MULLINS-SEKERKA AS THE WASSERSTEIN FLOW OF THE PERIMETER

Antonin Chambolle, Tim Laux

MULLINS-SEKERKA AS THE WASSERSTEIN FLOW OF THE PERIMETER

ANTONIN CHAMBOLLE AND TIM LAUX

Abstract. We prove the convergence of an implicit time discretization for the Mullins-Sekerka equation proposed in [F. Otto, Arch. Rational Mech. Anal. 141 (1998) 63–103]. Our simple argument shows that the limit satisfies the equation in a distributional sense as well as an optimal energy-dissipation relation. The proof combines simple arguments from optimal transport, gradient flows & minimizing movements, and basic geometric measure theory.

Keywords: Gradient flows, Wasserstein distance, sets of finite perimeter, Mullins-Sekerka, free boundary problems

Math. Subject Classification: 35A15, 35R37, 49Q20, 76D27, 90B06, 35R35

1. Introduction

The Mullins-Sekerka equation, see (5)–(8) below for its exact formulation, is a well-studied mathematical model which, among other phenomena, describes a Hele-Shaw cell: a viscous ferro-fluid is confined to a thin region between two parallel horizontal plates. Applying a strong magnetic field in the vertical direction leads to two opposing forces: (i) due to surface tension, the fluid wants to decrease its surface area; (ii) the probe becomes magnetized by the field and the particles repel each other due to the induced magnetic field. These two competing effects lead to the formation of intriguing patterns.

In this paper we construct weak solutions using an implicit time discretization proposed by F. Otto in [24]. Because of the gradient-flow structure of the equation, it is natural to consider minimizing movements, an implicit time discretization which comes as a sequence of variational problems [6]. The effective energy consists of two terms, (i) an attractive term due to surface tension, the total surface area of the lateral boundary of the region occupied by the fluid, and (ii) a nonlocal term due to the magnetic repulsion of the particles; see (3) below. In [24] it has been observed that the dissipation functional may be modelled by the Wasserstein distance, which arises in optimal mass transport; see (2). The Wasserstein distance plays a crucial role for many diffusion equations as was pointed out by Jordan, Kinderlehrer, and Otto in the seminal work [15], see also [14]. The resulting metric tensor $\int_E |\mathbf{u}|^2 \, dx$ defined on divergence-free vector fields $\mathbf{u}\colon E \to \mathbb{R}^d$ is less degenerate than the one of the mean curvature flow $\int_{\partial E} V^2 dS$ defined on normal velocities V, but more degenerate than the one of the two-phase Mullins-Sekerka problem $\int_{\mathbb{R}^d} |\mathbf{u}|^2 dx$, in which the ferro-liquid is assumed to be surrounded by another liquid of the same viscosity.

The main theorem of the present work is a refined version of the announced result [24, Theorem 1], for which a detailed proof was not provided. Our simple proof establishes the convergence of the approximations obtained from the minimizing movements scheme to a weak solution. We derive the Mullins-Sekerka equation (5)–(8) in a distributional form, and using De Giorgi's variational interpolations [2], we show that the limit satisfies an optimal energy-dissipation relation. The convergence of the energies as $h \to 0$, a well-known assumption known from the more difficult case of mean curvature flow [21], is *not* necessary in our case. In fact, our proof is much simpler and no regularity theory of almost minimal surfaces is needed. It may be expected that our solution concept satisfies a weak-strong uniqueness principle similar to the ones in the forthcoming works by Fischer, Hensel, Simon, and one of the authors for multiphase mean curvature flow [9], and for the simpler *two-phase* Mullins-Sekerka equation [10].

There has been continuous interest in the Mullins-Sekerka equation and similar gradient flows, so we only briefly point out some of the most relevant results related to the present work. Weak solutions to the two-phase Stefan problem have been constructed by Luckhaus [20]. In particular, Luckhaus discovered a hidden variational principle satisfied by his approximations, which allows to *verify* the convergence of the energies as $h \to 0$. Luckhaus and Sturzenhecker [21] constructed weak solutions of mean curvature flow and the two-phase Mullins-Sekerka equation *conditioned* on the convergence of the energies. Röger [25] was able to remove the assumption in the case of this two-phase Mullins-Sekerka equation by showing that the assumption may only be violated along flat parts of ∂E. In the case of mean curvature flow, the assumption can be verified in very particular cases, like convex sets [5], graphs [19], and mean convex sets [7]. For generalizations to the anisotropic case, which for mean curvarture flow has already been introduced by Almgren, Taylor, and Wang [1], we refer the interested reader to Garcke and Schaubeck [11] and Kraus [16]. A variant relevant for image denoising has been introduced by Carlier and Poon [4] who relax the constraint $\chi_E \in \{0, 1\}$, which leads to the total variation flow. However, it seems that the convergence can only be proven under an additional assumption on the density. Glasner [12] introduced a phase-field approximation to the one-phase Mullins-Sekerka equation and studied its convergence by formal asymptotic expansions. Recently, also the computationally efficient thresholding scheme by Merriman, Bence, and Osher [22, 23] has been reinterpreted as a minimizing movements scheme by Esedoğlu and Otto [8], which allowed one of the author together with Otto to prove conditional convergence results to multiphase mean curvature flow [18, 17]. Most recently, Jacobs, Kim, and Mészáros [13] introduced an interesting thresholding-type approximation for the Muskat problem and proved a similar (conditional) convergence result for their scheme.

