Lecture 4: Allen–Cahn. June 28, 2024

(Geometric Measure Theory)

We can now see the previous tools in action: while they were initially used to produce minimal submanifolds in an entirely intrinsic manner (the so-called *Almgren–Pitts theory*), a more recent approach is to use them in conjunction with the "level set viewpoint" that we mentioned at the beginning. We now explore this viewpoint, focusing for simplicity on the case of codimension one ($k = n - 1$), where we wish to obtain minimal hypersurfaces.

The idea is to view a hypersurface Σ^{n-1} as a level set (such as the zero set) of a map $u : M \to \mathbb{R}$. Since we want Σ to be critical for the area, it is natural to look for a map u critical for some energy defined on the set of maps.

We will look at the *Allen–Cahn* energy, which is a model of phase transitions. Namely, assuming from now on that (M^n, g) is closed, we consider

$$E_\varepsilon(u) := \int_M \left[\varepsilon |du|^2 + \frac{W(u)}{\varepsilon}\right],$$

where $W : \mathbb{R} \to [0, \infty)$ is a double-well potential (i.e., a smooth function whose graph is "W-shaped"). A standard choice is

$$W(s) := \frac{(1 - s^2)^2}{4}.$$

Note that this W vanishes precisely at ± 1 and is positive elsewhere; we will always assume that W has these properties. The typical maps $u : M \to \mathbb{R}$ that we have in mind have values in $[-1, 1]$ and M is mostly occupied by two regions A and B with

$$u \approx -1 \text{ on } A, \quad u \approx 1 \text{ on } B,$$

representing the two pure phases of a material. The two regions are separated by an interface of thickness $\simeq \varepsilon$, where the transition between the two phases happens. Thus, the parameter $\varepsilon > 0$ is dimensionally a length.

What we expect is that, in the static case of critical maps u, this interface resembles more and more a hypersurface as $\varepsilon \to 0$, and actually a minimal one, of area directly proportional to $E_\varepsilon(u)$ in the limit. Thus, the energy E_ε should approximate the area functional (for the purposes of the calculus of variations). This was confirmed by works of De Giorgi, Modica, Ilmanen, Hutchinson–Tonegawa, and others. The following is the main result of this theory; with a little abuse of notation, we write (u_ε) to denote any sequence of maps u_{ε_j}, associated to a sequence of parameters $\varepsilon_j \to 0$.

Theorem 4.1. *Given a sequence of maps $u_\varepsilon : M \to \mathbb{R}$, each critical for E_ε and satisfying an upper bound of the form*

$$E_\varepsilon(u) \le C$$

for some uniform $C > 0$ independent of ε, there exists a stationary integral $(n-1)$-varifold V such that the energy densities converge to c_W times the weight $|V|$, as well as $\{u_\varepsilon = 0\} \to \operatorname{spt}(|V|)$ in the Hausdorff topology, up to subsequences.

Here $c_W > 0$ is a constant depending only on the choice of the double-well potential W and the energy densities are the measures μ_ε given by

$$d\mu_\varepsilon := \Big[\varepsilon|du|^2 + \frac{W(u)}{\varepsilon}\Big]\, d\operatorname{vol}_g,$$

where vol_g is the volume measure of the ambient (M^n, g).

From heuristic dimensional analysis, it makes sense that $E_\varepsilon(u_\varepsilon)$ is an $(n-1)$-dimensional area. The way we normalize each term in the definition of E_ε can be seen to be natural also based on a heuristic computation: as above, the typical competitor u is more or less constant on two regions A and B, where both du and $W(u)$ are essentially zero; on the other hand, the interface has volume $\simeq \varepsilon$ and here $|du| \simeq \frac{1}{\varepsilon}$ (since u has a spatial room $\simeq \varepsilon$ to transition from -1 to 1), while $W(u) \simeq 1$, giving

$$E_\varepsilon(u) \simeq \varepsilon\Big[\varepsilon \cdot \frac{1}{\varepsilon^2} + \frac{1}{\varepsilon}\Big] \simeq 1.$$

Looking at E_ε rather than the $(n-1)$-dimensional area has two advantages:

- we are replacing the complicated "space of hypersurfaces" with the much simpler vector space of maps $M \to \mathbb{R}$;
- critical points of E_ε are C^∞-smooth.

The latter holds because, imposing that $\frac{d}{dt}E_\varepsilon(u + t\varphi)\Big|_{t=0} = 0$ for any smooth $\varphi : M \to \mathbb{R}$, we obtain

$$\int_M \Big[2\varepsilon\langle du, d\varphi\rangle + \frac{W'(u)\varphi}{\varepsilon}\Big] = 0,$$

and thus (integrating by parts) $\int_M[-2\varepsilon(\Delta u)\varphi + \varepsilon^{-1}W'(u)\varphi] = 0$; since φ is arbitrary, this gives the simple elliptic PDE

$$2\varepsilon^2\Delta u = W'(u), \quad \text{i.e.,} \quad 2\varepsilon^2\Delta u = (u^2 - 1)u$$

for the standard choice of W. This equation is nonlinear, but under reasonable assumptions (such as $u \in W^{1,2} \cap L^4$, a space that we will use later on) an integral version of the maximum principle immediately gives $|u| \le 1$. Hence, the right-hand side is bounded and by elliptic bootstrap we easily deduce $u \in C^\infty$; actually, we have uniform C^ℓ bounds for any given ℓ once we zoom in at scale ε.

