Ann. Inst. H. Poincaré C
Anal. Non Linéaire 41 (2024), 1527–1594
DOI 10.4171/AIHPC/129

On a degenerate elliptic problem arising in the least action principle for Rayleigh–Taylor subsolutions

Björn Gebhard, Jonas Hirsch, and József J. Kolumbán

Abstract. We address a degenerate elliptic variational problem arising in the application of the least action principle to averaged solutions of the inhomogeneous Euler equations in the Boussinesq approximation emanating from the horizontally flat Rayleigh–Taylor configuration. We give a detailed derivation of the functional starting from the differential inclusion associated with the Euler equations, i.e. the notion of an averaged solution is that of a subsolution in the context of convex integration, and illustrate how it is linked to the generalized least action principle introduced by Brenier [J. Amer. Math. Soc. 2 (1989), 225–255; in: New trends and results in mathematical description of fluid flows (2018), 53–75]. Concerning the investigation of the functional itself, we use a regular approximation in order to show the existence of a minimizer enjoying partial regularity, as well as other properties important for the construction of actual Euler solutions induced by the minimizer. Furthermore, we discuss to what extent such an application of the least action principle to subsolutions can serve as a selection criterion.

1. Introduction

Since its introduction in 2009 by De Lellis and Székelyhidi [18] in the context of fluid dynamics, the method of convex integration has been a powerful tool to show ill-posedness of initial value problems and to provide counterexamples to the conservation of physical quantities in a low enough regularity regime; we refer to [6,20] for recent surveys. Besides being an engine for counterexamples, due to their highly oscillatory nature, the solutions obtained by convex integration shortly after also began to be utilized to describe turbulent behavior in situations where a regular solution simply cannot exist due to irregular initial data. Examples include vortex sheets in the homogeneous two-dimensional Euler equations [36, 43], as well as the Muskat problem for the incompressible porous media equation [7–9, 15, 28, 32, 35, 38, 44], and the horizontally flat Rayleigh–Taylor instability in the inhomogeneous Euler equations [29, 30].

The existence of solutions emanating in the stated situations relies on a general convex integration theorem for the corresponding system, saying that a subsolution, which can be seen as an averaged solution, induces infinitely many turbulent solutions that are close in

Mathematics Subject Classification 2020: 35J70 (primary); 35J75, 76E17 (secondary).
Keywords: least action principle, degenerate elliptic variational problem, Euler equations, Boussinesq equation, Rayleigh–Taylor instability, subsolutions, generalized flows, convex integration.

a weak sense to the subsolution; see Theorem 2.1 below for example. Therefore, having such a theorem at hand, it remains to construct a suitable subsolution. However, typically there are plenty of admissible subsolutions emanating from the same initial data, i.e. also on the level of averaged motion, a vast number of evolutions differing, for instance, in the quantitative size of the induced turbulence of the solutions, are possible. One therefore has to choose a particular subsolution based on a meaningful selection criterion. Up to now, essentially two strategies have been used: reduction after some ansatzes to a hyperbolic conservation law and selection of the unique entropy solution [30, 43, 44], and in the case of the Euler equations short-time selection by means of maximal initial energy dissipation [29, 36].

In the case of the incompressible porous media equation with flat initial configuration, the former strategy has been applied by Székelyhidi [44] yielding a family of subsolutions out of which the one with "maximal mixing" coincides with the unique solution of a different relaxation given by Otto [39] based on gradient flows. Regarding nonflat initial interfaces, local-in-time subsolutions of different types have been constructed in [7,28,38], also in the only partially unstable case [9]. Properties of the subsolutions selected in the flat case by either of the above-mentioned relaxations have been incorporated in these constructions; see for instance [7, Remark 4.2].

The latter criterion, i.e. maximal initial energy dissipation in the case of the Euler equations, was motivated by the entropy rate admissibility criterion of Dafermos [16]. It has also been discussed in the context of solutions obtained by convex integration for the compressible Euler equations [12, 27].

Motivated by the search for a global-in-time selection criterion and the well-known fact that the Euler equations can formally be derived from the least action principle (see Section 2 for more details), the present article originates from the question of what happens if one imposes the least action principle on the level of subsolutions.

In the current paper we follow this question in the setting of the flat Rayleigh–Taylor instability modeled by the Euler equations in the Boussinesq approximation, while in general we believe that similar research can be extended to other models and configurations. The setting here has been chosen due to the presence of multiple symmetries (flat initial data, normalization of ρ to ± 1). In detail, we consider

$$\begin{aligned} \partial_t v + \operatorname{div}(v \otimes v) + \nabla p &= -\rho g A e_n, \\ \operatorname{div} v &= 0, \\ \partial_t \rho + \operatorname{div}(\rho v) &= 0, \end{aligned} \tag{1.1}$$

stated on $(0,T) \times \mathcal{D}$ with $T > 0$, $\mathcal{D} := (0,1)^{n-1} \times (-L, L) \subset \mathbb{R}^n$, and with initial data

$$\rho(0,x) = \operatorname{sign}(x_n), \quad v(0,x) = 0, \quad x \in \mathcal{D}. \tag{1.2}$$

On the boundary of $\mathcal{D}$ we set the usual no-penetration boundary condition,

$$v \cdot \nu = 0 \quad \text{on } \partial\mathcal{D} \times [0,T), \tag{1.3}$$

in order to complement the incompressibility condition. Here, ν denotes the exterior unit normal of $\partial\mathcal{D}$.

System (1.1), (1.2), (1.3) models two incompressible ideal fluids with homogeneous densities $0 < \rho_- < \rho_+$, with $\rho_+ - \rho_-$ small, initially at rest, and separated by a horizontally flat interface under the influence of gravity, i.e. in one of the most classical occurrences of the Rayleigh–Taylor instability. The unknowns are the normalized fluid density $\rho\colon [0,T) \times \mathcal{D} \to \mathbb{R}$, i.e. $\rho \in \{\pm 1\}$ a.e., the velocity field $v\colon [0,T) \times \mathcal{D} \to \mathbb{R}^n$, and the pressure of the fluid $p\colon [0,T) \times \mathcal{D} \to \mathbb{R}$. Furthermore, $e_n \in \mathbb{R}^n$ is used to denote the nth coordinate vector, $g > 0$ the gravitational constant, and $A := \frac{\rho_+ - \rho_-}{\rho_+ + \rho_-}$ is the Atwood number.

The article splits into two essentially independent parts. The first part, consisting of Section 2 and complemented by Appendices A, B, addresses the application of the least action principle. In Section 2 we first of all recall the relaxation of (1.1) seen as a differential inclusion and then illustrate how the least action principle imposed on a suitable class of one-dimensional subsolutions gives rise to a variational problem. In Appendix A we show how this problem relates to the relaxation of the least action principle by Brenier [2, 4] in terms of generalized flows. Appendix B contains details regarding a needed variation of the usual convex integration result stated in Theorem 2.1.

The variational problem derived in Section 2 is of the type

$$\text{minimize } \mathcal{A}(u) \text{ for } u \in X, \tag{1.4}$$

where the functional $\mathcal{A}$ is given by

$$\mathcal{A}(u) = \int_\Omega F(\nabla u) - V(x_2, u)\, dx \tag{1.5}$$

with $\Omega := (0,T) \times (-L, L) \subset \mathbb{R}^2$, $F\colon \mathbb{R}^2 \to [0, +\infty]$,

$$F(p) := \begin{cases} 0 & \text{if } p_1 = 0, \\ +\infty & \text{if } p_1 \neq 0, |p_2| \geq 1, \\ \dfrac{p_1^2}{2(1-p_2^2)} & \text{otherwise,} \end{cases} \tag{1.6}$$

and $V\colon [-L, L] \times \mathbb{R} \to \mathbb{R}$, $(x_2, z) \mapsto V(x_2, z)$ is a suitable nonlinear potential. To give an example, the reader may think of

$$V(x_2, z) = -gAz + \frac{3gA}{4L}(z - (|x_2| - L))^2. \tag{1.7}$$

The (affine) space of functions under consideration reads

$$X := \big\{u \in H^1(\Omega) : u(x_1, \pm L) = 0,\ u(0, x_2) = -u(T, x_2) = |x_2| - L\big\}. \tag{1.8}$$

The functional $\mathcal{A}$ is elliptic but degenerate in the sense that the minimal eigenvalue of $D^2F(p)$ vanishes for $p_1 = 0$, while the maximal eigenvalue of $D^2F(p)$ becomes infinite as $|p_2|$ approaches 1 from below. The functional will be further introduced and investigated in Sections 3–7, which form the second part of our paper. A reader only interested in the variational problem itself might directly jump to these sections, which are readable without the background given in Section 2. We would like to point out though that some aspects of our investigation, in particular in Section 7, stem from the usability of the minimizer as a subsolution for the Boussinesq system (1.1), (1.2), (1.3).

The two parts of our investigation are brought together again in the final Section 8, where we summarize our results and discuss some open problems. A rough version of our main theorem can be written here as follows.

Theorem 1.1. *Under suitable conditions on the potential V, problem (1.4) has a minimizer $u \in X \cap \mathscr{C}^0(\bar{\Omega})$ enjoying $\mathscr{C}^2$ regularity on a nonempty open set $\Omega' \subset \Omega$ of which every connected component is simply connected. On Ω' there holds $\partial_{x_1} u > 0$, $|\partial_{x_2} u| < 1$, while outside Ω' we have $\partial_{x_1} u = 0$, $|\partial_{x_2} u| \leq 1$ a.e. The gradient of u can be used to define a one-dimensional subsolution of (1.1), (1.2), (1.3) and it thus induces infinitely many solutions via convex integration.*

As stated, a more detailed version including statements on energy dissipation and attainment of initial and boundary data can be found in Section 8. There we also summarize the ansatzes we make and reflect upon our initial question regarding a selection criterion for subsolutions emanating from Rayleigh–Taylor initial data.

2. The action functional for subsolutions

The main point of this section is the derivation of the variational problem (1.4). We begin with a short review of related previous work before recalling the notion of a subsolution.

2.1. The Boussinesq system and previous results

For a sufficiently regular initial data system (1.1), (1.3) is locally well posed; see [10, 17, 23], where in [23] it also has been shown that finite-time singularity formation occurs for smooth data. We note however that the initial data of our interest (1.2), with ρ_0 being only essentially bounded, does not fall into the regularity classes considered in [10, 17, 23].

Contrary to local well-posedness for lower regularity classes the existence of infinitely many weak solutions can be shown by means of convex integration. Relying on the convex integration method for the homogeneous Euler equations from [18], the existence of infinitely many solutions with compact space-time support for (1.1) without the influence of gravity, i.e. $g = 0$, has been shown in [5]. Moreover, the paper [13] addresses system (1.1) under the additional influence of the Coriolis force in the momentum balance and dissipation in the continuity equation. For this dissipative Boussinesq system, it is shown that for a given initial density $\rho_0 \in \mathscr{C}^2 \cap L^\infty$, there exists an irregular initial velocity field v_0 inducing infinitely many solutions which are admissible in the sense that, for almost every positive time, the total energy of the solutions does not exceed the total initial energy. This is an important property both from a physical point of view and for the mathematical weak-strong uniqueness property; see [49] for an overview and [11] for the particular case of system (1.1).

In order to obtain existence of turbulently mixing solutions emanating from the actual Rayleigh–Taylor interface (1.2), Székelyhidi and the first and third authors have established in [30] the full relaxation of the inhomogeneous incompressible Euler equations (without the Boussinesq approximation) allowing them to construct admissible solutions to the corresponding initial value problem under the condition that the quotient of the two fluid densities satisfies

$$\frac{\rho_+}{\rho_-} \geq \left(\frac{4+2\sqrt{10}}{3}\right)^2.$$

This translates to an Atwood number $A \geq 0.845$, i.e. in the so-called "ultra-high" regime.

In view of this Atwood number condition, the first and third authors thereafter addressed in [29] the Euler equations in the Boussinesq approximation, which is applicable for low Atwood number. In the latter paper the full relaxation could also be explicitly given and, in contrast to [30], admissible turbulent solutions for (1.1), (1.2), (1.3) be constructed, respectively selected, without restrictions on the size of $A > 0$. Of course, the equations cannot be seen as a reasonable physical system for larger A. The selection of these subsolutions (cf. Definition 2.1 below) is based on imposing maximal initial energy dissipation in the class of one-dimensional self-similar subsolutions. Imposing these requirements leads to a variational problem for the self-similar density profile and the initial speed of the opening of the mixing zone. The problem could be solved explicitly, giving a possible small-time selection of subsolutions within the stated class.

While in [29] the first and third authors achieved the computation of the full relaxation of the Boussinesq system as a differential inclusion and the construction of first examples of subsolutions by means of the previously described short-time selection criterion, the present article in contrast picks up at the established relaxation and explores the utilization of the least action principle as a global-in-time selection criterion that is consistent with the underlying geometric structure of the Euler equations. Unlike [29], this approach for example allows us to tie connections between convex integration subsolutions and Brenier's 1989 relaxation [2]; see Appendix A for details. The resulting variational problem, and consequently its analysis, is entirely different to that for the self-similar subsolutions considered in [29].

In the following subsections we will derive this problem, i.e. problem (1.4), by applying the least action principle to the relaxation given in [29]. However, on the topic of connections between variational principles and convex integration subsolutions, we would first like to add a short comparison to another paper by Brenier [3].

The least action principle gives a way to derive the Euler equations. However, if in practice one would like to utilize the principle itself, it has the disadvantage that one has to specify not only the initial configuration but also a target or final configuration; see Section 2.6 for example. Brenier [3] designs a variational principle for the homogeneous

Euler equations ($\rho \equiv 1$, $g = 0$) that gets by solely with the initial condition, by adding the Euler equations themselves as a constraint. After reformulations he shows that the corresponding dual problem, which is a concave maximization problem, always admits a solution. Regarding consistency, he shows that smooth solutions to the Euler equations indeed give rise to maximizers of the dual problem.

Moreover, for a continuous initial velocity field V_0, Brenier also relates the optimal value of this concave problem to the infimum of a functional considered over all (non-strict) subsolutions emanating from V_0. These two values coincide [3, Theorem 2.6]. After some reformulations one can see that this functional is the kinetic energy and thus, in the homogeneous case, the action associated with the subsolutions. However, the analysis of [3] only states that the infimum of the "initial value subsolution action" equals the maximum of the designed dual functional for the Euler initial value problem. In particular, it does not give the existence of an optimal subsolution minimizing the action, or any strict subsolution emanating from V_0.

To the best of our knowledge, [3] is the only instance prior to the present article where an action functional for convex integration subsolutions of homogeneous or inhomogeneous Euler equations appears in some form.

2.2. Relaxation as a differential inclusion

We first of all rephrase (1.1) as a differential inclusion. As before, let $n \geq 2$, $L > 0$, $T > 0$, $\mathcal{D} = (0,1)^{n-1} \times (-L, L)$ and set $Z := \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_0^{n\times n} \times \mathbb{R}$, where $\mathcal{S}_0^{n\times n}$ denotes the set of symmetric trace-free matrices.

Consider the linear system

$$\begin{aligned} \partial_t v + \operatorname{div} \sigma + \nabla p &= -\rho g A e_n, \\ \operatorname{div} v &= 0, \\ \partial_t \rho + \operatorname{div} m &= 0, \end{aligned} \tag{2.1}$$

in $(0,T) \times \mathcal{D}$, with boundary conditions

$$m \cdot \nu = 0, \quad v \cdot \nu = 0 \tag{2.2}$$

on $(0,T) \times \partial\mathcal{D}$, as well as for given functions $e_0, e_1 \colon (0,T) \times \mathcal{D} \to \mathbb{R}$ with $e_0 \pm e_1 \geq 0$ the family of sets $K_{(t,x)} \subset Z$, $(t,x) \in (0,T) \times \mathcal{D}$ defined by $(\rho, v, m, \sigma, p) \in K_{(t,x)}$ if and only if

$$|\rho| = 1, \quad m = \rho v, \quad v \otimes v - \sigma = (e_0(t,x) + \rho e_1(t,x)) \operatorname{id}. \tag{2.3}$$

It is easy to see that if a tuple $z := (\rho, v, m, \sigma, p)$ of L^1_{loc} functions satisfies system (2.1), (2.2) distributionally and if there holds

$$z(t,x) \in K_{(t,x)} \tag{2.4}$$

for almost every $(t,x) \in (0,T) \times \mathcal{D}$, then (ρ, v) is a solution of the inviscid Boussinesq system with local energy density function

$$\begin{aligned} \mathcal{E}(t,x) &= \frac{1}{2}|v(t,x)|^2 + \rho(t,x) g A x_n \\ &= \frac{n}{2}(e_0(t,x) + \rho(t,x) e_1(t,x)) + \rho(t,x) g A x_n. \end{aligned} \tag{2.5}$$

Similarly in the other direction, if $(\rho, v) \in L^\infty((0,T)\times\mathcal{D}) \times L^2((0,T)\times\mathcal{D})$ solves (1.1), (1.3) and $\rho \in \{-1, 1\}$ a.e., then one can set $m = \rho v$, $\sigma = (v \otimes v)^\circ$ to see that there exist e_0, e_1, and a pressure, such that one obtains a solution to the differential inclusion (2.1), (2.2), (2.4) with $K_{(t,x)}$ defined with respect to the functions e_0, e_1.

Note that at this point there is no unique or canonical choice for e_0 and e_1. As can be seen from (2.5) they dictate the form of the local kinetic energy and should be continuous for the application of convex integration; cf. Theorem 2.1 below. The introduction of two such continuous functions in [29] instead of only one in the case of the homogeneous Euler equations [19] allows the kinetic energy of the solutions to oscillate along the oscillations of the density ρ. This additional flexibility turned out to be advantageous in setting up variational problems on the level of subsolutions and will also be exploited in Sections 2.3, 2.4 below.

For the relaxation of the Boussinesq system in the sense of differential inclusions, condition (2.4) is replaced by requiring the tuple $z(t,x)$ to take values in the interior of the convex hull $K^{co}_{(t,x)}$, or more generally of the Λ-convex hull $K^{\Lambda}_{(t,x)}$, instead. In the case of (2.1), (2.4) the two notions of convex hulls coincide and its interior is given by

$$\begin{aligned} U_{(t,x)} := \Big\{ z = (\rho, v, m, \sigma, p) : |\rho| < 1,\ & \tfrac{|m+v|^2}{n(\rho+1)^2} < e_0(t,x) + e_1(t,x), \\ & \tfrac{|m-v|^2}{n(\rho-1)^2} < e_0(t,x) - e_1(t,x), \\ & \lambda_{\max}\Big(\tfrac{(m-\rho v)\otimes(m-\rho v)}{1-\rho^2} + v \otimes v - \sigma \Big) < e_0(t,x) + \rho e_1(t,x) \Big\}; \end{aligned} \tag{2.6}$$

see [29, Proposition 3.6].

Definition 2.1. We say that $z = (\rho, v, m, \sigma, p)$ is a subsolution with respect to a pair of measurable functions e_0, e_1 iff $\rho \in L^\infty((0,T)\times\mathcal{D})$, $v, m \in L^2((0,T)\times\mathcal{D})$, $\sigma \in L^1((0,T)\times\mathcal{D})$, p is a distribution, z satisfies the linear system (2.1) with boundary and initial data (2.2), (1.2), there holds

$$e_0 + \rho e_1 \in L^1((0,T)\times\mathcal{D}), \quad (1-\rho^2) e_1 \in L^1((0,T)\times\mathcal{D}), \tag{2.7}$$

$$\rho e_1 \le 0 \text{ a.e. on } (0,T)\times\mathcal{D}, \tag{2.8}$$

and in addition there exists an open set $\mathcal{U} \subset (0,T)\times\mathcal{D}$ such that

(a) the functions $e_0, e_1, \rho, v, m, \sigma$ are continuous on $\mathcal{U}$, and $z(t,x) \in U_{(t,x)}$ for every $(t,x) \in \mathcal{U}$;

(b) $z(t,x) \in K_{(t,x)}$ for almost every $(t,x) \in ((0,T)\times\mathcal{D}) \setminus \mathcal{U}$.

We call the set $\mathcal{U}$ the mixing zone associated with z. Of course, the definition extends to initial conditions other than (1.2) compatible with $|\rho_0| = 1$ a.e. and $\operatorname{div} v_0 = 0$ in $\mathcal{D}$, $v_0 \cdot \nu = 0$ on $\partial\mathcal{D}$.

Note that condition (2.7) is a slight generalization of the one from [29] where it was assumed that e_0, e_1 are essentially bounded. This is done in view of the subsolutions we will obtain in this article for which it is not clear that even $e_0, e_1 \in L^1((0,T)\times\mathcal{D})$ holds true. However, condition (2.8) allows these less integrable subsolutions to still induce infinitely many weak solutions whose energy can be controlled. More precisely, there holds the following convex integration theorem, whose proof, i.e. the necessary modifications due to (2.7) instead of $e_0, e_1 \in L^\infty$, can be found in Appendix B.

Theorem 2.1. *Given an arbitrary error function $\delta\colon [0,T] \to \mathbb{R}$ with $\delta(0) = 0$ and $\delta(t) > 0$ for $t > 0$, and a subsolution z_{sub}, there exist infinitely many solutions $(\rho_{\mathrm{sol}}, v_{\mathrm{sol}})$ of (1.1), (1.2), (1.3) coinciding with $(\rho_{\mathrm{sub}}, v_{\mathrm{sub}})$ outside $\mathcal{U}$, and on $\mathcal{U}$ having the local energy density*

$$\begin{aligned}\mathcal{E}_{\mathrm{sol}}(t,x) &= \frac{n}{2}(e_0(t,x) + \rho_{\mathrm{sol}}(t,x)e_1(t,x)) + \rho_{\mathrm{sol}}(t,x)gAx_n \\ &= \frac{n}{2}(e_0(t,x) + \rho_{\mathrm{sub}}(t,x)e_1(t,x)) + \rho_{\mathrm{sub}}(t,x)gAx_n \\ &\quad + \mathcal{E}^1_\delta(t,x) + \mathcal{E}^2_\delta(t,x),\end{aligned} \tag{2.9}$$

with $\mathcal{E}^1_\delta := \frac{n}{2}e_1(\rho_{\mathrm{sol}} - \rho_{\mathrm{sub}})$, $\mathcal{E}^2_\delta := gAx_n(\rho_{\mathrm{sol}} - \rho_{\mathrm{sub}})$ satisfying

$$\left|\int_{\mathcal{D}} \mathcal{E}^i_\delta(t,x)\,dx\right| \le \delta(t) \quad \text{for a.e. } t \in (0,T),\ i = 1,2. \tag{2.10}$$

Moreover, the solutions are found arbitrarily close to $(\rho_{\mathrm{sub}}, v_{\mathrm{sub}})$ in the weak $L^2(\mathcal{U})$ topology.

Remark 2.2. Conditions (2.9), (2.10) allow us to conclude that the induced solutions are weakly admissible provided that the subsolution satisfies

$$\int_{\mathcal{D}} \frac{n}{2}(e_0(t,x) + \rho_{\mathrm{sub}}(t,x)e_1(t,x)) + \rho_{\mathrm{sub}}(t,x)gAx_n\,dx < \int_{\mathcal{D}} \rho_0(x)gAx_n\,dx \tag{2.11}$$

for a.e. $t \in (0,T)$. Here, the right-hand side is precisely the total initial energy associated with (1.2).

Remark 2.3. The solutions $(\rho_{\mathrm{sol}}, v_{\mathrm{sol}})$ given by Theorem 2.1 satisfy $|\rho_{\mathrm{sol}}(t,x)| = 1$ for a.e. $(t,x) \in (0,T)\times\mathcal{D}$. This is a consequence of (2.3), where this condition has been imposed as part of the differential inclusion consistent with the transport equation in (1.1) and the initial data ρ_0. Note however that the conclusion of Theorem 2.1 remains valid if $(\rho_{\mathrm{sub}}, v_{\mathrm{sub}})$ satisfies

(b′) $z(t,x) \in K_{(t,x)}$ or $\rho(t,x) \in (-1,1)$, $(v, m, \sigma, e_0, e_1)(t,x) = (0,0,0,0,0)$ for almost every $(t,x) \in ((0,T)\times\mathcal{D}) \setminus \mathcal{U}$, instead of Definition 2.1 (b). This is indeed true because the convex integration (cf. Appendix B) is carried out on $\mathcal{U}$ and the specified 0-state is already a solution to the Boussinesq system. The difference is that the induced solutions $(\rho_{\mathrm{sol}}, v_{\mathrm{sol}})$ may have regions of positive Lebesgue measure where the fluid is at rest, $v_{\mathrm{sol}} = 0$, but $|\rho_{\mathrm{sol}}| < 1$. We refer to this slightly relaxed notion of subsolutions as "subsolutions with mixed resting regions".

2.3. One-dimensional subsolutions

Since the initial data (1.2) that we consider only depends on x_n, one can consider subsolutions which are obtained by averaging solutions in all other spatial directions, i.e.

$$z(t, x_n) := \int_{[0,1]^{n-1}} z_{\mathrm{sol}}(t, x)\, d(x_1, \ldots, x_{n-1}).$$

One can then easily see that $\operatorname{div} v = 0$ and $v \cdot \nu = 0$ imply $v \equiv 0$, as well as $\partial_{x_j} m = 0$ for $j = 1, \ldots, n-1$ and $m \cdot \nu = 0$ imply $m_j = 0$ for $j = 1, \ldots, n-1$.

Definition 2.4. A subsolution $z = (\rho, v, m, \sigma, p)$ is called a one-dimensional subsolution provided $z(t, x)$ depends only on $(t, x_n) \in (0, T) \times (-L, L)$ and $v \equiv 0$, $m_j \equiv 0$, $j = 1, \ldots, n-1$.

Lemma 2.2. *Let z be a one-dimensional subsolution with respect to e_0, e_1. Then $\rho(t, x) = \rho(t, x_n)$ and $m(t, x) = m_n(t, x_n)e_n$ enjoy the following properties:*

(i) $\partial_t \rho + \partial_{x_n} m_n = 0$ *weakly,*

(ii) $m_n(\cdot, \pm L) = 0$, $\rho(0, \cdot) = \operatorname{sign}$ *weakly,*

(iii) *there exists $\mathcal{U}' \subset (0, T) \times (-L, L)$ open such that $\rho, m_n \in \mathcal{C}^0(\mathcal{U}')$, $|\rho| < 1$ on $\mathcal{U}'$,*

(iv) $|\rho| = 1$, $m_n = 0$ *a.e. outside* $\mathcal{U}'$,

(v) $\frac{m_n^2}{1-\rho^2} \in L^1(\mathcal{U}')$,

(vi) $n(e_0 + \rho e_1) > \frac{m_n^2}{1-\rho^2}$ *on* $\mathcal{U}'$.

Conversely, let $\rho, m_n \colon (0, T) \times (-L, L) \to \mathbb{R}$ be measurable functions satisfying properties (i)–(v) and $\tilde{e} \in \mathcal{C}^0(\mathcal{U}') \cap L^1(\mathcal{U}')$, $\tilde{e} > 0$. Then for suitable e_0, e_1 the pair (ρ, m_n) induces a one-dimensional subsolution z with $\rho(t, x) = \rho(t, x_n)$, $m(t, x) = m_n(t, x_n)e_n$, $\mathcal{U} := \{(t, x) : (t, x_n) \in \mathcal{U}'\}$ and kinetic energy density given by

$$\frac{n}{2}(e_0(t, x) + \rho(t, x)e_1(t, x)) = \frac{|m_n(t, x_n)|^2}{2(1 - \rho(t, x_n)^2)} + \tilde{e}(t, x_n) \tag{2.12}$$

for $(t, x) \in \mathcal{U}$.

Remark 2.5. The weak notion of solution in Lemma 2.2 (i), (ii) is understood in the sense that

$$\int_0^T \int_{-L}^{L} \rho \partial_t \varphi + m_n \partial_{x_n} \varphi \, dx_n\, dt + \int_{-L}^{L} \operatorname{sign}(x_n) \varphi(0, x_n)\, dx_n = 0 \tag{2.13}$$

for all $\varphi \in \mathcal{C}_c^\infty([0, T) \times [-L, L])$.

Remark 2.6. If (ρ, m_n) satisfies (i)–(iii), (v), but instead of (iv) only

$$m_n = 0, \quad |\rho| \leq 1 \text{ a.e. outside } \mathfrak{U}',$$

then (ρ, m_n) induces a one-dimensional subsolution with mixed resting regions as described in Remark 2.3.

Proof of Lemma 2.2. If z is a one-dimensional subsolution with respect to e_0, e_1 and mixing zone $\mathfrak{U}$, then properties (i)–(iv) clearly hold true for $\mathfrak{U}'$ being the projection of $\mathfrak{U}$. Moreover, (v) follows from (vi). Thus it remains to prove (vi).

By definition of a subsolution (cf. (2.6)), there holds

$$\frac{|m|^2}{n(\rho+1)^2} < e_0 + e_1, \quad \frac{|m|^2}{n(\rho-1)^2} < e_0 - e_1, \tag{2.14}$$

inside the mixing zone $\mathfrak{U}$. It follows that

$$\begin{aligned} n(e_0 + \rho e_1) &= n\Big(\frac{1+\rho}{2}(e_0+e_1) + \frac{1-\rho}{2}(e_0-e_1)\Big) \\ &> n\Big(\frac{1+\rho}{2}\frac{|m|^2}{n(\rho+1)^2} + \frac{1-\rho}{2}\frac{|m|^2}{n(\rho-1)^2}\Big) = \frac{m_n^2}{1-\rho^2}. \end{aligned}$$

Now let ρ, m_n, and $\tilde{e}$ be given as stated. We will define z and suitable e_0, e_1 in terms of these three functions. We set $\mathfrak{U} := \{(t,x) \in (0,T) \times \mathscr{D} : (t, x_n) \in \mathfrak{U}'\}$.

Since v has to be 0 throughout $(0,T) \times \mathscr{D}$ for a one-dimensional subsolution, we have to set $m = \rho v = 0$, $\sigma = (v \otimes v)^\circ = 0$, and $n(e_0 + \rho e_1) = |v|^2 = 0$, hence $e_0 = -\rho e_1$, outside the mixing zone $\mathfrak{U}$. Without loss of generality we set $e_0 = e_1 = 0$ on $((0,T) \times \mathscr{D}) \setminus \mathfrak{U}$. In consequence, for a.e. $(t,x) \notin \mathfrak{U}$ there holds $z(t,x) \in K_{(t,x)}$ with $K_{(t,x)}$ defined in (2.3) for $e_0 = e_1 = 0$.