The paper is organized as follows: In §2 we recall the minimizing movements scheme and state our main result Theorem 1, which will be proved in the following sections: §3 establishes the compactness; in §4 we recover the distributional equation for the limit and the optimal energy-dissipation relation; and §5 contains a simple nonlinear interpolation inequality and its proof.

Acknowledgement

The authors would like to warmly thank Prof. Felix Otto, Max Planck Institute (MPI), Leipzig, for mentioning this problem, and for numerous discussions and suggestions. This work was started during Mathematisches Forschungsinstitut Oberwolfach's (MFO) workshop #1746 and completed during MFO's workshop #1904, as well as visits to MPI Leipzig. The hospitality and support of these institutions is also acknowledged.

2. Statement of the main result

We recall the implicit time discretization introduced by F. Otto in [24]: Given a time-step size $h > 0$, and initial conditions $E_0 \subset \mathbb{R}^d$, for $n \geq 1$ find E_n solving

$$\min_E \Big\{ \frac{1}{2h} W_2^2(\chi_E, \chi_{E_{n-1}}) + P(E) + \int_E k * \chi_E \, dx \Big\}. \tag{1}$$

Here k is a non-negative, symmetric, and normalized convolution kernel $k \geq 0$, $k(-z) = k(z)$, and $\int k = 1$; $P(E) := \sup\{-\int_E \operatorname{div} \xi \, dx \colon \sup |\xi| \leq 1\}$ denotes the perimeter of $E \subset \mathbb{R}^d$; and

$$W_2^2(\chi_E, \chi_F) = \inf \int_E |x - T(x)|^2 dx = \min \iint |x - y|^2 d\gamma(x, y) \tag{2}$$

denotes the squared Wasserstein distance, where the infimum runs over all transport maps, i.e., volume preserving diffeomorphisms $T \colon \mathbb{R}^d \to \mathbb{R}^d$ such that $T_\sharp \chi_E = \chi_F$, and the minimum runs over all transport plans, i.e., finite measures γ in $\mathbb{R}^d \times \mathbb{R}^d$ with marginals $\chi_E(x)dx$ and $\chi_F(y)dy$. We denote

$$\mathcal{E}(E) := P(E) + \int_E k * \chi_E \, dx, \tag{3}$$

and let $E_h(t) := E_{[t/h]}$ for $t \geq 0$. Our standing assumption on the initial conditions is

$$P(E_0) < \infty \quad \text{and} \quad \int_{E_0} (1 + |x|^2) \, dx < \infty. \tag{4}$$

In particular $|E_0| < \infty$ and w.l.o.g. by scaling we may assume that $|E_0| = 1$.

The physical model is described by the following set of equations: The interface ∂E is transported by the fluid

$$V = \mathbf{u} \cdot (-\nu) \quad \text{on } \partial E \tag{5}$$

(throughout, ν denotes the *inner* normal to ∂E); the fluid is incompressible

$$\operatorname{div} \mathbf{u} = 0 \quad \text{in } E; \tag{6}$$

the flow is irrotational, i.e.,

$$\text{there exists } p \text{ such that } \mathbf{u} = -\nabla p \text{ in } E; \tag{7}$$

and

$$p = H - 2k * \chi_E \quad \text{on } \partial E. \tag{8}$$

Any smooth solution $E(t)$ is volume-preserving $\frac{d}{dt}|E| = 0$ and, more importantly, energy dissipating

$$\frac{d}{dt}\mathcal{E}(E) = -\int_E |\mathbf{u}|^2 \, dx \leq 0. \tag{9}$$

More precisely, the above set of equations have a gradient-flow structure. For more physical motivation, we refer to the introduction of [24] and the references therein.

The main result is the following construction of solutions.

Theorem 1. *Let $E_0 \subset \mathbb{R}^d$ be initial conditions satisfying (4) and let $E_h(t)$ be constructed as above. Then there exists a subsequence $h \downarrow 0$ and a one-parameter, continuous family of finite perimeter sets $E(t)$ and a vector field $\mathbf{u} \in L^2(\mathbb{R}^d \times (0,+\infty))$ such that*

$$\lim_{h\downarrow 0} \sup_{t\in[0,T]} |E_h(t)\triangle E(t)|\,dt = 0 \quad \textit{for all } T<+\infty$$

and

$$-\int_0^{+\infty}\int_{E(t)} (\partial_t\zeta + \mathbf{u}\cdot\nabla\zeta)\,dxdt = \int_{E_0}\zeta(0)\,dx \tag{10}$$

for all $\zeta \in C_0^\infty(\mathbb{R}^d\times[0,+\infty))$ and $E(t)$ satisfies the optimal energy dissipation rate

$$\mathcal{E}(E(T)) + \int_0^T\int_{E(t)} |\mathbf{u}(x,t)|^2\,dx\,dt \le \mathcal{E}(E_0) \quad \textit{for almost all } T>0. \tag{11}$$

