we have $|Du| \leq \frac{h}{R}$ on $B^n_\varrho(\xi_0)$, and hence by the fundamental theorem of calculus $|u - u(\xi_0)| \leq \varrho|Du| \leq R|Du| \leq h$. In other words, the graph of u within $B^n_\varrho(\xi_0)$ never escapes $C_{\varrho,h,T^n}(x_0)$ so the graph representation above can be refined to:

$$\text{graph}_{\,T^n} u\big|_{B^n_\varrho(\xi_0)} = \mathcal{U} \cap C_{\varrho,h,T^n}(x_0) \subseteq M^n \cap C_{\varrho,h,T^n}(x_0)$$

But graph $_{T^n} u\big|_{B^n_\varrho(\xi_0)}$ is certainly both relatively open and relatively closed in $M^n \cap C_{\varrho,h,T^n}$, and the result follows. $\square$

E.3. Arzelà-Ascoli on Submanifolds

It is time we applied our graph representation theory to prove an Arzelà-Ascoli type result for sequences of immersed n-dimensional submanifolds of $\mathbb{R}^{n+1}$. This theorem is used by our mean curvature flow approach to provide sufficient conditions for regularity. Our generic approach to graph representation theory allows us to proceed without requiring the compactness and embeddedness assumptions we had made for mean curvature flow. Without further ado:

THEOREM E.3.1 (Arzelà-Ascoli for Immersed Submanifolds). Let $\{M_j\}_{j=1}^{\infty}$ denote a sequence of n-dimensional immersed submanifolds in $\mathbb{R}^{n+1}$. Consider a sequence of points $x_j \in M_j$ with $x_j \to x_* \in \mathbb{R}^{n+1}$. Suppose that we are given a radius $\varrho > 0$ within which no M_j has a boundary, i.e.

$$\left(\overline{M_j} \setminus M_j\right) \cap B_\varrho(x_*) = \emptyset \quad \text{for all } j$$

and the following uniform curvature estimate applies:

$$\sup_{x \in M_j \cap B_\varrho(x_*)} |A|^2 \leq \frac{C_0}{\varrho^2} \quad \text{for all } j$$

Then there exists a constant $\sigma(\varrho, C_0) > 0$ and an affine space T^n containing x_* such that, after possibly passing to a subsequence, the submanifolds representing the connected component around x_j of M_j within a given cylinder, or

$$\mathscr{C}_{x_j}\left(M_j \cap C_{\sigma, \frac{2\varrho}{3}, T^n}(x_*)\right)$$

converge in $\mathcal{C}^{1,\alpha}$, $\alpha \in [0,1)$, as graphs to an n-dimensional $\mathcal{C}^{1,1}$ submanifold M_* containing x_* and tangent to T^n at that point. If we additionally have the higher order curvature estimates:

$$\sup_{x \in M_j \cap B_\varrho(x_*)} |\nabla^m A|^2 \leq \frac{C_m}{\varrho^{2(m+1)}} \quad \text{for all } m \geq 0 \text{ and all } j$$

then the convergence is in $\mathcal{C}^\infty$ to a smooth submanifold. The constant $\sigma(\varrho, C_0)$ above is:

$$\sigma(\varrho, C_0) = \varrho \, \min\left\{\frac{1}{6}, \frac{1}{8\sqrt{2}\sqrt{C_0}}\right\}$$

PROOF. Consider the sequence of normal vectors $\{\nu(x_j)\} \subset \mathbb{S}^n$ to our submanifolds M_j. Since $\mathbb{S}^n$ is compact, we may pass to a subsequence and assume that $\nu(x_j) \to \eta_* \in \mathbb{S}^n$. Let T^n denote the affine space, centered at x_*, that is orthogonal to η_*. Denote by $\sigma = \sigma(h, R, K, C)$ the constant from the previous theorem, (E.2.4), where in this case we are using $h = R = \frac{\varrho}{3}$, $K = \frac{1}{2}$ and $C = \frac{\sqrt{C_0}}{\varrho}$. This will amount to being twice the σ postulated in the statement of the theorem.