On the other hand, inside the mixing zone we of course set $\rho(t,x) = \rho(t,x_n)$, $m(t,x) = m_n(t,x)e_n$ and observe that for $z(t,x) \in U_{(t,x)}$ it remains to satisfy

$$\frac{|m|^2}{n(\rho+1)^2} < e_0(t,x) + e_1(t,x), \quad \frac{|m|^2}{n(\rho-1)^2} < e_0(t,x) - e_1(t,x),$$
$$\lambda_{\max}\Big(\frac{|m|^2}{1-\rho^2} e_n \otimes e_n - \sigma\Big) < e_0(t,x) + \rho e_1(t,x).$$

First of all we claim that the third inequality automatically holds true provided the first two inequalities are valid and we define $\sigma \in \mathcal{S}_0^{n\times n}$ such that

$$\frac{|m|^2}{1-\rho^2} e_n \otimes e_n - \sigma = \frac{|m|^2}{n(1-\rho^2)} \,\mathrm{id}\,.$$

Indeed, in that case, as just shown, there holds

$$\lambda_{\max}\Big(\frac{|m|^2}{1-\rho^2} e_n \otimes e_n - \sigma\Big) = \frac{|m|^2}{n(1-\rho^2)} < e_0 + \rho e_1.$$

Moreover, turning to the set of linear equations (2.1) one has

$$\partial_t v + \operatorname{div} \sigma + \rho g A e_n = (\partial_{x_n} \sigma_{nn} + \rho g A) e_n,$$

which can always be written as $-\nabla p$ by setting

$$p(t,x) := -\sigma_{nn}(t,x) - \int_0^{x_n} \rho(t, x_n')\, dx_n'\, gA.$$

The other two equations in (2.1) with correct initial and boundary data hold true by assumption.

In order to have a subsolution, it therefore only remains to find e_0, e_1 such that the two inequalities (2.14) are valid. It is easy to check that these conditions and (2.12) hold true for

$$e_0 := \frac{|m|^2(1+\rho^2)}{n(1-\rho^2)^2} + \tilde{e}, \quad e_1 := -\frac{2\rho|m|^2}{n(1-\rho^2)^2}.$$

Observing that conditions (2.7), (2.8) indeed hold true finishes the construction of the subsolution induced by ρ and m_n. ■

2.4. Applying the least action principle

In the classical case, the action functional consists of the difference between kinetic and potential energy and yields when minimized over suitable paths the equations of motion for the described mechanical system. Going back to Arnold it is well known that this principle can formally also be used to derive the Euler equations; see Appendix A and the references therein for more detail. We will now state it on the level of subsolutions.

The total potential energy of a one-dimensional subsolution $z = z_{\text{sub}}$ at time $t \in [0,T]$ is given by

$$E_{\text{pot}}(t) := \int_{\mathcal{D}} \rho(t,x) g A x_n \, dx = \int_{-L}^{L} \rho(t,x_n) g A x_n \, dx_n.$$

Note that, in view of Theorem 2.1, for any given error function $\delta(t)$ there exist solutions z_{sol} whose total potential energy $\int_{\mathcal{D}} \rho_{\text{sol}} g A x_n \, dx$ at time t is $\delta(t)$-close to $E_{\text{pot}}(t)$.

In a similar way, it follows from Theorem 2.1 that there exist solutions having at time t a total kinetic energy arbitrarily close to

$$\int_{\mathcal{D}} \frac{n}{2}(e_0(t,x) + \rho(t,x) e_1(t,x))\, dx.$$

Moreover, given ρ and m of a one-dimensional subsolution $z = z_{\text{sub}}$, Lemma 2.2 shows that pointwise on $\mathcal{U}$ there holds

$$\inf\left\{\tfrac{n}{2}(\tilde{e}_0 + \rho \tilde{e}_1) : \tilde{z} \text{ subsolution with respect to } \tilde{e}_0, \tilde{e}_1, \tilde{\rho} = \rho, \tilde{m} = m\right\} = \frac{|m|^2}{2(1-\rho^2)}.$$

Including this pointwise minimization over all possible subsolutions coinciding with z in the ρ and m components, we therefore define the total (least) kinetic energy of a one-dimensional subsolution z by

$$E_{\text{kin}}(t) := \int_{-L}^{L} \frac{|m(t, x_n)|^2}{2(1 - \rho(t, x_n)^2)}\, dx_n.$$

Here, the integrand is understood to be 0 when (t, x) is outside the mixing zone, i.e. where $|\rho| = 1$, $m = 0$. Indeed, outside the mixing zone there holds $e_0 + \rho e_1 = 0$ for any one-dimensional subsolution.

Note that the pointwise optimization above does not affect the potential energy $E_{\text{pot}}(t)$. It is therefore compatible with the following least action principle:

$$\text{minimize } \mathcal{A}_0(z) \text{ over one-dimensional subsolutions } z, \tag{2.15}$$

where the action is defined as

$$\begin{aligned} \mathcal{A}_0(z) &:= \int_0^T E_{\text{kin}}(t) - E_{\text{pot}}(t)\, dt \\ &= \int_0^T \int_{-L}^{L} \frac{m_n(t, x_n)^2}{2(1 - \rho(t, x_n)^2)} - \rho(t, x_n) g A x_n \, dx_n\, dt. \end{aligned} \tag{2.16}$$

As subsolutions relax the notion of solution for the Euler equations, the least action principle (2.15) can be seen as a generalization of the classical least action principle giving rise to the equations itself. We will show in Appendix A that this generalization is formally equivalent to the generalization of the least action principle given by Brenier in [2, 4].

Note that at this point we have not yet specified the final configuration for the subsolutions at the end time T, which is usually done in applications of the least action principle. This is postponed to Section 2.6.

2.5. Reformulation

Before continuing let us simplify our notation. First of all, from a one-dimensional subsolution z we keep only the information relevant for the action, that is, the density ρ and the last component of the momentum m, which we again denote by m.

Furthermore, let $\Omega := (0, T) \times (-L, L)$ and identify a point $(t, x_n) \in \Omega$ simply by $x = (x_1, x_2)$. That is, time is denoted now by x_1 and the last coordinate in the box $\mathcal{D}$ by x_2. Thus the action functional (2.16) can be written as

$$\mathcal{A}_0(\rho, m) = \int_\Omega \frac{m^2}{2(1 - \rho^2)} - \rho g A x_2\, dx = \int_\Omega F(m, \rho) - \rho g A x_2\, dx,$$

where F is defined in (1.6). Recall here that the kinetic energy density satisfies $\frac{m^2}{2(1-\rho^2)} = 0$ whenever $m = 0$.

As seen in Lemma 2.2, the pair $(m,\rho)\colon \Omega \to \mathbb{R}^2$ has to satisfy properties (i)–(v) of the said lemma in order to correspond to a one-dimensional subsolution in the sense of Definitions 2.1, 2.4. Some of these properties will be implemented directly into the variational formulation, while others will be shown a posteriori for an existing minimizer.

First of all, observe that property (v), i.e. the L^1-integrability of $\frac{m^2}{1-\rho^2}$, as well as $|\rho| \leq 1$ a.e. follows for (m,ρ) measurable with finite action $\mathcal{A}_0(\rho,m)$.

Next, property (i), i.e. the equation $\partial_{x_1}\rho + \partial_{x_2} m = 0$, will be encoded by introducing a stream function for the divergence-free vector field (ρ,m). If the action is finite we have $(\rho,m) \in L^2(\Omega;\mathbb{R}^2)$ and therefore find $u \in H^1(\Omega)$ with $m = -\partial_{x_1} u$, $\rho = \partial_{x_2} u$.

Moreover, in view of the needed initial and boundary data, i.e. property (ii), we require the stream function $u \in H^1(\Omega)$ to satisfy

$$u(0,x_2) = |x_2| - L, \;\; x_2 \in (-L,L), \quad u(x_1,\pm L) = 0, \;\; x_1 \in (0,T) \tag{2.17}$$

in the sense of traces. Indeed, one can easily check that for $\varphi \in \mathcal{C}_c^\infty([0,T) \times [-L,L])$ there holds

$$\begin{aligned}\int_\Omega \partial_{x_2} u \partial_{x_1}\varphi - \partial_{x_1} u \partial_{x_2}\varphi \, dx &= \int_{-L}^{L} (|x_2| - L)\partial_{x_2}\varphi(0,x_2)\, dx_2 \\ &= -\int_{-L}^{L} \operatorname{sign}(x_2)\varphi(0,x_2)\, dx_2.\end{aligned}$$

Hence (2.13) and therefore Lemma 2.2 (i), (ii) are satisfied. In Section 7.3 we will in fact show that the boundary and initial data are attained in a stronger sense. Moreover, our investigation will in addition show that $m(0,\cdot) = -\partial_{x_1} u(0,\cdot) = 0$ in this sense; see Lemma 7.5.

The remaining properties Lemma 2.2 (iii), (iv), as well as the admissibility of the total energy (cf. (2.11)), will be part of our investigation.

At this point the action functional in terms of a stream function u satisfying (2.17) can be written as

$$\mathcal{A}_0(u) = \int_\Omega F(\nabla u) + gAu \, dx, \tag{2.18}$$

i.e. we have arrived at (1.5) with $V(x_2,u) = -gAu$.

2.6. Final configuration

As mentioned above, the least action principle is formulated with respect to variations over a class of trajectories connecting a given initial and target configuration. While our initial configuration is clear, there are plenty of target configurations possible. In the present article we simply chose the stable interface configuration $-\rho_0$. This configuration has the overall least potential energy, thus also the overall least total energy provided the fluid is at rest, and therefore is a canonical candidate for the long-time limit of the system.

In terms of the introduced stream function u we therefore add

$$u(T,x_2) = L - |x_2|, \quad x_2 \in (-L,L) \tag{2.19}$$

to the list of requirements (2.17). Note that this implies that in Remark 2.5 the equation can be tested against $\varphi \in \mathcal{C}^\infty(\overline{\Omega})$, while adding $\int_{-L}^{L} \operatorname{sign}(x_n)\varphi(T, x_n)\,dx_n$ to the left-hand side of (2.13).

2.7. Energy dissipation

It is a well-known built-in feature of the least action principle that solutions conserve the total energy, which in our case at time x_1 reads

$$E_{\text{tot}}(x_1) := \int_{-L}^{L} F(\nabla u(x)) - gAu(x)\,dx_2. \tag{2.20}$$

Indeed, up to formally admitting the Euler–Lagrange equations $\operatorname{div}(\nabla F(\nabla u)) = gA$ of (2.18), here one also has

$$\begin{aligned}
\frac{d}{dx_1}E_{\text{tot}} &= \int_{-L}^{L} \nabla F(\nabla u)\cdot\nabla(\partial_{x_1}u) - gA\partial_{x_1}u\,dx_2 \\
&= \int_{-L}^{L} \partial_{p_1}F(\nabla u)\partial_{x_1}^2 u - \partial_{x_2}(\partial_{p_2}F(\nabla u))\partial_{x_1}u - gA\partial_{x_1}u\,dx_2 \\
&= \int_{-L}^{L} \frac{d}{dx_1}(\partial_{p_1}F(\nabla u)\partial_{x_1}u) - \operatorname{div}(F(\nabla u))\partial_{x_1}u - gA\partial_{x_1}u\,dx_2 \\
&= \int_{-L}^{L} \frac{d}{dx_1}(\partial_{p_1}F(\nabla u)\partial_{x_1}u) - 2gA\partial_{x_1}u\,dx_2 \\
&= \int_{-L}^{L} \frac{d}{dx_1}(2F(\nabla u)) - 2gA\partial_{x_1}u\,dx_2 = 2\frac{d}{dx_1}E_{\text{tot}}.
\end{aligned}$$

Hence the total energy is constant in time. This formal computation will be made rigorous in Section 7.2.

This is of course undesirable in the context of turbulent fluid dynamics where energy is anomalously dissipated. Note also that, as can be seen in (2.9), the energy of the associated solutions obtained via convex integration differs from the energy of the subsolution with a small margin of error. However, if the energy of the subsolution is conserved, it can happen a priori that due to this margin of error, the energy of the solution will increase.

There are plenty of modifications and extensions of the least action principle in order to include energy dissipation; see for instance also the discussion in [4].

In the present article we overcome the issue of energy conservation by introducing an additional nonlinear potential energy. That is, instead of $\mathcal{A}_0$ we consider (1.5), where now $V : [-L, L] \times \mathbb{R} \to \mathbb{R}$ has the form

$$V(x_2, z) = -gAz + f(x_2, z).$$

Through similar formal calculations to previously, one obtains that the total energy (2.20) now changes according to

$$\frac{d}{dx_1}\int_{-L}^{L} F(\nabla u) - gAu\,dx_2 = -\int_{-L}^{L} \partial_z f(x_2, u)\partial_{x_1}u\,dx_2. \tag{2.21}$$

We will show that $\partial_{x_1} u \geq 0$ (see Corollary 5.5), which means that the average momentum $m = -\partial_{x_1} u$ is negative. Thus, if $\partial_z f > 0$, then strict energy dissipation for the subsolution, and hence also for the associated solutions, is possible.

Integrating (2.21) in time one obtains that

$$E_{\mathrm{kin}}(x_1) + E_{\mathrm{pot}}(x_1) + \int_{-L}^{L} f(x_2, u(x))\, dx_2$$

is constant for minimizers of (1.5). Hence the dissipated kinetic and potential energy is absorbed in the new energy

$$E_f(x_1) := \int_{-L}^{L} f(x_2, u(x))\, dx_2.$$

Furthermore, note that in (2.21), only points where $\partial_{x_1} u > 0$ contribute to a dissipation, i.e. no points outside the mixing zone $\mathcal{U}'$. In terms of the induced solutions this means that the energy is only dissipated inside the mixing zone, where they are wildly oscillating.

Of course at this point there are plenty of choices for f possible. A specific example is given in Section 2.9 below, after we have introduced one more condition in the next Section 2.8. We refer also to the discussion in Section 8.

2.8. Initial and final energies

Formally taking the limits $x_1 \to 0$, $x_1 \to T$ we deduce that

$$E_f(T) - E_f(0) = -(E_{\mathrm{kin}}(T) - E_{\mathrm{kin}}(0) + E_{\mathrm{pot}}(T) - E_{\mathrm{pot}}(0)).$$

The initial and final potential energies can easily be computed for u satisfying (2.17), (2.19). There holds

$$-E_{\mathrm{pot}}(T) = E_{\mathrm{pot}}(0) = gAL^2.$$

Thus, requiring

$$E_f(T) - E_f(0) = 2gAL^2$$

renders the solutions to start and end with the same kinetic energy $E_{\mathrm{kin}}(0) = E_{\mathrm{kin}}(T)$, which in view of (1.2), (2.11) should be 0. Note that then the fluid is at rest also at the final time.

In order to achieve $E_{\mathrm{kin}}(0) = E_{\mathrm{kin}}(T) = 0$ as a consequence of minimizing the action functional (1.5) we chose f such that, for

$$s_V := \sup\left\{\int_{-L}^{L} V(x_2, \varphi)\, dx_2 : \varphi \in \mathcal{C}^0([-L, L]),\ |\varphi(x_2)| \leq L - |x_2|\right\},$$

there holds

$$s_V = \int_{-L}^{L} V(x_2, \varphi)\, dx_2 \quad \text{if and only if} \quad \varphi = \pm(L - |x_2|). \tag{2.22}$$

In fact, we will show in Section 7.4 that this condition allows us to conclude that the subsolution starts and ends with 0 kinetic energy at least as $T \to +\infty$. This can be seen similarly to the existence of heteroclinic orbits, for instance in pendulum equations.

2.9. Example

One of the simplest examples for f satisfying the requested properties is

$$f(x_2, z) = \frac{3gA}{4L}(z - (|x_2| - L))^2, \tag{2.23}$$

which is also stated in (1.7).

Indeed, the monotonicity with respect to z holds true for $z \geq |x_2| - L$, which in view of Corollary 5.5 turns out to be enough. Regarding (2.22) we first of all observe that

$$\int_{-L}^{L} V(x_2, L - |x_2|)\, dx_2 = \int_{-L}^{L} V(x_2, |x_2| - L) = gAL^2$$

by the choice of the constant $\frac{3gA}{4L}$ in (2.23). It remains to show that

$$\mathcal{V}(\varphi) := \int_{-L}^{L} V(x_2, \varphi(x_2))\, dx_2 < gAL^2$$

for any $\varphi\colon [-L, L] \to \mathbb{R}$ continuous, with $|\varphi(x_2)| \leq L - |x_2|$, but $\varphi(x_2)$ not identical to $L - |x_2|$ or $-(L - |x_2|)$. Let φ be such a function, i.e. there exists an open interval $I \subset (-L, L)$ on which $|\varphi(x_2)| < L - |x_2|$. Considering perturbations $\varphi + \varepsilon\psi$ with $\operatorname{supp}\psi \subset I$ and $|\varepsilon|$ small enough, one concludes that $\varepsilon \mapsto \mathcal{V}(\varphi + \varepsilon\psi)$ is a uniformly convex $\mathcal{C}^2$ function, and thus cannot have its supremum achieved in $\varepsilon = 0$.

We remark that for this specific example the new energy term absorbing kinetic and potential energies can be expressed in terms of the actual variables (ρ, m) (instead of the potential u) as

$$E_f(x_1) = \frac{3gA}{4L}\|u(x_1, \cdot) - u(0, \cdot)\|^2_{L^2(-L,L)} = \frac{3gA}{4L}\|\rho(x_1, \cdot) - \rho_0\|^2_{H^{-1}(-L,L)},$$

where $H^{-1}(-L, L)$ denotes the dual of $H^1_0(-L, L)$ with respect to the topology induced by the norm $\|\partial_{x_2}(\cdot)\|_{L^2(-L,L)}$.

3. A degenerate variational problem

We now turn to the investigation of problem (1.4). More precisely, we seek to minimize

$$\mathcal{A}(u) = \int_{\Omega} F(\nabla u) - V(x, u)\, dx \tag{3.1}$$

over the class of functions $u \in X$ with F defined in (1.6) and X given by (1.8).

The nonlinear potential $V\colon \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$, $(x, z) \mapsto V(x, z)$ is supposed to satisfy the following regularity condition:

$$\begin{gathered} V \text{ is 3 times differentiable with respect to } z, \\ \partial_z^k V\colon \overline{\Omega} \times \mathbb{R} \to \mathbb{R},\ k = 0, \dots, 3 \text{ are Lipschitz continuous and bounded.} \end{gathered} \tag{V$_{\text{reg}}$}$$

This condition will be assumed throughout the remainder of the article. Note that the example given in (1.7) and Section 2.9 satisfies ($\mathrm{V}_{\mathrm{reg}}$) when the stated V is extended smoothly outside $\bar{\Omega} \times [-L, L]$, such that V and its z-derivatives are globally bounded. We remark that the precise extension turns out to be irrelevant in view of Corollary 5.5.

Moreover, we will frequently also assume that the potential is autonomous with respect to x_1, i.e.

$$V(x, z) = V(x_2, z), \quad x \in \bar{\Omega},\ z \in \mathbb{R}, \tag{$\mathrm{V}_{\mathrm{aut}}$}$$

such that then (3.1) reduces to (1.5).

Two other conditions on V, as indicated in Sections 2.7, 2.8, will be introduced when needed, which is only in the very last part of our investigation in Section 7.4 when it comes to the interpretation of our minimizer as a subsolution for the Boussinesq system.

For now our main goal is to show the following existence and partial regularity result for the variational problem

$$\text{find } u \in X \text{ such that } \mathcal{A}(u) = \inf_{u \in X} \mathcal{A}(u), \tag{3.2}$$

as well as the associated energy balance in the autonomous case.

Theorem 3.1. *Suppose that* ($\mathrm{V}_{\mathrm{reg}}$) *holds. Then problem* (3.2) *with* $\mathcal{A}$ *defined in* (3.1) *has a solution* u *and there exists* $\Omega' \subset \Omega$ *open, nonempty such that* $u_{|\Omega'}$ *is of class* $\mathcal{C}^2$ *with* $\partial_{x_1} u \neq 0$, $|\partial_{x_2} u| < 1$ *on* Ω' *and* $\partial_{x_1} u(x) = 0$, $|\partial_{x_2} u(x)| \leq 1$ *for a.e.* $x \notin \Omega'$. *Moreover, if in addition* ($\mathrm{V}_{\mathrm{aut}}$) *holds true, then* $u \in \mathcal{C}^0(\bar{\Omega})$, $\partial_{x_1} u > 0$ *on* Ω' *and there holds*

$$\frac{d}{dx_1} \int_{-L}^{L} F(\nabla u(x)) + V(x_2, u(x))\, dx_2 = 0 \tag{3.3}$$

weakly on $(0, T)$.

The main difficulty lies in the degeneracy of the convex, lower semi-continuous integrand F. Indeed, denoting by $\Lambda(p)$, $\lambda(p)$ the maximal and minimal eigenvalues of $D^2 F(p)$ (cf. (4.1) with $\varepsilon = 0$), for $p_1 \neq 0$, $|p_2| < 1$ there holds

$$\Lambda(p) \to +\infty \text{ as } |p_2| \to 1 \quad \text{and} \quad \lambda(p) \to 0 \text{ as } p_1 \to 0.$$

Thus the problem degenerates on the nonconvex set $E := \{0\} \times [-1, 1] \cup \mathbb{R} \times \{\pm 1\}$, with indefinite behavior for $\det D^2 F(p)$ as $p \to (0, \pm 1)$.

This is in contrast with the prototype of degenerate problems, i.e. the p-Laplace problem with $F(\xi) = |\xi|^p$, where the ellipticity constants degenerate only at the single point $E = \{0\}$ and with definite behavior for all eigenvalues as $\xi \to 0$. In that case, $\mathcal{C}^{1,\alpha}$-regularity for minimizers is known; see [24, 34, 45–47].

Another proof for the regularity of p-harmonic maps is given by Wang [48], relying on a separation between degenerate and nondegenerate points and the fact that near the degenerate set ∇u is small anyway. Beyond the p-Laplacian, Colombo and Figalli [14] also applied a separation strategy based on ideas of [48] to problems that degenerate on a bounded convex set E having 0 in its interior. One of these "very degenerate" integrands is for example $F(\xi) = (|\xi| - 1)_+^p$, which arises in problems related to traffic congestion. The main theorem of [14] states that for this type of problem with $V(x,z) = V(x)$, the composition $\nabla F(\nabla u)$ is continuous. This extends the work by Santambrogio and Vespri [40] relying on two-dimensional methods for $E = B_1(0)$, to any dimension and a general convex bounded E with $0 \in \text{int}(E)$.

Near nondegenerate points, the proof of [14] is based on a compactness result for small solutions to elliptic equations in the spirit of Savin [41] which in the present work also renders one of the main ingredients in the proof of Theorem 3.1. However, the nature of the degeneracy of our F has not so far allowed us to conclude a global regularity result like the one in [14].

De Silva and Savin [21] considered a different type of degenerate variational problem arising, for instance, in questions related to limits of random surfaces. There the integrand F is a bounded function defined on the closure of a bounded two-dimensional polygon N and the set of degeneracy is given by the union of ∂N and finitely many points inside N. Apart from this union, F is smooth and strictly convex. There is no potential term, i.e. $V = 0$. In that setting, [21, Theorem 1.3] provides a partial regularity result for the unique minimizer u and characterizes the behavior at points where ∇u is not continuous. More precisely, every point of discontinuity is connected to the boundary of Ω along a straight segment perpendicular to one of the sides of N and on that segment u is affine linear. Besides the unboundedness of our F and our degeneracy set E, the absence of the maximum principle due to the nonlinear potential $V(x,z)$ prevents us from applying the methods of [21].

Still, in the following sense, points of discontinuity of ∇u are, for certain V, also in our case connected to $\partial\Omega$:

Lemma 3.2. *Suppose that V satisfies (V_{reg}), (V_{aut}), and $\partial_z^2 V(x_2, z) \geq 0$ for all $x_2, z \in [-L, L]$. Whenever $\Omega'' \subset \Omega$ is open with $\partial\Omega'' \subset \Omega'$, where Ω' is the set from Theorem 3.1, then there holds $\Omega'' \subset \Omega'$.*

For a further, more general overview on degenerate variational problems we refer to the survey [37].

The proof of Theorem 3.1 is carried out in Sections 4–7. We begin in Section 4.1 with the construction of regular approximations $\widehat{F}_\varepsilon$ for the degenerate integrand F. Here, some extra attention has to be paid (see Lemma 4.1 (iii)) in order to later conclude the energy balance (3.3). Having a family of regularized variational problems at hand, we deduce in Sections 4.2–4.4 the existence of regular minimizers u_ε enjoying corresponding ε-versions of the energy balance (3.3) or Lemma 3.2 for instance. After that, Section 5 deals with the limit $\varepsilon \to 0$. We prove Γ-convergence with respect to the weak H^1-topology and characterize in terms of the corresponding Young measure how strong convergence might fail. In particular, we deduce the existence of a minimizer u to the degenerate problem (3.2). Section 6 contains the proof of the partial regularity property which, as mentioned earlier, uses the compactness result of Savin [41]. Section 7 collects various additional properties of u, for example (3.3), Lemma 3.2, but also further properties that rely on the mentioned additional conditions on V. Finally, Section 8 contains a summary of all conditions on V together with the corresponding conclusions, as well as a discussion of open questions in the variational problem itself, but also regarding our application to the Boussinesq system.

4. A regular approximation

In order to deal with the possible singularity in the denominator of F and the degeneracy of $D^2F(p)$ when $p_1 = 0$, in this section we introduce the following regular approximation. For $\varepsilon > 0$, let $F_\varepsilon\colon \{p \in \mathbb{R}^2 : |p_2| < 1+\varepsilon\} \to \mathbb{R}$,

$$F_\varepsilon(p) = \frac{p_1^2 + \varepsilon^\theta}{2((1+\varepsilon)^2 - p_2^2)},$$

where $\theta \in (1,2)$ is a fixed constant. A quick calculation yields

$$\begin{aligned}
\nabla F_\varepsilon(p) &= \begin{pmatrix} \frac{p_1}{(1+\varepsilon)^2-p_2^2} & \frac{(p_1^2+\varepsilon^\theta)p_2}{((1+\varepsilon)^2-p_2^2)^2} \end{pmatrix}^{\mathsf{T}}, \\
D^2F_\varepsilon(p) &= \begin{pmatrix} \frac{1}{(1+\varepsilon)^2-p_2^2} & \frac{2p_1p_2}{((1+\varepsilon)^2-p_2^2)^2} \\ \frac{2p_1p_2}{((1+\varepsilon)^2-p_2^2)^2} & \frac{(p_1^2+\varepsilon^\theta)((1+\varepsilon)^2+3p_2^2)}{((1+\varepsilon)^2-p_2^2)^3} \end{pmatrix}, \qquad (4.1) \\
\det(D^2F_\varepsilon(p)) &= \frac{p_1^2}{((1+\varepsilon)^2-p_2^2)^3} + \frac{\varepsilon^\theta((1+\varepsilon)^2+3p_2^2)}{((1+\varepsilon)^2-p_2^2)^4}.
\end{aligned}$$

Hence F_ε is uniformly convex.

4.1. Global extension

Next, we define the compact sets

$$K^\varepsilon := \left\{p \in \mathbb{R}^2 : |p_1| \le \varepsilon^{-4\theta},\ |p_2| \le 1+\varepsilon-\varepsilon^{4\theta}\right\}$$

and extend $F_{\varepsilon|K^\varepsilon}$ in a uniformly elliptic way onto all of $\mathbb{R}^2$, with some additional properties. The stated uniform bound in (ii) for instance will be used to achieve Γ-convergence in Section 5, while property (iii) will be needed in order to conclude the energy balance in Section 7.

Lemma 4.1. *For every $\varepsilon \in (0,1)$ there exists a smooth extension $\widehat{F}_\varepsilon\colon \mathbb{R}^2 \to [0,\infty)$ of $F_{\varepsilon|K^\varepsilon}$ satisfying*

(i) $\lambda_\varepsilon \,\mathrm{id} \le D^2\widehat{F}_\varepsilon(p) \le \Lambda_\varepsilon \,\mathrm{id}$ *for all $p \in \mathbb{R}^2$ with some constants $0 < \lambda_\varepsilon, \Lambda_\varepsilon < \infty$,*

(ii) $\lambda_0 \,\mathrm{id} \le D^2\widehat{F}_\varepsilon(p)$ *for $|p_2| \ge 1+\varepsilon-\varepsilon^{4\theta}$ or $|p_1| \ge 1$ with a constant $\lambda_0 > 0$ independent of $\varepsilon \in (0,1)$,*

(iii) $-\varepsilon \leq \partial_{p_1} \widehat{F}_\varepsilon(p) p_1 \leq 3\widehat{F}_\varepsilon(p)$ *for all* $p \in \mathbb{R}^2$.

The proof of Lemma 4.1 relies on Lemmas 4.2, 4.3.

Lemma 4.2. *Let* $N_\varepsilon := \max F_{\varepsilon|K^\varepsilon} + 1$, $K_{N_\varepsilon} := F_\varepsilon^{-1}((-\infty, N_\varepsilon])$. *There exists a convex and globally Lipschitz extension* $\widetilde{F}_\varepsilon \colon \mathbb{R}^2 \to [0,\infty)$ *of* $F_{\varepsilon|K_{N_\varepsilon}}$ *satisfying*

$$0 \leq \partial_{p_1} \widetilde{F}_\varepsilon(p) p_1 \leq 2\widetilde{F}_\varepsilon(p) \tag{4.2}$$

for almost every $p \in \mathbb{R}^2$. *Moreover, on* $\partial K_{N_\varepsilon}$ *the extension* $\widetilde{F}_\varepsilon$ *is differentiable with* $\nabla \widetilde{F}_\varepsilon = \nabla F_\varepsilon$.

Proof. We abbreviate $N = N_\varepsilon$ and consider the smallest convex extension of $F_{\varepsilon|K_N}$ defined by $\widetilde{F}_\varepsilon \colon \mathbb{R}^2 \to \mathbb{R}$,

$$\widetilde{F}_\varepsilon(p) = \sup_{\tilde{p} \in \partial K_N} \{F_\varepsilon(\tilde{p}) + \nabla F_\varepsilon(\tilde{p}) \cdot (p - \tilde{p})\}$$

for $p \notin K_N$ and $\widetilde{F}_\varepsilon(p) = F_\varepsilon(p)$ for $p \in K_N$.