Furthermore, the measures $\mu^h := \delta_{\nu_{E_h(t)}} \otimes \left|\nabla\chi_{E_h(t)}\right| dt$ converge, $\mu^h \rightharpoonup \mu = \mu_t\,dt$, to some varifold μ, i.e., a measure on $(\tilde\nu,x,t) \in \mathbb{S}^{d-1}\times\mathbb{R}^d\times[0,+\infty)$, such that $\left|\nabla\chi_{E(t)}\right| dt \le \mu_t dt$ and

$$\begin{aligned}-\int_0^\infty\int_{E(t)} \mathbf{u}\cdot\xi\,dxdt = &\int_0^\infty\int\int (\operatorname{div}\xi - \tilde\nu\cdot D\xi\,\tilde\nu)\,d\mu_t(\tilde\nu,x)\,dt \\ &+ 2\int_0^\infty\int k*\chi_{E(t)}\,\xi\cdot\nu_{E(t)}\left|\nabla\chi_{E(t)}\right| dt\end{aligned} \tag{12}$$

for all $\xi \in C_0^\infty(\mathbb{R}^d\times(0,+\infty),\mathbb{R}^d)$ with $\operatorname{div}\xi = 0$ where $\nu_{E(t)} = \nu_{E(t)}(x) = \frac{\nabla\chi_{E(t)}}{|\nabla\chi_{E(t)}|}$ denotes the inner normal of $E(t)$; and the pair (μ,E) satisfies the optimal energy dissipation relation

$$\begin{aligned}&\mu_T(\mathbb{S}^{d-1},\mathbb{R}^d) + \int_{E(T)} k*\chi_{E(T)}\,dx + \frac12\int_0^T\int_{E(t)}|\mathbf{u}(x,t)|^2\,dx\,dt \\ &+\int_0^T\int\int(\operatorname{div}\xi - \tilde\nu\cdot D\xi\,\tilde\nu)\,d\mu_t(\tilde\nu,x)\,dt + 2\int_0^T\int k*\chi_{E(t)}\,\xi\cdot\nu\left|\nabla\chi_{E(t)}\right| dt \\ &\qquad -\frac12\int_0^T\int_{E(t)}|\xi|^2\,dx\,dt \le \mathcal{E}(E_0)\end{aligned} \tag{13}$$

for almost all $T<+\infty$ and all $\xi \in C_0^\infty(\mathbb{R}^d\times(0,+\infty),\mathbb{R}^d)$ with $\operatorname{div}\xi = 0$.

Remark 1. The set of equations derived in the theorem are indeed a weak form of the free boundary problem described earlier:

- The continuity equation (10) encodes (5) & (6) as well as the initial conditions E_0.
- Equation (12) encodes both (7) (since $\mathbf{u}$ is orthogonal to divergence-free fields in $E(t)$) and the balance law (8).
- The last three left-hand side terms involving the test vector field ξ in (13) can be viewed as a fractional Sobolev norm of $H - 2k*\chi_E$, and (13) is a type of De Giorgi inequality, which for a smooth gradient flow characterizes the solution.
- Note also that we may replace the sum of the first two left-hand side terms in (13) by the (smaller) energy $\mathcal{E}(E(T))$. If additionally

$$\lim_{h\downarrow 0}\int_0^T P(E_h(t))\,dt = \int_0^T P(E(t))\,dt,$$

we may replace the measure μ by the BV-version $\delta_{\nu_{E(t)}} \otimes \left|\nabla \chi_{E(t)}\right|$ in all terms appearing in (12) & (13).

- The optimal energy-dissipation rate (here in form of (11) or (13)) plays a crucial role in recent weak-strong uniqueness proofs and does not follow from the weak formulation (10).

In the following we write $A \lesssim B$ if there exists a generic constant $C = C(d)$ such that $A \leq C\,B$.

3. Compactness

Lemma 1 (Compactness). *Suppose E_0 satisfies (4) and let E_h be constructed by the scheme as above. Then*

$$W_2(\chi_{E_h(t)}, \chi_{E_h(s)}) \lesssim \mathcal{E}(E_0)^{\frac{1}{2}}(t-s)^{\frac{1}{2}} \tag{14}$$

and

$$|E_h(t) \triangle E_h(s)| \lesssim \mathcal{E}(E_0)^{\frac{3}{4}}(t-s)^{\frac{1}{4}} \tag{15}$$

for all $t > s \geq 0$ with $t - s \geq h$.

Therefore, there exists a subsequence $h \downarrow 0$ and a one-parameter family of finite perimeter sets $(E(t))_{t\geq 0}$ such that for any $T < +\infty$

$$\sup_{t\in[0,T]} |E_h(t) \triangle E(t)|\, dt \to 0 \qquad \text{as } h \downarrow 0. \tag{16}$$

Furthermore, the limit satisfies

$$W_2(\chi_{E(t)}, \chi_{E(s)}) \lesssim \mathcal{E}(E_0)^{\frac{1}{2}}(t-s)^{\frac{1}{2}} \tag{17}$$

and

$$|E(t) \triangle E(s)| \lesssim \mathcal{E}(E_0)^{\frac{3}{4}}(t-s)^{\frac{1}{4}} \tag{18}$$

for all $t > s \geq 0$.