Since $\nu(x_j) \to \eta_*$ and $x_j \to x_*$, without loss of generality we may assume that:

$$\nu(x_j) \cdot \eta_* > \frac{1}{2} \quad \text{and} \quad |x_j - x_*| < \frac{\varrho}{3} \quad \text{and} \quad |x_j - x_*| < \frac{\sigma}{2} \quad \text{for all } j$$

Let us now work at each submanifold M_j independently. Observe that $C_{\frac{\varrho}{3},\frac{\varrho}{3},T^n}(x_j) \subset B_\varrho(x_0)$, so all three conditions of (E.2.4)–no boundary, curvature estimate, gradient estimate–apply on this cylindrical domain. If ξ_j denotes the projection of x_j onto T^n, then the theorem guarantees that there exists a smooth function $u_j : B^n_\sigma(\xi_j) \to \mathbb{R}$ for which $|Du_j| \leq 1$ and:

$$\mathscr{C}_{x_j}\left(M_j \cap C_{\sigma,\frac{\varrho}{3},T^n}(x_j)\right) = \text{graph}_{\,T^n} u_j\big|_{B^n_\sigma(\xi_j)}$$

Let us denote the vertical component of x_j over T^n by ψ_j, so that $C_{\sigma,\frac{\varrho}{3},T^n}(x_j) = B^n_\sigma(\xi_j) \times B^1_{\frac{\varrho}{3}}(\psi_j)$. Since $|x_j - x_*| < \frac{\varrho}{3}$ and x_* has no vertical component with respect to T^n, the triangle inequality gives $B^1_{\frac{\varrho}{3}}(\psi_j) \subseteq B^1_{\frac{2\varrho}{3}}(0)$. Now:

$$C_{\sigma,\frac{\varrho}{3},T^n}(x_j) \subseteq B^n_\sigma(\xi_j) \times B^1_{\frac{2\varrho}{3}}(0)$$

is but a vertically enlarged cylinder. Since we're only interested in the connected component of x_j, we can accordingly enlarge our cylinder vertically in our graph representation result:

$$\mathscr{C}_{x_j}\left(M_j \cap \left(B^n_\sigma(\xi_j) \times B^1_{\frac{2\varrho}{3}}(0)\right)\right) = \text{graph}_{\,T^n} u_j\big|_{B^n_\sigma(\xi_j)}$$

Since $|x_j - x_*| < \frac{\sigma}{2}$, we also have that $B^n_{\frac{\sigma}{2}}(x_*) \subseteq B^n_\sigma(\xi_j)$ so we can horizontally shrink our cylinder per:

$$B^n_\sigma(\xi_j) \times B^1_{\frac{2\varrho}{3}}(0) \supseteq B^n_{\frac{\sigma}{2}}(x_*) \times B^1_{\frac{2\varrho}{3}}(0) = C_{\frac{\sigma}{2},\frac{2\varrho}{3},T^n}(x_*)$$

Shrinking our cylindrical domain horizontally obviously does not affect our graph representation, so in conclusion we see that:

$$\mathscr{C}_{x_j}\left(M_j \cap C_{\frac{\sigma}{2},\frac{2\varrho}{3},T^n}(x_*)\right) = \text{graph}_{\,T^n}\, u_j\big|_{B^n_{\frac{\sigma}{2}}(x_*)}$$

for every j, with u_j a smooth function on $B^n_{\frac{\sigma}{2}}(x_*)$ satisfying $|Du_j| \le 1$. Our curvature estimate lets us extend this to a second order derivative estimate according to $|D^2 u_j| \le \sqrt{1+|Du_j|^2}|A| \le \frac{\sqrt{2}C_0}{\varrho^2}$. We have constructed, then, a sequence $u_j : B^n_{\frac{\sigma}{2}}(x_*) \to \mathbb{R}$ such that:

$$u_j(\xi_j) \to 0 \quad \text{and} \quad |Du_j| \le 1 \quad \text{and} \quad |D^2 u_j| \le \frac{\sqrt{2}C_0}{\varrho^2} \quad \text{for every } j$$