4.1. **Solution in dimension one.** While we will see how the limit varifold V is constructed, we will not explain carefully how its integrality is obtained. We just mention that it comes from a blow-up analysis. Essentially, we expect the interface to resemble an ε-neighborhood of a hypersurface; on each "line" perpendicular to the hypersurface, we expect $u = u_\varepsilon$ to resemble several copies of the solution $U = U_\varepsilon : \mathbb{R} \to \mathbb{R}$ to the equation

$$2\varepsilon^2 U'' = W'(U), \quad \text{with } \lim_{x\to\pm\infty} U(x) = \pm 1,$$

which is just an ODE. Multiplying it by U', we obtain

$$\varepsilon^2[(U')^2]' = [W(U)]'.$$

Taking primitives, we get

$$\varepsilon^2(U')^2 = W(U)$$

(there is no additive constant since $U(x) \to \pm 1$ for $x \to \pm\infty$). This is saying that we have a perfect *equipartition of energy*, even in a pointwise sense: the two terms in the definition of E_ε are equal to each other at every point. We can check that $U' > 0$, so that $\varepsilon U' = \sqrt{W(U)}$ and thus

$$E_\varepsilon(U) = \int_{\mathbb{R}} [\varepsilon(U')^2 + \varepsilon^{-1}W(U)] = \int_{\mathbb{R}} 2\sqrt{W(U)}U' = \int_{-1}^{1} 2\sqrt{W(s)}\, ds.$$

Exercise 4.2. Show that the previous ODE, with boundary conditions $\lim_{x\to\pm\infty} U(x) = \pm 1$, has a unique solution up to translations in x, and that $U' > 0$ everywhere.

In higher dimension, we expect energy to be "quantized" in multiples of the energy of this one-dimensional heteroclinic solution. And indeed, the aforementioned constant is $c_W = \int_{-1}^{1} 2\sqrt{W(s)}\, ds$.

4.2. **Existence of nontrivial critical points.** The previous general result is definitely interesting, but it would be rather disappointing to end up with the trivial limit $V = 0$. Since $|V|$ is the limit of the energy densities, in order to obtain a limit $V \neq 0$ it suffices to show the following.

Proposition 4.3. *Given a compact Riemannian manifold (M^n, g) without boundary, for any $\varepsilon > 0$ small enough there exists a critical point u_ε of E_ε such that*

$$0 < c \leq E_\varepsilon(u_\varepsilon) \leq C,$$

for positive constants $0 < c \leq C$ independent of ε.

It can be shown that if M has positive (Ricci) curvature then no stable minimal hypersurface exists. This suggests to build critical points u_ε which are *unstable* instead.

Example 4.4. In fact, for the round sphere S^n, the construction outlined below will yield the equator S^{n-1} in the limit; this is an unstable critical point of the area since (while "random" perturbations of it will have larger area) shrinking it towards the north pole or the south pole makes the area smaller.

A very general idea to produce unstable critical points is to single out a *mountain pass* situation: given a Banach space X and a C^1 function $E : X \to \mathbb{R}$, we assume that there exists a threshold $\lambda \in \mathbb{R}$ and two points $a, b \in X$ with

$$E(a), E(b) < \lambda.$$

Moreover, we assume that these two points are separated by a "barrier" (like a chain of mountains) of height $\geq \lambda$: more precisely, for any continuous path $\gamma : [0,1] \to X$ with $\gamma(0) = a$ and $\gamma(1) = b$, we assume that there exists an intermediate time $t \in (0,1)$ such that $E(\gamma(t)) \geq \lambda$. Thus each path γ requires a certain "effort," which we can measure as the maximum height reached along it, namely

$$\max_{t \in [0,1]} E(\gamma(t)).$$

We look for an efficient path γ, namely we compute

$$\mu := \inf_{\gamma} \max_{t \in [0,1]} E(\gamma(t)) \geq \lambda,$$

where γ ranges among all continuous paths from a to b.

Our intuition suggests that, if such cheapest path exists, then the highest point along it is a saddle-type critical point. While the existence of this cheapest γ is not really guaranteed, the existence of a critical point at this precise energy level holds in great generality.

Lemma 4.5 (mountain pass lemma). *If $E : X \to \mathbb{R}$ satisfies the Palais–Smale condition (whenever we have a sequence of points (x_j) with $E(x_j)$ converging to some finite limit and $\|dE(x_j)\|_{X^*} \to 0$, there exists a subsequence converging strongly in X), then in the previous situation μ is the energy of a critical point.*

It is not hard to check that E_ε satisfies the Palais–Smale condition on the Banach space $X := W^{1,2}(M) \cap L^4(M)$ for the standard choice of W (we intersect with L^4 because of the growth at infinity of this double-well potential). It now remains to find such a mountain pass situation.

The geometric idea, in the limit $\varepsilon \to 0$, is to perform a *sweepout* of the manifold M. Roughly speaking, this means that we foliate it with (possibly singular) hypersurfaces Σ_t, parametrized for instance by the interval $[0,1]$, in such a way that Σ_t is a single point for $t = 0$

and $t = 1$. Thus, the area of Σ_t at the beginning and at the end is zero, while we expect to have an intermediate time such that Σ_t has substantially large area. For instance, a sweepout of the sphere is given by foliating it with parallels, taking $\Sigma_t := \{x \in S^n : x_{n+1} = 2t - 1\}$.

Clearly, even taking Σ_0 and Σ_1 to be different points, there exist sweepouts such that the area of Σ_t is very small for all times (for instance, Σ_t could be a single point, varying along a curve). To avoid this situation, we look for a topological condition on the sweepout, guaranteeing that it really wraps around M. In codimension one, there is a simple way out: we require that the region enclosed by Σ_t increases from $\emptyset$ to the whole M.

This is even easier to make rigorous at the ε level, where we have maps $u : M \to \mathbb{R}$: we simply look at continuous paths $(u^{(t)})_{t\in[0,1]} \subset X$ such that

$$u^{(0)} \equiv 1, \quad u^{(1)} \equiv -1.$$

Morally, the region $\{u^{(t)} < 0\}$ enclosed by the level set $\{u^{(t)} = 0\}$ is empty at time $t = 0$ and is the full ambient M at time $t = 1$; we expect to see an intermediate time where the two phases 1 and -1 coexist and the interface between them is large.