Proof. Using E_{n-1} as a competitor in (1) yields

$$\frac{1}{2h} W_2^2(\chi_{E_n}, \chi_{E_{n-1}}) + \mathcal{E}(E_n) \leq \mathcal{E}(E_{n-1}),$$

so that after summation in n and telescoping

$$\frac{h}{2} \sum_{n=n_0+1}^{n_1} \left(\frac{W_2(\chi_{E_n}, \chi_{E_{n-1}})}{h} \right)^2 + \mathcal{E}(E_{n_1}) \leq \mathcal{E}(E_{n_0}). \tag{19}$$

In particular, for any pair of integers $n_1 > n_0 \geq 0$, we have:

$$\begin{aligned} W_2(\chi_{E_{n_0}}, \chi_{E_{n_1}}) &\leq \sqrt{(n_1 - n_0)h} \left(\sum_{n=n_0+1}^{n_1} \frac{1}{h} W_2(\chi_{E_n}, \chi_{E_{n-1}})^2 \right)^{\frac{1}{2}} \\ &\leq \sqrt{(n_1 - n_0)h} \sqrt{2(\mathcal{E}(E_{n_0}) - \mathcal{E}(E_{n_1}))}, \end{aligned}$$

which implies (14).

The L^1 estimate (15) then follows from (14) in conjunction with the interpolation inequality in Corollary 1 and Jensen's inequality in the form of $W_1(\chi, \tilde{\chi}) \leq W_2(\chi, \tilde{\chi})$.

The energy estimate (19) also yields a uniform bound on the perimeter, so

$$\int |\chi_{E_h(t)}(x+z) - \chi_{E_h(t)}(x)|\, dx \leq |z| P(E_h(t)) \leq |z| \mathcal{E}(E_0),$$

i.e., we have a uniform modulus of continuity in space. Together with the uniform modulus of continuity in time (15), which is valid down to scales h, this allows us to apply the Riesz-Kolmogorov compactness theorem in $L^1([0,T]\times K)$ for any compact set $K\subset\mathbb{R}^d$ and any $T<+\infty$. A diagonal argument yields $\chi_{E_h}\to\chi_E$ in $L^1_{\mathrm{loc}}([0,\infty)\times\mathbb{R}^d)$. But since $\int|x|^2\chi_{E_h(t)}(x)dx<\infty$, which follows from (4) & (14), this implies the L^1-convergence globally in space, and locally in time. Eventually, an Ascoli-Arzelà type argument, together with the estimate in Corollary 1, allows to deduce the local uniform convergence in time, i.e., (16). The continuity estimates (17) & (18) then follow immediately. □

4. Convergence

Proof of Theorem 1.
Step 1: Construction of $\mathbf{u}$ *and verification of* (10). By Kantorovich duality

$$\frac{1}{2h}W_2^2(\chi_{E_n},\chi_{E_{n-1}})=\sup_{\phi(x)+\psi(y)\le\frac{|x-y|^2}{2h}}\int_{E_n}\phi(x)\,dx+\int_{E_{n-1}}\psi(y)\,dy.$$

This supremum is reached at (ϕ^n,ψ^n) such that

$$\Phi^n(x)=\frac{|x|^2}{2}-h\phi^n(x),\quad \Psi^n(y)=\frac{|y|^2}{2}-h\psi^n(y)$$

are convex conjugates and $(\nabla\Phi^n)_\sharp\chi_{E_n}=\chi_{E_{n-1}}$ is optimal in $W_2(\chi_{E_n},\chi_{E_{n-1}})$, see [3, Theorem 1.3] or [26, Theorem 2.12]. Hence

$$\frac{1}{2h}W_2^2(\chi_{E_n},\chi_{E_{n-1}})=h\int_{E_n}|\nabla\phi^n(x)|^2dx=h\int_{E_{n-1}}|\nabla\psi^n(y)|^2dy \tag{20}$$

and by (19)

$$h\sum_{n=n_0+1}^{n_1}\int_{E_n}|\nabla\phi^n(x)|^2dx\le\mathcal{E}(E_{n_0})-\mathcal{E}(E_{n_1})\le\mathcal{E}(E_0), \tag{21}$$

that is, if we set $\mathbf{u}^n:=\chi_{E_n}\nabla\phi^n$ and $\mathbf{u}_h(t):=\mathbf{u}^{[t/h]}$, then $\mathbf{u}_h$ is uniformly bounded in L^2. Let $\mathbf{u}=\mathbf{u}(x,t)$ be a weak limit. Since $(Id-h\nabla\phi^n)_\sharp\chi_{E_n}=\chi_{E_{n-1}}$, for $\eta(x,t)$ a smooth test function

$$\frac{1}{h}\int_{\mathbb{R}^d}(\chi_{E_n}-\chi_{E_{n-1}})\,\eta\,dx=\frac{1}{h}\int_{\mathbb{R}^d}\chi_{E_n}(x)\,(\eta(x)-\eta(x-h\nabla\phi^n(x)))\,dx.$$

Using Taylor's theorem in the form $\left|\eta(x)-\eta(x-h\xi)-h\xi\cdot\nabla\eta(x)\right|\le\frac{h^2}{2}|\xi|^2\sup_x|\nabla^2\eta|$, we can replace the right-hand side by

$$\int_{\mathbb{R}^d}\nabla\eta(x)\cdot\nabla\phi^n(x)\chi_{E_n}(x)dx$$

at the expense of the error

$$\frac{1}{h}\frac{h^2}{2}\sup|\nabla^2\eta|\int_{\mathbb{R}^d}|\nabla\phi^n|^2\chi_{E_n}\,dx=\sup|\nabla^2\eta|\frac{1}{2h}W_2^2(\chi_{E_n},\chi_{E_{n-1}}).$$

After integration in time, this error term vanishes as $h\downarrow 0$ because of (19) and we may pass to the limit in the time-integrated version of the above identity to obtain the continuity equation in form of (10).