By the classical Arzelà-Ascoli theorem, and after possibly passing to a subsequence, we see that there exists a $\mathcal{C}^{1,1}$ function $u_* : B^n_{\frac{\sigma}{2}}(x_*) \to \mathbb{R}$ to which our u_j converge in the Hölder spaces $\mathcal{C}^{1,\alpha}$, $\alpha \in [0,1)$. Furthermore we have $u_*(x_*) = \lim_j u_j(\xi_j) = 0$, so this is a graph function passing through x_* and hence represents an n-dimensional submanifold M_* containing x_*. The corresponding constant $\frac{\sigma}{2}$ is, then:

$$\frac{\sigma}{2} = \frac{\sigma\left(\frac{\varrho}{3},\frac{\varrho}{3},\frac{1}{2},\frac{\sqrt{C_0}}{\varrho}\right)}{2} = \frac{1}{2}\min\left\{\frac{\varrho}{3}, \frac{\frac{\varrho}{3}-\frac{1}{2}\frac{\varrho}{3}}{2\frac{\sqrt{C_0}}{\varrho}\left(\frac{\varrho^2}{9}+\frac{\varrho^2}{9}\right)^{1/2}}\right\} = \min\left\{\frac{\varrho}{6}, \frac{\frac{\varrho}{4}}{2\sqrt{2}\sqrt{C_0}}\right\}$$

If we also had uniform bounds on $|\nabla^m A|^2$, then of course they would inductively translate to uniform bounds on $|D^m u_j|$ for every m and j and hence our Arzelà-Ascoli convergence would actually be in $\mathcal{C}^\infty$. □

Bibliography

[AB11] Ben Andrews and Paul Bryan. Curvature bound for curve shortening flow via distance comparison and a direct proof of Grayson's theorem. *Journal für die reine und angewandte Mathematik*, 2011(653), 2011.

[And94] Ben Andrews. Contraction of convex hypersurfaces in Euclidean space. *Calculus of Variations and Partial Differential Equations*, 2(2), 1994.

[Bra78] Kenneth A. Brakke. *The Motion of a Surface by its Mean Curvature*, volume 20 of *Mathematical Notes*. Princeton University Press, University of Tokyo Press, 1978.

[CGG91] Yun-Gang Chen, Yoshikazu Giga, and Shun'ichi Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *The Journal of Differential Geometry*, 33(3), 1991.

[Coh94] Donald L. Cohn. *Measure Theory*. Birkhäuser, 1994.

[DeT83] Dennis M. DeTurck. Deforming metrics in the direction of their Ricci tensors. *The Journal of Differential Geometry*, 18(1), 1983.

[Eck04] Klaus Ecker. *Regularity Theory for Mean Curvature Flow*, volume 57 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, 2004.

[Eck11] Klaus Ecker. Partial regularity at the first singular time for hypersurfaces evolving by mean curvature. arXiv:1104.1212v7, 2011.

[EH91] Klaus Ecker and Gerhard Huisken. Interior estimates for hypersurfaces moving by mean curvature. *Inventiones Mathimaticae*, 105(1), 1991.

[Eid69] Samuil D. Eidel'man. *Parabolic Systems*. American Elsevier Publishing Company, 1969. Translated from Russian by Scripta Technica, London.

[ES91] Lawrence C. Evans and Joel Spruck. Motion of level sets by mean curvature I. *The Journal of Differential Geometry*, 33(3), 1991.

[ES92a] Lawrence C. Evans and Joel Spruck. Motion of level sets by mean curvature II. *Transactions of the American Mathematical Society*, 330(1), 1992.

