

# Convergence and non-convergence in a nonlocal gradient flow

Sangmin Park\* and Robert L. Pego†

Department of Mathematical Sciences  
Carnegie Mellon University  
Pittsburgh, PA 15213.

July 19, 2023

*Dedicated to Sir John Ball*

## Abstract

We study the asymptotic convergence of solutions of  $\partial_t u = -f(u) + \int f(u)$ , a nonlocal differential equation that is formally a gradient flow in a constant-mass subspace of  $L^2$  arising from simplified models of phase transitions. In case the solution takes finitely many values, we provide a new proof of stabilization that uses a Lojasiewicz-type gradient inequality near a degenerate curve of equilibria. Solutions with infinitely many values in general *need not* converge to equilibrium, however, which we demonstrate by providing counterexamples for piecewise linear and cubic functions  $f$ . Curiously, the exponential *rate* of convergence in the finite-value case can jump from order  $O(1)$  to arbitrarily small values upon perturbation of parameters.

*Keywords:* gradient flows, infinite-dimensional dissipative dynamical systems

*Mathematics Subject Classification:* 34D05, 35B40, 37L15

## 1 Introduction

Let  $(\Omega, \mathcal{F}, \nu)$  be a probability space, and assume  $f: \mathbb{R} \rightarrow \mathbb{R}$  is locally Lipschitz and piecewise  $C^1$ . This paper investigates the asymptotic behavior as  $t \rightarrow \infty$  of bounded solutions to the nonlocal differential equation

$$\partial_t u(x, t) = -f(u(x, t)) + \int_{\Omega} f(u(y, t)) d\nu(y), \quad x \in \Omega, \quad t \geq 0. \quad (1.1)$$

While our main concern involves general nonlinear functions  $f$ , key examples to bear in mind are nonmonotonic polynomials and piecewise-linear functions.

The initial-value problem for (1.1) is well-posed locally in time in  $B(\Omega)$ , the Banach space of bounded measurable functions on  $\Omega$  equipped with the supremum norm. The solution exists globally in time and remains uniformly bounded if, for example, the initial data  $u(x, 0) = u_0(x)$  lies in an interval  $[a, b]$  with the property that  $f(a) \leq f(s) \leq f(b)$  for all  $s \in [a, b]$ , as will be seen below.

The nonlocal term in (1.1) ensures that the mean of the solution is conserved, as

$$\frac{d}{dt} \int_{\Omega} u(t, x) d\nu(x) = \int_{\Omega} (-f(u(t, x)) + \bar{f}(t)) d\nu(x) = 0, \quad (1.2)$$

---

\*Email address: sangminp@andrew.cmu.edu

†Email address: rpego@cmu.edu

where

$$\bar{f}(t) = \int_{\Omega} f(u(x, t)) \, d\nu(x). \quad (1.3)$$

We can view (1.1) formally as the equation of  $L^2$ -gradient flow constrained by fixing the mean: Let  $F$  denote the antiderivative of  $f$  — i.e.,

$$F(x) = \int_0^x f(y) \, dy, \quad (1.4)$$

and define the energy  $\mathcal{E}$  by

$$\mathcal{E}(u) = \int_{\Omega} F(u(x)) \, d\nu(x) = \int_{\Omega} \int_0^{u(x)} f(y) \, dy \, d\nu(x). \quad (1.5)$$

Then the equation (1.1) can be written formally in the form

$$\partial_t u = -\mathcal{Q}\nabla\mathcal{E}(u), \quad (1.6)$$

where  $\nabla\mathcal{E}(u) = f \circ u$  is formally the  $L^2$ -gradient of  $\mathcal{E}$  at  $u$ , and  $\mathcal{Q}$  is the  $L^2$ -orthogonal projection on the space of functions with mean zero.

