for any set $\Delta \subset \Omega_0$ whose Lebesgue measure is less than 2δ. Thus, if we choose s_1 and t_1 above so close to s and t that $(s_1 - s) + (t - t_1) \leq \delta$, then (9) and (10) imply that

$$J_{s',t'}(v) \leq J_{s,t}(u) + 5\varepsilon \leq m_{s,t} + 6\varepsilon \tag{11}$$

for all pairs $(s', t') \in \Omega_0^2$ such that $s' \leq s \leq t \leq t'$ and $(s - s') + (t' - t) \leq \delta$.

Note that λ, and then δ, depend only on ε and Ω_0. Since the function $m_{s,t}$ is nondecreasing in t and nonincreasing in s, its continuity on $\Omega_0 \times \Omega_0$ follows from (11), which proves (5).

A direct consequence of our definitions is that

$$m = \inf \Big\{ \sum_{i=1}^{N-1} c(t_i) + \sum_{i=1}^{N} m_{t_{i-1},t_i} \Big\}, \tag{12}$$

where the infimum is taken over all finite sequences $\{t_i\}_{0 \leq i \leq N}$ such that $t_0 < t_1 < \cdots < t_N$ and t_0 and t_N are the endpoints of Ω. Moreover, we know that we can restrict to $N \leq C$, because c is bounded from below. If in addition c is lower semicontinuous (which means that $c(x) \leq \liminf_{k \to +\infty} c(x_k)$ whenever $\{x_k\}$ converges to x), then (5) allows us to find a finite sequence $\{t_i\}_{0 \leq i \leq N}$ such that

$$m = \sum_{i=1}^{N-1} c(t_i) + \sum_{i=1}^{N} m_{t_{i-1},t_i}.$$

Because of this, it is enough to study the individual functionals $J_{s,t}$, which is substantially easier than what happens in higher dimensions. In particular, the same convexity argument as in Section 3 gives the existence (and uniqueness) of minimizers for $J_{s,t}$ if $m_{s,t} < +\infty$ and under mild nondegeneracy conditions on the functions a and b (for instance, if they are locally bounded from below). The existence of minimizers for J follows.

Before we get to counterexamples, let us record that minimizers for $J_{s,t}$ satisfy a differential equation with Neumann conditions at the endpoints. To simplify the notation, let us only consider bounded intervals.

Lemma 13. *If b is of class C^1 and does not vanish on $[s,t]$, and if $u \in W^{1,2}_{\mathrm{loc}}(s,t)$ is such that $J_{s,t}(u) = m_{s,t} < +\infty$, then u is of class C^1 on $[s,t]$, $u'' \in L^2((s,t))$,*

$$b(x)u''(x) + b'(x)u'(x) = a(x)(u(x) - g(x)) \quad \text{on } (s,t), \qquad \text{and } u'(s) = u'(t) = 0. \tag{14}$$

First observe that $u' \in L^2(s,t)$, because $J_{s,t}(u) < +\infty$ and b is bounded from below. Similarly, $a(u-g) \in L^2(s,t)$, because a is bounded and $\int_s^t a(u-g)^2 \leq J_{s,t}(u) < +\infty$. Let us proceed as in Proposition 3.16, pick any smooth function φ (not necessarily compactly supported in (s,t) yet), and compare u with its competitors $u + \lambda\varphi$, $\lambda \in \mathbb{R}$. We get the scalar product condition

$$\int_s^t \big[a(x)(u(x) - g(x))\varphi(x) + b(x)u'(x)\varphi'(x)\big]\,dx = 0; \tag{15}$$

as in (3.18). When we take φ compactly supported in (s,t), we get that $(bu')' = a(u-g)$ in the sense of distributions. But this forces bu' to be the indefinite integral of $a(u-g)$ (which lies in $L^2(s,t)$, as we just noticed); see Exercise 59. In particular bu' is bounded, and hence also u', because b is bounded from below. Now $u' = b^{-1}(bu')$ and b^{-1} is C^1, so $u'' = b^{-1}(bu')' - b'b^{-2}bu' = b^{-1}a(u-g) - b'b^{-1}u'$, both in the sense of distributions and pointwise almost everywhere (see Exercise 59 again). The first part of (14) follows.