Step 2: De Giorgi's interpolation and argument for (11). De Giorgi's variational interpolation

$$\tilde{E}_h((n-1)h+t) \in \arg\min_E \Big\{ \frac{1}{2t} W_2^2(\chi_E, \chi_{E_{n-1}}) + \mathcal{E}(E) \Big\} \tag{22}$$

satisfies the identity

$$\frac{h}{2}\Big(\frac{W_2(\chi_{E_n},\chi_{E_{n-1}})}{h}\Big)^2 + \frac{1}{2}\int_{(n-1)h}^{nh} \Big(\frac{W_2(\chi_{\tilde{E}_h(t+(n-1)h)},\chi_{E_{n-1}})}{t}\Big)^2 dt \le \mathcal{E}(E_{n-1}) - \mathcal{E}(E_n). \tag{23}$$

Although the proof is contained—in a more general context— in [2, Theorem 3.1.4], we repeat it here for the reader's convenience.

W.l.o.g. we may assume $n=1$; for notational convenience we also drop the index h for this short argument. Defining momentarily

$$f(t) := \frac{1}{2t} W_2^2(\chi_{\tilde{E}(t)}, \chi_{E_0}) + \mathcal{E}(\tilde{E}(t))$$

to be the minimal value in the variational problem (22), we may compute for $s<t$, using the minimality of $\tilde{E}(t)$,

$$\begin{aligned} f(t) - f(s) &\le \frac{1}{2t} W_2^2(\chi_{\tilde{E}(s)}, \chi_{E_0}) + \mathcal{E}(E(s)) - \frac{1}{2s} W_2^2(\chi_{\tilde{E}(s)}, \chi_{E_0}) - \mathcal{E}(E(s)) \\ &= \frac{s-t}{2st} W_2^2(\chi_{\tilde{E}(s)}, \chi_{E_0}). \end{aligned}$$

Since $s<t$, this implies

$$\frac{f(t)-f(s)}{t-s} \le -\frac{1}{2st} W_2^2(\chi_{\tilde{E}(s)}, \chi_{E_0}) \to -\frac{1}{2t^2} W_2^2(\chi_{\tilde{E}(t)}, \chi_{E_0}) \quad \text{as } s \uparrow t.$$

The analogous reverse inequality may be obtained by using $s>t$ in the above argument with the roles of s and t interchanged. Hence f is locally Lipschitz in $(0,h]$ with

$$\frac{d}{dt} f(t) = -\frac{1}{2t^2} W_2^2(\chi_{\tilde{E}(t)}, \chi_{E_0})$$

for almost every $t \in (0,h)$. For $\varepsilon > 0$, integrating this inequality from $t=\varepsilon$ to $t=h$, and then using lower semicontinuity w.r.t. the L^1 convergence $E(\varepsilon) \to E_0$ yields (23).

Summing (23) over n from n_0+1 to n_1 and telescoping the right-hand side we obtain the sharp energy dissipation inequality:

$$\frac{h}{2}\sum_{n=n_0+1}^{n_1}\Big(\frac{W_2(\chi_{E_n},\chi_{E_{n-1}})}{h}\Big)^2 + \frac{1}{2}\int_{n_0 h}^{n_1 h} \Big(\frac{W_2(\chi_{\tilde{E}_h(t)},\chi_{E_h(t)})}{t-h[t/h]}\Big)^2 dt \le \mathcal{E}(E_{n_0}) - \mathcal{E}(E_{n_1}). \tag{24}$$

By (20), we have:

$$\frac{h}{2}\sum_{n=1}^{N}\Big(\frac{W_2(\chi_{E_n},\chi_{E_{n-1}})}{h}\Big)^2 = \frac{h}{2}\sum_{n=1}^{N}\int |\chi_{E_n}\nabla\phi^n|^2\,dx$$

which implies

$$\int_0^T \int_{E(t)} |\mathbf{u}(x,t)|^2\,dx\,dt \le \liminf_{h\to 0} \frac{h}{2}\sum_{n=1}^{N}\Big(\frac{W_2(\chi_{E_n},\chi_{E_{n-1}})}{h}\Big)^2 \tag{25}$$

since $\chi_{E_n}\nabla\phi^n = \mathbf{u}^n$ and $\mathbf{u}_h \rightharpoonup \mathbf{u}$ in L^2.