Proof of Proposition 4.3. To obtain an upper bound, we take a Morse function $f : M \to \mathbb{R}$, rescaled to take values in $[\frac{1}{4}, \frac{3}{4}]$. Moreover, we consider any smooth function $\chi : \mathbb{R} \to \mathbb{R}$ with $\chi(s) = -1$ for $s \le -1$ and $\chi(s) = 1$ for $s \ge 1$, and we let

$$\chi_\varepsilon^{(t)}(s) := \chi\Big(\frac{s-t}{\varepsilon}\Big).$$

In other words, $\chi_\varepsilon^{(t)}$ transitions from -1 to 1 on the interval $[t-\varepsilon, t+\varepsilon]$. We then take

$$u_\varepsilon^{(t)} := \chi_\varepsilon^{(t)} \circ f, \quad \text{for } t \in [0,1].$$

It is not hard to see that this is a continuous path in X with the desired endpoints and that

$$\max_{t\in[0,1]} E_\varepsilon(u_\varepsilon^{(t)}) \le C$$

for a constant C depending only on (M^n, g) (and f), for $\varepsilon > 0$ small enough.

To obtain a lower bound of the form

$$\max_{t\in[0,1]} E_\varepsilon(u^{(t)}) \ge c > 0$$

for *any* continuous path $(u^{(t)})$ with $u^{(0)} \equiv 1$ and $u^{(1)} \equiv -1$, we observe that

$$E_\varepsilon(u) \ge \int_M 2\sqrt{W(u)}|du| = \int_M |d(F(u))|$$

by Cauchy–Schwarz, where F is the primitive of $2\sqrt{W}$ satisfying $F(0) = 0$. Moreover,

$$\int_M F(u^{(0)}) = \mathcal{H}^n(M) \cdot F(1) > 0, \quad \int_M F(u^{(1)}) = \mathcal{H}^n(M) \cdot F(-1) < 0.$$

By continuity, there exists an intermediate time $t \in (0,1)$ such that

$$\int_M F(u^{(t)}) = 0.$$

Thus, at this specific t, Poincaré's inequality gives

$$E_\varepsilon(u^{(t)}) \geq \int_M |d(F(u^{(t)}))| \geq c' \int_M |F(u^{(t)})|,$$

where $c' > 0$ is a constant depending only on (M^n, g). Finally, since we can assume without loss of generality that $E_\varepsilon(u^{(t)}) \leq 1$ (as otherwise the desired lower bound is obvious), we have $\int_M W(u^{(t)}) \leq \varepsilon$. This immediately gives that $W(u^{(t)}) \leq \frac{2\varepsilon}{\mathcal{H}^n(M)}$ on at least half of M (in measure), giving $|u^{(t)}| \geq 1 - \delta_\varepsilon$ here (with an error $\delta_\varepsilon \to 0$ as $\varepsilon \to 0$). This gives a uniform lower bound on $\int_M |F(u^{(t)})|$, and hence on $E_\varepsilon(u^{(t)})$, as we wanted.

Finally, the mountain pass lemma (applied with $\lambda := c$) gives us a critical point $u_\varepsilon \in X$ with energy $E_\varepsilon(u_\varepsilon) \in [c, C]$. $\square$

4.3. **Stress-energy tensor and the limit varifold.** For each $\varepsilon > 0$ sufficiently small, we now have a critical point $u_\varepsilon : M \to \mathbb{R}$ for E_ε, with uniform bounds $c \leq E_\varepsilon(u_\varepsilon) \leq C$.

In order to extract a limiting (possibly singular) minimal hypersurface without "mass cancellation," we associate to each critical point u_ε a sort of diffuse varifold V_ε, with the goal of taking the limit of V_ε along a subsequence.

To motivate what follows, we observe that a k-varifold V in $\mathbb{R}^n$ whose disintegration is a Dirac mass δ_{P_x} centered at a k-plane P_x, at a.e. point $x \in \mathbb{R}^n$ (i.e., giving importance to just one plane for a.e. x), can be seen as a Radon measure with values into the set of matrices

$$A_k^n := \{S \in \mathbb{R}^{n\times n} \text{ symmetric s.t. } S^2 = S,\ \operatorname{tr} S = k\}.$$

Indeed, each symmetric matrix in this set gives the orthogonal projection onto a k-plane, and this gives a bijective correspondence between this set and the set of k-planes Gr_k^n. In other words, we identify the varifold V (with disintegration $dV(x,\cdot) = d|V|(x) \otimes \delta_{P_x}(\cdot)$) with the matrix-valued measure $\pi_{P_x}\, d|V|(x)$.

Note that, under this identification, the weight of the varifold is uniquely determined by V: it is just $\frac{\operatorname{tr}(V)}{k}$. Rectifiable varifolds are examples of such "univalued" varifolds.

Exercise 4.6. For a general varifold V on $\mathbb{R}^n$, check that stationarity is equivalent to

$$\int_{\mathbb{R}^n \times A_k^n} \langle \pi_P, \nabla X\rangle \, dV(x,P) = 0$$

for every vector field $X \in C^1_c$, where we use the usual Hilbert–Schmidt scalar product between matrices. For "univalued" k-varifolds V, show that V is stationary if and only if

$$\operatorname{div} V = 0$$

in the distributional sense, where we view V as a matrix-valued measure as above and we take the divergence of each row.

Remark 4.7. The same holds on a Riemannian manifold, using a measure with values into symmetric endomorphisms $S : T_xM \to T_xM$ (at each point $x \in M$) such that $S^2 = S$ and $\operatorname{tr} S = k$.

We have such a matrix-valued measure for each $\varepsilon > 0$: it is $T_\varepsilon \, d\operatorname{vol}_g$, with the so-called *stress-energy tensor*

$$T_\varepsilon := e_\varepsilon I - 2\varepsilon du_\varepsilon \otimes du_\varepsilon,$$

where $e_\varepsilon := \varepsilon |du_\varepsilon|^2 + \frac{W(u_\varepsilon)}{\varepsilon}$. It has zero divergence: exactly as stationarity of varifolds, this is seen by taking a smooth vector field X and by exploiting criticality of u_ε with respect to variations of the form

$$t \mapsto u_\varepsilon \circ \Phi^X_t,$$

called *inner variations*, where Φ^X_t is the flow of X (essentially, this moves the "interface" to its image under Φ^X_{-t}, since this is what happens to each level set).