Next, u' is continuous on $[s,t]$, because it is the indefinite integral of $u'' \in L^2$. In particular, $u'(s)$ and $u'(t)$ are well defined. To prove that $u'(t) = 0$, we apply (15) with a smooth nondecreasing function φ such that $\varphi(x) = 0$ for $x \leq t'$ and $\varphi(x) = 1$ for $x \geq t$, where $t' < t$ is very close to t. The contribution $\int_s^t a(x)(u(x) - g(x))\varphi(x)dx$ to (15) tends to zero as t' tends to t (because φ is bounded and $a(u-g) \in L^2$), while $\int_s^t b(x)u'(x)\varphi'(x)dx$ tends to $b(t)u'(t)$. This proves that $b(t)u'(t) = 0$; hence $u'(t) = 0$ because $b(t) > 0$. Of course $u'(s) = 0$ for similar reasons, and Lemma 13 follows. □

Example 16. Take $n = 1$, $\Omega = [-1,1]$, $a \equiv b \equiv c \equiv 1$, and $g_\lambda = \lambda \mathbf{1}_{[0,1]}$ for $\lambda \geq 0$. It is clear now that there are only two possible minimizers. The first one is given by $K = \{0\}$ and $u = g_\lambda$, and then $J(u,K) = 1$. The second one is for $K = \emptyset$ and $u = u_\lambda$, where u_λ is the only minimizer of $J_{-1,1}$ in (3). We could easily use Lemma 13 to compute u_λ, but the main point is that $u_\lambda = \lambda u_1$, so that $J(u_\lambda, \emptyset) = A\lambda^2$, where $A = \int [(u_1 - g_1)^2 + (u_1')^2]$ is a fixed positive constant.

For $\lambda^2 < A^{-1}$, the only minimizer for J is $(u_\lambda, \emptyset)$, while for $\lambda^2 > A^{-1}$ the only minimizer is $(g_\lambda, \{0\})$. And for $\lambda^2 = A^{-1}$, we have exactly two minimizers. Thus uniqueness and smooth dependence on parameters fail in a very simple way. The principle behind all our other examples will be the same.

b. Simple examples in dimension 2

Here we return to the usual Mumford-Shah functional in (2.13), with $a = b = c = 1$. We start with the product of a line segment with Example 16.

Example 17. (A white square above a black square.) This is taken from [DaSe5] but the example could be much older. We take $\Omega = (0,1) \times (-1,1) \subset \mathbb{R}^2$ and $g = g_\lambda = \lambda \mathbf{1}_{\Omega^+}$, where $\Omega^+ = (0,1) \times (0,1)$ and λ is a positive parameter. We shall denote by $L = (0,1) \times \{0\}$ the singular set of g_λ.

There are two obvious candidates to minimize J. The first one is the pair (g_λ, L) itself. The second one is $(\tilde{u}_\lambda, \emptyset)$, where $\tilde{u}_\lambda(x,y) = u_\lambda(y)$ and u_λ is the C^1-function on $(-1,1)$ for which

$$I = \int_{(-1,1)} \left|u_\lambda(y) - \lambda \mathbf{1}_{(0,1)}(y)\right|^2 + |u_\lambda'(y)|^2 \tag{18}$$

is minimal. Thus u_λ is the same as in Example 16. Set

$$m = \mathrm{Min}(J(g_\lambda, L), J(\tilde{u}_\lambda, \emptyset))\,; \tag{19}$$

we want to check that

$$J(u,K) \geq m \text{ for all } (u,K) \in \mathcal{A}. \tag{20}$$

So let $(u,K) \in \mathcal{A}$ be an admissible pair (see (2.1) for the definition). We may as well assume that u minimizes J with the given set K. This will not really change the proof; it will only make it a little more pleasant because we know that u is C^1 on $\Omega \setminus K$. Set

$$X = \Big\{x \in (0,1)\,;\, K \cap \big[\{x\} \times (-1,1)\big] = \emptyset\Big\}. \tag{21}$$