Due to this constrained gradient structure, the energy is dissipated along solutions of (1.1), with

$$\frac{d}{dt}\mathcal{E}(u) = \int_{\Omega} f(u)(\partial_t u) \, d\nu = \int_{\Omega} (f(u) - \bar{f}(t))(\partial_t u) \, d\nu = - \int_{\Omega} (\partial_t u)^2 \, d\nu.$$

Hence for a bounded solution the limit  $\mathcal{E}_{\infty} = \lim_{t \rightarrow \infty} \mathcal{E}(u(t))$  exists, and we have

$$\mathcal{E}_{\infty} + \int_0^{\infty} \int_{\Omega} (\partial_t u)^2 \, d\nu \, dt = \mathcal{E}(u_0). \quad (1.7)$$

By (1.1),  $\partial_t u$  is uniformly Lipschitz in  $t$ , so  $\int_{\Omega} (\partial_t u)^2 \, d\nu$  is as well, whence it follows that

$$\int_{\Omega} (\partial_t u)^2 \, d\nu \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (1.8)$$

Then it follows any limit point of the orbit  $\{u(\cdot, t)\}_{t \geq 0}$  (in the  $L^2$  sense) must be an equilibrium, a (possibly discontinuous) function  $\hat{u}$  such that  $f(\hat{u}(x))$  is a.e. a constant.

The main question that we resolve herein is this:

$$\text{Does } u(\cdot, t) \text{ necessarily converge to a } \textit{single} \text{ equilibrium as } t \rightarrow \infty? \quad (1.9)$$

It is well-known that solutions of gradient systems need not converge in general, even in  $\mathbb{R}^2$  [25, p. 13]. But in the paper [27], the second author proved that for solutions of (1.1) the answer is *yes, assuming* the initial data  $u_0$  has finite range, taking only finitely many values  $u_1^0, \dots, u_N^0$ . In that case (1.1) is equivalent to a finite-dimensional system for  $\mathbf{u}(t) = (u_1(t), \dots, u_N(t))$  in  $\mathbb{R}^N$ . In [27], the solution's  $\omega$ -limit set is shown to contain points in a normally hyperbolic curve of equilibria, and a theorem of Hale and Massat [13] is invoked to conclude convergence as  $t \rightarrow \infty$ .

As pointed out by Şengül [30], the theorem of Hale and Massat used in [27] was improved by Hale and Raugel [14], and this could also improve the convergence proof in [27] in the finite range case. One thing we provide in the present paper is a different and considerably simpler proof of convergence in the finite range case, based on a *gradient inequality* of the form

$$c\|\mathcal{E}(u) - \mathcal{E}(\hat{u})\|^{1/2} \leq \|\mathcal{Q}\nabla\mathcal{E}(u)\|, \quad (1.10)$$

which is proved valid for  $u$  on the orbit near a “regular” equilibrium  $\hat{u}$  in the  $\omega$ -limit set under the assumption that  $\bar{f}(t)$  fails to converge. The use of gradient inequalities to analyze convergence of gradient flows was pioneered by Łojasiewicz [21] and Simon [31], and has since expanded greatly in the the field of optimization [2] and in the analysis of dynamics in PDE [16]. The proof of such inequalities in general involves a deep study of objects such as subanalytic sets and o-minimal structures [5, 7]. But in our case, a proof based on simple Taylor approximation works, since we use (1.10) not for arbitrarily degenerate equilibria  $\hat{u}$ , but only for families of equilibria that, although they are not isolated, correspond to regular values of  $f$ .

Our main result, however, is that the general answer to the main question (1.9) is *no!* — It is possible that  $u$  fails to converge if  $u_0$  takes infinitely many values. We construct counterexamples to convergence in cases when  $f$  is piecewise-linear or a cubic polynomial, having an “N-shaped” graph. Our constructions are motivated by the observation that perturbations (arbitrarily small in  $L^2$ ) of certain degenerate unstable equilibria can cause the value of  $\bar{f}(t)$  to eventually drift a finite distance either up or down. An infinite number of such perturbations can then be superimposed to cause  $\bar{f}(t)$  to oscillate, slower and slower, with no limit.

## 1.1 Related works

Equation (1.1) is a simplified model for dissipative dynamics in a number of models of phase transitions that are related to each other. These include models of viscoelastic materials [1, 3, 26], models of formation of material microstructure [4, 10, 11], regularized forward-backward diffusion models [24], and shear flows in non-Newtonian fluids [22, 23]. Şengül has recently reviewed work on nonlinear viscoelastic models of strain rate type [30].