Following the same strategy as in Step 1, we can show that $\tilde{\mathbf{u}}_h := \chi_{\tilde{E}_h(t)} \nabla \tilde{\phi}_h(t)$ with $x-(t-h[t/h])\nabla\tilde{\phi}_h(x,t)$ optimal in $W_2(\chi_{\tilde{E}_h(t)}, \chi_{E_h(t)})$—after passage to a subsequence—weakly converges to the same limit $\mathbf{u} = w - \lim_{h\downarrow 0} \mathbf{u}_h$. In particular, as before,

$$\frac{1}{t-h[t/h]} W_2^2(\chi_{\tilde{E}_h(t)}, \chi_{E_h(t)}) = (t-h[t/h]) \int |\nabla \tilde{\phi}|^2 dx$$

so that after division by $(t-h[t/h])$ and integration in t

$$\frac{1}{2}\int_0^T \left(\frac{W_2(\chi_{\tilde{E}_h(t)}, \chi_{E_h(t)})}{t-h[t/h]}\right)^2 dt = \frac{1}{2}\int_0^T \int |\nabla\tilde{\phi}_h|^2 dx\, dt = \frac{1}{2}\int_0^T \int |\tilde{\mathbf{u}}_h|^2 dx\, dt,$$

which is again lower semi-continuous. This concludes the argument for (11).

Step 3: Derivation of (12). The Euler-Lagrange equation of the minimization problem (1) reads

$$-\int_{E_n} \nabla\phi^n \cdot \xi\, dx = \int \left(\operatorname{div}\xi - \nu_{E_n} \cdot D\xi \nu_{E_n} + 2k * \chi_{E_n}\xi \cdot \nu_{E_n}\right) |\nabla \chi_{E_n}| \tag{26}$$

for all smooth test vector fields ξ with $\operatorname{div}\xi = 0$, where $\nu_{E_n} = \frac{\nabla\chi_{E_n}}{|\nabla\chi_{E_n}|}$ denotes the inner normal.

The nonlocal term

$$\int_0^\infty \int 2k * \chi_{E_h(t)}\xi \cdot \nu_{E_h(t)} \left|\nabla\chi_{E_h(t)}\right| dt = \int_0^\infty \int 2k * \chi_{E_h}\xi \cdot \nabla\chi_{E_h}$$

converges since $\nabla\chi_{E_h} \to \nabla\chi_E$ weakly as measures, and since $k * \chi_{E_h(t)}$ converges uniformly, which follows from the strong L^1 convergence $\chi_{E_h} \to \chi_E$ and the observation that

$$\sup |k * \chi - k * \tilde{\chi}| \leq \int |\chi - \tilde{\chi}|\, dx$$

for any two characteristic functions $\chi, \tilde{\chi}$ for which only the integrability $\int k = 1$ and non-negativity $k \geq 0$ are needed.

Since the measures $\mu^h = \mu_t^h dt = \delta_{\nu_{E_h(t)}(x)} \otimes \left|\nabla\chi_{E_h(t)}\right| dt$ are bounded (with moreover $\mu^h(\mathbb{S}^{d-1} \times \mathbb{R}^d \times I) \leq P(E_0)|I|$ for any $I \subset (0,+\infty)$ measurable), by Banach-Alaoglu, they have a weak-$*$ limit $\mu = \mu_t dt$ (after passage to a subsequence). Hence we can identify the limit of the first right-hand side term in (26) as well:

$$\begin{aligned}\lim_{h\downarrow 0} \int_0^\infty \int (\operatorname{div}\xi - \nu_{E_h} \cdot D\xi\, \nu_{E_h})|\nabla\chi_{E_h}|\, dt \\ = \lim_{h\downarrow 0} \int_0^\infty \int\int (\operatorname{div}\xi - \tilde{\nu}\cdot D\xi\, \tilde{\nu})\, d\mu_t^h(\tilde{\nu}, x) dt \\ = \int_0^\infty \int\int (\operatorname{div}\xi - \tilde{\nu}\cdot D\xi\, \tilde{\nu})\, d\mu_t(\tilde{\nu}, x) dt\end{aligned}$$

for any test vector field $\xi \in C_0^\infty(\mathbb{R}^d \times (0,+\infty), \mathbb{R}^d)$.

Step 4: Proof of the optimal energy dissipation relation (13). The local slope of $\mathcal{E}$, defined via

$$|\partial\mathcal{E}(E)| := \limsup_{F\to E} \frac{(\mathcal{E}(E) - \mathcal{E}(F))_+}{W_2(\chi_E, \chi_F)},$$

where the convergence of the sets $F \to E$ is to be understood with respect to W_2, satisfies

$$|\partial\mathcal{E}|(\tilde{E}((n-1)h+t)) \leq \frac{W_2(\chi_{\tilde{E}((n-1)h+t)}, \chi_{E_{n-1}})}{t},$$

cf. [2, Lemma 3.1.3]. Applying this to (24) yields the sharp energy dissipation inequality

$$\frac{h}{2}\sum_{n=n_0+1}^{n_1}\left(\frac{W_2(\chi_{E_n},\chi_{E_{n-1}})}{h}\right)^2+\frac{1}{2}\int_{n_0h}^{n_1h}|\partial\mathcal{E}|^2\left(\tilde{E}_h(t)\right)dt \le \mathcal{E}(E_h(n_0h))-\mathcal{E}(E_h(n_1h)). \tag{27}$$

Our goal is to pass to the limit in (27). We have already done this for the metric term in Step 2. We are left with

$$\int_0^T\int(\operatorname{div}\xi-\tilde{\nu}\cdot D\xi\,\tilde{\nu})d\mu_t(\tilde{\nu},x)dt+\int_0^T\int 2k*\chi_{E(t)}\,\xi\cdot\nu_{E(t)}\,|\nabla\chi_{E(t)}|\,dt -\frac{1}{2}\int_0^T\int_{E(t)}|\xi|^2\,dx\,dt\le\liminf_{h\to0}\frac{1}{2}\int_0^T|\partial\mathcal{E}|^2(\tilde{E}(t))\,dt \tag{28}$$

for all test vector fields ξ with $\operatorname{div}\xi=0$.