[ES92b] Lawrence C. Evans and Joel Spruck. Motion of level sets by mean curvature III. *The Journal of Geometric Analysis*, 2(2), 1992.

[Fol95] Gerald B. Folland. *Introduction to Partial Differential Equations*. Princeton University Press, second edition, 1995.

[GH86] Michael E. Gage and Richard S. Hamilton. The heat equation shrinking convex plane curves. *The Journal of Differential Geometry*, 23(1), 1986.

[GP74] Victor Guillemin and Alan Pollack. *Differential Topology*. Prentice Hall, 1974.

[Gra87] Matthew A. Grayson. The heat equation shrinks embedded plane curves to round points. *The Journal of Differential Geometry*, 26(2), 1987.

[Ham82] Richard S. Hamilton. Three-manifolds with positive Ricci curvature. *The Journal of Differential Geometry*, 17(2), 1982.

[Ham95] Richard S. Hamilton. *Isoperimetric estimates for the curve shrinking flow in the plane*, volume 137 of *Annals of Mathematics Studies*, pages 201–222. Princeton University Press, 1995.

[HS11] Xiaoli Han and Jun Sun. An ε-regularity theorem for the mean curvature flow. arXiv:1102.4800v1, 2011.

[Hui84] Gerhard Huisken. Flow by mean curvature of convex surfaces into spheres. *Journal of Differential Geometry*, 20(1), 1984.

[Hui90] Gerhard Huisken. Asymptotic behavior for singularities of the mean curvature flow. *Journal of Differential Geometry*, 31(1), 1990.

[Hui98] Gerhard Huisken. A distance comparison principle for evolving curves. *Asian Journal of Mathematics*, 2(1), 1998.

[Ilm94] Tom Ilmanen. *Elliptic Regularization and Partial Regularity for Motion by Mean Curvature*, volume 108 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 1994.

[Ilm95] Tom Ilmanen. Singularities of mean curvature flow of surfaces. Preprint (available at: http://www.math.ethz.ch/ ilmanen/papers/pub.html), 1995.

[Lan95] Serge Lang. *Differential and Riemannian Manifolds*, volume 160 of *Graduate Texts in Mathematics*. Springer, third edition, 1995.

[Lee97] John M. Lee. *Riemannian Manifolds: An Introduction to Curvature*, volume 176 of *Graduate Texts in Mathematics*. Springer, 1997.

[Lie05] Gary M. Lieberman. *Second Order Parabolic Differential Equations*. World Scientific Publishing Company, 2005.

[LSU68] Olga A. Ladyzhenskaya, Vsevolod A. Solonnikov, and Nina N. Ural'tseva. *Linear and Quasi-Linear Equations of Parabolic Type*, volume 23 of *Translations of Mathematical Monographs*. American Mathematical Society, 1968.

[Sim83] Leon M. Simon. *Lectures on Geometric Measure Theory*, volume 3 of *Proceedings of the Centre for Mathematical Analysis, Australian National University*. Centre for Mathematical Analysis, Australian National University, 1983.

[Tay96a] Michael E. Taylor. *Partial Differential Equations I: Basic Theory*, volume 115 of *Applied Mathematical Sciences*. Springer, 1996.

[Tay96b] Michael E. Taylor. *Partial Differential Equations III: Nonlinear Equations*, volume 117 of *Applied Mathematical Sciences*. Springer, 1996.

[Whi00] Brian White. The size of the singular set in mean curvature flow of mean convex surfaces. *Journal of the American Mathematical Society*, 13(3), 2000.

[Whi03] Brian White. The nature of singularities in mean curvature flow of mean convex surfaces. *Journal of the American Mathematical Society*, 16, 2003.

[Whi05] Brian White. A local regularity theorem for mean curvature flow. *Annals of Mathematics*, 161(3), 2005.

[Zhu02] Xi-Ping Zhu. *Lectures on Mean Curvature Flows*, volume 32 of *AMS/IP Studies in Advanced Mathematics*. American Mathematical Society, International Press, 2002.