In order to ensure convergence of solutions in a problem of viscoelasticity, Andrews and Ball [1] introduced a hypothesis that they called a *nondegeneracy condition*, which works also for solutions of (1.1). To explain, suppose for simplicity that  $f$  is piecewise monotone, so that for  $z$  in any bounded set of  $\mathbb{R}$ , the equation  $f(z) = s$  has a finite number  $M = M(s)$  of roots  $z_1(s) < z_2(s) < \dots < z_M(s)$ , where  $M$  is piecewise continuous jumping a finite number of times. Then the nondegeneracy condition requires that no nonzero linear combination of  $z_1, \dots, z_M$  is constant on any common interval of definition. For counterexamples to convergence as constructed in this paper, it is important that the nondegeneracy condition be violated. This is indeed the case however if, e.g.,  $f$  is any piecewise linear function, or a nonmonotonic cubic polynomial (since then the sum of the roots  $z_1 + z_2 + z_3$  is constant).

In 2015, Ball and Şengül published an in-depth study [3] of an equation of the form exactly as in (1.1) in the context of quasistatic nonlinear viscoelasticity in one space dimension. In this context, the variable  $u$  represents the material strain and should remain positive. For the measure space  $\Omega = [0, 1]$  with Lebesgue measure (or any Borel-isomorphic space), they establish that (1.1) is well-posed in the positive cone of  $L^2(\Omega)$  when  $F$  is  $\lambda$ -convex (i.e.,  $F(u) + \frac{1}{2}\lambda u^2$  is convex) and  $f(u) \rightarrow -\infty$  as  $u \downarrow 0$ , by making use of a one-sided Lipschitz condition on  $f$  to obviate the problem that the Nemytskii operator  $u \mapsto f \circ u$  is not Lipschitz on  $L^2$ . Ball and Şengül then make rigorous the interpretation of these solutions as a gradient flow of  $\mathcal{E}$  in a constant-mass subset of  $L^2(\Omega)$ . Further, they prove the  $L^2$  compactness of positive orbits using monotone rearrangement and Helly’s theorem, and they improve the convergence analysis in the studies [1, 24] in several ways. They prove that solutions converge to equilibrium under a weakened nondegeneracy condition. For the cubic case  $f(u) = u^3 - u$  in particular, convergence is proved under the hypothesis that

$$\int_{\Omega} u(x, 0) \, d\nu(x) \neq 0. \quad (1.11)$$

A nearly contemporaneous study by Hilhorst *et al.* [18] was motivated by study of a singularly perturbed Allen-Cahn equation with mass conservation [29]. These authors studied existence and uniqueness of solutions of (1.1) taking values in  $L^\infty(\Omega)$  for multistable nonlinearities [18, Theorem 1.4], and proved stabilization for bistable nonlinearities  $f$  when the initial data have no flat portions [18, Theorem 1.6], having the

property that all level sets  $\{x \in \Omega \mid u(x, 0) = c\}$  have measure zero. Based on the asymptotic behavior of solution to of the nonlocal ODE, they study the generation of interfaces for solutions of the mass-conserved Allen-Cahn equation [19].

## 1.2 Discussion and plan

Gradient flows are generally important in many areas in mathematics, including in optimization for purposes such as training artificial neural networks [6,9,20,28] and improving methods of statistical sampling [12]. The Łojasiewicz gradient estimates provide a powerful tool to conclude convergence of finite-dimensional gradient flows with analytic and also nonsmooth subanalytic nonlinearities [5]. Simon's extensions have allowed the handling of some infinite-dimensional flows, particularly for partial differential equations of parabolic type in which the infinite-dimensional dynamics can be slaved to some finite-dimensional part by a kind of Lyapunov-Schmidt reduction [16]. More recently, Łojasiewicz-type inequalities have been extended to general metric spaces [17].

In light of these strong results from gradient-estimate theory, our counterexamples for solutions of (1.1) are puzzling insofar as they work for the simplest kinds of probability spaces and nonlinearities. For example, non-convergent solutions can be found on the one-dimensional domain  $\Omega = [0, 1]$  which are monotone in  $x$  and have compact trajectories in  $L^2$ , and which have finite-dimensional (actually one-dimensional)  $\omega$ -limit sets. Moreover, the nonlinear function  $f$  can be polynomial (cubic), both as a real function and as a Nemytskii operator on  $B(\Omega)$  (although the latter is not even once Fréchet differentiable on the space  $L^2(\Omega)$ ).