For $x \in X$, the function $y \to u(x,y)$ is C^1 on $(-1,1)$; then

$$\begin{aligned} &\int_{X \times (-1,1)} |u - g_\lambda|^2 + |\nabla u|^2 \\ &\geq \int_X \Big\{ \int_{(-1,1)} |u(x,y) - \lambda \mathbf{1}_{(0,1)}(y)|^2 + \Big|\frac{\partial u}{\partial y}(x,y)\Big|^2 dy \Big\} dx \geq H^1(X)\, I, \end{aligned} \tag{22}$$

where I is as in (18). On the other hand, $(0,1) \setminus X = \pi(K)$, where π is the orthogonal projection on the first axis. Since π is 1-Lipschitz, Remark 2.11 says that $H^1(K) \geq H^1((0,1) \setminus X) = 1 - H^1(X)$. We can combine this with (22) and get that

$$J(u,K) \geq \int_{X \times (-1,1)} \Big[|u - g_\lambda|^2 + |\nabla u|^2\Big] + H^1(K) \geq H^1(X)\, I + [1 - H^1(X)]. \tag{23}$$

Thus $J(u,K)$ is a convex combination of I and 1. Since $I = J(\tilde{u}_\lambda, \emptyset)$ and $1 = J(g_\lambda, L)$, (20) follows from (23) and (19).

With just a little more work, we could prove that if $J(u,K) = m$, then $(u,K) = (\tilde{u}_\lambda, \emptyset)$ or $(u,K) = (g_\lambda, L)$, modulo adding a set of measure 0 to K.

Thus at least one of our two favorite candidates is a minimizer for J. Since I is obviously proportional to λ^2, there is a value of λ for which $I = 1$, and for this value J has two minimizers. Of course one could object that this example is very special, and that we are using the two vertical sides of $\partial\Omega$, so we shall give two other ones.

Example 24. Take $g = \lambda \mathbf{1}_{B(0,1)}$, in the domain $\Omega = \mathbb{R}^2$ or $\Omega = B(0,R) \subset \mathbb{R}^2$, where $\lambda > 0$ and $R > 1$ are given constants.

It is tempting here to use the symmetry of the problem to compute things almost explicitly. By an averaging argument similar to the proof of (20) above (but radial), we can show that for each $(u,K) \in \mathcal{A}$ there is a radial competitor (u^*, K^*) such that $J(u^*, K^*) \leq J(u,K)$ (see [DaSe5], page 164, for details). Here radial means that K^* is a union of circles centered at the origin, and $u^*(\rho e^{i\theta})$ is just a function of ρ. Then the choice can be reduced to two competitors, as follows.

for all $\lambda > 0$. Thus, $\mu^* \in L^1_{\text{weak}}(\mathbb{R}^n)$. Note that this applies in particular to $d\mu = f(x)dx$, with $f \in L^1(\mathbb{R}^n, dx)$.

Hint: cover $\Omega_\lambda = \{x \in \mathbb{R}^n\,;\, \mu^*(x) > \lambda\}$ by balls $B(x,r)$ such that $\mu(B(x,r)) > \lambda r^n$, and apply Lemma 1.

Exercise 41 (Sard). Let $\Omega \subset \mathbb{R}^d$ be open, and $f : \Omega \to \mathbb{R}^n$ be of class C^1. Set $Z = \{x \in \Omega;\, Df(x) \text{ is not injective}\}$. We want to show that

$$H^d(f(Z)) = 0\,. \tag{42}$$

1. Let $x \in Z$ be given. Show that there is a constant C_x such that for each $\varepsilon > 0$ we can find r_0 such that for $r < r_0$, $f(B(x,r))$ can be covered by less than $C_x \varepsilon^{1-d}$ balls of radius εr. Hint: show it first with f replaced with its differential $Df(x)$ at x.
2. Show that for each choice of $x \in Z$, $\tau > 0$, and $\eta \in (0,1)$, we can find $r_x > 0$ such that $r_x < \tau$, $B(x, r_x) \subset \Omega$, and $f(B(x, 5r_x))$ can be covered by finitely many balls D_j, with $\sum_j (\text{diam}\, D_j)^d < \eta r_x^d$.
3. Pick $R > 0$ and cover $Z \cap B(0,R)$ by balls $B(x, r_x)$ as above. Then apply Lemma 1. Show that this yields

$$H^d_\tau(f(Z \cap B(0,R))) \leq C\eta |\Omega \cap B(0,R)|. \tag{43}$$

Conclude.