Exercise 4.8. Prove that $\operatorname{div} T_\varepsilon = 0$.

While T_ε is certainly symmetric, it is not even clear if it is positive semidefinite. We thus need to enlarge the previous set of matrices. We introduce the convex set

$$\tilde{A}^n_k := \{S \in \mathbb{R}^{n\times n} \text{ symmetric s.t. } -nI \le S \le I,\ \operatorname{tr} S \ge k\}.$$

Given a positive Radon measure μ, a *generalized k-dimensional varifold* with weight μ is a matrix-valued measure of the form

$$dV = S_x \, d\mu(x), \quad \text{with } S_x \in \tilde{A}^n_k.$$

Note that the weight must be specified, since it is no longer uniquely determined by V.

Remark 4.9. This is the "univalued" version of a more general notion, according to which a generalized varifold is a Radon measure on the bundle $\mathbb{R}^n \times \tilde{A}^n_k$ (and similarly on manifolds). However, we stick to this special univalued case, since the next rectifiability criterion would anyway require to replace such a varifold with its "collapsed" version (where, in the disintegration, we replace the probability measure above each point x with its center of mass, thus collapsing it to a single "generalized" plane, an element of $\tilde{A}^n_k$). Note that stationarity in the sense that $\int_{\mathbb{R}^n \times \tilde{A}^n_k} \langle S, \nabla X \rangle \, dV(x, S)$ is equivalent to $\operatorname{div} V' = 0$ for the "collapsed" varifold V'; in other words, stationarity sees only the center of mass of the disintegration.

In this setting, we have the following important result by Ambrosio–Soner.

Theorem 4.10 (Ambrosio–Soner rectifiability theorem). *Given a generalized k-dimensional varifold V with weight μ, assume that it is stationary (i.e., $\operatorname{div} V = 0$) or has locally bounded first variation (i.e., $\operatorname{div} V$ is itself a Radon measure) and that $\Theta^*_k(\mu, x) > 0$ for μ-a.e. x. Then V is (the matrix-valued measure given by) a rectifiable varifold. In particular, μ is k-rectifiable and, writing $dV = S_x \, d\mu$, the matrix S_x is the orthogonal projection onto the tangent plane to μ at x, for μ-a.e. x.*

We skip the proof, which involves checking that $S_x \in A^n_k$ a.e. and then relies on Allard's rectifiability theorem.

Remark 4.11. Conversely, Allard's rectifiability theorem follows from this result: given a varifold V, with disintegration $dV(x, P) = d|V|(x) \otimes \alpha_x(P)$, we consider the "collapsed" varifold V' as in the previous remark, obtained by taking the center of mass $S_x := \int_{G^n_k} \pi_P \, d\alpha_x(P)$ at each point. Note that $S_x \in \tilde{A}^n_k$ since $\pi_P \in A^n_k \subseteq \tilde{A}^n_k$ and $\tilde{A}^n_k$ is convex, and hence V', given by $dV'(x) = S_x \, d|V|(x)$, is a generalized varifold. Since V is stationary or has locally bounded first variation, the same holds for V'. Hence, by the Ambrosio–Soner criterion, $|V|$ is k-rectifiable and $S_x \in A^n_k$ for $|V|$-a.e. x. This implies that α_x was in fact a Dirac mass, since otherwise it is easy to check that its center of mass S_x would have kernel of dimension $< n - k$. This proves that $\alpha_x = \pi_{T_x|V|}$, so that V is a rectifiable varifold.

These results and observations also hold on a manifold (M^n, g). In our setting, we define V_ε as follows: the weight is the energy density, i.e.,

$$d\mu_\varepsilon := e_\varepsilon \, d\operatorname{vol}_g,$$

while $dV_\varepsilon := T_\varepsilon \, d\operatorname{vol}_g$ is our generalized varifold. Thus, we have $dV_\varepsilon = S_\varepsilon \, d\mu_\varepsilon$ with

$$S_\varepsilon := I - \frac{2\varepsilon du_\varepsilon \otimes du_\varepsilon}{e_\varepsilon}.$$

Remark 4.12. The fact that $S_\varepsilon \in \tilde{A}^n_k$, and in particular the lower bound on its trace, requires the fundamental inequality proved at the end of these notes. Strictly speaking, for general metrics g this lower bound holds only at the limit $\varepsilon \to 0$ (since Modica's inequality contains an error term if we do not assume $\operatorname{Ric} \geq 0$), and thus we do have a generalized varifold only in the limit.

Since V_ε is just a measure taking values in a closed convex set, we can extract a limit generalized varifold V_0 along a suitable subsequence $\varepsilon_j \to 0$. The weight μ_0 of V_0 is taken to be the limit of the weights μ_ε. In particular, μ_0 is not trivial. We want to apply this rectifiability criterion to conclude that V_0 is a rectifiable varifold. In fact, the ultimate result that one can prove is the following.

Theorem 4.13 (Hutchinson–Tonegawa structure theorem). *Given a sequence of critical points u_{ε_j} of E_{ε_j}, with $\varepsilon_j \to 0$ and*

$$\sup_j E_{\varepsilon_j}(u_j) < \infty,$$

the limit generalized varifold V_0 built as above (along a subsequence) is c_W times an integral stationary varifold. Moreover, $\operatorname{spt}|V_0|$ is the limit of the level sets $u_{\varepsilon_j}^{-1}(\lambda)$ for any fixed $\lambda \in (-1, 1)$ (along the same subsequence).