So despite the rather benign nature of nonlocally coupled integrodifferential equations from the point of view of nonlinear analysis, having a very regular nonlinear structure and having essentially finite-dimensional long-time dynamics appears insufficient to ensure gradient-flow convergence. For finite-dimensional flows, solutions converge, but our constructions indicate that the rate of convergence can be arbitrarily slow, even for fixed nonlinearity and fixed dimension as small as 3. The appearance of arbitrarily slow rates of convergence is a curious phenomenon, in fact—it happens by perturbation from a situation in which the rate of convergence is  $O(1)$  and a Łojasiewicz inequality applies.

Our non-convergent examples are all non-generic and highly unstable. To emphasize how delicate non-convergence has to be for the cubic nonlinearity, we present the following criterion that is necessary (but far from sufficient) for non-convergence, which shows that non-convergence is far more unlikely to arrange than the codimension-1 necessary condition  $\int_{\Omega} u = 0$  from (1.11) might suggest.

**Proposition 1** (Unstable nature of non-convergence). *Let  $f(u) = u^3 - u$ , and suppose  $u(\cdot, t)$  is a bounded solution of (1.1) that fails to converge in  $L^2$  to a limit as  $t \rightarrow \infty$ . Then  $\int_{\Omega} u(x, 0) d\nu(x) = 0$ , and moreover, there exists  $c$  such that the three sets, consisting of all  $x \in \Omega$  where  $u(x, 0) = c$ , where  $u(x, 0) > c$ , and where  $u(x, 0) < c$  respectively, each have measure exactly equal to  $\frac{1}{3}$ .*

The plan of this paper is as follows. We develop a few basic properties of solutions of (1.1) in Section 2, regarding well-posedness, the relative preservation of order at different values of  $x$ , and invariant sets for solutions (a kind of maximum principle). In Section 3 we re-prove long-time convergence for solutions with finite range, in a simpler way than in [27] using gradient estimates. Our construction of non-convergent solutions for piecewise-linear bistable  $f$  appears in Section 4. Subsection 4.6 contains an  $L^2$  gradient inequality that is valid in this case (Lemma 7) which is curiously similar to the one used to prove convergence in the finite-range case with arbitrary nonlinearity (Lemma 3). In Section 5 we construct non-convergent examples for cubic  $f$ , and also complete the proof of Proposition 1. Finally we discuss in Section 6 a phenomenon of instability of convergence rates under perturbation around degenerate equilibria. For suitable three-valued initial data, parameter perturbations of order  $O(\varepsilon)$  leads to slow exponential convergence at rate  $O(\varepsilon)$ , whereas a rate of order  $O(1)$  is observed when  $\varepsilon = 0$ .

## 2 Basic properties of solutions

We begin our analysis with a brief discussion of the well-posedness of the initial value problem for (1.1), and some basic properties that solutions have regarding preservation of order and positively invariant sets.

We choose to work with solutions taking values  $u(\cdot, t)$  in the space of bounded measurable functions  $B(\Omega)$ , as it is convenient to interpret them as pointwise satisfying the integrodifferential equation in (1.1), without having to take the trouble of selecting representatives from equivalence classes as was done in [24] for elements of  $C([0, T], L^\infty(\Omega))$ . Local-time well-posedness (existence, uniqueness, and continuous dependence on initial data) follows by the standard Picard iteration method. This use of  $B(\Omega)$  makes well-posedness and the study of pointwise properties rather easy, as solutions  $u(x, t)$  are  $C^1$  in  $t$  for every  $x$ , but some other things become more difficult. E.g., even in case  $\Omega = [0, 1]$  with Lebesgue measure, it does not seem easy to determine whether, say, measurable monotone reordering is possible pointwise everywhere for all initial data.

We will make considerable use of the pointwise properties that solutions enjoy according to the two following results. The second one is similar to results observed in [18, Lemma 2.5], and previously for viscous diffusion equations in [24, Proposition 2.7].