Given a set E with finite measure, any smooth divergence free vector field ξ provides a one-parameter family of candidates for the lim sup in the definition of the local slope $|\partial\mathcal{E}|(E)$ at E via the inner variations $\partial_s\chi_{E_s}+\xi\cdot\nabla\chi_{E_s}=0$. Therefore

$$\begin{aligned}\frac{1}{2}|\partial\mathcal{E}|^2(E)&\ge\lim_{s\to0}\frac{1}{2}\left(\frac{\frac{1}{s}\left(\mathcal{E}(E)-\mathcal{E}(E_s)\right)_+}{\frac{1}{s}W_2(\chi_E,\chi_{E_s})}\right)^2\\&\ge\lim_{s\to0}\frac{1}{s}\left(\mathcal{E}(E)-\mathcal{E}(E_s)\right)_+-\frac{1}{2s^2}W_2^2(\chi_E,\chi_{E_s})\end{aligned}$$

On the one hand, since ξ is divergence free, it generates one particular volume-preserving flow from E to E_s. More precisely, the rescaled field $s\xi$ solves $\partial_{s'}\chi_{E_{s'}}+\operatorname{div}(s\xi\chi_{E_{s'}})=0$ and transports E to E_s in one unit of time and hence provides a particular candidate for the minimum problem in W_2:

$$W_2^2(\chi_E,\chi_{E_s})\le\int_0^1\int_{E_{s'}}|s\xi|^2\,dx\,ds'=s^2\int_E|\xi|^2dx+o(s^2).$$

On the other hand, we have:

$$\begin{aligned}(\mathcal{E}(E_s)-\mathcal{E}(E))_+\ge&\,\mathcal{E}(E_s)-\mathcal{E}(E)\\=&\,s\frac{d}{ds}\Big|_{s=0}\mathcal{E}(E_s)+o(s)\\=&\,s\int(\operatorname{div}\xi-\nu\cdot D\xi\,\nu)\,d\mu_t(x,\tilde{\nu})\\&+2s\int k*\chi_E\,\xi\cdot\nu_{E(t)}\,|\nabla\chi_E|+o(s).\end{aligned}$$

Therefore:

$$|\partial\mathcal{E}|(E(t))\ge\int(\operatorname{div}\xi-\nu\cdot D\xi\,\nu)\,d\mu_t(x,\tilde{\nu}) +2\int k*\chi_E\xi\cdot\nu_E\,|\nabla\chi_E|-\frac{1}{2}\int_E|\xi|^2dx, \tag{29}$$

and taking the limit $h\to0$ yields (28). □

5. An interpolation inequality

Lemma 2. *There exists $C > 0$ such that for $u, v \in BV(\mathbb{R}^d)$ with $\int(|u| + |v|)|x|dx < +\infty$, one has*

$$\|u - v\|_{L^1}^2 \le C\|u - v\|_{W^{-1,1}}|D(u - v)|(\mathbb{R}^d).$$

Proof. We first consider, f, g smooth with compact support, a symmetric mollifier ρ, and $\sigma > 0$. We have

$$\int fg\,dx = \int g\,(\rho_\sigma * f)\,dx + \int g\,(f - \rho_\sigma * f)\,dx.$$

For the first integral we use $|\nabla(\rho_\sigma * g)| \le (\|\nabla\rho\|_{L^1}/\sigma)\|g\|_{L^\infty}$, hence by symmetry of ρ:

$$\int g\,(\rho_\sigma * f)\,dx = \int f\,(\rho_\sigma * g)\,dx \le \frac{C_1}{\sigma}\|f\|_{W^{-1,1}}\|g\|_{L^\infty},$$

with $C_1 = \|\nabla\rho\|_{L^1}$. For the second integral, we write that for $f \in BV(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$,

$$\int |f - \rho_\sigma * f|dx = \int \left|\int\int_0^\sigma \rho(\xi)\xi\cdot\nabla f(x - t\xi)\,dt\,d\xi\right| dx \le C_2\sigma\|\nabla f\|_{L^1},$$

where $C_2 = \int |\xi|\rho(\xi)d\xi$. We deduce that

$$\int g(f - \rho_\sigma * f)\,dx \le C_2\sigma\|g\|_\infty|Df|(\mathbb{R}^d).$$

Hence for any $\sigma > 0$

$$\int f\,g\,dx \le \|g\|_{L^\infty}\left(\frac{C_1}{\sigma}\|f\|_{W^{-1,1}} + C_2\sigma|D(u - v)|(\mathbb{R}^d)\right)$$

so that (minimizing the right-hand side in $\sigma > 0$)

$$\int f\,g\,dx \le 2\|g\|_{L^\infty}\sqrt{C_1C_2\|f\|_{W^{-1,1}}|Df|(\mathbb{R}^d)}.$$