Remark 4.14. We could have also taken a limit in the sense of currents. The 1-form $2\sqrt{W(u_\varepsilon)}\, du_\varepsilon$ is closed, and hence (if M is oriented) we can associate to it an $(n-1)$-current Γ_ε with no boundary, given by

$$\langle \Gamma_\varepsilon, \omega \rangle := \int_M 2\sqrt{W(u_\varepsilon)}\, du_\varepsilon \wedge \omega.$$

It has mass $\mathbb{M}(\Gamma_\varepsilon) \le E_\varepsilon(u_\varepsilon)$ by Cauchy–Schwarz (with equality in dimension one; it can be shown that this is asymptotically an equality also when $n > 1$). Hence, we can extract a limit normal current Γ_0. Since its weight $|\Gamma_0| \le \mu_0$ and, as we will show, μ_0 is rectifiable, also Γ_0 is rectifiable. Actually, it can be shown to have density c_W a.e. (these facts can be established in a quicker way using BV functions and showing that $\frac{\Gamma_0}{c_W}$ bounds a finite perimeter set). The drawback is that we could have $\Gamma_0 = 0$ a priori, due to some cancellation. However, these currents can be used to show the so-called Γ-*convergence* of E_ε to the $(n-1)$-dimensional area.

4.4. **Monotonicity and consequences.** In order to apply the previous rectifiability criterion to the limiting generalized varifold V_0, we need to check that μ_0 has positive (upper) density at μ_0-a.e. point. The next result essentially says that if the energy on a ball $B_r(p)$ is much smaller than r^{n-1} (which lower bounds the area of a minimal hypersurface $\Sigma \ni p$ in the same ball, by the monotonicity formula for minimal submanifolds), then no interface is forming near p. We write $u = u_\varepsilon$ for simplicity.

Proposition 4.15 (clearing-out). *Given $\eta > 0$, there exists a constant $\delta > 0$, depending only on η, W, and (M^n, g), such that if $E_\varepsilon(u, B_r(p)) \le \delta r^{n-1}$ and $\varepsilon \le r/2 \le \delta$ then $|u| \in (1-\eta, 1+\eta)$ on $B_{r/2}(p)$.*

In particular, assuming that $W''(\pm 1) > 0$, on the smaller ball $B_{r/2}(p)$ we compute that

$$\Delta(W(u)) = W''(u)|\nabla u|^2 + \frac{W'(u)^2}{2\varepsilon^2} \geq \frac{W'(u)^2}{2\varepsilon^2} \geq \frac{c}{\varepsilon^2} W(u)$$

for some constant $c > 0$, since $|W'(u)| \simeq |u - (\pm 1)| \simeq \sqrt{W(u)}$ for $u \approx \pm 1$. In the sequel, we assume that (M^n, g) is locally Euclidean for simplicity, just to avoid dealing with error terms. Calling $f(\sigma)$ the spherical average of $W(u)$ on the sphere $\partial B_\sigma(p)$, we get

$$\begin{aligned} f'(\sigma) &= \frac{1}{\mathcal{H}^{n-1}(\partial B_\sigma(p))} \int_{\partial B_\sigma(p)} \partial_\nu (W(u)) \\ &= \frac{1}{\mathcal{H}^{n-1}(\partial B_\sigma(p))} \int_{B_\sigma(p)} \Delta(W(u)) \\ &\geq \frac{c}{\varepsilon^2 \sigma^{n-1}} \int_{B_\sigma(p)} W(u) \end{aligned}$$

by the divergence theorem (for a different $c > 0$), for all $\sigma < \frac{r}{2}$.

By the mean value theorem, since

$$\int_{r/4}^{r/2} \rho^{n-1} f(\rho)\, d\rho \leq C \int_0^{r/2} \int_{\partial B_\sigma(p)} W(u)\, d\sigma \leq C \int_{B_{r/2}(p)} W(u) \leq C\varepsilon E_\varepsilon(u, B_{r/2}(p)),$$

we can find an intermediate radius $\rho \in (\frac{r}{4}, \frac{r}{2})$ such that

$$f(\rho) \leq \frac{4^n C\varepsilon}{r^n} E_\varepsilon(u, B_{r/2}(p)).$$

Since $\int_{r/8}^{\rho} f'(\sigma)\, d\sigma \leq f(\rho)$ and $\rho - \frac{r}{4} \geq \frac{r}{8}$, we can find another radius $\sigma \in (\frac{r}{8}, \sigma)$ such that

$$f'(\sigma) \leq \frac{8}{r} f(\rho) \leq \frac{C\varepsilon}{r^{n+1}} E_\varepsilon(u, B_{r/2}(p)),$$

for a different $C > 0$. Thanks to the previous bound, we arrive at

$$\frac{c}{\varepsilon^2 \sigma^{n-1}} \int_{B_\sigma(p)} W(u) \leq \frac{C\varepsilon}{r^{n+1}} E_\varepsilon(u, B_{r/2}(p)).$$

Since σ is comparable with r, this gives

$$\int_{B_{r/8}(p)} W(u) \leq \int_{B_\sigma(p)} W(u) \leq \frac{C\varepsilon^3}{r^2} E_\varepsilon(u, B_r(p)).$$

As a consequence, we arrive at the following result.

Theorem 4.16 (rectifiability of the limit varifold). *We have $\Theta_{n-1}(\mu_0, p) > 0$ at all points $p \in \operatorname{spt}(\mu_0)$. In particular, the previous rectifiability criterion applies: μ_0 is an $(n-1)$-rectifiable measure.*

We write $\Theta_{n-1}(\mu_0, p)$ since the density exists at every p, as a consequence of a monotonicity formula proved later on.

Proof. We assume for simplicity that the ambient (M^n, g) is flat. In fact, the proof will show that $\Theta_{n-1}(\mu_0, p) \geq c$ for some threshold $c > 0$ independent of $p \in \operatorname{spt}(\mu_0)$. Assume that $\Theta_{n-1}(\mu_0, p)$ is very small. Then there exists a (small) radius $r > 0$ such that $\mu_0(\bar{B}_r(p))$ is much smaller than r^{n-1}, implying that the same eventually holds for $\mu_\varepsilon(B_r(p)) = E_\varepsilon(u_\varepsilon, B_r(p))$, for ε along our subsequence.