A priori, $\widetilde{F}_\varepsilon$ as a convex function is only locally Lipschitz, but since the subdifferential of $\widetilde{F}_\varepsilon$ at $p \notin \operatorname{int}(K_N)$ is given by

$$\partial \widetilde{F}_\varepsilon(p) = \left\{\nabla F_\varepsilon(\bar{p}) : \bar{p} \in \partial K_N,\ \widetilde{F}_\varepsilon(p) = F_\varepsilon(\bar{p}) + \nabla F_\varepsilon(\bar{p}) \cdot (p - \bar{p})\right\}^{\mathrm{co}} \tag{4.3}$$

(cf. [50, Theorem 2.4.18]) and K_N is compact, it follows that $\|\nabla \widetilde{F}_\varepsilon\|_{L^\infty(\mathbb{R}^2)}$ is finite.

Moreover, the strict convexity of F_ε and (4.3) imply

$$\partial \widetilde{F}_\varepsilon(p) = \{\nabla F_\varepsilon(p)\} \quad \text{for all } p \in \partial K_{N_\varepsilon}.$$

Next, knowing that $0 \leq \partial_{p_1} F_\varepsilon(p) p_1 \leq 2F_\varepsilon(p)$, $p \in \mathbb{R}^2$, $|p_2| < 1 + \varepsilon$ holds true, let us prove (4.2) whenever $\widetilde{F}_\varepsilon$ is differentiable at $p \in \mathbb{R}^2$.

In order to do this let us suppose that the supremum is achieved at some $\bar{p} = \bar{p}(p) \in \partial K_N$ and let us write $F_\varepsilon = \frac{1}{2} e^g$, with

$$g(p) = g_\varepsilon(p) = \log(p_1^2 + \varepsilon^\theta) - \log((1+\varepsilon)^2 - p_2^2).$$

We obtain that

$$\widetilde{F}_\varepsilon(p) = N(1 + \nabla g(\bar{p}) \cdot (p - \bar{p})) \tag{4.4}$$

and under the assumption of differentiability that $\nabla \widetilde{F}_\varepsilon(p) = N \nabla g(\bar{p})$.

Since $\bar{p} \in \partial K_N = \{F_\varepsilon = N\}$, we have

$$\bar{p}_1^2 + \varepsilon^\theta - 2N((1+\varepsilon)^2 - \bar{p}_2^2) = 0, \tag{4.5}$$

and hence we may rewrite

$$\nabla g(\bar{p}) \cdot (p - \bar{p}) = \frac{\bar{p}_1 p_1 + 2N \bar{p}_2 p_2 - (2N(1+\varepsilon)^2 - \varepsilon^\theta)}{N((1+\varepsilon)^2 - \bar{p}_2^2)}. \tag{4.6}$$

We first of all observe that $\partial_{p_1} \widetilde{F}_\varepsilon(p) p_1 = N \partial_{p_1} g(\bar{p}) p_1 \geq 0$ as otherwise $(-\bar{p}_1, \bar{p}_2)$ would be a better choice for the supremum.

For the upper bound we distinguish two cases:

Case (a)*:* $p_2 - \bar{p}_2 < 0$. Since $\bar{p}$ maximizes (4.6) under the constraint (4.5), there exists a Lagrange multiplier $\lambda \in \mathbb{R}$ such that there holds

$$\begin{aligned} \frac{p_1}{N((1+\varepsilon)^2 - \bar{p}_2^2)} - 2\lambda \bar{p}_1 = 0, \\ \frac{2Np_2}{N((1+\varepsilon)^2 - \bar{p}_2^2)} + \nabla g(\bar{p}) \cdot (p - \bar{p}) \frac{2N\bar{p}_2}{N((1+\varepsilon)^2 - \bar{p}_2^2)} - 4N\lambda \bar{p}_2 = 0. \end{aligned} \tag{4.7}$$

We may assume without loss of generality that neither $\bar{p}_1$ nor $\bar{p}_2$ is zero. Indeed, if $\bar{p}_1 = 0$, from (4.7) it follows that $p_1 = 0$, and hence the inequality that we want to prove follows trivially. If $\bar{p}_2 = 0$, from (4.7) once more it follows that $p_2 = 0$, which contradicts $p_2 - \bar{p}_2 < 0$.

Expressing λ from both equations and simplifying leads to

$$\nabla g(\bar{p}) \cdot (p - \bar{p}) = \frac{p_1}{\bar{p}_1} - \frac{p_2}{\bar{p}_2}. \tag{4.8}$$

As described we need to estimate $\partial_{p_1} g(\bar{p}) p_1$, which is given by

$$\partial_{p_1} g(\bar{p}) p_1 = 2\frac{\bar{p}_1 p_1}{\bar{p}_1^2 + \varepsilon^\theta} \le 2\frac{p_1}{\bar{p}_1} = 2\nabla g(\bar{p}) \cdot (p - \bar{p}) + 2\frac{p_2}{\bar{p}_2} \le 2\nabla g(\bar{p}) \cdot (p - \bar{p}) + 2,$$

where we have used (4.8) and the fact that we are in the case $\frac{p_2}{\bar{p}_2} < 1$. Hence, using (4.4), it follows that

$$\partial_{p_1} \widetilde{F}_\varepsilon(p) p_1 = N\partial_{p_1} g(\bar{p}) p_1 \le 2N(\nabla g(\bar{p}) \cdot (p - \bar{p}) + 1) = 2\widetilde{F}_\varepsilon(p),$$

which is the desired inequality.

Case (b)*:* $p_2 - \bar{p}_2 \ge 0$. Observe that since $p_2 \mapsto F_\varepsilon(p)$ is even, we may assume without loss of generality that $p_2 \ge 0$. In (4.6), replacing the denominator by $\frac{1}{2}(\bar{p}_1^2 + \varepsilon^\theta)$ via (4.5) it follows that $\bar{p}_2 \ge 0$, as otherwise $(\bar{p}_1, -\bar{p}_2)$ would be better. This however implies that

$$\partial_{p_2} g(\bar{p}) = \frac{2\bar{p}_2}{(1+\varepsilon)^2 - \bar{p}_2^2} \ge 0.$$

We may then write

$$\begin{aligned} \partial_{p_1} g(\bar{p}) p_1 &\le \nabla g(\bar{p}) \cdot (p - \bar{p}) + \partial_{p_1} g(\bar{p}) \bar{p}_1 \\ &= \nabla g(\bar{p}) \cdot (p - \bar{p}) + 2\frac{\bar{p}_1^2}{\bar{p}_1^2 + \varepsilon^\theta} \\ &\le \nabla g(\bar{p}) \cdot (p - \bar{p}) + 2 \le 2\nabla g(\bar{p}) \cdot (p - \bar{p}) + 2, \end{aligned}$$

from where we conclude as in the previous case. This finishes the proof of the lemma. ∎

Lemma 4.3. *Let $\varphi_\eta \in \mathcal{C}_c^\infty(B_\eta(0))$, $\eta > 0$ be a standard symmetric mollifier and $\widetilde{F}_\varepsilon$ from Lemma 4.2. The convolution $\widetilde{F}_\varepsilon^\eta := \varphi_\eta * \widetilde{F}_\varepsilon$ satisfies $\widetilde{F}_\varepsilon^\eta \geq \widetilde{F}_\varepsilon$ and*

$$-c_\varepsilon \eta \leq \partial_{p_1} \widetilde{F}_\varepsilon^\eta(p) p_1 \leq 2\widetilde{F}_\varepsilon^\eta(p) + c_\varepsilon \eta, \tag{4.9}$$

where $c_\varepsilon > 0$ depends only on ε. Moreover, there exists $\Lambda_{\varepsilon,\eta} > 0$ with

$$0 \leq D^2 \widetilde{F}_\varepsilon^\eta(p) \leq \Lambda_{\varepsilon,\eta}\, \mathrm{id}, \quad p \in \mathbb{R}^2. \tag{4.10}$$

Proof. Denoting by $c_\varepsilon > 0$ the global Lipschitz constant of $\widetilde{F}_\varepsilon$ and using (4.2) one concludes

$$\begin{aligned}\partial_{p_1} \widetilde{F}_\varepsilon^\eta(p) p_1 &= \int_{\mathbb{R}^2} \varphi_\eta(q) \partial_{p_1} \widetilde{F}_\varepsilon(p-q)(p_1 - q_1)\, dq + \int_{\mathbb{R}^2} \varphi_\eta(q) \partial_{p_1} \widetilde{F}_\varepsilon(p-q) q_1\, dq \\ &\leq 2\widetilde{F}_\varepsilon^\eta(p) + c_\varepsilon \eta\end{aligned}$$

and $\partial_{p_1} \widetilde{F}_\varepsilon^\eta(p) p_1 \geq -c_\varepsilon \eta$.

Moreover, since $\widetilde{F}_\varepsilon^\eta$ is a convex smooth function there holds

$$0 \leq v^\intercal D^2 \widetilde{F}_\varepsilon^\eta(p) v = \int_{\mathbb{R}^2} (\nabla\varphi(p-q)\cdot v)(\nabla \widetilde{F}_\varepsilon(q)\cdot v)\, dq \leq \tilde{c}_\varepsilon \int_{\mathbb{R}^2} |\nabla \varphi_\eta(q)|\, dq |v|^2$$

for any $v \in \mathbb{R}^2$. ■

Proof of Lemma 4.1. We will now construct the extension $\widehat{F}_\varepsilon$. Let N_ε, K_{N_ε}, $\widetilde{F}_\varepsilon$ and $\widetilde{F}_\varepsilon^\eta$, $\eta > 0$ be as in Lemmas 4.2 and 4.3 respectively. Furthermore, consider Q_ε, $\check{F}_\varepsilon$, $\check{F}_\varepsilon^\eta : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$Q_\varepsilon(p) := \frac{p_1^2 + 2N_\varepsilon p_2^2}{2N_\varepsilon(1+\varepsilon)^2 - \varepsilon^\theta},$$

$$\check{F}_\varepsilon(p) := \widetilde{F}_\varepsilon(p) + \frac{N_\varepsilon}{4}(Q_\varepsilon(p) - 1), \quad \check{F}_\varepsilon^\eta(p) := \widetilde{F}_\varepsilon^\eta(p) + \frac{N_\varepsilon}{4}(Q_\varepsilon(p) - 1).$$

Note that $Q_\varepsilon = 1$ on $\partial K_{N_\varepsilon}$ by the definition of K_{N_ε}.

We claim that there exist constants $C = C_\varepsilon > 0$ and $\delta' = \delta'_\varepsilon > 0$, such that for any $\delta \in (0, \delta')$ and $q \in \mathbb{R}^2$ with $\operatorname{dist}(q, \partial K_{N_\varepsilon}) = \delta$ there holds

$$F_\varepsilon(q) - \check{F}_\varepsilon(q) \begin{cases} \leq -C\delta, & q \notin K_{N_\varepsilon}, \\ \geq C\delta, & q \in K_{N_\varepsilon}. \end{cases} \tag{4.11}$$

For $q \in K_{N_\varepsilon}$, where also $\check{F}_\varepsilon$ is smooth, this is a straightforward consequence of the fact that on $\partial K_{N_\varepsilon}$ the functions F_ε, $\check{F}_\varepsilon$ coincide, while their gradients are related via $\nabla \check{F}_\varepsilon = \nabla F_\varepsilon + \frac{N_\varepsilon}{4} \nabla Q_\varepsilon$; cf. Lemma 4.2. For $q \notin K_{N_\varepsilon}$ we can argue by convexity instead. Indeed, let $p \in \partial K_{N_\varepsilon}$ and $q = p + \delta \frac{\nabla Q_\varepsilon(p)}{|\nabla Q_\varepsilon(p)|}$. Note that any $q \notin K_{N_\varepsilon}$ with $\operatorname{dist}(q, \partial K_{N_\varepsilon}) = \delta$ can be written like this. There holds

$$\check{F}_\varepsilon(q) - F_\varepsilon(p) \geq (\nabla \check{F}_\varepsilon(p) - \nabla F_\varepsilon(p)) \cdot (q - p) + o(\delta) = \frac{N_\varepsilon}{4} |\nabla Q_\varepsilon(p)| \delta + o(\delta),$$

with an error uniform in $p \in \partial K_{N_\varepsilon}$. Thus (4.11) follows.

Next we fix $\delta \in (0, \delta')$ small, such that $B_\delta(K_{N_\varepsilon}) \subset\subset \{p \in \mathbb{R}^2 : |p_2| < 1 + \varepsilon\}$, $\operatorname{dist}(K^\varepsilon, B_\delta(\partial K_{N_\varepsilon})) > 0$ and such that

$$\frac{3}{4} N_\varepsilon \leq \check{F}_\varepsilon(p) \text{ for all } p \notin K_{N_\varepsilon} \text{ and all } p \in K_{N_\varepsilon} \text{ with } \operatorname{dist}(p, \partial K_{N_\varepsilon}) \leq \delta. \tag{4.12}$$

We then choose the convolution scale $\eta > 0$ such that $\check{F}_\varepsilon^\eta$ satisfies

$$F_\varepsilon(q) - \check{F}_\varepsilon^\eta(q) \begin{cases} \leq -C\delta/2, & q \notin K_{N_\varepsilon},\ \operatorname{dist}(q, \partial K_{N_\varepsilon}) = \delta, \\ \geq C\delta/2, & q \in K_{N_\varepsilon},\ \operatorname{dist}(q, \partial K_{N_\varepsilon}) = \delta. \end{cases} \tag{4.13}$$

Finally, we define the extension $\widehat{F}_\varepsilon : \mathbb{R}^2 \to \mathbb{R}$ by

$$\widehat{F}_\varepsilon(p) := \begin{cases} F_\varepsilon(p), & p \in K_{N_\varepsilon},\ \operatorname{dist}(p, \partial K_{N_\varepsilon}) > \delta, \\ \max_{\tilde{\eta}}(F_\varepsilon(p), \check{F}_\varepsilon^\eta(p)), & \operatorname{dist}(p, \partial K_{N_\varepsilon}) \leq \delta, \\ \check{F}_\varepsilon^\eta(p), & p \notin K_{N_\varepsilon},\ \operatorname{dist}(p, \partial K_{N_\varepsilon}) > \delta. \end{cases}$$

Here $\max_{\tilde{\eta}}$, $\tilde{\eta} > 0$ is a sufficiently sharp convolution of the maximum of two numbers, that is,

$$\max_{\tilde{\eta}}(y_1, y_2) := \int_{\mathbb{R}^2} \varphi_{\tilde{\eta}}(y_1 - a_1, y_2 - a_2) \max\{a_1, a_2\}\, da.$$

By this definition and (4.13) we see that $\widehat{F}_\varepsilon$ is indeed a smooth function for any $\tilde{\eta} > 0$ chosen smaller than $C\delta/4$. Also, $\widehat{F}_{\varepsilon|K^\varepsilon} = F_{\varepsilon|K^\varepsilon}$ by the choice of δ. It thus remains to verify properties (i)–(iii).

We begin with (iii): F_ε clearly satisfies (iii) with a factor 2 on the right-hand side. Moreover, the definition of Q_ε and (4.9), (4.12) imply

$$-c_\varepsilon \eta \leq \partial_{p_1} \check{F}_\varepsilon^\eta(p) p_1 \leq 2\check{F}_\varepsilon^\eta(p) + \frac{N_\varepsilon}{2} + c_\varepsilon \eta \leq \Big(2 + \frac{2}{3}\Big) \check{F}_\varepsilon^\eta(p) + c_\varepsilon \eta$$

for all $p \notin K_{N_\varepsilon}$ and all $p \in K_{N_\varepsilon}$ with $\operatorname{dist}(p, \partial K_{N_\varepsilon}) \leq \delta$. Hence, by shrinking η further, we obtain that $\check{F}_\varepsilon^\eta(p)$ satisfies (iii) with a factor $2 + \frac{3}{4}$ on the right-hand side for the said points p.

It remains to observe that in the transition zone $\operatorname{dist}(p, \partial K_{N_\varepsilon}) \leq \delta$, the gradient of $\widehat{F}_\varepsilon$ is given by the convex combination

$$\begin{aligned} \nabla \widehat{F}_\varepsilon(p) &= \nabla F_\varepsilon(p) \int_{\{a_1 > a_2\}} \varphi_{\tilde{\eta}}(F_\varepsilon(p) - a_1, \check{F}_\varepsilon^\eta(p) - a_2)\, da \\ &\quad + \nabla \check{F}_\varepsilon^\eta(p) \int_{\{a_1 < a_2\}} \varphi_{\tilde{\eta}}(F_\varepsilon(p) - a_1, \check{F}_\varepsilon^\eta(p) - a_2)\, da \\ &=: \lambda(p) \nabla F_\varepsilon(p) + (1 - \lambda(p)) \nabla \check{F}_\varepsilon^\eta(p), \end{aligned}$$

while the values satisfy

$$\big| \widehat{F}_\varepsilon(p) - \big(\lambda(p) F_\varepsilon(p) + (1 - \lambda(p)) \check{F}_\varepsilon^\eta(p)\big) \big| \leq \tilde{\eta}.$$

Thus by shrinking $\tilde{\eta}$ we deduce property (iii) for $\widehat{F}_\varepsilon$.

Next we turn to (ii). Regarding $\check{F}_\varepsilon^\eta(p)$ we have

$$\lambda_{\min}(D^2\check{F}_\varepsilon^\eta(p)) \geq \frac{N_\varepsilon}{4}\lambda_{\min}(D^2 Q_\varepsilon(p)) = \frac{N_\varepsilon}{4N_\varepsilon(1+\varepsilon)^2 - 2\varepsilon^\theta} \geq \frac{1}{16}.$$

Regarding the original F_ε, observe that $1+\varepsilon > |p_2| \geq 1+\varepsilon-\varepsilon^{4\theta}$ implies

$$F_\varepsilon(p) \geq \frac{\varepsilon^\theta}{2\cdot 4\varepsilon^{4\theta}} \geq \frac{1}{8}. \tag{4.14}$$

The same estimate also holds true for $p \in \mathbb{R}^2$, $|p_2| < 1+\varepsilon$ with $|p_1| \geq 1$.

For $p \in \mathbb{R}^2$, $|p_2| < 1+\varepsilon$ such that (4.14) holds true, we abbreviate the quantity $\sigma := (1+\varepsilon)^2 - p_2^2 \leq 4$ and estimate by means of (4.1) the minimal eigenvalue as follows:

$$\begin{aligned}\lambda_{\min}(D^2F_\varepsilon(p)) &\geq \frac{\det(D^2F_\varepsilon(p))}{\operatorname{tr}(D^2F_\varepsilon(p))} = \frac{p_1^2 + \sigma^{-1}\varepsilon^\theta((1+\varepsilon)^2+3p_2^2)}{\sigma^2 + (p_1^2+\varepsilon^\theta)((1+\varepsilon)^2+3p_2^2)} \\ &\geq \frac{4^{-1}p_1^2 + 4^{-1}\varepsilon^\theta}{4\sigma + 16(p_1^2+\varepsilon^\theta)} = \frac{1}{8F_\varepsilon(p)^{-1}+64} \geq \frac{1}{128}.\end{aligned}$$

Thus we set $\lambda_0 = \frac{1}{128}$ and it remains to check the behavior of the minimal eigenvalue in the transition zone $\operatorname{dist}(p, \partial K_{N_\varepsilon}) \leq \delta$. From what we have seen it follows that for both functions $f_1 := F_\varepsilon$, $f_2 := \check{F}_\varepsilon^\eta$ there holds

$$f_i(p) - a_i \geq f_i(p_0) - a_i + \nabla f_i(p_0)\cdot(p-p_0) + \frac{\lambda_0}{2}|p-p_0|^2$$

for all points $p, p_0 \in \mathbb{R}^2$ which are δ-close to $\partial K_{N_\varepsilon}$ and $a_i \in \mathbb{R}$. Therefore,

$$\begin{aligned}\hat{F}_\varepsilon(p) &= \int_{A:=\{f_1(p_0)-a_1>f_2(p_0)-a_2\}} \varphi_{\tilde\eta}(a)\max\{f_1(p)-a_1, f_2(p)-a_2\}\,da \\ &\quad + \int_{\substack{B:=\{f_1(p_0)-a_1\\<f_2(p_0)-a_2\}}} \varphi_{\tilde\eta}(a)\max\{f_1(p)-a_1, f_2(p)-a_2\}\,da \\ &\geq \int_A \varphi_{\tilde\eta}(a)\Big(f_1(p_0)-a_1+\nabla f_1(p_0)\cdot(p-p_0)+\frac{\lambda_0}{2}|p-p_0|^2\Big)\,da \\ &\quad + \int_B \varphi_{\tilde\eta}(a)\Big(f_2(p_0)-a_2+\nabla f_2(p_0)\cdot(p-p_0)+\frac{\lambda_0}{2}|p-p_0|^2\Big)\,da \\ &= \hat{F}_\varepsilon(p_0) + \nabla\hat{F}_\varepsilon(p_0)\cdot(p-p_0) + \frac{\lambda_0}{2}|p-p_0|^2,\end{aligned}$$

which shows (ii).

Finally, property (i) follows in a similar way by observing that F_ε and $\check{F}_\varepsilon^\eta$ are uniformly elliptic with ε-dependent bounds; cf. also (4.10). ∎

4.2. The regularized variational problem

Our regular approximation of the action functional $\mathcal{A}(u)$ defined in (3.1) then reads

$$\mathcal{A}_\varepsilon(u) := \int_\Omega \widehat{F}_\varepsilon(\nabla u) - V(x,u)\,dx$$

with $\widehat{F}_\varepsilon$ from Lemma 4.1.

Moreover, we will not only use an approximation of the integrand, but also introduce ε-dependent boundary data, which is used for the a priori bounds in Section 4.3. We set

$$X_\varepsilon := \big\{u \in H^1(\Omega) : u(\cdot, \pm L) = 0,\ u(0,\cdot) = -U_\varepsilon,\ u(T,\cdot) = U_\varepsilon\big\},$$

where $U_\varepsilon\colon [-L, L] \to \mathbb{R}$,

$$U_\varepsilon(x_2) := L - |x_2| + \frac{\varepsilon^\beta}{2L}(L^2 - x_2^2)$$

for some fixed constant $\beta \in (1, 3-\theta)$. The boundary data is again attained in the trace sense. Also observe that $\mathcal{A}_\varepsilon$ is well defined on all of $H^1(\Omega)$ due to the uniform ellipticity of $\widehat{F}_\varepsilon$. We then consider the corresponding regularized minimization problem

$$\text{find } u_\varepsilon \in X_\varepsilon \text{ such that } \mathcal{A}_\varepsilon(u_\varepsilon) = \inf\nolimits_{u \in X_\varepsilon} \mathcal{A}_\varepsilon(u). \tag{4.15}$$

The uniform ellipticity of the approximations allows us to conclude the existence of sufficiently smooth minimizers in a standard way.

Lemma 4.4. *Problem (4.15) admits a solution. Every solution belongs to $\mathcal{C}^0(\overline{\Omega})$, as well as $\mathcal{C}^{2,\alpha}(K)$ for any compact K contained in $\overline{\Omega}$ and having a positive distance to* $\{(0, \pm L), (T, \pm L), (0, 0), (T, 0)\}$.

Proof. Lemma 4.5 below in particular shows that $\inf_{X_\varepsilon} \mathcal{A}_\varepsilon$ is finite. The existence of a minimizer $u_\varepsilon \in X_\varepsilon$ of $\mathcal{A}_\varepsilon$ then follows by the boundedness of V and the uniform convexity of $\widehat{F}_\varepsilon$ (e.g. [25, Chapter 8.2]). Due to the ellipticity condition Lemma 4.1 (i) and ($\mathrm{V}_{\mathrm{reg}}$), the regularity follows in the classical manner, e.g. from $u_\varepsilon \in H^1$ to $u_\varepsilon \in H^2$ to $\nabla u_\varepsilon \in \mathcal{C}^{0,\alpha}$ to $u_\varepsilon \in \mathcal{C}^{2,\alpha}$; see [25, 31]. The points excluded are the points where either the boundary or the boundary data lacks the necessary smoothness. ∎

Lemma 4.5. *There holds*

$$\sup_{\varepsilon \in (0,1)} \inf_{u \in X_\varepsilon} \mathcal{A}_\varepsilon(u) < \infty.$$

Proof. Define $w_\varepsilon\colon \Omega \to \mathbb{R}$, $w_\varepsilon(x) = -U_\varepsilon(x_2)\cos(\pi x_1/T)$. Then $w_\varepsilon \in X_\varepsilon$ and

$$|\partial_{x_1} w_\varepsilon| \le \frac{2L\pi}{T}, \quad |\partial_{x_2} w_\varepsilon| \le 1 + \varepsilon^\beta.$$

Hence, for sufficiently small $\varepsilon > 0$, we have $\nabla w^\varepsilon(x) \in K^\varepsilon$ and therefore

$$\begin{aligned}\int_\Omega \widehat{F}_\varepsilon(\nabla w_\varepsilon(x))\,dx &= \int_\Omega F_\varepsilon(\nabla w_\varepsilon(x))\,dx \\ &= \int_\Omega \frac{U_\varepsilon(x_2)^2 \sin^2(\pi x_1/T)^2\pi^2/T^2 + \varepsilon^\theta}{(1+\varepsilon)^2 - U_\varepsilon'(x_2)^2\cos^2(\pi x_1/T)}\,dx \\ &\le \frac{2L\pi^2 U_\varepsilon(0)^2}{T(1+\varepsilon)^2} + \frac{2LT\varepsilon^\theta}{(1+\varepsilon)^2-(1+\varepsilon^\beta)^2}.\end{aligned}$$

The right-hand side is uniformly bounded as $\varepsilon \to 0$, since the exponents θ, β are both bigger than 1.

The uniform boundedness of $\mathcal{A}_\varepsilon(w_\varepsilon)$ follows since V is uniformly bounded due to condition (V_{reg}). ■

4.3. A priori bounds

Next we establish some first a priori bounds for solutions of the regularized problem (4.15). Let u_ε be such a solution. In view of Lemma 4.4, u_ε is a classical, and therefore in particular also a viscosity, solution of the associated Euler–Lagrange equation

$$\operatorname{div}(\nabla \widehat{F}_\varepsilon(\nabla u)) + \partial_z V(x,u) = D^2\widehat{F}_\varepsilon(\nabla u) : D^2 u + \partial_z V(x,u) = 0. \tag{4.16}$$

We quickly recall the notion of being a solution in viscosity sense.

Definition 4.1. A viscosity subsolution of (4.16) is a continuous function $u\colon \Omega \to \mathbb{R}$, such that whenever u is touched at a point $x_0 \in \Omega$ from above by a function $\varphi \in \mathcal{C}^2(B_\delta(x_0))$, then

$$D^2\widehat{F}_\varepsilon(\nabla\varphi(x_0)) : D^2\varphi(x_0) + \partial_z V(x_0,\varphi(x_0)) \ge 0.$$

If the above inequality is strict in every such situation we say that u is a strict viscosity subsolution. The notion of a (strict) supersolution is defined analogously, and a viscosity solution is both a viscosity sub- and supersolution.

Lemma 4.6. *For $\varepsilon > 0$ sufficiently small, the functions $\Omega \ni x \mapsto U_\varepsilon(x_2) + c \in \mathbb{R}$, $c \in \mathbb{R}$ are strict viscosity supersolutions. The corresponding functions induced by $-U_\varepsilon$ are strict viscosity subsolutions.*

Proof. Consider U_ε as an x_1-independent function defined on Ω. For $x_0 \in \Omega$ with $x_{0,2} \neq 0$ there holds

$$\begin{aligned}D^2\widehat{F}_\varepsilon(\nabla U_\varepsilon(x_0)) : \nabla^2 U_\varepsilon(x_0) &= -\frac{\varepsilon^\beta}{L}\partial^2_{p_2} F_\varepsilon(\nabla U_\varepsilon(x_0)) \\ &\le \frac{-\varepsilon^{\beta+\theta}}{L\big((1+\varepsilon)^2 - (\operatorname{sign}(x_{0,2}) + \frac{\varepsilon^\beta}{L}x_{0,2})^2\big)^3} \le -C\varepsilon^{\beta+\theta-3}\end{aligned}$$

for a suitable constant $C > 0$ independent of x_0. By the choice of the exponents $\theta \in (1, 2)$, $\beta \in (1, 3 - \theta)$ and the boundedness of $\partial_z V$ (cf. ($\mathrm{V}_{\mathrm{reg}}$)), we deduce that

$$D^2 \hat{F}_\varepsilon(\nabla U_\varepsilon(x_0)) : \nabla^2 U_\varepsilon(x_0) + \partial_z V(x_0, U_\varepsilon(x_0) + c) < 0$$

provided $\varepsilon \in (0, \varepsilon_0)$ for some $\varepsilon_0 > 0$ small enough and independent of x_0 and c. The same inequality holds true for any $\mathcal{C}^2$ function φ touching $U_\varepsilon + c$ from below at x_0. Note that $U_\varepsilon + c$ can only be touched by a $\mathcal{C}^2$ function from below at points with $x_{0,2} \neq 0$. Thus $U_\varepsilon + c$ is a strict viscosity supersolution. Similarly, one concludes that $-U_\varepsilon + c$ is a strict viscosity subsolution. ∎

Corollary 4.7. *For $\varepsilon > 0$ sufficiently small, any minimizer $u_\varepsilon \in X_\varepsilon$ of $\mathcal{A}_\varepsilon$ satisfies $|u_\varepsilon(x)| \leq U_\varepsilon(x_2)$ for all $x \in \bar{\Omega}$. Moreover, the inequality is strict for $x \in \Omega$.*

Proof. Consider again U_ε as a function defined on all of $\bar{\Omega}$. Assuming to the contrary that u_ε touches $U_\varepsilon + c$ for some $c \geq 0$ from below at a point $x_0 \in \Omega$ with $x_{0,2} \neq 0$ (touching at a point with $x_{0,2} = 0$ again is not possible) directly gives a contradiction with the fact that $U_\varepsilon + c$ is a strict viscosity supersolution and $u_\varepsilon \in \mathcal{C}^2(\Omega)$ is a solution. Hence $u_\varepsilon < U_\varepsilon$ on Ω. An analogous statement regarding the comparison of u_ε with $-U_\varepsilon$ is obtained similarly. ∎

4.4. Autonomous potentials

In this section, under condition ($\mathrm{V}_{\mathrm{aut}}$) we conclude the positiveness of $\partial_{x_1} u_\varepsilon$ and the existence of a first integral, which is the total energy. Moreover, we lay the basis for Lemma 3.2.