Choosing g a mollification of sign f and passing to the limit it follows that

$$\int |f|\,dx \le 2\sqrt{C_1C_2\|f\|_{W^{-1,1}}|Df|(\mathbb{R}^d)}.$$

This extends to $f \in BV(\mathbb{R}^d)$ such that $\|f\|_{-1,1} = \sup_{|\nabla g|<1}\int fg\,dx < +\infty$ and $\int |x||f(x)|dx < +\infty$, by approximation. □

Corollary 1. *For any sets $E, F \subset \mathbb{R}^d$ with finite perimeter and with $\int_{E\cup F}|x|dx < +\infty$*

$$|E\triangle F| \lesssim \sqrt{P(E) + P(F)}\sqrt{W_1(\chi_E, \chi_F)}.$$

References

[1] Fred Almgren, Jean E. Taylor, and Lihe Wang. Curvature-driven flows: a variational approach. *SIAM Journal on Control and Optimization*, 31(2):387–438, 1993.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Birkhäuser, 2008.

[3] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44(4):375–417, 1991.

[4] Guillaume Carlier and Clarice Poon. On the total variation Wasserstein gradient flow and the TV-JKO scheme. *ESAIM: COCV, Forthcoming article*, forthcoming.

[5] Vicent Caselles and Antonin Chambolle. Anisotropic curvature-driven flow of convex sets. *Nonlinear Analysis: Theory, Methods & Applications*, 65(8):1547–1577, 2006.

[6] Ennio De Giorgi. Movimenti minimizzanti. In *Aspetti e problemi della Matematica oggi, Proc. of Conference held in Lecce*, 1992.
[7] Guido De Philippis and Tim Laux. Implicit time discretization for the mean curvature flow of mean convex sets. `arXiv:1806.02716[math.AP]`, 2018.
[8] Selim Esedoğlu and Felix Otto. Threshold dynamics for networks with arbitrary surface tensions. *Communications on Pure and Applied Mathematics*, 68(5):808–864, 2015.
[9] Julian Fischer, Sebastian Hensel, Tim Laux, and Thilo M. Simon. Weak-strong uniqueness for multiphase mean curvature flow. In preparation.
[10] Julian Fischer, Sebastian Hensel, Tim Laux, and Thilo M. Simon. Weak-strong uniqueness for the Mullins-Sekerka equation. In preparation.
[11] Harald Garcke and Stefan Schaubeck. Existence of weak solutions for the Stefan problem with anisotropic Gibbs-Thomson law. *Advances in Mathematical Sciences and Applications*, 21(1):255–283, 2011.
[12] Karl Glasner. A diffuse interface approach to Hele-Shaw flow. *Nonlinearity*, 16(1):49, 2002.
[13] Matt Jacobs, Inwon Kim, and Alpár R Mészáros. Weak solutions to the muskat problem with surface tension via optimal transport. `arXiv:1905.05370[math.AP]`, 2019.
[14] Richard Jordan, David Kinderlehrer, and Felix Otto. Free energy and the Fokker-Planck equation. *Phys. D*, 107(2-4):265–271, 1997. Landscape paradigms in physics and biology (Los Alamos, NM, 1996).
[15] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
[16] Christiane Kraus. The degenerate and non-degenerate Stefan problem with inhomogeneous and anisotropic Gibbs-Thomson law. *European Journal of Applied Mathematics*, 22(5):393–422, 2011.
[17] Tim Laux and Felix Otto. Convergence of the thresholding scheme for multi-phase mean-curvature flow. *Calculus of Variations and Partial Differential Equations*, 55(5):129, 2016.
[18] Tim Laux and Felix Otto. Brakke's inequality for the thresholding scheme. `arXiv:1708.03071[math.AP]`, 2017.
[19] Philippe Logaritsch. *An Obstacle Problem for Mean Curvature Flow*. PhD thesis, University of Leipzig, 2016.
[20] Stephan Luckhaus. Solutions for the two-phase Stefan problem with the Gibbs-Thomson law for the melting temperature. In *Fundamental Contributions to the Continuum Theory of Evolving Phase Interfaces in Solids*, pages 317–327. Springer, 1999.
[21] Stephan Luckhaus and Thomas Sturzenhecker. Implicit time discretization for the mean curvature flow equation. *Calculus of Variations and Partial Differential Equations*, 3(2):253–271, 1995.
[22] B. Merriman, J. K. Bence, and S. J. Osher. Diffusion generated motion by mean curvature. In J. E. Taylor, editor, *"Computational Crystal Growers Workshop"*, Selected Lectures in Math. Amer. Math. Soc., Providence, Rhode Island, 1992.
[23] Barry Merriman, James K. Bence, and Stanley J. Osher. Motion of multiple junctions: A level set approach. *Journal of Computational Physics*, 112(2):334–363, 1994.
[24] Felix Otto. Dynamics of labyrinthine pattern formation in magnetic fluids: A mean-field theory. *Archive for Rational Mechanics and Analysis*, 141(1):63–103, 1998.
[25] Matthias Röger. Existence of weak solutions for the Mullins–Sekerka Flow. *SIAM Journal on Mathematical Analysis*, 37(1):291–301, 2005.
[26] Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

A.C.: CMAP, Ecole Polytechnique, CNRS, 91128 Palaiseau, France
Email address: `antonin.chambolle@cmap.polytechnique.fr`

T.L.: Department of Mathematics, University of California, Berkeley, CA 94720-3840 USA
Email address: `tim.laux@math.berkeley.edu`