Writing $u = u_\varepsilon$, we can then apply clearing-out, and thus the integral bound on $W(u)$ that we deduced from it, as well as Modica's inequality $\varepsilon|du|^2 \leq \frac{W(u)}{\varepsilon}$ from the next section, to conclude that

$$E_\varepsilon(u, B_{r/8}(p)) \leq \int_{B_{r/8}(p)} \frac{2W(u)}{\varepsilon} \leq \frac{C\varepsilon^2}{r^2} E_\varepsilon(u, B_r(p)).$$

Since $r > 0$ is fixed and $\limsup_{\varepsilon\to 0} E_\varepsilon(u, B_r(p)) \leq \mu_0(\bar{B}_r(p))$ is finite, the right-hand side converges to zero. Hence,

$$\mu_0(B_{r/8}(p)) \leq \liminf_{\varepsilon\to 0} E_\varepsilon(u, B_{r/8}(p)) = 0.$$

This shows that $\mu_0(B_{r/8}(p)) = 0$, contradicting the assumption that $p \in \operatorname{spt}(\mu_0)$. □

The previous clearing-out result is in turn a simple consequence of a *monotonicity formula*, which mimics the one for minimal hypersurfaces (recall that the energy density is expected to concentrate along a minimal hypersurface).

Proposition 4.17 (monotonicity formula). *If (M^n, g) is flat (i.e., locally Euclidean), then*

$$r \mapsto \frac{E_\varepsilon(u, B_r(p))}{r^{n-1}}$$

is increasing for any given center $p \in M$, as r varies in the interval $(0, \operatorname{inj}(M))$. For general closed Riemannian manifolds, the slightly perturbed function $e^{Cr}\frac{E_\varepsilon(u,B_r(p))}{r^{n-1}}$ is increasing for ε, r small enough (where C depends only on (M^n, g)).

Remark 4.18. Since $e^{Cr} \to 1$ as $r \to 0$, the perturbing factor e^{Cr} usually does not affect the arguments exploiting the monotonicity formula.

Proof of Proposition 4.15 assuming Proposition 4.17. We sketch a proof in the flat case, for simplicity. Since $E_\varepsilon(u, B_r(p)) \leq \delta r^{n-1}$ and $\varepsilon \leq r/2$, by monotonicity we know that

$$E_\varepsilon(u, B_\varepsilon(p)) \leq \delta\varepsilon^{n-1}.$$

As computed in the next section, we have $|u| < 1$ and, as a consequence, $\Delta|du|^2 \geq -C\varepsilon^{-2}|du|^2$. Using standard techniques such as Moser iteration, it is not hard to conclude that we have a pointwise bound of the form

$$|du|^2 \leq C\varepsilon^{-2}$$

on a smaller ball $B_{cr}(p)$ (this is best seen by rescaling $B_\varepsilon(p)$ to the unit ball, so that after rescaling we have a function $\tilde{u}$ with the energy bound $\int_{B_1} |d\tilde{u}|^2 \leq \delta \leq 1$ and $\Delta|d\tilde{u}|^2 \geq -C|d\tilde{u}|^2$).

In particular, since W is Lipschitz on $[-1,1]$, we also have $|d(W(u))| \leq C\varepsilon^{-1}$. Now if $|u| \notin (1-\eta, 1+\eta)$ at p, so that here $W(u) \geq \tilde{\eta}$ for some fixed $\tilde{\eta} > 0$, by this Lipschitz bound we would have $W(u) \geq \frac{\tilde{\eta}}{2}$ on a smaller ball $B_{c\tilde{\eta}\varepsilon}(p)$, for a different $c > 0$. But then

$$E_\varepsilon(u, B_\varepsilon(p)) \geq \int_{B_{c\tilde{\eta}\varepsilon}(p)} \frac{W(u)}{\varepsilon} \geq \frac{\tilde{\eta}}{2\varepsilon}\mathcal{H}^n(B_{c\tilde{\eta}\varepsilon}(p)) = \frac{\tilde{\eta} \cdot \omega_n (c\tilde{\eta})^n}{2}\varepsilon^{n-1}.$$

This contradicts the fact that $E_\varepsilon(u, B_\varepsilon(p)) \leq \delta\varepsilon^{n-1}$ once we choose δ small enough. We proved the claim at the center p, but it is easy to see that a bound $E_\varepsilon(u, B_\varepsilon(q)) \leq 2^{n-1}\delta\varepsilon^{n-1}$ also holds for nearby center points $q \in B_{r/2}(p)$, so that the conclusion also holds at these points. $\square$

Recalling how we obtained the monotonicity formula for stationary varifolds, we can try to get it in the present setting by exploiting our generalized varifold V_ε. Assuming for simplicity that (M^n, g) is flat, mimicking the computation used for stationary varifolds we get the following:

$$\int_{B_r(p)} (ne_\varepsilon - 2\varepsilon|du|^2) = r\int_{\partial B_r(p)} (e_\varepsilon - 2\varepsilon|\partial_\nu u|^2) \leq r\int_{\partial B_r(p)} e_\varepsilon.$$

Exercise 4.19. Show the previous formula.

This formula immediately implies the fact that $\frac{E_\varepsilon(u, B_r(p))}{r^{n-2}}$ is increasing, since $ne_\varepsilon - 2\varepsilon|du|^2 \geq (n-2)e_\varepsilon$ and thus the formula implies that

$$(n-2)\int_{B_r(p)} e_\varepsilon \leq r\int_{\partial B_r(p)} e_\varepsilon,$$

which rearranges to

$$\frac{d}{dr}\frac{\int_{B_r(p)} e_\varepsilon}{r^{n-2}} \geq 0.$$

However, this is not enough for us. This unnatural power in the denominator comes from the Dirichlet term $|du|^2$ in the energy.

Remark 4.20. In fact, in other variational settings such as the study of harmonic maps between manifolds, where the energy contains just this term, this becomes the natural power: for instance, $x \mapsto \frac{x}{|x|}$ is a harmonic map from $\mathbb{R}^3$ to S^2 and its Dirichlet energy on $B_r(0)$ is a multiple of $r = r^{n-2}$.