Lemma 4.8. *Assume that V in addition to ($\mathrm{V}_{\mathrm{reg}}$) satisfies ($\mathrm{V}_{\mathrm{aut}}$) and let u_ε be a solution of (4.15). Then $\partial_{x_1} u_\varepsilon \geq 0$ on Ω for $\varepsilon > 0$ sufficiently small.*

Proof. Consider the function $w := u_\varepsilon + U_\varepsilon$, which is nonnegative by Corollary 4.7. As before, U_ε is considered here as an x_1-independent function defined on $\bar{\Omega}$. For every $x \in \Omega$, $x_2 \neq 0$, Lemma 4.6 implies

$$\operatorname{div}(A(x)\nabla w(x)) + c(x)w(x) < 0,$$

where

$$A(x) := \int_0^1 D^2 \hat{F}_\varepsilon(-\nabla U_\varepsilon + s\nabla w)\, ds, \quad c(x) := \int_0^1 \partial_z^2 V(x, -U_\varepsilon + sw)\, ds.$$

Splitting the zeroth-order term into $c = c^+ - c^-$ with $c^+, c^- \geq 0$ and neglecting $c^+ w \geq 0$, the Hopf maximum principle implies that $\partial_{x_1} u_\varepsilon(0, x_2) = \partial_{x_1} w(0, x_2) > 0$ for $x_2 \in (-L, L)$, $x_2 \neq 0$. Similarly one sees that $\partial_{x_1} u_\varepsilon(T, x_2) > 0$ for $x_2 \in (-L, L)$, $x_2 \neq 0$.

Next we claim that $(\partial_{x_1} u_\varepsilon)^- \in H_0^1(\Omega)$. Indeed, the difference quotients

$$v_h(x) := \frac{u_\varepsilon(x_1 + h, x_2) - u_\varepsilon(x)}{h}$$

satisfy $v_h^- \in H_0^1(\Omega_h)$, $\Omega_h := (0, T-h) \times (-L, L)$ by Corollary 4.7. We therefore can use v_h^- as a test function for the equation

$$\operatorname{div}(B(x)\nabla v_h(x)) + d(x)v_h(x) = 0, \quad x \in \Omega_h, \tag{4.17}$$

where

$$\begin{aligned} B(x) &:= \int_0^1 D^2 \widehat{F}_\varepsilon(\nabla u_\varepsilon(x) + sh\nabla v_h(x))\, ds, \\ d(x) &:= \int_0^1 \partial_z^2 V(x, u_\varepsilon(x) + shv_h(x))\, ds. \end{aligned}$$

Note that the potential V is autonomous with respect to x_1 by assumption ($\mathrm{V}_{\mathrm{aut}}$). Now testing (4.17) with v_h^- we obtain

$$\begin{aligned} 0 &= \int_{\{v_h<0\}} B(x)\nabla v_h \cdot \nabla v_h^- - d(x)v_h v_h^-\, dx \\ &= \int_{\Omega_h} B(x)\nabla v_h^- \cdot \nabla v_h^- - d(x)(v_h^-)^2\, dx. \end{aligned}$$

Hence the uniform ellipticity of $D^2\widehat{F}_\varepsilon$ (for fixed ε), the boundedness of $\partial_z^2 V$ and $\partial_{x_1} u_\varepsilon \in L^2(\Omega)$ imply a bound on $\|\nabla v_h^-\|_{L^2(\Omega_h)}$, or $\|\nabla v_h^-\|_{L^2(\Omega)}$ when extending ∇v_h^- by 0 outside Ω_h. Now, by the regularity of u_ε in Ω, there holds $\nabla v_h^-(x) \to \nabla(\partial_{x_1} u_\varepsilon)^-(x)$ for all $x \in \Omega$ as $h \to 0$, which together with the L^2 bound for ∇v_h implies $\nabla(\partial_{x_1} u_\varepsilon)^- \in L^2(\Omega)$. Therefore, $(\partial_{x_1} u_\varepsilon)^- \in H_0^1(\Omega)$.

Now using $\psi := (\partial_{x_1} u_\varepsilon)^-$ as a test function for the differentiated equation

$$\operatorname{div}(C(x)\nabla \partial_{x_1} u_\varepsilon(x)) + e(x)\partial_{x_1} u_\varepsilon(x) = 0, \quad x \in \Omega,$$

where this time $C(x) := D^2\widehat{F}_\varepsilon(\nabla u_\varepsilon(x))$, $e(x) = \partial_z^2 V(x, u_\varepsilon(x))$, we deduce

$$0 = \int_\Omega C(x)\nabla\psi \cdot \nabla\psi - e(x)\psi^2\, dx.$$

On the other hand, since u_ε is a minimizer of $\mathcal{A}_\varepsilon$, we also have for any $\phi \in H_0^1(\Omega)$ that

$$0 \le \frac{d^2}{ds^2}_{|s=0} \mathcal{A}_\varepsilon(u_\varepsilon + s\phi) = \int_\Omega C(x)\nabla\phi \cdot \nabla\phi - e(x)\phi^2\, dx.$$

Hence, if we assume $\psi \neq 0$, then the first eigenvalue of the self-adjoint operator $\mathcal{L}\phi := -\operatorname{div}(C(x)\nabla\phi) - e(x)\phi$ is 0 and ψ is in the associated eigenspace. However, the eigenspace associated with the first eigenvalue is one-dimensional and spanned by a function which is positive a.e. on Ω. This contradicts the fact that ψ is vanishing in a neighborhood of $(0, L/2)$, due to $\partial_{x_1} u_\varepsilon(0, L/2) > 0$ and the continuity of $\partial_{x_1} u_\varepsilon$ at that point. In consequence, $(\partial_{x_1} u_\varepsilon)^- = \psi = 0$. ■

Lemma 4.9. *Under the additional assumption* ($\mathrm{V_{aut}}$) *the quantity*

$$\int_{-L}^{L} \partial_{p_1} \widehat{F}_\varepsilon(\nabla u_\varepsilon(x))\partial_{x_1} u_\varepsilon(x) - \widehat{F}_\varepsilon(\nabla u_\varepsilon(x)) + V(x_2, u_\varepsilon(x))\, dx_2 \tag{4.18}$$

is independent of $x_1 \in (0, T)$ *for any solution* u_ε *of* (4.15).

Proof. Let us denote the function in the integral (4.18) by $\widehat{H}_\varepsilon(x)$. Using the Euler–Lagrange equation (4.16) and ($\mathrm{V_{aut}}$) one easily checks that inside Ω there holds

$$\begin{aligned}&\partial_{x_1} \widehat{H}_\varepsilon + \partial_{x_2}(\partial_{p_2} \widehat{F}_\varepsilon(\nabla u_\varepsilon)\partial_{x_1} u_\varepsilon)\\ &\quad = \operatorname{div}(\nabla_p \widehat{F}_\varepsilon(\nabla u_\varepsilon)\partial_{x_1} u_\varepsilon) + \partial_z V(\cdot, u_\varepsilon)\partial_{x_1} u_\varepsilon - \partial_{x_1}(\widehat{F}_\varepsilon(\nabla u_\varepsilon)) = 0.\end{aligned}$$

Integration and $\partial_{x_1} u_\varepsilon = 0$ on $(0, T) \times \{\pm L\}$ imply the stated conservation. ∎

Lemma 4.10. *Suppose* ($\mathrm{V_{aut}}$) *and* $\partial_z^2 V(x_2, z) \geq 0$ *for* $|x_2| \leq L$, $|z| \leq L + 1$*; then there holds the following one-sided maximum principle for* $\partial_{x_1} u_\varepsilon$*:*

$$\inf_{\Omega'} \partial_{x_1} u_\varepsilon = \inf_{\partial\Omega'} \partial_{x_1} u_\varepsilon$$

for all $\Omega' \subset\subset \Omega$.

Proof. Differentiation, the imposed convexity of V with respect to z, Corollary 4.7, and Lemma 4.8 show that $w_\varepsilon := \partial_{x_1} u_\varepsilon$ indeed satisfies

$$-\operatorname{div}(\nabla^2 \widehat{F}_\varepsilon(\nabla u_\varepsilon)\nabla w_\varepsilon) = \partial_z^2 V(x_2, u_\varepsilon) w_\varepsilon \geq 0.$$ ∎

5. Γ-convergence and Young measure representation

In this section we will show that the found regular minimizers converge to a minimizer of the corresponding unperturbed variational problems.

Proposition 5.1. *Let* u_ε, $\varepsilon > 0$ *be a solution of* (4.15)*; then there exists a solution* u *of the variational problem* (3.2) *such that* $u_\varepsilon \rightharpoonup u$ *in* $H^1(\Omega)$ *along a subsequence.*

The proof of Proposition 5.1 is contained in the proof of Proposition 5.4 below, which by means of the Young measure representation will also characterize where and how strong convergence can fail.

5.1. The recovery sequence

Recall that X consists of all $u \in H^1(\Omega)$ which satisfy $\|\partial_{x_2} u\|_{L^\infty(\Omega)} \leq 1$ and agree in the trace sense with

$$\widehat{U}_0(x) := \left(\frac{2x_1}{T} - 1\right) U_0(x_2) \tag{5.1}$$

on $\partial\Omega$, where $U_0(x_2) = L - |x_2|$.

Lemma 5.2. *For every $u \in X$ with $\mathcal{A}(u) < \infty$ there exists a sequence $u^\varepsilon \in X_\varepsilon$ with $u^\varepsilon \to u$ in $H^1(\Omega)$ and*

$$\limsup_{\varepsilon\to 0} \mathcal{A}_\varepsilon(u^\varepsilon) \le \mathcal{A}(u).$$

Proof. We first define an $H^1(\Omega)$ function $\tilde{u}^\delta$ by setting

$$\tilde{u}^\delta(x) := \begin{cases} -\cos(x_1)U_0(x_2), & x_1 \in (0,\delta), \\ \cos(\delta)u\Big(\frac{T}{T-2\delta}(x_1-\delta), x_2\Big), & x_1 \in (\delta, T-\delta), \\ \cos(T-x_1)U_0(x_2), & x_1 \in (T-\delta, T). \end{cases}$$

It is easy to see that $\tilde{u}^\delta$ actually belongs to X and that $\tilde{u}^\delta \to u$ in $H^1(\Omega)$ as $\delta \to 0$.

Next we improve the integrability of $\partial_{x_1}\tilde{u}^\delta$ from L^2 to L^∞ in order to be able to evaluate the extension $\widehat{F}_\varepsilon$ on the set K^ε where it agrees with the old F_ε. In order to do this, fix $\delta > 0$ and extend $\tilde{u}^\delta$ onto $(-\delta, T+\delta) \times (-L, L)$ by setting

$$\tilde{u}^\delta(x) = \begin{cases} -(2-\cos(x_1))U_0(x_2), & x_1 \in (-\delta, 0), \\ (2-\cos(T-x_1))U_0(x_2), & x_1 \in (T, T+\delta). \end{cases} \tag{5.2}$$

Due to this symmetric extension we can now convolute $\tilde{u}^\delta$ in the x_1-direction and conserve the boundary data, i.e. for $x \in \Omega$ we set

$$\tilde{u}^{\eta,\delta}(x) := (\varphi_\eta *_1 \tilde{u}^\delta)(x) = \int_{\mathbb{R}} \varphi_\eta(x_1 - s)\tilde{u}^\delta(s, x_2)\, ds$$

with a symmetric one-dimensional kernel φ_η of scale $\eta \in (0, \delta/2)$.

Observe that $\tilde{u}^{\eta,\delta}$ indeed agrees with $\widehat{U}_0$ on $\partial\Omega$ and that for a.e. $x \in \Omega$ there holds

$$|\partial_{x_1}\tilde{u}^{\eta,\delta}(x)| \le c\eta^{-1}\|\tilde{u}^\delta\|_{L^\infty}, \quad |\partial_{x_2}\tilde{u}^{\eta,\delta}(x)| \le 2-\cos(\delta) = 1 + \frac{\delta^2}{2} + o(\delta^2)$$

with a constant $c > 0$ depending only on the kernel φ_1. Since $\|\tilde{u}^\delta\|_{L^\infty}$ is bounded (as a bounded extension of a function in X), we will pick $\eta(\varepsilon) = \varepsilon^\theta$ and $\delta(\varepsilon) = \varepsilon$ in order to satisfy $\nabla\tilde{u}^{\eta(\varepsilon),\delta(\varepsilon)} \in K^\varepsilon$ a.e. for sufficiently small ε.

It remains to adapt the boundary data. Therefore, we finally set

$$u^\varepsilon(x) := \tilde{u}^{\varepsilon^\theta,\varepsilon}(x) + \widehat{U}_\varepsilon(x) - \widehat{U}_0(x),$$

where in analogy with (5.1) the function $\widehat{U}_\varepsilon\colon \Omega \to \mathbb{R}$ is defined as

$$\widehat{U}_\varepsilon(x) = \Big(\frac{2x_1}{T} - 1\Big)U_\varepsilon(x_2) = \widehat{U}_0(x) + \Big(\frac{2x_1}{T} - 1\Big)\frac{\varepsilon^\beta}{2L}(L^2 - x_2^2). \tag{5.3}$$

By our construction it is clear that $u^\varepsilon \in X_\varepsilon$ and that still $\nabla u^\varepsilon(x) \in K^\varepsilon$ for a.e. $x \in \Omega$ and $\varepsilon > 0$ small enough. Hence $\widehat{F}_\varepsilon \circ \nabla u^\varepsilon = F_\varepsilon \circ \nabla u^\varepsilon$ a.e.

Next we will show that $u^\varepsilon \to u$ in $H^1(\Omega)$. Clearly, $u^\varepsilon - \tilde{u}^{\varepsilon^\theta,\varepsilon} \to 0$ in $H^1(\Omega)$ as $\varepsilon \to 0$. In order to see that $\tilde{u}^{\varepsilon^\theta,\varepsilon} \to u$, let $\tilde{u} \in H^1((-1, T+1) \times (-L, L))$ be the extension of u with the values given in (5.2) for $\tilde{u}^\delta$. Then

$$\begin{aligned}\|\tilde{u}^{\varepsilon^\theta,\varepsilon} - u\|_{H^1(\Omega)} &\le \|\varphi_{\varepsilon^\theta} *_1 (\tilde{u}^\varepsilon - \tilde{u})\|_{H^1(\Omega)} + \|\varphi_{\varepsilon^\theta} *_1 \tilde{u} - \tilde{u}\|_{H^1(\Omega)} \\ &\le \|\tilde{u}^\varepsilon - \tilde{u}\|_{H^1(\Omega)} + \|\varphi_{\varepsilon^\theta} *_1 \tilde{u} - \tilde{u}\|_{H^1(\Omega)} \to 0.\end{aligned}$$

Finally, it remains to look at the values of the action functionals. The just shown convergence $u^\varepsilon \to u$ in $H^1(\Omega)$ and condition ($\mathrm{V}_{\mathrm{reg}}$) easily imply

$$\int_\Omega V(x, \tilde{u}^\varepsilon(x))\, dx \to \int_\Omega V(x, u(x))\, dx.$$

Concerning $\widehat{F}_\varepsilon$ we first of all observe that

$$\begin{aligned}\limsup_{\varepsilon\to 0} \int_\Omega \widehat{F}_\varepsilon(\nabla u^\varepsilon)\, dx &= \limsup_{\varepsilon\to 0} \int_\Omega F_\varepsilon(\nabla u^\varepsilon)\, dx \\ &\le \limsup_{\varepsilon\to 0} \int_\Omega \frac{(\partial_{x_1}\tilde{u}^{\varepsilon^\theta,\varepsilon})^2}{2(1+\varepsilon-(\partial_{x_2}\tilde{u}^{\varepsilon^\theta,\varepsilon})^2)}\, dx,\end{aligned}$$

where we have used that $u^\varepsilon - \tilde{u}^{\varepsilon^\theta,\varepsilon} = O(\varepsilon^\beta)$ in $W^{1,\infty}(\Omega)$ as $\varepsilon \to 0$ and $\theta, \beta > 1$. The function $\widetilde{F}_\varepsilon\colon \{p \in \mathbb{R}^2 : |p_2| < 1+\varepsilon\} \to \mathbb{R}$,

$$\widetilde{F}_\varepsilon(p) = \frac{p_1^2}{2(1+\varepsilon-p_2^2)},$$

is also convex. Hence Jensen's inequality implies

$$\begin{aligned}\int_\Omega \widetilde{F}_\varepsilon(\nabla \tilde{u}^{\varepsilon^\theta,\varepsilon})\, dx &= \int_\Omega \widetilde{F}_\varepsilon(\varphi_{\varepsilon^\theta} *_1 \nabla\tilde{u}^\varepsilon)\, dx \le \int_\Omega \varphi_{\varepsilon^\theta} *_1 (\widetilde{F}_\varepsilon \circ \nabla\tilde{u}^\varepsilon)\, dx \\ &= \int_D \varphi_1(s)\widetilde{F}_\varepsilon(\nabla\tilde{u}^\varepsilon(x_1 + s\varepsilon^\theta, x_2))\, d(s,x),\end{aligned}$$

where $D := (-1,1)\times\Omega$. We now split the domain of integration into the following sets:

$$D_1 = \big\{(s,x) \in D : x_1 + s\varepsilon^\theta < 0\big\}, \quad D_2 = \big\{(s,x) \in D : 0 < x_1 + s\varepsilon^\theta < \varepsilon\big\},$$
$$D_3 = \big\{(s,x) \in D : \varepsilon < x_1 + s\varepsilon^\theta < T - \varepsilon\big\},$$
$$D_4 = \big\{(s,x) \in D : T-\varepsilon < x_1 + s\varepsilon^\theta < T\big\}, \quad D_5 = \big\{(s,x) \in D : x_1 + s\varepsilon^\theta > T\big\}$$

and estimate the corresponding integrals $I_1, \dots, I_5$. There holds

$$\begin{aligned}I_1 &= \int_{D_1} \varphi_1(s) \frac{\sin^2(x_1+s\varepsilon^\theta)U_0(x_2)^2}{2\big(1+\varepsilon-(2-\cos(x_1+s\varepsilon^\theta))^2\big)}\, d(s,x) \\ &\le L^3\|\varphi_1\|_{L^\infty} \int_{-1}^1 \int_0^{\varepsilon^\theta} \frac{1}{1+\frac{\varepsilon-4+4\cos(x_1+s\varepsilon^\theta)}{\sin^2(x_1+s\varepsilon^\theta)}}\, dx_1\, ds \le 2\varepsilon^\theta L^3 \|\varphi_1\|_{L^\infty} \frac{1}{1+\frac{\varepsilon}{2}}.\end{aligned}$$

Hence $I_1 \to 0$ as $\varepsilon \to 0$. A similar reasoning also yields $I_2 + I_4 + I_5 \to 0$ as $\varepsilon \to 0$.

For I_3 we use the transformation $x_1 \mapsto y_1(s, x_1) = \frac{T}{T-2\varepsilon}(x_1 + s\varepsilon^\theta - \varepsilon)$ in order to see that

$$I_3 = \int_{-1}^{1} \varphi_1(s) \int_{\varepsilon - s\varepsilon^\theta}^{T-\varepsilon-s\varepsilon^\theta} \int_{-L}^{L} \frac{(\frac{T}{T-2\varepsilon})^2 \big(\partial_{x_1} u(y_1(s,x_1), x_2)\big)^2}{2\big(\frac{1+\varepsilon}{\cos^2(\varepsilon)} - \big(\partial_{x_2} u(y_1(s,x_1), x_2)\big)^2\big)} \, dx_2 \, dx_1 \, ds$$

$$\leq \frac{T}{T-2\varepsilon} \int_0^T \int_{-L}^{L} \widetilde{F}_\varepsilon(\nabla u(y_1, x_2)) \, dx_2 \, dy_1 \to \int_\Omega F(\nabla u) \, dx$$

by means of monotone convergence, since $F(\nabla u) \in L^1(\Omega)$ by assumption.

Altogether we therefore have shown

$$\limsup_{\varepsilon \to 0} \int_\Omega \widehat{F}_\varepsilon(\nabla u^\varepsilon) - V(x, u^\varepsilon) \, dx \leq \int_\Omega F(\nabla u) - V(x, u) \, dx. \qquad \blacksquare$$

5.2. Young measures

Before stating a general convergence result, which contains the weak lower semicontinuity for the Γ-convergence, we quickly recall the notion of generalized Young measures; see e.g. [33].

Let S denote the unit sphere of $\mathbb{V} = \mathbb{R}^m$ and Ω be a bounded domain in $\mathbb{R}^l$ with $\partial\Omega$ having 0 Lebesgue measure. Every weakly converging sequence $U_k \rightharpoonup U$ in $L^q(\Omega, \mathbb{V})$ induces a q-Young measure $\boldsymbol{\nu} = ((\nu_x)_{x\in\Omega}, \lambda, (\nu_x^\infty)_{x\in\overline{\Omega}})$. That is, $(\nu_x)_{x\in\Omega}$ is a Lebesgue measurable family of probability measures on $\mathbb{V}$ (oscillation measure), λ is a positive measure on $\overline{\Omega}$ (concentration measure), and $(\nu_x^\infty)_{x\in\overline{\Omega}}$ is a λ-measurable family of probability measures on S (concentration angle measure), such that for all q-admissible integrands Φ there holds

$$\int_\Omega \Phi(x, U_k(x)) \, dx \to \int_\Omega \int_{\mathbb{V}} \Phi(x, v) \, d\nu_x(v) \, dx + \int_{\overline{\Omega}} \int_S \Phi^{q,\infty}(x, v) \, d\nu_x^\infty(v) \, d\lambda(x).$$

A continuous function $\Phi: \overline{\Omega} \times \mathbb{V} \to \mathbb{R}$ is q-admissible provided the q-recession function

$$\Phi^{q,\infty}(x, v) := \lim_{t\to\infty} \frac{\Phi(x, tv)}{t^q}$$

exists, is finite, and the convergence is locally uniform with respect to $(x, v) \in \overline{\Omega} \times (\mathbb{V} \setminus \{0\})$.

In terms of the Young measure, the weak limit U of U_k is represented as the barycenter of ν_x, i.e.

$$U(x) = \int_{\mathbb{V}} v \, d\nu_x(v) \quad \text{for a.e. } x \in \Omega.$$

Strong convergence $U_k \to U$ in $L^q(\Omega, \mathbb{V})$ can equivalently be characterized by the absence of concentration and oscillation, i.e. $\lambda = 0$ and $\nu_x = \delta_{U(x)}$ for a.e. $x \in \Omega$.

5.3. A general convergence result

While the existence of a recovery sequence has been a construction specific to our particular problem, we now investigate a more general class of variational problems and their approximations. The more general setting in this subsection is the following.

Let $0 \in C \subset \mathbb{R}^m$ be open, convex and $f : \bar{C} \to [0, \infty]$ be a convex lower semi-continuous function. We suppose further that there exist convex functions $f_k : \mathbb{R}^m \to [0, \infty)$ with

$$f_k(0) = 0, \quad f_k(p) \geq c|p|^q - d \text{ for all } p \in \mathbb{R}^m, \tag{5.4}$$

for some constants $c, d > 0$ and $q \in (1, \infty)$ independent of k, and

$$f_k \to f \text{ uniformly on any compact subset of } C \text{ as } k \to \infty, \tag{5.5}$$

$$f_k(p) \to +\infty \text{ uniformly on any set with positive distance to } \bar{C}. \tag{5.6}$$

Observe that these conditions imply $f(0) = 0$ and that f is continuous and finite on C.

We then consider the functionals

$$\mathcal{F}_k(U) := \int_\Omega f_k(U(x))\,dx, \quad \mathcal{F}(U) := \int_\Omega f(U(x))\,dx,$$

where $U \in L^q(\Omega; \mathbb{R}^m)$, $\Omega \subset \mathbb{R}^l$ open, bounded with $\partial\Omega$ being a Lebesgue null set.

Lemma 5.3. *Suppose $U_k \rightharpoonup U$ in $L^q(\Omega; \mathbb{R}^m)$ and let $\boldsymbol{\nu} = ((\nu_x)_{x \in \Omega}, \lambda, (\nu_x^\infty)_{x \in \bar{\Omega}})$ be the associated q-Young measure. Assume that $\sup_k \mathcal{F}_k(U_k) < \infty$. Then*

(i) *there holds* $\operatorname{supp} \nu_x \subset \bar{C}$ *for a.e.* $x \in \Omega$,

(ii) *there holds* $\mathcal{F}(U) \leq \liminf_{k \to +\infty} \mathcal{F}_k(U_k)$,

(iii) *and if there holds equality in* (ii)*, then $\lambda = 0$, and for a.e. $x \in \Omega$ there holds*

$$\operatorname{supp} \nu_x \cap \{U(x) + \mathbb{R}w\} = \{U(x)\},$$

whenever $w \in \mathbb{R}^m$ satisfies

$$f(U(x) + sw) - f(U(x)) > sz_0 \cdot w \tag{5.7}$$

for a subgradient $z_0 \in \mathbb{R}^m$ of f at $U(x)$ and any $s \neq 0$ with $U(x) + sw \in \bar{C}$.

Part (i) implies $U(x) \in \bar{C}$ for a.e. $x \in \Omega$. Part (ii) states the weak lower semi-continuity, except that the functionals are changing along the sequence. Part (iii) shows that, if the values of the functional actually converge, then there is no concentration ($\lambda = 0$). Furthermore, strict convexity of the integrand f in direction w excludes oscillations in that direction. In consequence, if f is strictly convex (in all directions) on C, $U(x) \in C$ for a.e. $x \in \Omega$, then $U_k \to U$ strongly in $L^q(\Omega; \mathbb{R}^m)$. Note here that a subgradient z_0 of f at $U(x)$, i.e. a vector z_0 for which (5.7) holds with sw replaced by any $\tilde{w} \in \bar{C} - U(x)$ and $>$ replaced by $\geq$, always exists when $U(x) \in C$. In part (iii) the existence of a corresponding subgradient z_0 is part of the condition.

Proof of Lemma 5.3. *Part* (i). For $\delta > 0$ define $K_\delta := \{p \in \mathbb{R}^m : \mathrm{dist}(p, \overline{C}) \geq \delta\}$. The uniform boundedness of $\mathcal{F}_k(U_k)$ and condition (5.6) imply that

$$\left|\{x \in \Omega : U_k(x) \in K_\delta\}\right| \to 0$$

as $k \to \infty$, where $|\cdot|$ denotes the l-dimensional Lebesgue measure. Now for any $\Phi \in \mathcal{C}^0(\mathbb{R}^m)$ with $\Phi = 0$ on $\mathbb{R}^m \setminus K_\delta$ and $0 \leq \Phi \leq 1$ on K_δ, there holds

$$0 = \lim_{k\to\infty} \left|\{x \in \Omega : U_k(x) \in K_\delta\}\right| \geq \lim_{k\to\infty} \int_\Omega \Phi(U_k(x))\,dx = \int_\Omega \int_{\mathbb{R}^m} \Phi(v)\,d\nu_x(v)\,dx$$

by the definition of the Young measure. Hence $\operatorname{supp} \nu_x \subset \mathbb{R}^m \setminus K_\delta$ for a.e. $x \in \Omega$, for all $\delta > 0$.

Part (ii). Let $C_j \subset \mathbb{R}^m$, $j \in \mathbb{N}$ be compact convex sets with

$$0 \in C_j \subset C_{j+1} \subset\subset C, \quad \bigcup_{j\in\mathbb{N}} C_j = C$$

and define $T_j : \mathbb{R}^m \to C_j$ as the radial retraction onto C_j, that is $T_j(p) = r_j(p)p$, where $r_j(p) := \inf\{r > 0 : r^{-1}p \in C_j\}$. Note that r_j is convex. Thus T_j is continuous.

The convexity of f_k and $f_k(0) = 0$ (cf. (5.4)) imply that

$$f_k(T_j(p)) = f_k\big(r_j(p)p + (1 - r_j(p))0\big) \leq r_j(p) f_k(p) + 0 \leq f_k(p) \tag{5.8}$$

for every $k, j \in \mathbb{N}$, $p \in \mathbb{R}^m$.

In consequence, the functions $g_{k,j}, g_j : \mathbb{R}^m \to [0, \infty)$,

$$g_{k,j}(p) := \max\big\{f_k(T_j(p)), \tfrac{c}{2}|p|^q - d\big\}, \quad g_j(p) := \max\big\{f(T_j(p)), \tfrac{c}{2}|p|^q - d\big\}$$

with the constants c, d, q taken from (5.4), are also continuous, and there holds

$$g_{k,j}(p) \leq f_k(p) \quad \text{for any } k, j \in \mathbb{N},\ p \in \mathbb{R}^m. \tag{5.9}$$

By condition (5.5) we also see that $g_{k,j} \to g_j$ uniformly on all of $\mathbb{R}^m$ as $k \to \infty$.

Furthermore, the functions g_j are q-admissible (autonomous) integrands. Indeed, $f_{|C_j}$ is bounded by the compactness of C_j and therefore the q-regression function reads

$$g_j^{q,\infty}(p) = \lim_{t\to\infty} t^{-q} g_j(tp) = \frac{c}{2}|p|^q,$$

while the convergence is locally uniform with respect to $p \in \mathbb{R}^m \setminus \{0\}$.

By (5.9), the uniform convergence $g_{k,j} \to g_j$, and the definition of the Young measure associated with $(U_k)_{k\in\mathbb{N}}$, we therefore obtain for any $j \in \mathbb{N}$,

$$\begin{aligned}
\liminf_{k\to\infty} \mathcal{F}_k(U_k) &\geq \liminf_{k\to\infty} \int_\Omega g_{k,j}(U_k(x))\,dx \\
&\geq \liminf_{k\to\infty} \int_\Omega g_j(U_k(x))\,dx + \liminf_{k\to\infty} \int_\Omega g_{k,j}(U_k(x)) - g_j(U_k(x))\,dx \\
&= \int_\Omega \int_{\overline{C}} g_j(v)\,d\nu_x(v)\,dx + \frac{c}{2}\lambda(\overline{\Omega}) + 0.
\end{aligned}$$

In the last step we also used (i).

Now, similarly to (5.8), one can check that the sequence g_j is monotone increasing. Thus by monotone convergence, assumption (5.4), and the lower semi-continuity of f we conclude

$$\liminf_{k\to\infty} \mathcal{F}_k(U_k) \ge \int_\Omega \int_{\bar{C}} \lim_{j\to\infty} g_j(v)\, d\nu_x(v)\, dx + \frac{c}{2}\lambda(\bar{C})$$
$$\ge \int_\Omega \int_{\bar{C}} f(v)\, d\nu_x(v)\, dx + \frac{c}{2}\lambda(\bar{C}).$$

Therefore, Jensen's inequality finally shows that

$$\liminf_{k\to\infty} \mathcal{F}_k(U_k) \ge \int_\Omega f(U(x))\, dx + \frac{c}{2}\lambda(\bar{\Omega}) \ge \mathcal{F}(U). \tag{5.10}$$

This finishes the proof of part (ii).