34 Local Hausdorff convergence of sets

The main goal of this section is to give a natural (and of course very classical) definition of convergence for closed subsets of an open set $\Omega \subset \mathbb{R}^n$. The main point will be to make sure that we can extract convergent subsequences from any given sequence of sets in Ω. Note that for $\Omega = \mathbb{R}^n$, say, the usual Hausdorff distance between sets is too large for this; what we need is something like convergence for the Hausdorff metric on every compact subset.

So let $\Omega \subset \mathbb{R}^n$ be a given open set, and let $\{H_m\}_{m \geq 0}$ be an exhaustion of Ω by compact subsets. This means that each H_m is compact, $H_m \subset \text{interior}(H_{m+1})$ for $m \geq 0$, and $\Omega = \bigcup_m H_m$. When $\Omega = \mathbb{R}^n$, a good choice is $H_m = \overline{B}(0, 2^m)$. We set

$$d_m(A,B) = \sup\{\text{dist}(x,B)\,;\, x \in A \cap H_m\} + \sup\{\text{dist}(y,A)\,;\, y \in B \cap H_m\} \tag{1}$$

for $A, B \subset \Omega$ and $m \geq 0$. We use the convention that $\text{dist}(x,B) = +\infty$ when $B = \emptyset$, but $\sup\{\text{dist}(x,B)\,; x \in A \cap H_m\} = 0$ when $A \cap H_m = \emptyset$. Thus $d_m(A,B) = 0$ when $A \cap H_m = B \cap H_m = \emptyset$, but $d_m(A,\emptyset) = +\infty$ if $A \cap H_m \neq \emptyset$.

We shall find it more convenient here to use $d_m(A,B)$, rather than the Hausdorff distance between $A \cap H_m$ and $B \cap H_m$. Recall that the Hausdorff distance between A and B is

$$d_{\mathcal{H}}(A,B) = \sup_{x \in A} \operatorname{dist}(x,B) + \sup_{y \in B} \operatorname{dist}(y,A). \tag{2}$$

Here it could happen that $d_m(A,B) << d_{\mathcal{H}}(A \cap H_m, B \cap H_m)$, for instance because points of $A \cap H_m$ can be close to B, even if B does not meet H_m. Thus our functions d_m are a little less sensitive to boundary effects. Also, they have the nice feature that

$$d_m(A,B) \leq d_l(A,B) \quad \text{for } 0 \leq m \leq l, \tag{3}$$

because $H_m \subset H_l$.

Definition 4. *Let $\{A_k\}$ be a sequence of subsets of Ω, and $B \subset \Omega$. We say that $\{A_k\}$ converges to B if*

$$\lim_{k \to +\infty} d_m(A_k, B) = 0 \quad \textit{for all } m \geq 0. \tag{5}$$

This is the notion of convergence that we shall systematically use in the following sections. A few comments may be useful here. The reader may be worried because we did not restrict to closed sets. This is not really needed, but anyway $\{A_k\}$ converges to B if and only if the sequence $\{\overline{A}_k\}$ of closures in Ω converges to B. Also, convergence to B is equivalent to convergence to $\overline{B}$, and if we want to have uniqueness of the limit, we can require that it be closed. [See Exercise 15.] In the later sections, we shall always apply the definition to sequences of closed sets, and take a closed limit.

It is fairly easy to check that our notion of convergence does not depend on the choice of exhaustion $\{H_m\}$. See Exercise 17.

We now come to the main point of this section.

Proposition 6. *For each sequence $\{A_k\}$ of subsets of Ω, we can find a subsequence of $\{A_k\}$ that converges to some closed subset of Ω.*

This is a very standard exercise on complete boundedness, Cauchy sequences, and the diagonal process. We nonetheless give the proof for the convenience of the reader. Let us say that $\{A_k\}$ is a Cauchy sequence if

$$\lim_{k,l \to +\infty} d_m(A_k, A_l) = 0 \quad \text{for every } m \geq 0. \tag{7}$$

It is easy to see that if $\{A_k\}$ converges to some $B \subset \Omega$, it is a Cauchy sequence.