**Lemma 1** (Preservation of order). *Let  $u$  solve the nonlocal ODE (1.1). If  $u(x, 0) < u(y, 0)$ , then for all  $t > 0$  we have  $u(x, t) < u(y, t)$ . Further, equality is also preserved.*

*Proof.* This is a simple consequence of the fact that if we regard  $\bar{f}(t)$  as given, then  $u(x, t)$  and  $u(y, t)$  satisfy the same scalar ODE with locally Lipschitz nonlinearity.  $\square$

We call a set  $S \subset \mathbb{R}$  *positively invariant* for (1.1) if the condition  $u(x, 0) \in S$  for all  $x \in \Omega$  implies that  $u(x, t) \in S$  for all  $x \in \Omega$  and  $t > 0$ . For a given solution  $u$ , we call a set  $\hat{S} \subset \mathbb{R}$  *pointwise stable* if  $u(\hat{x}, 0) \in \hat{S}$  implies  $u(\hat{x}, t) \in \hat{S}$  for all  $t > 0$ , for any (particular)  $\hat{x} \in \Omega$ .

**Lemma 2.** (i) (*Positively invariant sets*) *Let  $[a, b]$  be a closed interval such that*

$$f(a) \leq f(s) \leq f(b) \quad \text{for all } s \in [a, b].$$

*Then  $[a, b]$  is positively invariant.*

(ii) (*Pointwise stable subsets*) *If further  $[\hat{a}, \hat{b}] \subset [a, b]$  with  $f(\hat{a}) = f(a)$  and  $f(\hat{b}) = f(b)$ , then  $[\hat{a}, \hat{b}]$  is pointwise stable for any solution with  $u(x, 0) \in [a, b]$  for all  $x \in \Omega$ .*

*Proof.* Let  $u(0, x) \in [a, b]$  for all  $x \in \Omega$ . If  $f(u(\cdot, 0))$  is a.e. constant, then  $\bar{f}(0) \in [f(a), f(b)]$  is this same constant. So  $u$  is at equilibrium a.e., and trivially the invariance properties in parts (i) and (ii) hold.

Suppose  $f(u(\cdot, 0))$  is not a.e. constant. Then  $u$  is not a.e. at equilibrium, and  $\bar{f}(0) \in (f(a), f(b))$ . Define

$$t_* = \inf\{t > 0 : \bar{f}(t) \in \{f(a), f(b)\}\}.$$

This is the first exit time of  $\bar{f}(t)$  from the interval  $(f(a), f(b))$ . By continuity of  $\bar{f}$ , we know  $t_* > 0$ .

Next note that for any  $x \in \Omega$  and  $t \in [0, t_*)$ ,

$$-f(u(x, t)) + f(a) < -f(u(x, t)) + \bar{f}(t) = \partial_t u(x, t) < -f(u(x, t)) + f(b).$$

By consequence,  $\partial_t u(x, t)$  is positive if  $u(x, t) = a$  (or  $\hat{a}$ ) and negative if  $u(x, t) = b$  (or  $\hat{b}$ ). It follows  $u(x, t) \in (a, b)$  for all  $t \in (0, t_*)$ , and all  $x$ . Moreover if  $u(x, 0)$  is in  $[\hat{a}, \hat{b}]$  then  $u(x, t)$  remains there for all  $t \in [0, t_*)$ .

Now we claim  $t_* = \infty$ . If  $t_* < \infty$ , then by continuity  $u(x, t_*) \in [a, b]$  and  $f(a) \leq f(u(x, t_*)) \leq f(b)$  for all  $x$ . But then  $f(u(x, t_*))$  must a.e. equal  $f(a)$  if  $\bar{f}(t_*) = f(a)$ , and must a.e. equal  $f(b)$  if  $\bar{f}(t_*) = f(b)$ . This contradicts our hypothesis and establishes  $t_* = \infty$ . The invariance properties follow.  $\square$

By this result, if  $[a, b]$  is an interval with the property stated and the initial data  $u(x, 0)$  belong to this interval, then the solution to (1.1) exists globally with  $u(x, t) \in [a, b]$  for all  $t \geq 0$  and all  $x \in \Omega$ .