Part (iii). We immediately see that equality in (5.10) implies $\lambda = 0$. Going one step back there also has to hold

$$\int_{\bar{C}} f(v)\, d\nu_x(v) = f(U(x)) \tag{5.11}$$

for a.e. $x \in \Omega$. Now we fix such a point x and suppose that $f(U(x)) < \infty$, $w \in \mathbb{R}^m \setminus \{0\}$, $z_0 \in \partial f(U(x))$ (the subdifferential of f at $U(x)$) satisfy (5.7) for any $s \neq 0$ with $U(x) + sw \in \bar{C}$.

Assume to the contrary of the statement that $\operatorname{supp} \nu_x \cap \{U(x) + \mathbb{R}w\} \neq \{U(x)\}$, which means that there exists $s_0 \neq 0$ with

$$\nu_x\big(B_r(U(x) + s_0 w)\big) > 0 \tag{5.12}$$

for all $r > 0$. By the properties of f and (5.7) with $s = s_0$ we can pick a radius $r_0 > 0$, such that

$$f(v) - f(U(x)) > z_0 \cdot (v - U(x)) \tag{5.13}$$

for all $v \in B := \bar{C} \cap B_{r_0}(U(x) + s_0 w)$. Combining (5.11), (5.12), (5.13), and the fact that z_0 is a subgradient yields the contradiction

$$\begin{aligned} f(U(x)) &= \int_{\bar{C}\cap B} f(v)\, d\nu_x(v) + \int_{\bar{C}\setminus B} f(v)\, d\nu_x(v) \\ &> \int_{\bar{C}\cap B} f(U(x)) + z_0 \cdot (v - U(x))\, d\nu_x(v) + \int_{\bar{C}\setminus B} f(v)\, d\nu_x(v) \\ &\ge \int_{\bar{C}} f(U(x)) + z_0 \cdot (v - U(x))\, d\nu_x(v) \\ &= f(U(x)). \end{aligned}$$

Hence $\operatorname{supp} \nu_x \cap \{U(x) + \mathbb{R}w\} = \{U(x)\}$. ■

5.4. Our case

Applying Lemmas 5.2, 5.3 to our case we obtain the following extended version of Proposition 5.1.

Proposition 5.4. *Let u_ε, $\varepsilon > 0$ be a solution of (4.15). Then there exists a solution u of the variational problem (3.2), such that $u_\varepsilon \rightharpoonup u$ in $H^1(\Omega)$, $\mathcal{A}_\varepsilon(u_\varepsilon) \to \mathcal{A}(u)$, and*

$$\int_\Omega \widehat{F}_\varepsilon(\nabla u_\varepsilon(x))\,dx \to \int_\Omega F(\nabla u(x))\,dx \tag{5.14}$$

along a subsequence. Moreover, in terms of the Young measure representation of the weak limit of ∇u_ε there holds $\boldsymbol{\nu} = ((\nu_x)_{x\in\Omega}, 0, 0)$ with

$$\nu_x(v) = \begin{cases} \delta_{\nabla u(x)}(v), & \text{for a.e. } x \in \Omega \text{ with } \partial_{x_1} u(x) \neq 0,\\ \delta_{\partial_{x_1} u(x)}(v_1) \otimes \hat{\nu}_x(v_2) & \text{for a.e. } x \in \Omega \text{ with } \partial_{x_1} u(x) = 0, \end{cases}$$

where $v \in \mathbb{R}^2$ and $\hat{\nu}_x$ is a probability measure on $\mathbb{R}$ with support in $[-1, 1]$.

Proof. Let $u_\varepsilon \in X_\varepsilon$ be a minimizer of $\mathcal{A}_\varepsilon(u)$. By Lemma 4.5 the minimal values $\mathcal{A}_\varepsilon(u_\varepsilon)$, $\varepsilon \in (0, 1)$ are bounded. Using the uniform convexity of the extension $\widehat{F}_\varepsilon$ outside the compact set $\{p \in \mathbb{R}^2 : |p_1| \le 1,\ |p_2| \le 1 + \varepsilon - \varepsilon^{4\theta}\}$ (cf. Lemma 4.1), we conclude

$$\widehat{F}_\varepsilon(p) \ge c_1|p|^2 - c_2 \tag{5.15}$$

for all $p \in \mathbb{R}^2$ and some constants $c_1, c_2 > 0$ independent of $\varepsilon \in (0, 1)$. Condition ($\mathrm{V}_{\mathrm{reg}}$) therefore yields

$$\|\nabla u_\varepsilon\|^2_{L^2(\Omega)} \le c_3 \mathcal{A}_\varepsilon(u_\varepsilon) + c_4$$

with ε-independent constants $c_3, c_4 > 0$. In consequence, there exists a subsequence $(u_{\varepsilon_k})_k$, as well as $u \in H^1(\Omega)$ with $u_{\varepsilon_k} \rightharpoonup u$ in $H^1(\Omega)$. Since $u_\varepsilon - \widehat{U}_\varepsilon \in H^1_0(\Omega)$ with $\widehat{U}_\varepsilon$ defined in (5.3) we conclude that u coincides with $\widehat{U}_0$ on $\partial\Omega$ in the trace sense. In order to see that $u \in X^D$ it therefore remains to show $\|\partial_{x_2} u\|_{L^\infty(\Omega)} \le 1$. For this we will rely on Lemma 5.3.

It is easily seen (cf. (5.15)) that the set $C := \{p \in \mathbb{R}^2 : |p_2| < 1\}$, the approximating functions $f_k : \mathbb{R}^2 \to \mathbb{R}$, $f_k(p) := \widehat{F}_{1/k}(p) - \widehat{F}_{1/k}(0)$, $p \in \mathbb{R}^2$, and the limiting function $f : \overline{C} \to \mathbb{R}$, $f(p) := F(p)$ satisfy the conditions postulated in Section 5.3 with exponent $q = 2$. We therefore can apply Lemma 5.3 with

$$\mathcal{F}_k(U) = \int_\Omega f_k(U(x))\,dx, \quad \mathcal{F}(U) = \int_\Omega f(U(x))\,dx,$$

and $U_k := \nabla u_{\varepsilon_k}$. Observe that $\mathcal{F}_k(U_k)$ is indeed uniformly bounded due to the boundedness of $\mathcal{A}_\varepsilon(u_\varepsilon)$ and condition ($\mathrm{V}_{\mathrm{reg}}$). In consequence, Lemma 5.3 tells us that if $\boldsymbol{\nu} = ((\nu_x)_{x\in\Omega}, \lambda, (\nu_x^\infty)_{x\in\overline{\Omega}})$ denotes the Young measure representation for $(U_k)_k$, then $\operatorname{supp} \nu_x \subset \{p \in \mathbb{R}^2 : |p_2| \le 1\}$ and

$$\liminf_{k\to\infty} \int_\Omega \widehat{F}_{\varepsilon_k}(\nabla u_{\varepsilon_k}) - \widehat{F}_{\varepsilon_k}(0)\,dx \ge \int_\Omega F(\nabla u)\,dx.$$

It follows that $|\partial_{x_2} u| \leq 1$ a.e. and

$$\liminf_{k\to\infty} \mathcal{A}_{\varepsilon_k}(u_{\varepsilon_k}) \geq \mathcal{A}(u),$$

since without loss of generality $V(\cdot, u_{\varepsilon_k}) \to V(\cdot, u)$ in $L^1(\Omega)$ by ($\mathrm{V}_{\mathrm{reg}}$).

In particular, $u \in X$ and $\mathcal{A}(u) < \infty$. Hence the recovery sequence Lemma 5.2 on one hand shows that $\mathcal{A}(u) \leq \mathcal{A}(\tilde{u})$ for all $\tilde{u} \in X$, i.e. u is a minimizer of (3.2), and on the other hand it shows that

$$\lim_{k\to\infty} \mathcal{A}_{\varepsilon_k}(u_{\varepsilon_k}) = \mathcal{A}(u),$$

which enables us to utilize part (iii) of Lemma 5.3. Doing this we first of all see that $(\nabla u_{\varepsilon_k})_k$ does not concentrate, i.e. $\lambda = 0$, $\nu_x^\infty = 0$, $x \in \overline{\Omega}$. Next, since

$$F(p + sw) - F(p) - s\nabla F(p) \cdot w > 0$$

for any triple $(p, w, s) \in \{p \in \mathbb{R}^2 : |p_2| < 1\} \times (\mathbb{R}^2 \setminus \{0\}) \times (\mathbb{R} \setminus \{0\})$ with $p_1 \neq 0$ or $w_1 \neq 0$ (cf. (4.1) for $\varepsilon = 0$), there also holds $\operatorname{supp} \nu_x = \{\nabla u(x)\}$ for a.e. $x \in \Omega$ with $\partial_{x_1} u(x) \neq 0$, $|\partial_{x_2} u(x)| < 1$, and $\operatorname{supp} \nu_x \subset \{(0, v_2) : |v_2| \leq 1\}$ for a.e. $x \in \Omega$ with $\partial_{x_1} u(x) = 0$. Note here that 0 is a subgradient of F, or rather $f : \overline{C} \to \mathbb{R}$, at $(0, \pm 1)$. Moreover, due to the fact that the set $\{x \in \Omega : \partial_{x_1} u(x) \neq 0,\ |\partial_{x_2} u(x)| = 1\}$ must be of zero measure, otherwise $\mathcal{A}(u)$ would be infinite, we can also in the first case simply say $\operatorname{supp} \nu_x = \{\nabla u(x)\}$ for a.e. $x \in \Omega$ with $\partial_{x_1} u(x) \neq 0$. This finishes the proof of Proposition 5.4. ■

Proposition 5.4, Corollary 4.7, and Lemma 4.8 directly imply the following bounds.

Corollary 5.5. *The minimizer u satisfies $|u(x)| \leq L - |x_2|$ for a.e. $x \in \Omega$. If in addition ($\mathrm{V}_{\mathrm{aut}}$) holds true, then $\partial_{x_1} u \geq 0$ a.e.*

Remark 5.1. In view of the bound on $|u|$, problem (3.2) can, as the degenerate variational problem arising in the study of random surfaces in [21], be written as an obstacle problem:

$$\min_{|x_2|-L\leq u\leq L-|x_2|} \mathcal{A}(u) \quad (+\text{ boundary conditions}).$$

5.5. Autonomous potentials

In the case of $\partial_{x_1} u_\varepsilon \geq 0$ we can extend the list of convergences as $\varepsilon \to 0$ as follows.

Proposition 5.6. *Let u_ε, u be as in Proposition 5.4 and suppose that in addition there holds ($\mathrm{V}_{\mathrm{aut}}$). Then $u_\varepsilon \to u$ uniformly on $\overline{\Omega}$. In particular, the minimizer u is continuous.*

The proof is a direct consequence of Lemma 4.8, as said, and Lemma 5.7 below.

Lemma 5.7. *For any $u \in \mathcal{C}^0(\overline{\Omega}) \cap \mathcal{C}^1(\Omega)$ with $\partial_{x_1} u \geq 0$ there holds*

$$\operatorname{osc}(u; B_r(z_0)) \lesssim \frac{\|\nabla u\|_{L^2(\Omega)}}{\sqrt{|\log r|}},$$

for any $r \in (0,1)$ *and* $z_0 \in \mathbb{R}^2$ *such that* $B_{\sqrt{r}}(z_0) \subset \Omega$*, as well as*

$$\operatorname{osc}(u; B_r(z_0) \cap \bar{\Omega}) \lesssim \frac{\|\nabla u\|_{L^2(\Omega)}}{\sqrt{|\log r|}} + \operatorname{osc}(u; B_{\sqrt{r}}(z_0) \cap \partial\Omega),$$

for $r \in (0,1)$ *and arbitrary* $z_0 \in \mathbb{R}^2$*.*

The proof relies on the classical Courant–Lebesgue lemma:

Lemma 5.8 (Courant–Lebesgue). *Let* $u \in \mathcal{C}^1(\Omega)$*. For* $z \in \mathbb{R}^2$ *denote the length of* $u(\partial B_r(z) \cap \Omega)$ *by* $L(r) := \int_{\partial B_r(z)\cap\Omega} |\partial_\tau u|$*, where* $\partial_\tau u$ *is the tangential derivative. Then there holds*

$$\int_0^\infty \frac{L(r)^2}{r}\, dr \le 2\pi \|\nabla u\|^2_{L^2(\Omega)},$$

and consequently

$$\min_{a<r<b} L(r)^2 \le \frac{2\pi \|\nabla u\|^2_{L^2(\Omega)}}{\log(b/a)}.$$

Proof. Using the parametrization $\theta \mapsto u(z + re^{i\theta})$ of $\partial B_r(z)$, one has

$$\begin{aligned} L(r) &= r \int_{\{\theta : z + re^{i\theta} \in \Omega\}} |\partial_\tau u(z + re^{i\theta})|\, d\theta \\ &\le r \left(2\pi \int_{\{\theta : z + re^{i\theta} \in \Omega\}} |\nabla u(z + re^{i\theta})|^2\, d\theta \right)^{1/2}, \end{aligned}$$

and hence

$$\int_0^\infty \frac{L(r)^2}{r}\, dr \le 2\pi \int_0^\infty \int_{\{\theta : z + re^{i\theta} \in \Omega\}} |\nabla u(z + re^{i\theta})|^2\, r\, d\theta\, dr = 2\pi \|\nabla u\|^2_{L^2(\Omega)}. \quad \blacksquare$$

Proof of Lemma 5.7. In either of the two cases it follows from Courant–Lebesgue that there exists $\rho > 0$ such that $r < \rho < \sqrt{r}$ and

$$L(\rho)^2 \le \frac{4\pi \|\nabla u\|^2_{L^2(\Omega)}}{|\log(r)|}.$$

If now $B_{\sqrt{r}}(z_0) \subset \Omega$ and $z_1, z_2 \in B_r(z_0)$, we denote by $z_i^{\pm}$ the associated boundary points such that

$$\partial B_\rho(z_0) \cap \{z_i + \mathbb{R}e_1\} = \{z_i^-, z_i^+\}.$$

Hence we may write

$$u(z_1) - u(z_2) \le u(z_1^+) - u(z_2^-) \le L(\rho) \le \frac{\sqrt{4\pi}\|\nabla u\|_{L^2(\Omega)}}{\sqrt{|\log(r)|}},$$

where for the first inequality we use $\partial_{x_1} u \ge 0$. This concludes the proof of the first part of the lemma.

For the second part, given $z_1, z_2 \in B_{\sqrt{r}}(z_0) \cap \overline{\Omega}$, we denote by $z_i^\pm$ the boundary points now given by

$$\partial(B_\rho(z_0) \cap \Omega) \cap \{z_i + \mathbb{R}e_1\} = \{z_i^-, z_i^+\}.$$

We may once more write

$$\begin{aligned} u(z_1) - u(z_2) &\le u(z_1^+) - u(z_2^-) \le L(\rho) + \operatorname{osc}(u; B_\rho(z_0) \cap \partial\Omega) \\ &\le \frac{\sqrt{4\pi}\|\nabla u\|_{L^2(\Omega)}}{\sqrt{|\log(r)|}} + \operatorname{osc}(u; B_{\sqrt{r}}(z_0) \cap \partial\Omega), \end{aligned}$$

which finishes the proof of the lemma. ∎

6. Partial regularity near good points

In this section we will apply the result of Savin [41] to the minimizers $(u_\varepsilon)_\varepsilon$, in order to obtain partial regularity for the limit u. Throughout the section we mean by $(u_\varepsilon)_\varepsilon$ the subsequence from Proposition 5.4 converging to a minimizer u of problem (3.2).

Let

$$G := \{p \in \mathbb{R}^2 : p_1 \neq 0,\ |p_2| < 1\},$$

and note that $D^2F(p)$ is nondegenerate and positive definite for $p \in G$. Therefore, we would like to show that whenever ∇u takes values in this "good" set, one may deduce higher regularity of u via the Euler–Lagrange equation. However, the set of points $x \in \Omega$ for which $\nabla u(x) \in G$ is a priori not open, hence we will need to adapt our argument and use the Euler–Lagrange equations associated with the approximation u_ε in order to deduce the openness of this set, and hence the regularity. The main result of this section is the following proposition, which allows us to directly conclude the partial regularity stated in Theorem 3.1.

Proposition 6.1. *For any $p_0 \in G$ there exists $\delta > 0$, $R_0 > 0$ such that whenever $B_r(x_0) \subset \Omega$, $r \in (0, R_0)$, and*

$$⨍_{B_r(x_0)} |\nabla u - p_0|^2 < \delta,$$

then $u \in \mathscr{C}^{2,\alpha}(B_{r/2}(x_0))$ for some $\alpha \in (0,1)$.

The idea behind the proof is the following. Let $x_0 \in \Omega$ and $r > 0$ be such that $B_r(x_0) \subset \Omega$, and let $u_0 \in \mathbb{R}$, $p_0 \in G$. For $\varepsilon > 0$ we know that $u_\varepsilon \colon B_r(x_0) \to \mathbb{R}$ is a viscosity solution (cf. Definition 4.1) to the Euler–Lagrange equation

$$D^2\widehat{F}_\varepsilon(\nabla u_\varepsilon) : D^2 u_\varepsilon + \partial_z V(x, u_\varepsilon) = 0.$$

We will show that after rescaling

$$u_\varepsilon(x) = u_0 + p_0 \cdot (x - x_0) + r v_\varepsilon((x - x_0)/r), \tag{6.1}$$

which implies that $v_\varepsilon\colon B_1(0) \to \mathbb{R}$ is a (viscosity) solution of

$$D^2 \widehat{F}_\varepsilon(p_0 + \nabla v_\varepsilon) : D^2 v_\varepsilon + r\partial_z V(x_0 + rx, u_0 + rp_0 \cdot x + rv_\varepsilon) = 0, \tag{6.2}$$

and some further technical manipulations, we can apply the following regularity result of Savin. Recall that $\mathcal{S}^{n\times n}$ denotes the set of symmetric $n \times n$ matrices. We also set $B_1 := B_1(0) \subset \mathbb{R}^n$.

Theorem 6.2 (Savin [41, Theorem 1.3]). *Let $\mathcal{F}\colon \mathcal{S}^{n\times n} \times \mathbb{R}^n \times \mathbb{R} \times B_1 \to \mathbb{R}$, $(M, p, z, x) \mapsto \mathcal{F}(M, p, z, x)$ be a measurable map and K, $\bar\delta > 0$, $\Lambda \geq \lambda > 0$ constants satisfying*

(H1) *if $M, N \in \mathcal{S}^{n\times n}$, $N \geq 0$, $|p| \leq \bar\delta$, $|z| \leq \bar\delta$, $x \in B_1$, then*

$$\mathcal{F}(M + N, p, z, x) \geq \mathcal{F}(M, p, z, x),$$

(H2) *if $M, N \in \mathcal{S}^{n\times n}$, $N \geq 0$, $\|M\| \leq \bar\delta$, $\|N\| \leq \bar\delta$, $|p| \leq \bar\delta$, $|z| \leq \bar\delta$, $x \in B_1$, then*

$$\Lambda\|N\| \geq \mathcal{F}(M + N, p, z, x) - \mathcal{F}(M, p, z, x) \geq \lambda\|N\|,$$

(H4) *$\mathcal{F}(0, 0, 0, x) = 0$, and in the $\bar\delta$-neighborhood of $\{(0, 0, 0, x) : x \in B_1\}$ the map $\mathcal{F}$ is of class $\mathcal{C}^2$ and there holds the uniform bound $\|D^2\mathcal{F}\| \leq K$.*

Then there exists a constant $c_1 > 0$ depending only on K, $\bar\delta$, Λ, λ such that if the function $u\colon B_1 \to \mathbb{R}$ is a viscosity solution of $\mathcal{F}(D^2u, \nabla u, u, x) = 0$ with $\|u\|_{L^\infty(B_1)} \leq c_1$, then $u \in \mathcal{C}^{2,\alpha}(B_{1/2})$ and $\|u\|_{\mathcal{C}^{2,\alpha}(B_{1/2})} \leq \bar\delta$.

Remark 6.1. We will apply Theorem 6.2 to maps $\mathcal{F}$ that are of class $\mathcal{C}^2$ with respect to M, p, and z, whereas they are only Hölder continuous with respect to x. That is, instead of (H4) we have

(H4′) $\mathcal{F}(0, 0, 0, x) = 0$, and in the $\bar\delta$-neighborhood of $\{(0, 0, 0, x) : x \in B_1\}$ the derivatives $D^2_{(M,p,z)}\mathcal{F}$ exist, are continuous, and $\|D^2_{(M,p,z)}\mathcal{F}\| \leq K$. Moreover, $\|\mathcal{F}(M, p, z, \cdot)\|_{\mathcal{C}^{0,\beta}(\bar B_1)} \leq K$ for some $\beta \in (0, 1)$ and any (M, p, z) in the said $\bar\delta$-neighborhood.

However, the proof in [41, Sections 3 and 4] shows that the conclusions of Theorem 6.2 remain valid for any $\alpha < \beta$.

In the next two lemmas we show that the rescaled functions $v_\varepsilon\colon B_1 \to \mathbb{R}$ introduced in (6.1) satisfy the needed L^∞-bound for a suitable choice of the constant $u_0 \in \mathbb{R}$.

Lemma 6.3. *Let $x_0 \in \Omega$, $r > 0$ such that $B_r(x_0) \subset \Omega$, $p_0 \in G$. There holds*

$$\lim_{\varepsilon\to 0} \fint_{B_r(x_0)} |\nabla u_\varepsilon - p_0|^2 \, dx \leq \Big(1 + \frac{1}{(p_{0,1})^2}\Big) \fint_{B_r(x_0)} |\nabla u - p_0|^2 \, dx.$$

Proof. Using the associated Young measure $\boldsymbol{\nu} = ((\nu_x)_{x\in\Omega}, 0, 0)$ given in Proposition 5.4, and in particular the fact that the concentration measure λ vanishes, which allows us to approximate the noncontinuous indicator function of $B_r(x_0)$ by continuous integrands, there holds

$$
\begin{aligned}
\lim_{\varepsilon\to 0} &\fint_{B_r(x_0)} |\nabla u_\varepsilon - p_0|^2\,dx \\
&= \fint_{B_r(x_0)} \int_{\mathbb{R}^2} |v - p_0|^2\,d\nu_x(v)\,dx \\
&= \fint_{\substack{B_r(x_0)\cap\\ \{\partial_{x_1}u=0\}}} \int_{-1}^{1} |(0,v_2) - p_0|^2\,d\hat{\nu}_x(v_2)\,dx + \fint_{\substack{B_r(x_0)\cap\\ \{\partial_{x_1}u\neq 0\}}} |\nabla u - p_0|^2\,dx \\
&= \fint_{\substack{B_r(x_0)\cap\\ \{\partial_{x_1}u=0\}}} \int_{-1}^{1} |v_2 - p_{0,2}|^2 - |\partial_{x_2}u(x) - p_{0,2}|^2\,d\hat{\nu}_x(v_2)\,dx \\
&\quad + \fint_{B_r(x_0)} |\nabla u - p_0|^2\,dx.
\end{aligned}
$$

We further estimate

$$
\begin{aligned}
&\fint_{\substack{B_r(x_0)\cap\\ \{\partial_{x_1}u=0\}}} \int_{-1}^{1} |v_2 - p_{0,2}|^2 - |\partial_{x_2}u(x) - p_{0,2}|^2\,d\hat{\nu}_x(v_2)\,dx \\
&\quad = \fint_{\substack{B_r(x_0)\cap\\ \{\partial_{x_1}u=0\}}} \int_{-1}^{1} |v_2|^2 - |\partial_{x_2}u(x)|^2\,d\hat{\nu}_x(v_2)\,dx \le \fint_{\substack{B_r(x_0)\cap\\ \{\partial_{x_1}u=0\}}} dx \\
&\quad = \frac{1}{(p_{0,1})^2} \fint_{\substack{B_r(x_0)\cap\\ \{\partial_{x_1}u=0\}}} |\partial_{x_1}u(x) - p_{0,1}|^2\,dx \le \frac{1}{(p_{0,1})^2} \fint_{B_r(x_0)} |\nabla u - p_0|^2\,dx,
\end{aligned}
$$

which concludes the proof of the lemma. ∎

Lemma 6.4. *Let $p_0 \in G$, $x_0 \in \Omega$, $r > 0$ with $B_r(x_0) \subset \Omega$, and*

$$
|p_{0,1}| - 4r\|\partial_z V\|_{L^\infty(\Omega\times\mathbb{R})} > 0. \tag{6.3}
$$

There exists $\varepsilon_0 \in (0,1)$ such that for any $\varepsilon \in (0,\varepsilon_0)$ there exist $u_0 \in \mathbb{R}$ and $r_1 \in (r/2, r)$, such that the rescaled functions $v_\varepsilon\colon B_1 \to \mathbb{R}$,

$$
v_\varepsilon(x) := \frac{u_\varepsilon(x_0 + r_1 x) - u_0}{r_1} - p_0 \cdot x
$$

satisfy

$$
\|v_\varepsilon\|_{L^\infty(B_1)} \le C\left(\fint_{B_r(x_0)} |\nabla u - p_0|^2\,dx\right)^{\frac{1}{2}} + 2r\|\partial_z V\|_{L^\infty(\Omega\times\mathbb{R})} \tag{6.4}
$$

for a constant $C > 0$ depending only on $(p_{0,1})^{-1}$.

Proof. For p_0 and $B_r(x_0)$ as stated we set

$$
E := \fint_{B_r(x_0)} |\nabla u - p_0|^2\,dx.
$$

We assume $E > 0$, otherwise our main goal, Proposition 6.1, is trivial. By Lemma 6.3 there exists $\varepsilon_0 > 0$ small enough such that

$$\fint_{B_r(x_0)} |\nabla u_\varepsilon - p_0|^2 \, dx < C_0 E,$$

for $\varepsilon \in (0, \varepsilon_0)$, where $C_0 := 2 + \frac{1}{(p_{0,1})^2}$. Moreover, in view of condition (6.3) we can, after shrinking $\varepsilon_0 > 0$, also assume that for $\varepsilon \in (0, \varepsilon_0)$ the extension $\widehat{F}_\varepsilon$ coincides with the original approximation F_ε in a neighborhood of the segment

$$I := \left[p_0 - 4r\|\partial_z V\|_\infty e_1, \, p_0 + 4r\|\partial_z V\|_\infty e_1\right], \tag{6.5}$$

which is compactly contained in the good set G.

It follows that there exists $r_1 \in (r/2, r)$ such that

$$\frac{r}{2} \int_{\partial B_{r_1}(x_0)} |\nabla u_\varepsilon - p_0|^2 \, dS \le \int_{B_r(x_0)} |\nabla u_\varepsilon - p_0|^2 \, dx < C_0 E r^2 \pi,$$

hence

$$\frac{1}{2\pi r_1} \int_{\partial B_{r_1}(x_0)} |\nabla u_\varepsilon - p_0|^2 \, dS < 2C_0 E.$$

Then, defining v_ε as stated with $u_0 := \fint_{\partial B_{r_1}(x_0)} u_\varepsilon \, dS$ and r_1 as chosen before, one obtains that

$$\fint_{\partial B_1} |\nabla v_\varepsilon|^2 \, dS < 2C_0 E.$$

By Morrey's inequality for instance, we get

$$\|v_\varepsilon\|_{L^\infty(\partial B_1)} \lesssim \sqrt{E}, \tag{6.6}$$

with a proportionality constant depending only on $(p_{0,1})^{-1}$.

Now define

$$A_0 := \|\partial_z V\|_\infty ((1 + \varepsilon_0)^2 - (p_{0,2})^2),$$

and note that for any $\varepsilon \in (0, \varepsilon_0)$, since $\partial_1^2 \widehat{F}_\varepsilon(p_0) = \partial_1^2 F_\varepsilon(p_0) = \frac{1}{(1+\varepsilon)^2 - (p_{0,2})^2}$, we have

$$\partial_1^2 \widehat{F}_\varepsilon(p_0) A_0 > \|\partial_z V\|_\infty. \tag{6.7}$$

Finally, we want to show that

$$\|v_\varepsilon\|_{L^\infty(B_1)} \le \|v_\varepsilon\|_{L^\infty(\partial B_1)} + \frac{r_1 A_0}{2}. \tag{6.8}$$

We argue by contradiction: suppose that

$$\eta := \max_{B_1} \left\{ v_\varepsilon - \|v_\varepsilon\|_{L^\infty(\partial B_1)} - \frac{r_1 A_0}{2}(1 - x_1^2) \right\} > 0.$$

Then the function $\bar{\phi}\colon B_1 \to \mathbb{R}$,

$$\bar{\phi}(x) := \eta + \|v_\varepsilon\|_{L^\infty(\partial B_1)} + \frac{r_1 A_0}{2}(1 - x_1^2),$$

touches v_ε from above at some point $\bar{x} \in B_1$. Since v_ε is a viscosity solution to (6.2), there must hold

$$\begin{aligned} 0 &\le D^2\widehat{F}_\varepsilon(p_0 + \nabla\bar{\phi}(\bar{x})) : D^2\bar{\phi}(\bar{x}) + r_1\partial_z V(x_0 + r_1\bar{x}, u_0 + r_1 p_0 \cdot \bar{x} + r\bar{\phi}(\bar{x})) \\ &\le -\partial_1^2\widehat{F}_\varepsilon(p_0 - r_1A_0\bar{x}_1 e_1)r_1A_0 + r_1\|\partial_z V\|_\infty = r_1(-\partial_{p_1}^2 F_\varepsilon(p_0)A_0 + \|\partial_z V\|_\infty), \end{aligned}$$

which contradicts (6.7). Note here that in the last step we have made use of the fact that $p_0 - r_1A_0\bar{x}_1e_1$ lies on the segment I defined in (6.5).

A similar contradiction is obtained if one assumes that

$$\min_{B_1}\left\{v_\varepsilon + \|v_\varepsilon\|_{L^\infty(\partial B_1)} + \frac{r_1A_0}{2}(1 - x_1^2)\right\} < 0.$$

Therefore, combining (6.6) and (6.8) we obtain the inequality stated in (6.4). ■

The next lemma will help us to set up a family of functionals which satisfies the conditions of Theorem 6.2.

Lemma 6.5. *Let $p_0 \in G$ and $\alpha \in (0,1)$. There exist $\varepsilon_1 \in (0,1)$ and $r_2 > 0$ such that for any $u_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^2$, $r \in (0, r_2)$ with $B_r(x_0) \subset \Omega$, $\varepsilon \in [0, \varepsilon_1]$, the boundary value problem*

$$\begin{cases} D^2\widehat{F}_\varepsilon(p_0 + \nabla\phi) : \nabla^2\phi + r\partial_z V(x_0 + rx, u_0 + rp_0\cdot x + r\phi) = 0 & \text{in } B_1, \\ \phi = 0 & \text{on } \partial B_1, \end{cases} \tag{6.9}$$

has a $\mathcal{C}^{2,\alpha}$ solution $\phi_r^{\varepsilon,x_0,u_0,p_0}$ satisfying $\|\phi_r^{\varepsilon,x_0,u_0,p_0}\|_{\mathcal{C}^{2,\alpha}(\bar{B}_1)} \to 0$ as $r \to 0$ uniformly in $(\varepsilon, x_0, u_0) \in [0,\varepsilon_1] \times K$ for any $K \subset\subset \Omega \times \mathbb{R}$.