Lemma 8. *Every Cauchy sequence has a limit.*

Let $\{A_k\}$ be a Cauchy sequence, and set

$$B = \{x \in \Omega \, ; \lim_{k \to +\infty} \operatorname{dist}(x, A_k) = 0\}. \tag{9}$$

Note that B is closed. We want to check that $\{A_k\}$ converges to B. Let $m \geq 0$ be given. First note that the functions $f_k = \operatorname{dist}(\cdot, A_k)$ are 1-Lipschitz, so they are equicontinuous on H_m. Since they converge to 0 on B, the convergence is uniform on $B \cap H_m$. [See Exercise 22.] Hence

$$\lim_{k \to +\infty} \Big\{ \sup \{\operatorname{dist}(x, A_k) \, ; \, x \in B \cap H_m\} \Big\} = 0. \tag{10}$$

Note that

$$B = \{x \in \Omega \, ; \liminf_{k \to +\infty} \operatorname{dist}(x, A_k) = 0\}, \tag{11}$$

because $\{A_k\}$ is a Cauchy sequence. Indeed, if x lies in this last set and ε is small, we can find arbitrarily large values of k for which A_k meets $B(x, \varepsilon)$. Then all A_l, l large enough, meet $B(x, 2\varepsilon)$, because $d_m(A_k, A_l) \leq \varepsilon$, where we choose m so large that $B(x, \varepsilon) \subset H_m$. The other inclusion is trivial. Next we want to check that

$$\lim_{k \to +\infty} \Big\{ \sup \{\operatorname{dist}(x, B) \, ; \, x \in A_k \cap H_m\} \Big\} = 0 \tag{12}$$

for all $m \geq 0$. Let ε be given, so small that $B(x, 2\varepsilon) \subset H_{m+1}$ for $x \in H_m$. Let k_0 be such that $d_{m+1}(A_k, A_l) \leq \varepsilon$ for $k, l \geq k_0$. If $k \geq k_0$ and $x \in A_k \cap H_m$, then $\overline{B}(x, \varepsilon)$ meets A_l for all $l \geq k$. Thus $\overline{B}(x, \varepsilon)$ contains points of adherence of the sequence $\{A_l\}$. These points lie in B, by (11). This proves (12), and the lemma follows because of (10). □

Remark 13. If $\{A_k\}$ converges to B and B is closed in Ω, then B is given by both formulae (9) and (11). This follows from our proof of Lemma 8 and the uniqueness of the limit (given that it is closed). See Exercise 15 for this last point.

Now we can prove Proposition 6. Let $\{A_k\}$ be any sequence of sets in Ω. It is enough to show that we can extract a Cauchy subsequence.

For each $m \geq 0$ and $N \geq 0$, cover H_m by finitely many balls $B_{m,N,s}$, $s \in S(m, N)$, that are centered on H_m, contained in H_{m+1}, and with radii smaller than 2^{-N}. Then define functions $g_{m,N,s}$ on $\mathbb{N}$ by $g_{m,N,s}(k) = 1$ if $B_{m,N,s}$ meets A_k, and $g_{m,N,s}(k) = 0$ otherwise. We can use the diagonal process to extract a subsequence $\{k_i\}$ such that every $\{g_{m,N,s}(k_i)\}_i$ has a limit $L_{m,N,s}$ at infinity. Let us check that $\{A_{k_i}\}$ converges.

Let m and N be given, and let i_0 be so large that $g_{m,N,s}(k_i) = L_{m,N,s}$ for $i \geq i_0$ and $s \in S(m, N)$. Let $i, j \geq i_0$ and $x \in H_m \cap A_{k_i}$ be given. Then x lies in some $B_{m,N,s}$, and $g_{m,N,s}(k_i) = 1$ by definition. Then $g_{m,N,s}(k_j) = 1$ also, and $B_{m,N,s}$ meets A_{k_j}. Hence $\operatorname{dist}(x, A_{k_j}) \leq 2^{-N+1}$. The same reasoning can be done with $x \in A_{k_j}$, and altogether $d_m(A_{k_i}, A_{k_j}) \leq 2^{-N+1}$.

Since this can be done for all m and N, $\{A_{k_i}\}$ is a Cauchy sequence, as desired. Proposition 6 follows. □