### 3 The case of finite range: convergence via gradient inequalities

Let  $u$  take finitely many values  $u_j$  on sets  $\Omega_j \subset \Omega$  of measure  $\mu_j$ ,  $j = 1, \dots, N$ , with  $\sum_k \mu_k = 1$ . Our equation is then equivalent to the following system in  $\mathbb{R}^N$ :

$$\frac{d}{dt}u_j(t) = -f(u_j(t)) + \bar{f}(t), \quad \bar{f}(t) = \sum_k \mu_k f(u_k(t)). \quad (3.1)$$

We define a reduced energy for vectors  $\mathbf{u} = (u_j) \in \mathbb{R}^N$  by

$$E(\mathbf{u}) = \sum_k \mu_k F(u_k).$$

With respect to the reduced inner product  $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_k \mu_k u_k v_k$ , the gradient  $\nabla E(\mathbf{u}) = (f(u_j))$ , and we may write the system above in the vector form

$$\frac{d\mathbf{u}}{dt} = -Q\nabla E(\mathbf{u}(t)), \quad Q\mathbf{v} = \mathbf{v} - \mathbb{1}\langle \mathbb{1}, \mathbf{v} \rangle = \left( v_j - \sum_k \mu_k v_k \right).$$

Here  $Q$  is the orthogonal projection on the subspace perpendicular to  $\mathbb{1} = (1, \dots, 1)$ .

Recall that we assume  $f$  is locally Lipschitz and piecewise  $C^1$ . Our goal in this section is to provide a simplified proof of the following theorem from [27].

**Theorem 1.** *If  $\mathbf{u}: [0, \infty) \rightarrow \mathbb{R}^N$  is a bounded solution of (3.1), then  $\lim_{t \rightarrow \infty} \mathbf{u}(t)$  exists.*

Our simplified proof avoids a spectral analysis of curves of equilibria of (3.1) and the use of the Hale-Massat theorem. Instead we rely on the gradient inequality contained in the following lemma. Its proof involves a simple Taylor approximation argument near ‘‘regular equilibria,’’ which stands in contrast to general Lojasiewicz inequalities valid near arbitrary equilibria for energies that are analytic, semi-algebraic, or more generally definable in an o-minimal structure [2].

We recall as in [27] that by Sard’s theorem, the set of *regular values* of  $f$  in any bounded interval of  $\mathbb{R}$  is open and dense. If  $\hat{s}$  is a regular value of  $f$ , then the equation  $f(z) = s$  has a finite number of solutions  $z_i(s)$  at which  $f'(z_i(s)) \neq 0$ , for all  $s$  in some neighborhood  $\hat{J}$  of  $\hat{s}$ . We will call  $\hat{\mathbf{u}} \in \mathbb{R}^N$  a *regular equilibrium* for (3.1) if  $\hat{s} = f(\hat{u}_j)$  is independent of  $j$  and is a regular value of  $f$ . In this case, then for each  $j$  there exists  $i(j)$  such that  $\hat{u}_j = z_{i(j)}(\hat{s})$ . We define  $\phi(s) = (z_{i(j)}(s))$  for  $s \in \hat{J}$ ; then  $s \mapsto \phi(s)$  is a curve of regular equilibria and  $\phi(\hat{s}) = \hat{\mathbf{u}}$ .

**Lemma 3.** *Let  $\hat{\mathbf{u}} \in \mathbb{R}^N$  be a regular equilibrium for (3.1) as above. Then in some neighborhood of  $\hat{\mathbf{u}}$ , all equilibria of (3.1) have the form  $\phi(s)$  for some  $s \in \hat{J}$ , and all states  $\mathbf{u}$  satisfy the gradient inequality*

$$c|E(\mathbf{u}) - E(\phi(s)) - s\langle \mathbb{1}, \mathbf{u} - \phi(s) \rangle| \leq \|Q\nabla E(\mathbf{u})\|^2, \quad s = \sum_j \mu_j f(u_j),$$

for some  $c > 0$  independent of  $\mathbf{u}$ .