Proof. Let $p_0 \in G$ and $\alpha \in (0,1)$ be fixed. We first of all pick $\varepsilon_1 > 0$ and $\eta_0 > 0$ such that $\overline{B_{\eta_0}(p_0)} \subset G$ and $\widehat{F}_\varepsilon(p_0 + p) = F_\varepsilon(p_0 + p)$ for all $\varepsilon \in (0, \varepsilon_1]$, $|p| \le \eta_0$.

Next let

$$\mathcal{B} := \left\{\phi \in \mathcal{C}^{2,\alpha}(\bar{B}_1) : \phi_{|\partial B_1} = 0,\ \|\phi\|_{\mathcal{C}^{2,\alpha}(\bar{B}_1)} < \eta_0\right\}$$

and consider for $R > 0$ the family of maps $\mathfrak{F}_{p_0}^a\colon [0,R) \times \mathcal{B} \to \mathcal{C}^{0,\alpha}(\bar{B}_1)$,

$$\mathfrak{F}_{p_0}^a(r,\phi)(x) = D^2F_\varepsilon(p_0 + \nabla\phi(x)) : \nabla^2\phi(x) + r\partial_z V(x_0 + rx, u_0 + rp_0\cdot x + r\phi(x)),$$

where a is an abbreviation for the tuple of parameters $a := (\varepsilon, x_0, u_0)$ satisfying $\varepsilon \in [0,\varepsilon_1]$, $x_0 \in \Omega$, $\operatorname{dist}(x_0, \partial\Omega) > R$, $u_0 \in \mathbb{R}$. Observe that $\mathfrak{F}_{p_0}^a$ is well defined and that for $\varepsilon > 0$ the equation $\mathfrak{F}_{p_0}^a(r,\phi) = 0$ holds true if and only if ϕ solves the boundary value problem (6.9).

Since F_ε is smooth on the closure of $B_{\eta_0}(p_0)$ and by (V_{reg}), one can check that $\mathfrak{F}_{p_0}^a$ is continuous and Fréchet-differentiable with respect to ϕ, and that the corresponding derivative $D_\phi\mathfrak{F}_{p_0}^a\colon [0,\infty) \times \mathcal{B} \to \mathcal{L}(\{\psi \in \mathcal{C}^{2,\alpha}(\bar{B}_1) : \psi_{|\partial B_1} = 0\}; \mathcal{C}^{0,\alpha}(\bar{B}_1))$ is continuous. For later use we point out that not only is each $D_\phi \mathfrak{F}^a_{p_0}$ continuous as a function of (r, ϕ), but also that the joint function

$$[0, \varepsilon_1] \times \{\operatorname{dist}(x, \partial\Omega) > R\} \times \mathbb{R} \times [0, R) \times \mathcal{B} \ni (\varepsilon, x_0, u_0, r, \phi) \mapsto D_\phi \mathfrak{F}^{(\varepsilon, x_0, u_0)}_{p_0}(r, \phi)$$

is continuous. The same is true for $\mathfrak{F}_{p_0}$ itself.

Moreover, there clearly holds $\mathfrak{F}^a_{p_0}(0,0) = 0$ for any considered parameter triple a and $\mathcal{L}^a_{p_0} := D_\phi \mathfrak{F}^a_{p_0}(0,0)\colon \{\psi \in \mathcal{C}^{2,\alpha}(\bar{B}_1)\colon \psi_{|\partial B_1} = 0\} \to \mathcal{C}^{0,\alpha}(\bar{B}_1)$ is given by

$$\mathcal{L}^a_{p_0}[\psi](x) = D^2 F_\varepsilon(p_0) : \nabla^2 \psi(x).$$

By Schauder theory (cf. [31]) we see that $\mathcal{L}^a_{p_0}$ is an isomorphism with

$$\|(\mathcal{L}^a_{p_0})^{-1}[\theta]\|_{\mathcal{C}^{2,\alpha}(\bar{B}_1)} \le C \|\theta\|_{\mathcal{C}^{0,\alpha}(\bar{B}_1)} \tag{6.10}$$

for all $\theta \in \mathcal{C}^{0,\alpha}(\bar{B}_1)$ and a constant $C > 0$ depending only on the ellipticity constants of $D^2 F_\varepsilon(p_0)$. Therefore, since $p_0 \in G$ is a good point, this constant can be chosen independently of $\varepsilon \in [0, \varepsilon_1]$. Thus (6.10) holds with a constant $C > 0$ independent of the considered parameter triple $a = (\varepsilon, x_0, u_0)$.

Hence, by using a quantitative version of the implicit function theorem, we conclude the existence of $r_2 \in (0, R)$ depending solely on p_0, as well as for every a a continuous family $[0, r_2) \to \mathcal{B}$, $r \mapsto \phi^a_r$ with $\phi^a_0 = 0$ satisfying $\mathfrak{F}^a_{p_0}(r, \phi^a_r) = 0$, $r \in [0, r_2)$.

Now it only remains to show that the convergence $\phi^a_r \to 0$ in $\mathcal{C}^{2,\alpha}(\bar{B}_1)$ as $r \to 0$ is in fact uniform in $a = (\varepsilon, x_0, u_0)$ for $\varepsilon \in [0, \varepsilon_1]$ and (x_0, u_0) taken from a compact subset of $\Omega \times \mathbb{R}$. First of all, note that there exists $R > 0$ such that the above-defined map $\mathfrak{F}^a_{p_0}$ is well defined for any $(x_0, u_0) \in K$.

The uniform convergence then follows from the fact that, as observed earlier, the maps $\mathfrak{F}_{p_0}$, $D_\phi \mathfrak{F}_{p_0}$ are continuous as functions of $(\varepsilon, x_0, u_0, r, \phi)$, such that the implicit function theorem also provides us with the continuity of the joint map $[0, \varepsilon_1] \times \{\operatorname{dist}(x, \partial\Omega) \ge R\} \times \mathbb{R} \times [0, r_2) \to \mathcal{B}$,

$$(\varepsilon, x_0, u_0, r) \mapsto \phi_r^{(\varepsilon, x_0, u_0)}.$$

Now the stated uniform convergence is a direct consequence of the compactness of the set $[0, \varepsilon_1] \times K \times \{0\}$. ∎

Proof of Proposition 6.1. We fix $p_0 \in G$. For $\varepsilon \in (0, \varepsilon_1)$, $r \in (0, r_2)$, $x_0 \in \mathbb{R}^2$, $B_r(x_0) \subset \Omega$, and $u_0 \in \mathbb{R}$ as in Lemma 6.5, we consider the family of nonlinear maps $\mathcal{F}_r^{\varepsilon, x_0, u_0}\colon \mathcal{S}^{2\times 2} \times \mathbb{R}^2 \times \mathbb{R} \times B_1 \to \mathbb{R}$ defined by

$$\begin{aligned}\mathcal{F}_r^{\varepsilon, x_0, u_0}(M, p, z, x) &:= D^2 \hat{F}_\varepsilon(p_0 + p + \nabla \phi^a_r(x)) : (M + D^2 \phi^a_r(x)) \\ &\quad + r \partial_z V(x_0 + rx, u_0 + r p_0 \cdot x + rz + r\phi^a_r(x)).\end{aligned}$$

Here we have denoted by $\phi^a_r := \phi_r^{\varepsilon, x_0, u_0, p_0}$ the $\mathcal{C}^{2,\alpha}$ solution of (6.9) provided by Lemma 6.5.

We will now show that conditions (H1), (H2), and (H4) from Theorem 6.2, or rather condition (H4′) from Remark 6.1, are satisfied for $\mathcal{F}_r^{\varepsilon,x_0,u_0}$. Indeed, condition (H1) in fact even holds globally, while (H2) holds true with a constant $\bar{\delta} > 0$ independent of r, ε, x_0, u_0 provided (x_0, u_0) is restricted to a compact subset of $\Omega \times \mathbb{R}$ and r, ε are chosen small enough. That this is possible is a consequence of p_0 being a good point and Lemma 6.5.

Furthermore, we see that $\mathcal{F}_r^{\varepsilon,x_0,u_0}(0,0,0,x) = 0$ by (6.9) and the partial second derivative $D^2_{(M,p,z)} F_r^{\varepsilon,x_0,u_0}$ is bounded on a neighborhood of $\{(0,0,0,x) : x \in B_1\}$. The size of the neighborhood and the bound can be chosen with the same (in)-dependencies as $\bar{\delta} > 0$ above in (H2). Moreover, in view of ($\mathrm{V}_{\mathrm{reg}}$) and Lemma 6.5 we see that any $\mathcal{F}_r^{\varepsilon,x_0,u_0}$ is Hölder continuous with any exponent $\beta \in (0,1)$, where for fixed β the corresponding Hölder norm can again be assumed to be bounded uniformly in r, ε, x_0, u_0 for (x_0, u_0) from a compact set, and r, ε small enough. We therefore also have property (H4′).

We may then apply Theorem 6.2 to $\mathcal{F}_r^{\varepsilon,x_0,u_0}$ and obtain a constant $c_1 > 0$ independent of ε, r small and (x_0, u_0) from a compact subset of $\Omega \times \mathbb{R}$ having the property that any viscosity solution $v: B_1 \to \mathbb{R}$ of $\mathcal{F}_r^{\varepsilon,x_0,u_0} = 0$ with $\|v\|_{L^\infty(B_1)} \le c_1$ belongs to $\mathcal{C}^{2,\alpha}(B_{1/2})$ and $\|v\|_{\mathcal{C}^{2,\alpha}(B_{1/2})} \le \bar{\delta}$.

Let us now also fix $x_0 \in \Omega$ and consider for $\varepsilon, r > 0$ small, such that the above conclusions hold true, as well as the conclusions of Lemma 6.4, the function $w_\varepsilon: B_1 \to \mathbb{R}$,

$$w_\varepsilon(x) := v_\varepsilon(x) - \phi_{r_1}^{\varepsilon,x_0,u_0,p_0}(x) = \frac{u_\varepsilon(x_0 + r_1 x) - u_0}{r_1} - p_0 \cdot x - \phi_{r_1}^{\varepsilon,x_0,u_0,p_0}(x),$$

where $r_1 \in (r/2, r)$, $u_0 \in \mathbb{R}$ are given by Lemma 6.4.

Then w_ε satisfies $\mathcal{F}_{r_1}^{\varepsilon,x_0,u_0}(D^2 w_\varepsilon, \nabla w_\varepsilon, w_\varepsilon, x) = 0$ and by Lemma 6.4 there holds

$$\begin{aligned}\|w_\varepsilon\|_{L^\infty(B_1)} \le C \left(\fint_{B_r(x_0)} |\nabla u - p_0|^2 \, dx \right)^{\frac{1}{2}} + 2r \|\partial_z V\|_{L^\infty(\Omega \times \mathbb{R})} \\ + \|\phi_{r_1}^{\varepsilon,x_0,u_0,p_0}\|_{L^\infty(B_1)}\end{aligned}$$

for a constant $C > 0$ depending only on p_0. In view of Lemma 6.5 we therefore reach

$$\|w_\varepsilon\|_{L^\infty(B_1)} \le c_1$$

by assuming that $r > 0$ and $\fint_{B_r(x_0)} |\nabla u - p_0|^2 \, dx$ are small enough. Therefore, we may conclude that $(w_\varepsilon)_\varepsilon$ is bounded in $\mathcal{C}^{2,\alpha}(B_{1/2})$.

It then follows that $(u_\varepsilon)_\varepsilon$ is bounded in $\mathcal{C}^{2,\alpha}(B_{r/2}(x_0))$, and hence converges to the limit u in $\mathcal{C}^{2,\alpha-}(B_{r/2}(x_0))$. This finishes the proof of Proposition 6.1. ∎

7. Further properties

Here we collect some additional properties for our minimizer that are important in relation to the role of ∇u as a subsolution to the Boussinesq equation. Throughout this section we again consider $(u_\varepsilon)_\varepsilon$ and u as in Proposition 5.4.

7.1. Topology of Ω'

We begin by noting that Lemma 3.2 is a direct consequence of the one-sided maximum principle in Lemma 4.10 and the uniform convergence in Proposition 5.6.

7.2. Energy balance

The Young measure representation in Proposition 5.1 allows us to pass to the limit in the energy balance in Lemma 4.9.

Lemma 7.1. *The measure $(\partial_{p_1}\widehat{F}_\varepsilon(\nabla u_\varepsilon(x))\partial_{x_1}u_\varepsilon(x) - \widehat{F}_\varepsilon(\nabla u_\varepsilon(x)))\,dx$ converges weakly to $F(\nabla u(x))\,dx$. In particular, if ($\mathrm{V}_{\mathrm{aut}}$) holds true, then*

$$\frac{d}{dx_1}\int_{-L}^{L} F(\nabla u(x)) + V(x_2, u(x))\,dx_2 = 0 \tag{7.1}$$

in the weak sense.

Proof. Let $\psi\colon \overline{\Omega} \to \mathbb{R}$ be an arbitrary continuous, bounded function and set $G_\varepsilon(p) := \partial_{p_1}\widehat{F}_\varepsilon(p)p_1 - \widehat{F}_\varepsilon(p)$, as well as

$$I_\varepsilon := \int_\Omega G_\varepsilon(\nabla u_\varepsilon(x))\psi(x)\,dx, \quad I_\varepsilon^1 := \int_\Omega G_\varepsilon(\nabla u_\varepsilon(x))\psi(x)\eta(\nabla u_\varepsilon(x))\,dx,$$
$$I_\varepsilon^2 := I_\varepsilon - I_\varepsilon^1 = \int_\Omega G_\varepsilon(\nabla u_\varepsilon(x))\psi(x)(1-\eta(\nabla u_\varepsilon(x)))\,dx,$$

where $\eta\colon \mathbb{R}^2 \to [0,1]$ is continuous with support compactly contained in the open strip $\{p \in \mathbb{R}^2 : |p_2| < 1\}$.

By Proposition 5.4 and the convergence of $G_\varepsilon(p)$ to $\partial_{p_1}F(p)p_1 - F(p) = F(p)$, which is uniform on the support of η, it follows that

$$\lim_{\varepsilon\to 0} I_\varepsilon^1 = \int_\Omega F(\nabla u)\eta(\nabla u)\psi\,dx.$$

For I_ε^2 we use Lemma 4.1 (iii) in order to estimate

$$|I_\varepsilon^2| \le \|\psi\|_{L^\infty(\Omega)}\int_\Omega (\varepsilon + 4\widehat{F}_\varepsilon(\nabla u_\varepsilon))(1-\eta(\nabla u_\varepsilon))\,dx$$
$$\to 4\|\psi\|_{L^\infty(\Omega)}\int_\Omega F(\nabla u)(1-\eta(\nabla u))\,dx$$

as $\varepsilon \to 0$. Indeed, the convergence of the integral can be seen by splitting it up, the use of the convergence of the total kinetic energies (5.14), and the same argument as above for the convergence of I_ε^1. We conclude

$$\left| I_\varepsilon - \int_\Omega F(\nabla u)\eta(\nabla u)\psi\,dx\right| \le o(1) + C\int_\Omega F(\nabla u)(1-\eta(\nabla u))\,dx$$

as $\varepsilon \to 0$, with a constant $C > 0$ depending on ψ, but not on η.

Now taking a sequence $(\eta_j)_j$ converging pointwise to 1 on $\{p \in \mathbb{R}^2 : |p_2| \leq 1\}$ we deduce that

$$\lim_{\varepsilon \to 0} I_\varepsilon = \int_\Omega F(\nabla u)\psi \, dx.$$

Recall here that $F(p) = 0$ for $p_1 = 0$, $|p_2| = 1$ and that the measure of the set $|\{x \in \Omega : \partial_{x_1} u(x) \neq 0,\ |\partial_{x_2} u(x)| = 1\}|$ is 0 as otherwise $F(\nabla u)$ would not be integrable. Thus we have shown the stated weak convergence.

Now, if (V_{aut}) holds true, it easily follows from Lemma 4.9 that

$$\int_0^T \varphi'(x_1) \int_{-L}^L F(\nabla u(x)) + V(x_2, u(x)) \, dx_2 \, dx_1 = 0$$

for all $\varphi \in \mathcal{C}_c^1(0, T)$. ■

Note that at this point all the claims of Theorem 3.1 have been shown.

7.3. Stronger attainment of boundary data

As already discussed in Sections 2.5, 2.6, considering functions $u \in X$ implies that $\rho := \partial_{x_2} u$ and $m := -\partial_{x_1} u$ satisfy

$$\int_\Omega \rho \partial_{x_1} \varphi + m \partial_{x_2} \varphi \, dx + \int_{-L}^L \operatorname{sign}(x_2)(\varphi(0, x_2) + \varphi(T, x_2)) \, dx_2 = 0 \tag{7.2}$$

for all $\varphi \in H^1(\Omega)$. Concerning the boundary data for m and the initial and final data for ρ, we can conclude from this certain weak convergences; see for instance Lemma 7.2 below. The goal of this subsection is to improve these weak convergences to strong convergences. This is the statement of Lemmas 7.3, 7.4. Moreover, the energy balance allows us to also conclude that m attains 0 initial and final data; see Lemma 7.5. This information has not been encoded in the function space X, not even in a weaker form.

We begin with the claimed weak convergence of ρ near $\{x_1 = 0\}$. A similar statement holds true near $\{x_1 = T\}$.

Lemma 7.2. *For $a > 0$, let $v_a(x) := u(ax_1, x_2) - (|x_2| - L)$. Then there holds $\partial_{x_2} v_a \overset{*}{\rightharpoonup} 0$ in $L^\infty((0,1) \times (-L, L))$ as $a \to 0$.*

Proof. Let $\psi \in \mathcal{C}^1([0,1] \times [-L, L])$ and define for $a \in (0, T)$ the function $\varphi_a : \overline{\Omega} \to \mathbb{R}$,

$$\varphi_a(x) = -\int_{x_1/a}^1 \psi(x_1', x_2) \, dx_1'$$

for $x_1 \in [0, a)$ and $\varphi_a(x) = 0$ for $x_1 \geq a$. Then using (7.2) one computes

$$\begin{aligned} \int_0^1 \int_{-L}^L \partial_{x_2} v_a(x) \psi(x) \, dx &= \int_\Omega (\partial_{x_2} u(x) - \operatorname{sign}(x_2)) \partial_{x_1} \varphi_a(x) \, dx \\ &= \int_\Omega \partial_{x_1} u(x) \partial_{x_2} \varphi_a(x) \, dx \to 0 \quad \text{as } a \to 0. \end{aligned}$$

The general case $\psi \in L^1((0,1) \times (-L, L))$ follows by observing that $(\partial_{x_2} v_a)_a$ is bounded in $L^\infty((0,1) \times (-L, L))$ (cf. Proposition 5.4) and approximation. ■

This weak convergence can easily be improved.

Lemma 7.3. *For v_a as in Lemma 7.2 there holds $\partial_{x_2} v_a \to 0$ in $L^1((0,1)\times(-L,L))$ as $a \to 0$.*

Proof. Recalling that $|\partial_{x_2} u| \le 1$ a.e. one may write

$$\begin{aligned}\|\partial_{x_2} v_a\|_{L^1((0,1)\times(-L,L))} &= \int_0^1 \int_{-L}^{L} |\partial_{x_2} u(ax_1, x_2) - \operatorname{sign}(x_2)|\,dx \\ &= \int_0^1 \int_0^L (\mathbb{1}_{\{x_2<0\}}(x) - \mathbb{1}_{\{x_2>0\}}(x))\partial_{x_2} v_a(x)\,dx \to 0,\end{aligned}$$

by the previous weak convergence and the fact that the indicators are in L^1. ■

For a similar strong convergence of $\partial_{x_1} u$ near $\{x_2 = \pm L\}$ we argue directly.

Lemma 7.4. *Let ($\mathrm{V_{aut}}$) be satisfied, such that $\partial_{x_1} u \ge 0$. There holds the convergence $\|\partial_{x_1} u(\cdot, -L + b\cdot)\|_{L^1((0,T)\times(0,1))} \to 0$ as $b \to 0$.*

Proof. Using $\partial_{x_1} u \ge 0$, $u(\cdot, x_2) \in H^1(0,T)$ for a.e. $x_2 \in (-L, L)$, as well as the continuity of u (cf. Proposition 5.6), we observe that for any $b > 0$ there holds

$$\begin{aligned}\|\partial_{x_1} u(\cdot, -L + b\cdot)\|_{L^1((0,T)\times(0,1))} &= \int_0^T \int_0^1 \partial_{x_1} u(x_1, -L + bx_2)\,dx \\ &= \frac{1}{b}\int_0^T \int_{-L}^{-L+b} \partial_{x_1} u(x)\,dx \\ &= \frac{1}{b}\int_{-L}^{-L+b} u(T, x_2) - u(0, x_2)\,dx_2 \\ &= \frac{1}{b}\int_{-L}^{-L+b} 2(L + x_2)\,dx_2 \le \frac{1}{b} 2\int_{-L}^{-L+b} b\,dx_2 = 2b,\end{aligned}$$

hence the claim follows. ■

Utilizing the energy balance from Lemma 7.1 we in addition obtain a corresponding strong convergence of $\partial_{x_1} u$ near $\{x_1 = 0\}$ and $\{x_1 = T\}$.

Lemma 7.5. *Let ($\mathrm{V_{aut}}$) be satisfied. For the minimizer u there holds $\partial_{x_1} u(a\cdot, \cdot) \to 0$ in $L^1((0,1)\times(-L,L))$ as $a \to 0^+$.*

Proof. We may write

$$\begin{aligned}&\int_0^1 \int_{-L}^{L} |\partial_{x_1} u(ax_1, x_2)|\,dx \\ &\quad= \frac{1}{a}\int_0^a \int_{-L}^{L} \partial_{x_1} u(x)\,dx = \frac{1}{a}\int_0^a \int_{-L}^{L} \sqrt{2F(\nabla u(x))}\sqrt{1 - \partial_{x_2} u(x)^2}\,dx \\ &\quad\le \left(\frac{1}{a}\int_0^a \int_{-L}^{L} 2F(\nabla u(x))\,dx\right)^{1/2} \left(\frac{1}{a}\int_0^a \int_{-L}^{L} 1 - \partial_{x_2} u(x)^2\,dx\right)^{1/2}.\end{aligned}$$

Denoting by $E_0 \in \mathbb{R}$ the constant total energy value given by the balance (7.1), the first factor can be estimated against $(2E_0 + 2\|V\|_{L^\infty(\Omega\times\mathbb{R})})^{1/2}$, while for the second factor we have

$$\frac{1}{a}\int_0^a \int_{-L}^{L} 1 - \partial_{x_2}u(x)^2\,dx = \frac{1}{a}\int_0^a \int_{-L}^{L} (\mathrm{sign}(x_2) - \partial_{x_2}u(x))(\mathrm{sign}(x_2) + \partial_{x_2}u(x))\,dx$$
$$\leq 2\|\partial_{x_2}v_a\|_{L^1((0,1)\times(-L,L))} \to 0$$

as $a \to 0$ by Lemma 7.3. ■

Comparing Lemmas 7.3, 7.5 with [26, Theorem 5.7] one can say that the functions

$$x_1 \mapsto \int_{-L}^{L} |\partial_{x_2}u(x) - \mathrm{sign}(x_2)|\,dx_2, \quad x_1 \mapsto \int_{-L}^{L} |\partial_{x_1}u(x)|\,dx_2,$$

which are a priori only in $L^\infty(0,T)$, resp. $L^2(0,T)$, have a trace at $x_1 = 0$ and $x_1 = T$.

7.4. Admissibility

We will now discuss the actual energy balance of the Boussinesq subsolution induced by the minimizer u. Recall that the one-dimensional subsolution $(\rho, m) := (\partial_{x_2}u, -\partial_{x_1}u)$ is called weakly admissible provided (2.11), i.e.

$$E_{\mathrm{tot}}(x_1) := \int_{-L}^{L} F(\nabla u(x)) - gAu(x)\,dx_2 < \int_{-L}^{L} \rho_0(x_2)gAx_2\,dx_2 \tag{7.3}$$

holds true for a.e. $x_1 \in (0,T)$.

As indicated in Section 2.7 we in fact will have a monotone decay of the total energy on $(0,T)$ provided that, in addition to ($\mathrm{V}_{\mathrm{reg}}$), V also satisfies

$$\begin{aligned} &V \text{ has the form } V(x,z) = -gAz + f(x_2,z) \\ &\text{with } \partial_z f(x_2,z) > 0 \text{ whenever } |z| < L - |x_2|. \end{aligned} \tag{$\mathrm{V}_{\mathrm{dis}}$}$$

Note that ($\mathrm{V}_{\mathrm{dis}}$) contains ($\mathrm{V}_{\mathrm{aut}}$).

As a direct consequence of Lemma 7.1 and the fact that our minimizer u satisfies $|x_2| - L \leq u(x) \leq L - |x_2|$ for all $x \in \overline{\Omega}$ (cf. Corollary 5.5 and Proposition 5.6), we indeed deduce the energy balance (2.21) on the open interval $(0,T)$.

Corollary 7.6. *If V satisfies ($\mathrm{V}_{\mathrm{reg}}$) and ($\mathrm{V}_{\mathrm{dis}}$), then the sum of kinetic and potential energy (given only by the gravity potential) satisfies*

$$\frac{d}{dx_1}\int_{-L}^{L} F(\nabla u) - gAu\,dx_2 = -\int_{-L}^{L} \partial_z f(x_2,u)\partial_{x_1}u\,dx_2 \leq 0$$

weakly on $(0,T)$.

Thus we have strict dissipation of the total energy on any time interval $I \subset (0,T)$ on which $((L-|x_2|)^2 - u^2)\partial_{x_1} u$ is not essentially vanishing, i.e. u has to be different from the initial and final configurations and the momentum $-\partial_{x_1} u$ has to be strictly negative.

Let us now turn to the behavior of the energy as $x_1 \to 0$ or $x_1 \to T$. Regarding the potential energy Proposition 5.6 implies that

$$\lim_{x_1\to 0} E_{\mathrm{pot}}(x_1) := \lim_{x_1\to 0}\int_{-L}^{L} -gAu(x)\,dx_2 = \int_{-L}^{L} -gA(|x_2|-L)\,dx_2$$
$$= \int_{-L}^{L} \rho_0(x_2) gAx_2\,dx_2,$$

and similarly, including the dissipated energy,

$$\lim_{x_1\to 0}\int_{-L}^{L} V(x_2,u(x))\,dx_2 = \int_{-L}^{L} V(x_2,|x_2|-L)\,dx_2.$$

The corresponding limits also exist at $x_1 = T$.

Due to the continuity of the potential energy $E_{\mathrm{pot}}(t)$ at $t=0$ one sees that weak admissibility (7.3) requires

$$\operatorname*{ess\,lim}_{x_1\to 0} E_{\mathrm{kin}}(x_1) := \operatorname*{ess\,lim}_{x_1\to 0}\int_{-L}^{L} F(\nabla u(x))\,dx_2 = 0. \tag{7.4}$$

Although the initial momentum vanishes in the sense of Lemma 7.5 we can a priori not conclude (7.4). However, we will show that for suitable V the possible initial jump in kinetic energy becomes arbitrarily small when the variational problem is considered over a longer and longer time interval.

For that we no longer consider the final time $T>0$ as a fixed constant and indicate the T-dependency in our variational problem by writing Ω_T, $\mathcal{A}_T$, X_T instead of Ω, $\mathcal{A}$, X.

Moreover, as in Section 2.8 we define

$$s_V := \sup\left\{\textstyle\int_{-L}^{L} V(x_2,\varphi)\,dx_2 : \varphi \in \mathcal{C}^0([-L,L]),\ |\varphi(x_2)| \le L-|x_2|\right\}$$

and assume in addition to ($\mathrm{V}_{\mathrm{reg}}$), ($\mathrm{V}_{\mathrm{dis}}$),

$$s_V = 0 \quad \text{and} \quad s_V = \int_{-L}^{L} V(x_2,\varphi)\,dx_2 \ \text{ if and only if } \varphi = \pm(L-|x_2|). \tag{$\mathrm{V}_{\mathrm{sup}}$}$$

Note that assuming s_V to be 0 is not a restriction as one can always shift V by a fixed constant without changing the variational problem.

Lemma 7.7. *Assume* ($\mathrm{V}_{\mathrm{reg}}$), ($\mathrm{V}_{\mathrm{dis}}$), ($\mathrm{V}_{\mathrm{sup}}$). *Let* u_T, $T>0$ *be a minimizer of* $\mathcal{A}_T$ *over* X_T *given by Proposition* 5.4. *For* $T \ge 1$ *there holds*

$$\operatorname*{ess\,lim}_{x_1\to T}\int_{-L}^{L} F(\nabla u_T(x))\,dx_2 = \operatorname*{ess\,lim}_{x_1\to 0}\int_{-L}^{L} F(\nabla u_T(x))\,dx_2 \le \frac{\mathcal{A}_1(u_1)}{T}.$$

Proof. In view of Lemma 7.1 there exists a constant $c_T \in \mathbb{R}$ such that

$$\int_{-L}^{L} F(\nabla u_T(x)) + V(x_2, u_T(x))\, dx_2 = c_T$$

for a.e. $x_1 \in (0, T)$. By (V_{sup}) we therefore have

$$\operatorname*{ess\,lim}_{x_1 \to 0} \int_{-L}^{L} F(\nabla u_T(x))\, dx_2 = c_T - \lim_{x_1 \to 0} \int_{-L}^{L} V(x_2, u_T(x))\, dx_2 = c_T.$$

Similarly, we also conclude that the essential limit of the kinetic energy as $x_1 \to T$ is given by c_T. This shows the stated equality between the two limits. Note also that $c_T \geq 0$.