However, quite miraculously, in our setting we are able to show that there is a tradeoff between the two terms appearing in the energy E_ε: in the next section we will prove the bound

$$\varepsilon|du|^2 \le \frac{W(u)}{\varepsilon},$$

due to Modica. This bound allows to improve the exponent $n-2$ to the expected one $n-1$.

Exercise 4.21. Show that this bound implies Proposition 4.17 in the flat case.

The previous results imply that the limit generalized varifold V_0 satisfies the density assumption $\Theta_{n-1}(\mu_0,\cdot) > 0$ on $\operatorname{spt}(\mu_0)$ in the statement of the Ambrosio–Soner rectifiability criterion. Moreover, it is stationary since it is a limit of stationary generalized varifolds. Hence, we conclude that it is a genuine rectifiable varifold.

Remark 4.22. To conclude the discussion, let us hint at how to obtain integrality of V_0 (up to a multiplicative constant $c_W > 0$). Since it is rectifiable, at a.e. point x its blow-up is an $(n-1)$-plane P with constant multiplicity θ. In order to show that $\theta \in c_W \cdot \mathbb{N}$, by rescaling each u_ε (thus slightly changing the parameter ε in the PDE) and performing a diagonal argument, we can assume that this plane with multiplicity θ is the limit varifold. Since θ is constant, the multiplicity formula for E_ε is almost saturated if we center it at any point $x \in P$. Since in its derivation we threw away the positive term $2\varepsilon|\partial_\nu u_\varepsilon|^2$ on spherical shells, it is easy to conclude that $\varepsilon|\partial_v u_\varepsilon|^2 \to 0$ in L^1_{loc} for any direction $v \in P$. Essentially, this says that u_ε is "almost one-dimensional," since its partial derivatives along P are "small." We then look at generic lines ℓ perpendicular to P and, on each line ℓ, we check that u_ε resembles several copies of the standard heteroclinic solution (scaled by ε), separated by long intervals where $|u_\varepsilon| \approx 1$, where no energy concentrates. Thus, the energy along each line ℓ is close to an integer multiple of c_W. The fact that this multiple is roughly the same as ℓ varies relies on Allard's strong constancy lemma mentioned in the previous lecture.

4.5. **Two proofs of Modica's bound.** We will now show two possible ways of proving the fundamental inequality

$$\varepsilon|du|^2 \le \frac{W(u)}{\varepsilon},$$

assuming that $\mathrm{Ric} \ge 0$, i.e., (M^n, g) has nonnegative Ricci curvature (in general, with some additional work, one can obtain a slightly weaker bound of the form $\varepsilon|du|^2 \le \frac{W(u)}{\varepsilon} + C$, with C independent of $\varepsilon \in (0,1)$; this is enough to obtain an almost-monotonicity and carry out the preceding analysis, again with little extra work).

First proof. We start with Modica's original argument, assuming actually $\mathrm{Ric} > 0$. Let

$$\xi := \varepsilon |du|^2 - \frac{W(u)}{\varepsilon},$$

the so-called *discrepancy*, and assume by contradiction that it attains a positive maximum at a point $p \in M$.

In the Euclidean space we would have $\Delta |du|^2 = 2\langle d\Delta u, du\rangle + 2|\operatorname{Hess} u|^2$ for any smooth function u. On a Riemannian manifold, there is a correcting term involving the Ricci curvature: specifically, Bochner's formula tells us that

$$\Delta |du|^2 = 2\langle d\Delta u, du\rangle + 2|\operatorname{Hess} u|^2 + 2\,\mathrm{Ric}(\nabla u, \nabla u).$$

In our setting, since $2\Delta u = \varepsilon^{-2} W'(u)$, we obtain

$$\Delta |du|^2 = \frac{W''(u)}{\varepsilon^2} |\nabla u|^2 + 2|\operatorname{Hess} u|^2 + 2\,\mathrm{Ric}(\nabla u, \nabla u).$$

As for $W(u)$, we have

$$\Delta W(u) = \operatorname{div}(W'(u)\nabla u) = W''(u)|\nabla u|^2 + W'(u)\Delta u = W''(u)|\nabla u|^2 + \frac{W'(u)^2}{2\varepsilon^2}.$$

Hence, when we compute $\Delta \xi$, the terms involving $W''(u)$ cancel, giving

$$\Delta \xi = 2\varepsilon |\operatorname{Hess} u|^2 + 2\varepsilon\,\mathrm{Ric}(\nabla u, \nabla u) - \frac{W'(u)^2}{2\varepsilon^3}.$$

Since p is a maximum point for ξ, we must have $\Delta \xi(p) \leq 0$. Also, since we are assuming that $\xi(p) > 0$, we have $\varepsilon |\nabla u|^2(p) = \varepsilon |du|^2(p) > \frac{W(u)}{\varepsilon}(p) \geq 0$, and hence $\nabla u(p) \neq 0$, so that $\mathrm{Ric}(\nabla u, \nabla u) > 0$ at p. Thus, we arrive at

$$0 > 2\varepsilon |\operatorname{Hess} u|^2 - \frac{W'(u)^2}{2\varepsilon^3} \quad \text{at } p,$$

or equivalently

$$|W'(u)| > 2\varepsilon^2 |\operatorname{Hess} u| \quad \text{at } p.$$

Moreover, since p is a maximum point for ξ, we also have $d\xi(p) = 0$. We compute that this is the same as

$$2\varepsilon \operatorname{Hess} u[\nabla u, \cdot] = \frac{W'(u)}{\varepsilon}\, du \quad \text{at } p.$$

Taking absolute values, we obtain

$$|W'(u)||\nabla u| = 2\varepsilon^2 |\operatorname{Hess} u[\nabla u, \cdot]| \leq 2\varepsilon^2 |\operatorname{Hess} u||\nabla u| \quad \text{at } p.$$

Canceling $|\nabla u|(p) > 0$, we arrive at the opposite inequality $|W'(u)| \leq 2\varepsilon^2 |\operatorname{Hess} u|$, a contradiction. $\square$

We now present an alternative proof, which works assuming that $\sqrt{W}$ is a smooth, strictly concave function on $(-1,1)$ (in the sense that the second derivative of $\sqrt{W}$ is negative). Note that this holds for the standard choice $W(s) := \frac{(1-s^2)^2}{4}$. This proof has the minor advantage of working immediately also in the case $\mathrm{Ric} \geq 0$ and, most importantly, it avoids exploiting the first order optimality at a maximum point of the discrepancy, making it conceptually simpler. This is the route followed to achieve an analogous bound in codimension two, involving critical points of the abelian Yang–Mills–Higgs energy (in the work by Pigati–Stern).