Next let $T' > T$ and define $v \in X_{T'}$ by setting

$$v(x) = \begin{cases} u_T(x), & x_1 \in (0, T), \\ L - |x_2|, & x_1 \in (T, T'). \end{cases}$$

Due to (V_{sup}) and $F(0, \pm 1) = 0$ one deduces

$$\mathcal{A}_{T'}(u_{T'}) \leq \mathcal{A}_{T'}(v) = \mathcal{A}_T(u_T).$$

Thus (V_{sup}) and Corollary 5.5 imply

$$0 \leq Tc_T \leq Tc_T - 2 \int_{\Omega_T} V(x_2, u_T(x))\, dx = \mathcal{A}_T(u_T) \leq \mathcal{A}_1(u_1)$$

and the statement follows. ■

8. Summary and further questions

Let us first of all repeat the full list of used requirements regarding the nonlinear potential $V(x, z)$ and formulate an extended version of Theorem 3.1. There is

$$\partial_z^k V : \overline{\Omega} \times \mathbb{R} \to \mathbb{R},\ k = 0, 1, 2, 3 \text{ exist, are Lipschitz and bounded,} \tag{V_{reg}}$$

$$V(x, z) = V(x_2, z), \quad x \in \overline{\Omega},\ z \in \mathbb{R}, \tag{V_{aut}}$$

$$\partial_z^2 V(x, z) \geq 0, \quad x \in \overline{\Omega},\ z \in [-L - 1, L + 1], \tag{V_{con}}$$

$$V(x, z) = -gAz + f(x_2, z) \text{ with } \partial_z f(x_2, z) > 0 \text{ whenever } |z| < L - |x_2|, \tag{V_{dis}}$$

$$s_V = 0 \quad \text{and} \quad s_V = \int_{-L}^{L} V(x_2, \varphi)\, dx_2 \text{ if and only if } \varphi = \pm(L - |x_2|), \tag{V_{sup}}$$

where

$$s_V := \sup\left\{\textstyle\int_{-L}^{L} V(x_2, \varphi)\, dx_2 : \varphi \in \mathcal{C}^0([-L, L]),\ |\varphi(x_2)| \leq L - |x_2|\right\}.$$

We have seen (cf. Section 2.9) that a suitable extension of

$$V(x,z) = -gAz + \frac{3gA}{4L}(z - (|x_2| - L))^2 + \text{const.}$$

satisfies all of the stated conditions. Summarizing the statements of Sections 4–7 there holds the following extension of Theorem 3.1:

Theorem 8.1. *Suppose that V satisfies all five conditions (V_{reg})–(V_{sup}). Then problem (3.2) with $\mathcal{A}$ defined in (3.1) has a solution u with the following properties:*

- (a) *u is continuous on $\overline{\Omega}$ with $|u(x)| \leq L - |x_2|$ and $\partial_{x_1} u \geq 0$, $|\partial_{x_2} u| \leq 1$ a.e.,*
- (b) *there exists $\Omega' \subset \Omega$ open, nonempty, and such that every connected component of Ω' is simply connected, on which u is of class $\mathcal{C}^2$ with $\partial_{x_1} u > 0$, $|\partial_{x_2} u| < 1$, while $\partial_{x_1} u(x) = 0$ for a.e. $x \notin \Omega'$,*
- (c) *$\partial_{x_1} u(\cdot, \pm L) = \partial_{x_1} u(0,\cdot) = \partial_{x_1} u(T,\cdot) = 0$, $\partial_{x_2} u(0,x_2) = -\partial_{x_2} u(T,x_2) = \operatorname{sign}(x_2)$ in the sense specified in Lemmas 7.3, 7.4, 7.5,*
- (d) *on $(0,T)$ the balance*

$$\frac{d}{dx_1}\int_{-L}^{L} F(\nabla u) - gAu \, dx_2 = -\int_{-L}^{L} \partial_z f(x_2,u)\partial_{x_1} u \, dx_2 \leq 0$$

  holds in a weak sense,
- (e) *while at $x_1 = 0$, $x_1 = T$ there holds*

$$\operatorname*{ess\,lim}_{x_1 \to T} \int_{-L}^{L} F(\nabla u(x)) \, dx_2 = \operatorname*{ess\,lim}_{x_1 \to 0} \int_{-L}^{L} F(\nabla u(x)) \, dx_2 \leq \frac{c_1}{T}$$

  for all $T \geq 1$ and a constant $c_1 > 0$ (specified in Lemma 7.7).

8.1. Use of ∇u as a Boussinesq subsolution

Let u be the minimizer from Theorem 8.1 with partial regularity set Ω' and set ρ, m_n: $(0,T) \times (-L,L) \to \mathbb{R}$,

$$\rho(t,x_n) = \partial_{x_2} u(t,x_n), \quad m_n(t,x_n) = -\partial_{x_1} u(t,x_n).$$

We indeed see that ρ, m_n, and $\mathcal{U}' := \Omega'$ satisfy Lemma 2.2 (i)–(iii), (v), while (iv) is relaxed to $m_n = 0$, $|\rho| \leq 1$ a.e. outside Ω'. In consequence, ∇u induces a subsolution with mixed resting regions (cf. Remark 2.6) and therefore via Theorem 2.1 and Remark 2.3 infinitely many solutions (ρ_{sol}, v_{sol}) to the Boussinesq system (1.1), (1.2), (1.3) that are turbulently mixing on $\mathcal{U}'$, and of which the above density ρ, momentum $m = m_n e_n$, and velocity $v \equiv 0$ can be seen as horizontally averaged quantities.

At this point we have proven Theorem 1.1. Let us however state some consequences of Theorem 8.1 for the induced subsolution and associated solutions. We conclude that

- the average momentum m is directed downwards ($m_n \leq 0$),

- outside the mixing zone $\mathcal{U}'$, the fluid is at rest ($v_{\text{sol}} = 0$), but our investigation does not allow us to conclude that the density is in one of the two initial phases ($|\rho_{\text{sol}}| < 1$ not excluded),
- the resting regions, with or without $\rho_{\text{sol}} \in \{\pm 1\}$, cannot be surrounded by the mixing zone $\mathcal{U}'$,
- besides the initial and boundary conditions for (1.1), m also vanishes in a certain trace sense as $t \to 0$, $t \to T$, and ρ approaches the stable interface configuration $-\rho_0(x)$ as $t \to T$,
- the total energy of the subsolution

$$E_{\text{sub}}(t) := \int_{-L}^{L} \frac{m_n(t, x_n)^2}{2(1 - \rho(t, x_n)^2)} + \rho(t, x_n) g A x_n \, dx_n$$

  might jump upwards from $\int_{-L}^{L} \rho_0(x_n) g A x_n \, dx_n$ at $t = 0$, and then monotonically decays on $(0, T]$ to $-\int_{-L}^{L} \rho_0(x_n) g A x_n \, dx_n$ with a reversed jump at $t = T$,
- the heights of the initial and final energy jumps vanish as the considered time interval $(0, T)$ becomes unbounded.

We recall that induced solutions can be found with total energy $E_{\text{tot}}(t)$ in an arbitrary $\delta(t)$-neighborhood of $E_{\text{sub}}(t)$; cf. Theorem 2.1.

8.2. Open questions

We discuss here some further questions regarding the variational problem (3.2), properties of the induced subsolutions, and the modeling in general.

Starting with the list of the previous subsection it would be of interest to see whether, under suitable conditions on V, the possibility of mixed resting regions can be excluded, i.e. whether in Theorem 8.1 (b) one could have $\partial_{x_1} u = 0$ *and* $|\partial_{x_2} u| = 1$ a.e. outside Ω'. An analogous property, for instance, holds true in the setting of De Silva and Savin [21]; cf. Section 3 and Remark 5.1.

Other questions for problem (3.2) address uniqueness of minimizers (this property holds true in [21]), global regularity, for instance comparable to the result of Colombo and Figalli [14], and any further information regarding the partial regularity set Ω' which corresponds to the turbulent mixing zone of the induced solutions. Of particular interest in applications is the growth of this zone in time.

On a larger scale of questions, we recall that our investigation was motivated by the search for global-in-time selection criteria for subsolutions of the Euler equations. Here, we first of all point out that the derivation of the nondissipative action functional $\mathcal{A}_0(u)$ in Sections 2.2–2.5 relies on almost no ansatzes besides the imposition of the least action principle itself. The only a priori unjustified choice made is that the kinetic energy density of the solutions $(\rho_{\text{sol}}, v_{\text{sol}})$ in the turbulent zone $\mathcal{U}$ satisfies $|v_{\text{sol}}|^2 \in \mathcal{C}^0(\mathcal{U}) + \rho_{\text{sol}} \mathcal{C}^0(\mathcal{U})$; cf. (2.5). This choice was made in [29] and is, at least in the here-considered one-dimensional initial configuration, a posteriori backed up by the fact that the functional $\mathcal{A}_0$ can also be derived from a different point of view, avoiding the notion of subsolutions at all; see Appendix A.

However, after the derivation of $\mathcal{A}_0$, in Section 2.7 we introduced the nonlinear potential $V(x_2, z) = -gAu + f(x_2, z)$ that allowed us to have energy dissipation (up to the initial jump controlled by T^{-1}) while staying within the variational framework of the least action principle. Note that, in terms of the subsolution components ρ and m_n, the associated Euler–Lagrange equations (formal on all of $(0,T)\times(-L,L)$, rigorous on Ω') are given by

$$\partial_t\Big(\frac{m_n}{1-\rho^2}\Big) - \partial_{x_n}\Big(\frac{m_n^2\rho}{(1-\rho^2)^2}\Big) - \Psi[\rho, m_n] = -gA,$$

where Ψ is the nonlocal operator

$$\Psi[\rho, m_n](t, x_n) = \partial_z f\Big(x_n, \int_{-L}^{x_n} \rho(t,s)\, ds\Big).$$

Besides the here-stated properties ($\mathrm{V}_{\mathrm{reg}}$)–($\mathrm{V}_{\mathrm{sup}}$), a further investigation and justification concerning suitable choices of f, or more generally, of a different type of relation $\Psi[\rho, m_n]$ consistent with energy dissipation, remains open.

A. Relation to Brenier's generalized least action principle

We quickly recall the least action principle, Brenier's generalization of it, and thereafter focus on a special one-dimensional problem leading to a functional formally equivalent to our $\mathcal{A}_0$ derived in Section 2.

A.1. The least action principle

Let $\mathcal{D} \subset \mathbb{R}^n$ be a compact domain, $T > 0$, $U\colon (0,T)\times\mathcal{D} \to \mathbb{R}$ be a given potential, and $\rho_0\colon \mathcal{D} \to (0,\infty)$ an initial mass distribution. It is well known, originating in the work of Arnold [1], that the Euler equations

$$\begin{aligned} \partial_t(\rho v) + \operatorname{div}(\rho v \otimes v) + \nabla p &= -\rho \nabla U, \\ \operatorname{div} v &= 0, \\ \partial_t \rho + \operatorname{div}(\rho v) &= 0, \end{aligned} \tag{A.1}$$

can formally be derived by minimizing the action functional

$$\mathcal{A}(g) = \int_0^T \int_{\mathcal{D}} \rho_0(x)\Big(\frac{1}{2}|\partial_t g(t,x)|^2 - U(t, g(t,x))\Big)\, dx\, dt$$

over trajectories $t \mapsto g(t,\cdot)$ in the manifold of volume-preserving diffeomorphisms $\mathcal{D} \to \mathcal{D}$ connecting a given initial and end state, say $g(0,\cdot) = \mathrm{id}$, $g(T,\cdot) = h$. Assuming the existence of a regular minimizer g and an associated Lagrange multiplier $p\colon (0,T)\times\mathcal{D} \to$

$\mathbb{R}$, one derives that the tuple (ρ, v, p), where v is the velocity field inducing the flow g, i.e. $\partial_t g(t,x) = v(t, g(t,x))$, and ρ is the corresponding transported density distribution, i.e. $\rho(t, g(t,x)) = \rho_0(x)$, is a solution of (A.1) with initial mass distribution ρ_0. For more detail we refer to [4] and the references therein.

Rigorous existence results of minimizers for suitable target diffeomorphisms h not too far away from the identity are due to Ebin and Marsden [22]. However, for a general h it also has been shown by Shnirel'man [42] that there does not need to be a solution in the classical sense described above.

In order to overcome this, Brenier [2] introduced the aforementioned generalization of the least action principle, which allowed him to conclude the existence of a solution given the existence of at least one competitor with finite action.

A.2. Relaxation via generalized flows

Let us recall that Brenier's generalized action functional associated with (A.1) is defined as

$$\mathcal{A}(\mu) := \int_{\Omega(\mathcal{D})} \rho_0(\omega(0)) \left(\int_0^T \frac{1}{2} |\omega'(t)|^2 - U(t, \omega(t))\, dt \right) \mu(d\omega),$$

where $\Omega(\mathcal{D}) := \{[0,T] \ni t \mapsto \omega(t) \in \mathcal{D}\} = \mathcal{D}^{[0,T]}$ is equipped with the product topology, hence compact, and μ is a generalized flow, namely a regular Borel probability measure on $\Omega(\mathcal{D})$ satisfying the incompressibility constraint

$$\int_{\Omega(\mathcal{D})} f(\omega(t))\, \mu(d\omega) = \fint_{\mathcal{D}} f(x)\, dx, \quad \text{for all } f \in \mathcal{C}(\mathcal{D}), t \in [0,T], \tag{A.2}$$

as well as the initial and final data constraint

$$\int_{\Omega(\mathcal{D})} f(\omega(0), \omega(T))\, \mu(d\omega) = \fint_{\mathcal{D}} f(x, h(x))\, dx, \quad \text{for all } f \in \mathcal{C}(\mathcal{D}^2). \tag{A.3}$$

Here, h again denotes the target configuration, which in this setting is only required to be a measure-preserving map $(\mathcal{D}, dx) \to (\mathcal{D}, dx)$, and not necessarily a diffeomorphism.

It was shown in [2] that if $\inf_\mu \mathcal{A}(\mu) < +\infty$, then there exists a minimizer, and furthermore, if system (A.1) has a solution (enjoying certain properties) then the minimizer corresponds to the flow associated with the fluid velocity of the solution; see [2, Theorems 3.2, 5.1] for the precise statements.

Through slight modifications one sees that the same two properties remain valid for the generalized action functional corresponding to the here-investigated Euler system in the Boussinesq approximation (1.1), which reads

$$\mathcal{A}(\mu) := \int_{\Omega(\mathcal{D})} \int_0^T \left(\frac{1}{2} |\omega'(t)|^2 - \rho(0, \omega(0))\, gA\omega_n(t) \right) dt\, \mu(d\omega), \tag{A.4}$$

with $\rho(0, \cdot) = \operatorname{sign}$ being the normalized initial density distribution.

A.3. One-dimensional two-phase flows

For the particular Rayleigh–Taylor situation we consider $\mathcal{D} = [0, 1]^{n-1} \times [-L, L]$, $\rho(0, x) = \operatorname{sign}(x_n)$, and the target transformation $h \colon \mathcal{D} \to \mathcal{D}$, $h(x', x_n) = (x', h_n(x_n))$, where

$$h_n(x_n) = \begin{cases} x_n - L, & x_n \geq 0, \\ x_n + L, & x_n < 0. \end{cases}$$

That is,, h swaps the upper half $\mathcal{D}_+$ of the container with the lower half $\mathcal{D}_-$. Note that in fact h even prescribes a particle-by-particle exchange of the two halves.

This situation (without the first $n-1$ dimensions and without the potential) appears in Brenier's revisitation of the least action principle [4] as one of the examples for generalized incompressible flows; see [4, Section 4.3].

Also here, i.e. with potential term, it follows formally (ignoring the conditions on p and T) from [2, Theorem 5.1] that the minimizer of (A.4) is given by a one-dimensional two-phase flow provided the associated vector fields satisfy the corresponding Euler–Lagrange equation.

More precisely, a one-dimensional two-phase flow is a generalized flow of the type

$$\mu(d\omega) = \big(\mu_+(0, x)\delta_{G_+(\cdot,x)}(\omega) + \mu_-(0, x)\delta_{G_-(\cdot,x)}(\omega)\big)\, dx,$$

where

- the $\delta_{G_\pm(\cdot,x)}$ denote Dirac measures on $\Omega(\mathcal{D})$,
- $G_\pm(t, x) = (x', g_\pm(t, x_n))$ denote actual flows of two one-dimensional vector fields $V_\pm(t, x) = (0, v_\pm(t, x_n))$, i.e.

$$\partial_t g_\pm(t, x_n) = v_\pm(t, g_\pm(t, x_n)),$$
$$g_\pm(0, x_n) = x_n, \quad g_\pm(T, x_n) = x_n \mp L,$$

  and the maps $g_\pm(t, \cdot)$ are understood as orientation-preserving diffeomorphisms $\mathbb{R} \to \mathbb{R}$ with the property that

$$g_+(t, [0, L]) \subset [-L, L], \quad g_-(t, [-L, 0]) \subset [-L, L], \quad t \in [0, T], \tag{A.5}$$

- the functions $\mu_\pm \colon [0, T] \times \mathcal{D} \to \mathbb{R}$ indicate the two phases initially given by

$$\mu_\pm(0, x) = |\mathcal{D}|^{-1} \mathbb{1}_{\mathcal{D}_\pm}(x) = (2L)^{-1} \mathbb{1}_{\mathcal{D}_\pm}(x)$$

  and obeying

$$\mu_+ + \mu_- = (2L)^{-1}, \quad \partial_t \mu_\pm + \partial_{x_n}(\mu_\pm v_\pm) = 0. \tag{A.6}$$

  Note that the continuity equations imply

$$\mu_\pm(t, g_\pm(t, x_n))\partial_{x_n} g_\pm(t, x_n) = \mu_\pm(0, x_n). \tag{A.7}$$

One can then check using (A.5), (A.6), (A.7) that such a two-phase flow indeed satisfies the incompressibility constraint (A.2). Condition (A.3) is directly stated. Moreover, it follows from (A.5), (A.6) that the average of the velocities satisfies

$$\mu_- v_- + \mu_+ v_+ = 0 \quad \text{on } [0,T] \times \mathcal{D}. \tag{A.8}$$

As indicated above it follows formally from [2, Theorem 5.1] that such a two-phase flow minimizes (A.4) provided $v_\pm$ satisfy

$$\begin{aligned} \partial_t v_+ + \partial_{x_n}\Big(\frac{1}{2}v_+^2 + p\Big) &= -gA \quad \text{for } x_n \in g_+(t,[0,L]), \\ \partial_t v_- + \partial_{x_n}\Big(\frac{1}{2}v_-^2 + p\Big) &= gA \qquad \text{for } x_n \in g_-(t,[-L,0]), \end{aligned} \tag{A.9}$$

with a pressure function $p\colon [0,T] \times [-L,L] \to \mathbb{R}$, $(t,x_n) \mapsto p(t,x_n)$ independent of the sign $\pm$.

The generalized action of a two-phase flow with initial density $\rho(0,x) = \operatorname{sign}(x_n)$ transformed to Eulerian coordinates using (A.6), (A.7) reads

$$\mathcal{A}(\mu) = \int_0^T \int_{-L}^L \frac{1}{2}(\mu_+ v_+^2 + \mu_- v_-^2) - gA(\mu_+ - \mu_-)x_n \, dx_n \, dt. \tag{A.10}$$

As a side remark we mention that as in [4] the action (A.10) for a two-phase flow can be written solely in terms of the flow of one of the phases, i.e.

$$\begin{aligned} \mathcal{A}(\mu) = \frac{1}{2}\int_0^T \int_0^L &\partial_t g_+(t,x_n)^2\Big(1 + \frac{2L}{\partial_{x_n} g_+(t,x_n) - 2L}\Big)\, dx_n \, dt \\ &- 2gA \int_0^T \int_0^L g_+(t,x_n)\, dx_n \, dt. \end{aligned} \tag{A.11}$$

The computations for the kinetic energy are not original to us: this is precisely Brenier's example in [4, Section 4.3]. The stated form of the potential energy easily follows from the incompressibility condition (A.2) applied to the odd map $f(x) = x_n$. Also, here one can check that condition (A.9) is precisely the (formal) Euler–Lagrange equation of (A.11).

A.4. Comparison

In order to see how the two-phase action (A.10) corresponds to the functional (2.16) derived for subsolutions we define $\rho, m\colon [0,T] \times [-L,L] \to \mathbb{R}$ by

$$\begin{aligned} \rho &:= 4L\mu_+ - 1 = 1 - 4L\mu_-, \\ m &:= 4L\mu_+ v_+ = -4L\mu_- v_-, \end{aligned}$$

and observe that (A.8) implies

$$\begin{aligned} \mu_- v_-^2 + \mu_+ v_+^2 &= -\frac{v_- v_+}{2L} = \frac{m^2}{2L(1-\rho^2)}, \\ \mu_+ - \mu_- &= 2\mu_+ - \frac{1}{2L} = \frac{\rho}{2L}. \end{aligned}$$

Thus (A.10) becomes

$$\mathcal{A}(\mu) = \frac{1}{2L}\int_0^T \int_{-L}^L \frac{m^2}{2(1-\rho^2)} - \rho g A x_n \, dx_n \, dt.$$

Finally, observe that the tuple (ρ, m) satisfies

$$\partial_t \rho + \partial_{x_n} m = 4L(\partial_t \mu_+ + \partial_{x_n}(\mu_+ v_+)) = 0$$

by (A.6). Thus we arrive at (2.16) and Lemma 2.2 (i).

Regarding Lemma 2.2 (ii) and (2.19) we point out that the conditions for initial and final data for ρ are built into the two-phase-flow framework by specifying $\mu_\pm(0,\cdot)$ and the target diffeomorphism h. As mentioned earlier, the specification of h is even stronger than requiring $\rho(T,\cdot) = -\rho_0$ via (2.19) as it corresponds to a particle-by-particle exchange of the two fluids.

Next we will convince ourselves that $m(t,\pm L) = 0$ holds true as well. Indeed, a generalized flow $\mu(d\omega)$ has to be a measure on the path space $\Omega(\mathcal{D})$. In terms of a two-phase flow this is ensured by (A.5). Now, if $g_+(t,x_n) = L$ for some $x_n \in [0,L]$, then $\partial_t g_+(t,x_n) \leq 0$, and thus $v_+(t,L) \leq 0$. On the other hand, if $L \notin g_+(t,[0,L])$, then $\mu_+(t,L) = 0$. We therefore conclude $m(t,L) = 4L\mu_+(t,L)v_+(t,L) \leq 0$, and similarly $m(t,-L) \geq 0$. Now, by means of (A.7), we compute

$$\int_{\mathcal{D}} \rho(t,x)\, dx = 4L \int_{\mathcal{D}} \mu_+(t,x)\, dx - 2L = 0,$$

and therefore using (A.6) it follows that

$$0 = \frac{d}{dt}\int_{\mathcal{D}} \rho(t,x)\, dx = -(m(t,L) - m(t,-L)) \geq 0.$$

Thus $m(t,\pm L) = 0$.

Following from here the same reformulation as in Section 2.5, we therefore have shown that the variational problem (1.4) considered in this article can also be derived from Brenier's generalization of the least action principle, instead of subsolutions. The relations are summarized in Figure 1.

B. Regarding the convex integration

In this section we prove Theorem 2.1. Let z_{sub} be a subsolution with respect to e_0, e_1 and with mixing zone $\mathcal{U}$ and $\delta\colon [0,T] \to \mathbb{R}$ be continuous with $\delta(0) = 0$, $\delta(t) > 0$, for $t > 0$. We also define the set of functions

$$\mathcal{F} := \left\{\tfrac{n}{2} e_1, (t,x) \mapsto gAx_n\right\}$$

and take open sets $V_j \subset\subset V_{j+1} \subset \mathcal{U}$ with $\bigcup_{j\geq 1} V_j = \mathcal{U}$, $|\partial V_j| = 0$ for $j \geq 1$. We convex integrate recursively as follows:

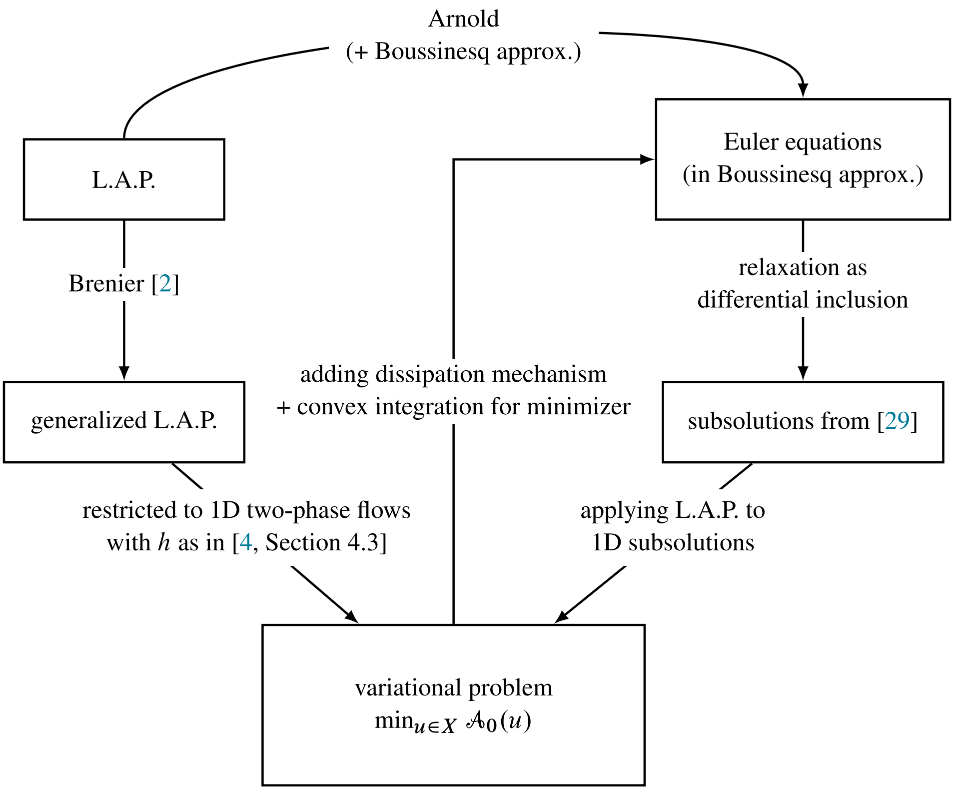

Figure 1. Relation between relaxations of the least action principle (L.A.P.) and Euler equations.

Step 1: Initiation. Let X_0^1 be the set of tuples $z = (\rho, v, m, \sigma, p)$ satisfying

- $(\rho, v, m, \sigma) \in (L^\infty \times L^2 \times L^2 \times L^1)((0,T) \times \mathcal{D})$,
- p is a distribution on $(0,T) \times \mathcal{D}$,
- z satisfies the linear system (2.1) with (1.2), (2.2),
- (ρ, v, m, σ) is continuous on V_1 and $z(t,x) \in U_{(t,x)}$ for $(t,x) \in V_1$,
- $z = z_{\text{sub}}$ a.e. in $(0,T) \times \mathcal{D} \setminus V_1$,
- $\exists\, C(z) \in (0,1)$ with
$$\left| \int_{\mathcal{D}} f(\rho - \rho_{\text{sub}})\, dx \right| \leq \frac{C(z)}{2} \delta(t)$$
for all $t \in [0,T]$ and all $f \in \mathcal{F}$.

Since e_0 and e_1 are continuous on $\overline{V_1}$, it follows that the set

$$\tilde{X}_0^1 := \left\{ (\rho, v|_{V_1}, m|_{V_1}, \sigma|_{V_1}) : z \in X_0^1 \text{ for some } p \right\}$$

is bounded in $L^\infty((0,T);L^2(\mathcal{D}))\times L^2(V_1;\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_0^{n\times n})$. Moreover, as in [29, Remark 2.4] it follows that for any such z there holds $\rho\in\mathcal{C}^0([0,T];L^2_w(\mathcal{D}))$.

Now let $B^{(1)}\subset L^2(\mathcal{D})$, $B^{(2)}\subset L^2(V_1;\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_0^{n\times n})$ be bounded balls such that

$$\rho(t,\cdot)\in B^{(1)},\quad \text{for all } t\in[0,T],\quad (v|_{V_1},m|_{V_1},\sigma|_{V_1})\in B^{(2)},$$

for any $(\rho,v|_{V_1},m|_{V_1},\sigma|_{V_1})\in\widetilde{X}_0^1$. Furthermore, let us denote by $d^{(1)}$ and $d^{(2)}$ the metrizations of the respective weak-L^2 topologies on these balls, and set

$$d_{X^1}(y,y'):=\max\Big\{\sup_{t\in[0,T]}d^{(1)}(\rho(t,\cdot),\rho'(t,\cdot)),\\ d^{(2)}((v|_{V_1},m|_{V_1},\sigma|_{V_1}),(v'|_{V_1},m'|_{V_1},\sigma'|_{V_1}))\Big\},$$

for $y,y'\in\widetilde{X}_0^1$. We define X^1 as the closure of $\widetilde{X}_0^1$ with respect to d_{X^1}, such that (X^1,d_{X^1}) is a complete metric space.

Proceeding as in [29] one obtains that the functional

$$I_1(y):=\int_{V_1}|y(t,x)|^2\,d(t,x)$$

is Baire-1 on X^1 and that

$$J_1(y):=\int_{V_1}d(y(t,x),\pi(K_{(t,x)}))\,d(t,x)$$

is continuous with respect to the strong-L^2 topology on X^1. Here π is the canonical projection from $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_0^{n\times n}\times\mathbb{R}$ to $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_0^{n\times n}$, eliminating the pressure component.

One may then use an argument based on a perturbation lemma (see for instance [29, Lemma 3.13]) to show that $J_1^{-1}(0)$ is residual in X^1. Hence, for every $\varepsilon_1>0$, we find $y_1\in J_1^{-1}(0)$, which after augmentation by a suitable $p^{(1)}$ gives a subsolution $z_{\text{sub}}^{(1)}$, ε_1-close to $(\rho_{\text{sub}},v_{\text{sub}},m_{\text{sub}},\sigma_{\text{sub}})$. More precisely, there holds

- $z_{\text{sub}}^{(1)}$ solves the linear system (2.1), (1.2), (2.2),
- $z_{\text{sub}}^{(1)}=z_{\text{sub}}$ outside V_1,
- $z_{\text{sub}}^{(1)}(t,x)\in K_{(t,x)}$ for a.e. $(t,x)\in V_1$,
- $d_{X^1}\big((\rho_{\text{sub}}^{(1)},v_{\text{sub}}^{(1)},m_{\text{sub}}^{(1)},\sigma_{\text{sub}}^{(1)}),(\rho_{\text{sub}},v_{\text{sub}},m_{\text{sub}},\sigma_{\text{sub}})\big)<\varepsilon_1$,
- $\left|\int_{\mathcal{D}}f(\rho_{\text{sub}}^{(1)}-\rho_{\text{sub}})\,dx\right|\le\frac{1}{2}\delta(t)$ for all $t\in[0,T]$, $f\in\mathcal{F}$.

That is, the mixing zone of $z_{\text{sub}}^{(1)}$ is given by $\mathcal{U}\setminus\overline{V_1}$.

Step 2: Recursion. Suppose that $j\ge 1$ and that there exists a subsolution $z_{\text{sub}}^{(j)}$ with mixing zone $\mathcal{U}\setminus\overline{V_j}$, i.e.

- $z_{\text{sub}}^{(j)}$ solves the linear system (2.1), (1.2), (2.2),

- $z_{\mathrm{sub}}^{(j)} = z_{\mathrm{sub}}$ outside V_j,
- $z_{\mathrm{sub}}^{(j)} \in K_{(t,x)}$ for a.e. $(t,x) \in V_j$,
- $d_{X^j}\big((\rho_{\mathrm{sub}}^{(j)}, v_{\mathrm{sub}}^{(j)}, m_{\mathrm{sub}}^{(j)}, \sigma_{\mathrm{sub}}^{(j)}), (\rho_{\mathrm{sub}}^{(j-1)}, v_{\mathrm{sub}}^{(j-1)}, m_{\mathrm{sub}}^{(j-1)}, \sigma_{\mathrm{sub}}^{(j-1)})\big) < \varepsilon_j$,
- $\big|\int_{\mathcal{D}} f(\rho_{\mathrm{sub}}^{(j)} - \rho_{\mathrm{sub}}^{(j-1)})\, dx\big| \le \frac{1}{2^j}\delta(t)$ for all $t \in [0,T]$, $f \in \mathcal{F}$.

Here, the spaces (X^j, d_{X^j}) are defined recursively by first saying that a tuple $z = (\rho, v, m, \sigma, p)$ belongs to X_0^{j+1} if and only if

- $(\rho, v, m, \sigma) \in (L^\infty \times L^2 \times L^2 \times L^1)((0,T) \times \mathcal{D})$,
- p is a distribution on $(0,T) \times \mathcal{D}$,
- z satisfies the linear system (2.1) with (1.2), (2.2),
- (ρ, v, m, σ) is continuous on $W_{j+1} := V_{j+1} \setminus \overline{V_j}$ and $z(t,x) \in U_{(t,x)}$ for $(t,x) \in W_{j+1}$,
- $z = z_{\mathrm{sub}}^{(j)}$ a.e. in $(0,T) \times \mathcal{D} \setminus W_{j+1}$,
- $\exists\, C(z) \in (0,1)$ with
$$\left|\int_{\mathcal{D}} f(\rho - \rho_{\mathrm{sub}}^{(j)})\, dx\right| \le \frac{C(z)}{2^{j+1}}\delta(t)$$
  for all $t \in [0,T]$ and all $f \in \mathcal{F}$,

and then constructing an appropriate completion $(X^{j+1}, d_{X^{j+1}})$ as in Step 1. Note that the whole space X^{j+1} depends on the previously chosen jth-order subsolution $z_{\mathrm{sub}}^{(j)}$.

As in Step 1, relying on the functionals $I_{j+1}, J_{j+1} \colon X^{j+1} \to \mathbb{R}$,

$$I_{j+1}(y) := \int_{W_{j+1}} |y(t,x)|^2\, d(t,x), \quad J_{j+1}(y) := \int_{W_{j+1}} d(y(t,x), \pi(K_{(t,x)}))\, d(t,x),$$

we may conclude the existence of subsolutions $z_{\mathrm{sub}}^{(j+1)}$ satisfying the properties listed at the beginning of Step 2 with j replaced by $j+1$ for any given $\varepsilon_{j+1} > 0$.

Step 3: Conclusion. In this manner we may construct (infinitely many) sequences $\{z_{\mathrm{sub}}^{(j)}\}_{j\ge1}$ which further satisfy

$$z_{\mathrm{sub}}^{(k)}|_{V_j} = z_{\mathrm{sub}}^{(j)}|_{V_j}, \quad \text{for } k \ge j \ge 1. \tag{B.1}$$

For simplicity of notation we also set $z_{\mathrm{sub}}^{(0)} := z_{\mathrm{sub}}$. Let us show that any such sequence converges to a solution.

First of all, we claim that there exists a dimensional constant $C = C(n) > 0$ such that

$$|m_{\mathrm{sub}}^{(j)}|^2 + |v_{\mathrm{sub}}^{(j)}|^2 + |\sigma_{\mathrm{sub}}^{(j)}| \le C(e_0 + \rho_{\mathrm{sub}}^{(j)} e_1) \quad \text{a.e. on } (0,T) \times \mathcal{D}. \tag{B.2}$$

Indeed, outside the corresponding mixing zone this is clear by the definition of $K_{(t,x)}$, whereas inside the mixing zone the inequalities involved in the definition of $U_{(t,x)}$ imply

$$\begin{aligned} |m|^2 + |v|^2 &= \frac{|m+v|^2}{2} + \frac{|m-v|^2}{2} < \frac{|m+v|^2}{1+\rho} + \frac{|m-v|^2}{1-\rho} \\ &< n(e_0+e_1)(1+\rho) + n(e_0-e_1)(1-\rho) = 2n(e_0 + \rho e_1), \end{aligned}$$

as well as

$$\begin{aligned}
-\lambda_{\min}(\sigma) = \lambda_{\max}(-\sigma) &< e_0 + \rho e_1 + \lambda_{\max}\Big(-\frac{v\otimes v - \rho(v\otimes m + m\otimes v) + m\otimes m}{1-\rho^2}\Big) \\
&= e_0 + \rho e_1 + \lambda_{\max}\Big(-\frac{(v+m)\otimes(v+m)}{2(1+\rho)} - \frac{(v-m)\otimes(v-m)}{2(1-\rho)}\Big) \\
&\le e_0 + \rho e_1.
\end{aligned}$$

Note that the bound on $-\lambda_{\min}(\sigma) = |\lambda_{\min}(\sigma)|$ is enough since σ is trace-free.

As a consequence we can conclude the following equi-integrability.

Lemma B.1. *For any $\varepsilon > 0$ there exists $\delta > 0$ such that for any measurable set $S \subset (0,T)\times\mathcal{D}$ with $|S| \le \delta$ there holds*

$$\int_S |\rho_{\text{sub}}^{(j)}| + |m_{\text{sub}}^{(j)}|^2 + |v_{\text{sub}}^{(j)}|^2 + |\sigma_{\text{sub}}^{(j)}|\,dx\,dt \le \varepsilon, \quad \textit{for all } j \ge 1.$$

Proof. The equi-integrability of the ρ-component is clear due to the uniform L^∞-bound.

Let us now show the equi-integrability of the σ-component. Let $\varepsilon > 0$ and $j_0 \ge 1$ be such that

$$2^{-j_0+1}\int_0^T \delta(t)\,dt \le \varepsilon \quad \text{and} \quad \int_{\mathcal{U}\setminus V_{j_0}} e_0 + \rho_{\text{sub}} e_1\,dx\,dt \le \varepsilon. \tag{B.3}$$

Then there exists $\delta > 0$ such that for any measurable set $S \subset (0,T)\times\mathcal{D}$ with $|S| \le \delta$ there holds

$$\int_S |\sigma_{\text{sub}}|\,dx\,dt \le \varepsilon \quad \text{and} \quad \int_S |\sigma_{\text{sub}}^{(j)}|\,dx\,dt \le \varepsilon \text{ for all } j \le j_0,$$

due to the equi-integrability of a finite family of integrable functions.

For $j > j_0$, using $f = \frac{n}{2}e_1 \in \mathcal{F}$ and (B.2), (B.3) we estimate

$$\begin{aligned}
\int_S |\sigma_{\text{sub}}^{(j)}|\,dx\,dt &= \int_{S\setminus V_j} |\sigma_{\text{sub}}|\,dx\,dt + \int_{S\cap(V_j\setminus V_{j_0})} |\sigma_{\text{sub}}^{(j)}|\,dx\,dt + \int_{S\cap V_{j_0}} |\sigma_{\text{sub}}^{(j_0)}|\,dx\,dt \\
&\le 2\varepsilon + \int_{\mathcal{U}\setminus V_{j_0}} |\sigma_{\text{sub}}^{(j)}|\,dx\,dt \le 2\varepsilon + C\int_{\mathcal{U}\setminus V_{j_0}} e_0 + \rho_{\text{sub}}^{(j)} e_1\,dx\,dt \\
&= 2\varepsilon + C\Big(\int_{\mathcal{U}\setminus V_{j_0}} e_0 + \rho_{\text{sub}} e_1\,dx\,dt + \frac{2}{n}\int_0^T\int_{\mathcal{D}} f(\rho_{\text{sub}}^{(j)} - \rho_{\text{sub}})\,dx\,dt\Big) \\
&\le 2\varepsilon + C\varepsilon + \frac{2C}{n}\sum_{k=j_0}^{j-1}\int_0^T \Big|\int_{\mathcal{D}} f(\rho_{\text{sub}}^{(k+1)} - \rho_{\text{sub}}^{(k)})\,dx\Big|\,dt \\
&\le 2\varepsilon + C\varepsilon + \frac{2C}{n}\sum_{k=j_0}^{+\infty} 2^{-(k+1)}\int_0^T \delta(t)\,dt \le \widetilde{C}(n)\varepsilon.
\end{aligned}$$

The proof for the m- and v-components is the same. ∎

The previous lemma and (B.1) allow us to check that $(\rho_{\mathrm{sub}}^{(j)}, v_{\mathrm{sub}}^{(j)}, m_{\mathrm{sub}}^{(j)}, \sigma_{\mathrm{sub}}^{(j)})_j$ is a Cauchy sequence in $L^1 \times L^2 \times L^2 \times L^1$ and thus converging to a limit (ρ, v, m, σ) solving the linear system (2.1), (1.2), (2.2) with some distributional pressure p.

Furthermore, outside $\mathcal{U}$ we have $z(t,x) = z_{\mathrm{sub}}(t,x) \in K_{(t,x)}$ a.e., while for every $(t,x) \in \mathcal{U}$ there exists $j \geq 1$ such that $(t,x) \in V_j$, and hence $z(t,x) = z_{\mathrm{sub}}^{(j)}(t,x) \in K_{(t,x)}$ a.e. in $\mathcal{U}$. So z is a solution to the Boussinesq system with $|\rho| = 1$ a.e. and its energy density is given by

$$\mathcal{E}(t,x) = \frac{n}{2}(e_0(t,x) + \rho(t,x)e_1(t,x)) + \rho(t,x)gAx_n.$$

Next we will show (2.10), i.e. we will show that

$$F(t) := \int_{\mathcal{D}} f(\rho - \rho_{\mathrm{sub}})\,dx$$

first of all is well defined and moreover satisfies $|F(t)| \leq \delta(t)$ for a.e. $t \in (0,T)$ and every $f \in \mathcal{F}$. In order to do this we define

$$F_j(t) := \int_{\mathcal{D}} f(\rho_{\mathrm{sub}}^{(j)} - \rho_{\mathrm{sub}})\,dx.$$

Since $\rho_{\mathrm{sub}}, \rho_{\mathrm{sub}}^{(j)} \in \mathcal{C}^0([0,T]; L^2_w(\mathcal{D}))$, and we have that $\rho_{\mathrm{sub}}^{(j)} - \rho_{\mathrm{sub}}$ is supported in V_j, where f is continuous, it follows that each F_j is continuous. On the other hand, for $j > j' \geq 1$ we may estimate

$$|F_j(t) - F_{j'}(t)| \leq \sum_{k=j'}^{j-1} \left| \int_{\mathcal{D}} f(\rho_{\mathrm{sub}}^{(k+1)} - \rho_{\mathrm{sub}}^{(k)})\,dx \right| \leq \sum_{k \geq j'} 2^{-(k+1)}\delta(t),$$

hence $\{F_j\}_{j\geq 1}$ is Cauchy in $\mathcal{C}^0([0,T])$, since δ was assumed to be continuous. Thus, $F_j \to \tilde{F}$ uniformly for some $\tilde{F}$, which satisfies $|\tilde{F}(t)| \leq \delta(t)$ for all $t \in [0,T]$.

In order to conclude $\tilde{F}(t) = F(t)$ for a.e. $t \in (0,T)$, we will show that $f(\rho - \rho_{\mathrm{sub}}) \in L^1((0,T) \times \mathcal{D})$. This is of course clear in the case $f = gAx_n$, but not for $f = \frac{n}{2}e_1$, as e_1 might be not integrable. Nonetheless, the claimed integrability for $f(\rho - \rho_{\mathrm{sub}})$ follows from the next lemma and monotone convergence.

Lemma B.2. *There holds*

$$\sup_j \int_0^T \int_{\mathcal{D}} |e_1(\rho_{\mathrm{sub}}^{(j)} - \rho_{\mathrm{sub}})|\,dx\,dt < \infty.$$

Proof. We abbreviate $\rho^j := \rho_{\mathrm{sub}}^{(j)}$ and $\bar{\rho} := \rho_{\mathrm{sub}}$. First of all note that if $|\rho^j| = 1$, which is the case almost everywhere on V_j, then there holds the following equivalence:

$$|\rho^j - \bar{\rho}| \leq 2(1 - \bar{\rho}^2) \quad \text{if and only if } \rho^j\bar{\rho} \geq -\frac{1}{2}.$$

We therefore estimate

$$\begin{aligned}\int_0^T \int_{\mathcal{D}} |e_1||\rho^j - \bar\rho| \, dx \, dt &= \int_{V_j} |e_1||\rho^j - \bar\rho| \, dx \, dt \\ &\le \int_{V_j \cap \{\rho^j \bar\rho < -\frac{1}{2}\}} |e_1||\rho^j - \bar\rho| \, dx \, dt \\ &\quad + 2 \int_{V_j \cap \{\rho^j \bar\rho \ge -\frac{1}{2}\}} |e_1|(1 - \bar\rho^2) \, dx \, dt.\end{aligned}$$

Now, the second term is bounded by $2\|e_1(1-\bar\rho^2)\|_{L^1((0,T)\times\mathcal{D})}$, which is finite by assumption (2.7).

Regarding the first term, observe that if $\rho^j \bar\rho < 0$, then $e_1(\rho^j - \bar\rho) \ge 0$ by condition (2.8). We therefore have

$$\begin{aligned}\int_{V_j \cap \{\rho^j \bar\rho < -\frac{1}{2}\}} |e_1||\rho^j - \bar\rho| \, dx \, dt &= \int_{\{\rho^j \bar\rho < -\frac{1}{2}\}} e_1(\rho^j - \bar\rho) \, dx \, dt \\ &= \int_0^T \int_{\mathcal{D}} e_1(\rho^j - \bar\rho) \, dx \, dt \\ &\quad - \int_{\{\rho^j \bar\rho \ge -\frac{1}{2}\}} e_1(\rho^j - \bar\rho) \, dx \, dt \\ &\le \int_0^T \delta(t) \, dt + 2\|e_1(1-\bar\rho^2)\|_{L^1((0,T)\times\mathcal{D})}.\end{aligned}$$

This finishes the proof of the lemma. ■

In consequence, $F(t)$ is well defined for a.e. $t \in (0,T)$ and using dominated convergence, one obtains that for any $\varphi \in L^\infty(0,T)$ there holds

$$\begin{aligned}\int_0^T \varphi(t)\tilde F(t) \, dt &= \lim_{j\to+\infty} \int_0^T \int_{\mathcal{D}} \varphi(t) f(\rho_{\mathrm{sub}}^{(j)} - \rho_{\mathrm{sub}}) \, dx \, dt \\ &= \lim_{j\to+\infty} \int_{V_j} \varphi(t) f(\rho - \rho_{\mathrm{sub}}) \, dx \, dt = \int_0^T \varphi(t) F(t) \, dt.\end{aligned}$$

Thus $\tilde F = F$ a.e. and therefore $|F(t)| \le \delta(t)$ for a.e. $t \in (0,T)$.

It remains to show the existence of a sequence of such solutions $(\rho_k, v_k)_k$ converging to $(\rho_{\mathrm{sub}}, v_{\mathrm{sub}})$ weakly in $L^2((0,T)\times\mathcal{D})$. Indeed, Step 2 and what we have shown so far in Step 3 allows us to have solutions $(\rho_k, v_k, m_k, \sigma_k)$, $k \ge 1$ that are generated by sequences of subsolutions $(\rho_{\mathrm{sub},k}^{(j)}, v_{\mathrm{sub},k}^{(j)}, m_{\mathrm{sub},k}^{(j)}, \sigma_{\mathrm{sub},k}^{(j)})_j$ such that for each fixed $j \ge 1$ there holds

$$(\rho_{\mathrm{sub},k}^{(j)}, v_{\mathrm{sub},k}^{(j)}, m_{\mathrm{sub},k}^{(j)}, \sigma_{\mathrm{sub},k}^{(j)}) \rightharpoonup (\rho_{\mathrm{sub}}, v_{\mathrm{sub}}, m_{\mathrm{sub}}, \sigma_{\mathrm{sub}})$$

weakly in $L^2((0,T)\times\mathcal{D})$.

For the ρ- and v-components this weak convergence extends as follows. Let $\varepsilon > 0$, $\varphi \in L^2((0,T) \times \mathcal{D})$ and pick a fixed $j \geq 1$ such that $\int_{\mathcal{U}\setminus V_j} |\varphi| < \varepsilon$. Then

$$\begin{aligned}\int_0^T \int_{\mathcal{D}} \varphi(\rho_k - \rho_{\mathrm{sub}})\,dx\,dt &= \int_{\mathcal{U}\setminus V_j} \varphi(\rho_k - \rho_{\mathrm{sub}})\,dx\,dt \\ &\quad + \int_0^T \int_{\mathcal{D}} \varphi(\rho_{\mathrm{sub},k}^{(j)} - \rho_{\mathrm{sub}})\,dx\,dt \\ &\leq 2\varepsilon + o(1)\end{aligned}$$

as $k \to \infty$. Thus $\rho_k \rightharpoonup \rho_{\mathrm{sub}}$ as $k \to \infty$.

The convergence $v_k \rightharpoonup v_{\mathrm{sub}}$ follows similarly, since for fixed $j \geq 1$ there holds

$$\begin{aligned}\int_0^T \int_{\mathcal{D}} \varphi(v_k - v_{\mathrm{sub}})\,dx\,dt &= \int_{\mathcal{U}\setminus V_j} \varphi(v_k - v_{\mathrm{sub}})\,dx\,dt + o(1) \\ &\leq \|\varphi\|_{L^2(\mathcal{U}\setminus V_j)}(\|v_k\|_{L^2((0,T)\times\mathcal{D})} + \|v_{\mathrm{sub}}\|_{L^2((0,T)\times\mathcal{D})}) + o(1),\end{aligned}$$

and

$$\begin{aligned}\|v_k\|^2_{L^2((0,T)\times\mathcal{D})} &= \int_0^T \int_{\mathcal{D}} n(e_0 + \rho_k e_1)\,dx\,dt \\ &\leq n\|e_0 + \rho_{\mathrm{sub}} e_1\|_{L^1((0,T)\times\mathcal{D})} + 2\int_0^T \delta(t)\,dt.\end{aligned}$$

Thus we have shown the convex integration Theorem 2.1.

Funding. J.J.K. has received support from the Ministry of Innovation and Technology of Hungary, from the National Research, Development and Innovation Fund (NKFIH), financed under the TKP2021 funding scheme, project no. BME-NVA-02. B.G. has been partially supported by the María Zambrano grant CA6/RSUE/2022-00097, the Severo Ochoa Programme for Centres of Excellence grant CEX2019-000904-S funded by MCIN/AEI/10.13039/501100011033, the grant PI2021-124-195NB-C32 funded by MCIN/AEI/10.13039/501100011033, and the ERC Advanced Grant 834728.

References

[1] V. Arnold, Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)* **16** (1966), no. fasc. 1, 319–361 Zbl 0148.45301 MR 0202082

[2] Y. Brenier, The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Amer. Math. Soc.* **2** (1989), no. 2, 225–255 Zbl 0697.76030 MR 0969419

[3] Y. Brenier, The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem. *Comm. Math. Phys.* **364** (2018), no. 2, 579–605 Zbl 1410.35102 MR 3869437

[4] Y. Brenier, Some concepts of generalized and approximate solutions in ideal incompressible fluid mechanics related to the least action principle. In *New trends and results in mathematical description of fluid flows*, pp. 53–75, Nečas Center Ser., Birkhäuser/Springer, Cham, 2018 Zbl 1418.35301 MR 3838269

[5] A. C. Bronzi, M. C. Lopes Filho, and H. J. Nussenzveig Lopes, Wild solutions for 2D incompressible ideal flow with passive tracer. *Commun. Math. Sci.* **13** (2015), no. 5, 1333–1343 Zbl 1326.35258 MR 3344429

[6] T. Buckmaster and V. Vicol, Convex integration constructions in hydrodynamics. *Bull. Amer. Math. Soc. (N.S.)* **58** (2021), no. 1, 1–44 Zbl 1461.35186 MR 4188806

[7] A. Castro, D. Córdoba, and D. Faraco, Mixing solutions for the Muskat problem. *Invent. Math.* **226** (2021), no. 1, 251–348 Zbl 1491.35331 MR 4309495

[8] Á. Castro, D. Faraco, and F. Mengual, Degraded mixing solutions for the Muskat problem. *Calc. Var. Partial Differential Equations* **58** (2019), no. 2, article no. 58 Zbl 1412.35246 MR 3921353

[9] Á. Castro, D. Faraco, and F. Mengual, Localized mixing zone for Muskat bubbles and turned interfaces. *Ann. PDE* **8** (2022), no. 1, article no. 7 Zbl 1492.35218 MR 4406890

[10] D. Chae and H.-S. Nam, Local existence and blow-up criterion for the Boussinesq equations. *Proc. Roy. Soc. Edinburgh Sect. A* **127** (1997), no. 5, 935–946 Zbl 0882.35096 MR 1475638

[11] F. Cheng, On the dissipative solutions for the inviscid Boussinesq equations. *AIMS Math.* **5** (2020), no. 4, 2869–2876 Zbl 1484.35317 MR 4146067

[12] E. Chiodaroli and O. Kreml, On the energy dissipation rate of solutions to the compressible isentropic Euler system. *Arch. Ration. Mech. Anal.* **214** (2014), no. 3, 1019–1049 Zbl 1304.35516 MR 3269641

[13] E. Chiodaroli and M. Michálek, Existence and non-uniqueness of global weak solutions to inviscid primitive and Boussinesq equations. *Comm. Math. Phys.* **353** (2017), no. 3, 1201–1216 Zbl 1373.35238 MR 3652488

[14] M. Colombo and A. Figalli, Regularity results for very degenerate elliptic equations. *J. Math. Pures Appl. (9)* **101** (2014), no. 1, 94–117 Zbl 1282.35175 MR 3133426

[15] D. Cordoba, D. Faraco, and F. Gancedo, Lack of uniqueness for weak solutions of the incompressible porous media equation. *Arch. Ration. Mech. Anal.* **200** (2011), no. 3, 725–746 Zbl 1241.35156 MR 2796131

[16] C. M. Dafermos, The entropy rate admissibility criterion for solutions of hyperbolic conservation laws. *J. Differential Equations* **14** (1973), 202–212 Zbl 0262.35038 MR 0328368

[17] R. Danchin, Remarks on the lifespan of the solutions to some models of incompressible fluid mechanics. *Proc. Amer. Math. Soc.* **141** (2013), no. 6, 1979–1993 Zbl 1283.35080 MR 3034425

[18] C. De Lellis and L. Székelyhidi, Jr., The Euler equations as a differential inclusion. *Ann. of Math. (2)* **170** (2009), no. 3, 1417–1436 Zbl 1350.35146 MR 2600877

[19] C. De Lellis and L. Székelyhidi, Jr., On admissibility criteria for weak solutions of the Euler equations. *Arch. Ration. Mech. Anal.* **195** (2010), no. 1, 225–260 Zbl 1192.35138 MR 2564474

[20] C. De Lellis and L. Székelyhidi, Jr., Weak stability and closure in turbulence. *Philos. Trans. Roy. Soc. A* **380** (2022), no. 2218, article no. 20210091 MR 4395517

[21] D. De Silva and O. Savin, Minimizers of convex functionals arising in random surfaces. *Duke Math. J.* **151** (2010), no. 3, 487–532 Zbl 1204.35080 MR 2605868

[22] D. G. Ebin and J. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math. (2)* **92** (1970), 102–163 Zbl 0211.57401 MR 0271984

[23] T. M. Elgindi and I.-J. Jeong, Finite-time singularity formation for strong solutions to the Boussinesq system. *Ann. PDE* **6** (2020), no. 1, article no. 5 Zbl 1462.35287 MR 4098032

[24] L. C. Evans, A new proof of local $C^{1,\alpha}$ regularity for solutions of certain degenerate elliptic p.d.e. *J. Differential Equations* **45** (1982), no. 3, 356–373 Zbl 0508.35036 MR 0672713

[25] L. C. Evans, *Partial differential equations*. 2nd edn., Grad. Stud. Math. 19, American Mathematical Society, Providence, RI, 2010 Zbl 1194.35001 MR 2597943

[26] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*. Revised edn., Textb. Math., CRC Press, Boca Raton, FL, 2015 Zbl 1310.28001 MR 3409135

[27] E. Feireisl, Maximal dissipation and well-posedness for the compressible Euler system. *J. Math. Fluid Mech.* **16** (2014), no. 3, 447–461 Zbl 1308.35190 MR 3247361

[28] C. Förster and L. Székelyhidi, Jr., Piecewise constant subsolutions for the Muskat problem. *Comm. Math. Phys.* **363** (2018), no. 3, 1051–1080 Zbl 1446.76159 MR 3858828

[29] B. Gebhard and J. J. Kolumbán, Relaxation of the Boussinesq system and applications to the Rayleigh–Taylor instability. *NoDEA Nonlinear Differential Equations Appl.* **29** (2022), no. 1, article no. 7 Zbl 1497.35368 MR 4353502

[30] B. Gebhard, J. J. Kolumbán, and L. Székelyhidi, A new approach to the Rayleigh–Taylor instability. *Arch. Ration. Mech. Anal.* **241** (2021), no. 3, 1243–1280 Zbl 1479.35643 MR 4284526

[31] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*. Classics in Mathematics, Springer, Berlin, 2001 Zbl 1042.35002 MR 1814364

[32] L. Hitruhin and S. Lindberg, Lamination convex hull of stationary incompressible porous media equations. *SIAM J. Math. Anal.* **53** (2021), no. 1, 491–508 Zbl 1458.35333 MR 4202508

[33] J. Kristensen and B. Raiţă, An introduction to generalized Young measures. Lecture Notes 45, Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig (2020)

[34] J. L. Lewis, Regularity of the derivatives of solutions to certain degenerate elliptic equations. *Indiana Univ. Math. J.* **32** (1983), no. 6, 849–858 Zbl 0554.35048 MR 0721568

[35] F. Mengual, H-principle for the 2-dimensional incompressible porous media equation with viscosity jump. *Anal. PDE* **15** (2022), no. 2, 429–476 Zbl 1490.35340 MR 4409883

[36] F. Mengual and L. Székelyhidi, Jr., Dissipative Euler flows for vortex sheet initial data without distinguished sign. *Comm. Pure Appl. Math.* **76** (2023), no. 1, 163–221 Zbl 07789785 MR 4544797

[37] C. Mooney, Hilbert's 19th problem revisited. *Boll. Unione Mat. Ital.* **15** (2022), no. 4, 483–501 Zbl 1500.35196 MR 4498802

[38] F. Noisette and L. Székelyhidi, Jr., Mixing solutions for the Muskat problem with variable speed. *J. Evol. Equ.* **21** (2021), no. 3, 3289–3312 Zbl 1502.35123 MR 4350274

[39] F. Otto, Evolution of microstructure in unstable porous media flow: a relaxational approach. *Comm. Pure Appl. Math.* **52** (1999), no. 7, 873–915 Zbl 0929.76136 MR 1682800

[40] F. Santambrogio and V. Vespri, Continuity in two dimensions for a very degenerate elliptic equation. *Nonlinear Anal.* **73** (2010), no. 12, 3832–3841 Zbl 1202.35107 MR 2728558

[41] O. Savin, Small perturbation solutions for elliptic equations. *Comm. Partial Differential Equations* **32** (2007), no. 4-6, 557–578 Zbl 1221.35154 MR 2334822

[42] A. I. Shnirel'man, The geometry of the group of diffeomorphisms and the dynamics of an ideal incompressible fluid. *Mat. Sb. (N.S.)* **128(170)** (1985), no. 1, 82–109, 144 Zbl 0725.58005 MR 0805697

[43] L. Székelyhidi, Weak solutions to the incompressible Euler equations with vortex sheet initial data. *C. R. Math. Acad. Sci. Paris* **349** (2011), no. 19-20, 1063–1066 Zbl 1230.35093 MR 2842999

[44] L. Székelyhidi, Jr., Relaxation of the incompressible porous media equation. *Ann. Sci. Éc. Norm. Supér. (4)* **45** (2012), no. 3, 491–509 Zbl 1256.35073 MR 3014484

[45] P. Tolksdorf, Regularity for a more general class of quasilinear elliptic equations. *J. Differential Equations* **51** (1984), no. 1, 126–150 Zbl 0488.35017 MR 0727034

[46] K. Uhlenbeck, Regularity for a class of non-linear elliptic systems. *Acta Math.* **138** (1977), no. 3-4, 219–240 Zbl 0372.35030 MR 0474389

[47] N. N. Ural'ceva, Degenerate quasilinear elliptic systems. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* **7** (1968), 184–222 Zbl 0199.42502 MR 0244628

[48] L. Wang, Compactness methods for certain degenerate elliptic equations. *J. Differential Equations* **107** (1994), no. 2, 341–350 Zbl 0792.35067 MR 1264526

[49] E. Wiedemann, Weak-strong uniqueness in fluid dynamics. In *Partial differential equations in fluid mechanics*, pp. 289–326, London Math. Soc. Lecture Note Ser. 452, Cambridge University Press, Cambridge, 2018 Zbl 1408.35158 MR 3838055

[50] C. Zălinescu, *Convex analysis in general vector spaces*. World Scientific, River Edge, NJ, 2002 Zbl 1023.46003 MR 1921556

Received 19 January 2023; revised 2 February 2024; accepted 22 April 2024.

Björn Gebhard
Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid; ICMAT, C. Nicolás Cabrera 13-15, 28049 Madrid, Spain; gebhard.bjorn@uam.es

Jonas Hirsch
Mathematisches Institut, Universität Leipzig, Augustusplatz 10, 04109 Leipzig, Germany; jonas.hirsch@math.uni-leipzig.de

József J. Kolumbán
Department of Analysis and Operations Research, Budapest University of Technology and Economics, Műegyetem rkp. 3, 1111 Budapest; HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda utca 13-15, 1053 Budapest, Hungary; jkolumban@math.bme.hu