Exercise 4.23. Check that this assumption is equivalent to $W'' < \frac{(W')^2}{2W}$ on $(-1,1)$.

We also assume that $W'(s) > 0$ for $s > 1$ and $W'(s) < 0$ for $s < -1$, which guarantees that a solution u satisfies $|u| \leq 1$ at each point (indeed, at the maximum point p for u we have $0 \geq 2\varepsilon^2 \Delta u = W'(u)$, which is a contradiction if $u(p) > 1$; a similar argument shows that $u \geq -1$). Actually, the maximum principle implies that $|u| < 1$ for a nonconstant solution: indeed, the function $v := 1 - u$ satisfies

$$\Delta v = -\frac{W'(u)}{2\varepsilon^2} = cv,$$

where $c := h(u)$ and $h(s)$ is a smooth function extending $-\frac{W'(s)}{2\varepsilon^2(1-s)}$ (note that $W'(1) = 0$). Since $v \geq 0$, by the maximum principle we have either $v > 0$ (everywhere) or $v \equiv 0$ (the same proof shows that, in general, two solutions $u \leq \tilde{u}$ to a second-order nonlinear elliptic PDE of the form $\Delta u = f(u)$ cannot touch, unless they coincide; in this case we are taking $\tilde{u} := 1$). This shows that $u < 1$ and similarly we can check that $u > -1$.

Second proof. This time we look at the modified discrepancy

$$\xi := \sqrt{\varepsilon}|du| - \frac{\sqrt{W(u)}}{\sqrt{\varepsilon}}.$$

Again, the goal is to show that $\xi \leq 0$, so we assume by contradiction that $\xi(p) > 0$ at a maximum point p. Using Bochner's formula and Cauchy–Schwarz, it is easy to check that

$$\Delta|du| \geq \left\langle d\Delta u, \frac{du}{|du|} \right\rangle = \frac{W''(u)}{2\varepsilon^2}|du|$$

on the region $\{|du| > 0\}$ (in particular at p), assuming $\mathrm{Ric} \geq 0$. Moreover, for $\sqrt{W(u)}$ we compute

$$\Delta\sqrt{W(u)} = \operatorname{div}\Big(\frac{W'}{2\sqrt{W}}\nabla u\Big) = \Big(\frac{W''}{2\sqrt{W}} - \frac{(W')^2}{4\sqrt{W^3}}\Big)|\nabla u|^2 + \frac{(W')^2}{4\varepsilon^2\sqrt{W}},$$

writing W in place of $W(u)$ for simplicity (and similarly for W', W'').

We thus obtain

$$\begin{aligned}\sqrt{\varepsilon}\Delta\xi &\geq \frac{W''}{2\varepsilon}|du| - \frac{(W')^2}{4\varepsilon^2\sqrt{W}} - \Big(\frac{W''}{2\sqrt{W}} - \frac{(W')^2}{4\sqrt{W^3}}\Big)|\nabla u|^2 \\ &= \frac{(W')^2}{4\sqrt{W^3}}\Big(|\nabla u|^2 - \frac{W}{\varepsilon^2}\Big) - \frac{W''}{2\sqrt{W}}|\nabla u|\Big(|\nabla u| - \frac{\sqrt{W}}{\varepsilon}\Big).\end{aligned}$$

If the two expressions between parentheses were the same multiple of ξ, we would immediately obtain an inequality of the form $\Delta\xi \geq c\xi$, with $c >$ (thanks to the assumption that $\sqrt{W}$ is concave), concluding the proof. However, they are slightly different: the first one is in fact a multiple of the first definition of discrepancy given above.

We thus proceed in a slightly more careful way, working at the point p from now on. Since $|\nabla u| - \frac{\sqrt{W}}{\varepsilon} > 0$ and $W'' < \frac{(W')^2}{2W}$, we have

$$\sqrt{\varepsilon}\Delta\xi > \frac{(W')^2}{4\sqrt{W^3}}\Big(|\nabla u|^2 - \frac{W}{\varepsilon^2}\Big) - \frac{(W')^2}{4\sqrt{W^3}}|\nabla u|\Big(|\nabla u| - \frac{\sqrt{W}}{\varepsilon}\Big)$$

at p. Now two terms cancel, leaving us with

$$\sqrt{\varepsilon}\Delta\xi > \frac{(W')^2}{4\sqrt{W^3}}\cdot\frac{\sqrt{W}}{\varepsilon}\Big(|\nabla u| - \frac{\sqrt{W}}{\varepsilon}\Big) \geq 0.$$

This is again a contradiction, since $\Delta\xi(p) \leq 0$. □

Useful references:

- for the mountain pass lemma: Section 8.1 of Ambrosetti–Malchiodi, *Nonlinear analysis and semilinear elliptic problems* (where the statement is a special case of the one given in these notes, but the same proof works);
- for the structure theorem on the limit varifold: Hutchinson–Tonegawa, *Convergence of phase interfaces in the van der Waals–Cahn–Hilliard theory*;
- for its geometric application to construct minimal hypersurfaces in a closed ambient: Guaraco, *Min-max for phase transitions and the existence of embedded minimal hypersurfaces*;
- for the Ambrosio–Soner rectifiability result: Section 3 of Ambrosio–Soner, *A measure theoretic approach to higher codimension mean curvature flow* (the definition of the set $\tilde{A}^n_k$ of "generalized planes" is slightly different from the one given here, but the same proof works);
- for analogous results in codimension two: Pigati–Stern, *Minimal submanifolds from the abelian Higgs model.*