*Proof.* For any equilibrium  $\mathbf{u}^{\text{eq}}$  in a small enough neighborhood  $\mathcal{N}$  of  $\hat{\mathbf{u}}$ ,  $s = f(u_j^{\text{eq}})$  is independent of  $j$  and near  $\hat{s}$ , so necessarily  $u_j^{\text{eq}} = z_{i(j)}(s)$  by the inverse function theorem. Taking  $\mathcal{N}$  smaller if necessary, for any  $\mathbf{u} \in \mathcal{N}$  we may let

$$s = \sum_j \mu_j f(u_j), \quad \mathbf{v} = \mathbf{u} - \phi(s),$$

and we may find constants  $0 < \underline{\lambda} < \bar{\lambda} < \infty$  such that  $\underline{\lambda} < |f'(u_j)| < \bar{\lambda}$  for all  $\mathbf{u} \in \mathcal{N}$  and all  $j$ . By Taylor's theorem we may write

$$F(u_j) = F(\phi_j(s)) + f(\phi_j(s))v_j + \frac{1}{2}\ell_j(\mathbf{u})v_j^2, \quad f(u_j) = f(\phi_j(s)) + \hat{\ell}_j(\mathbf{u})v_j, \quad (3.2)$$

where

$$\ell_j(\mathbf{u}) = 2 \int_0^1 f'(\phi_j(s) + rv)(1-r) dr, \quad \hat{\ell}_j(\mathbf{u}) = \int_0^1 f'(\phi_j(s) + rv) dr. \quad (3.3)$$

The bounds  $\underline{\lambda} < |\ell_j(\mathbf{u})|, |\hat{\ell}_j(\mathbf{u})| < \bar{\lambda}$  hold for all  $\mathbf{u} \in \mathcal{N}$ . Then since  $s = f(\phi_j(s))$  we have

$$\begin{aligned} E(\mathbf{u}) - E(\phi(s)) &= \sum_j \mu_j \left( F(u_j) - F(\phi_j(s)) \right) = \sum_j \mu_j \left( f(\phi_j(s))v_j + \frac{1}{2}\ell_j(\mathbf{u})v_j^2 \right) \\ &= s \sum_j \mu_j v_j + \frac{1}{2} \sum_j \mu_j \ell_j(\mathbf{u})v_j^2. \end{aligned}$$

Since also  $s = \sum_k \mu_k f(u_k)$ , we find

$$Q\nabla E(\mathbf{u})_j = f(u_j) - \sum_k \mu_k f(u_k) = f(u_j) - f(\phi_j(s)) = \hat{\ell}_j(\mathbf{u})v_j,$$

hence  $\|Q\nabla E(\mathbf{u})\|^2 = \sum_j \mu_j \hat{\ell}_j(\mathbf{u})^2 v_j^2$ . Evidently we have the estimates

$$\left| \sum_j \mu_j \ell_j(\mathbf{u})v_j^2 \right| \leq \bar{\lambda} \sum_j \mu_j v_j^2 \leq \frac{\bar{\lambda}}{\underline{\lambda}^2} \sum_j \mu_j \hat{\ell}_j(\mathbf{u})^2 v_j^2,$$

whence the result claimed in the Lemma follows with  $c = \frac{1}{2}\underline{\lambda}^2/\bar{\lambda}$ .  $\square$

The next (and main) step in the proof of Theorem 1 is to show that  $\bar{f}(t)$  converges. This is as in [27], but now the proof is much simpler.

**Lemma 4.** *If  $\mathbf{u}: \mathbb{R}^N \times [0, \infty)$  is a bounded solution of (3.1), then  $\lim_{t \rightarrow \infty} \bar{f}(t)$  exists.*

*Proof.* Suppose not. Then the interval  $(\liminf \bar{f}, \limsup \bar{f})$  is nonempty and strictly contains some interval  $\hat{J}$  of regular values of  $f$ , by Sard's theorem as above. Fixing some  $\hat{s} \in \hat{J}$ , using the compactness of the orbit we can find a sequence  $t_n \rightarrow \infty$  such that  $\bar{f}(t_n) = \hat{s}$  and  $\mathbf{u}(t_n)$  converges to some regular equilibrium  $\hat{\mathbf{u}} \in \omega(\mathbf{u})$ . Then because  $\omega(\mathbf{u})$  is connected and  $\hat{\mathbf{u}}$  cannot be isolated in  $\omega(\mathbf{u})$ , by taking  $\hat{J}$  smaller and on one side of  $\hat{s}$  if necessary, the curve of equilibria  $\{\phi(s) : s \in \hat{J}\}$  provided by the Lemma will be entirely contained in  $\omega(\mathbf{u})$ .

By consequence, we infer that for all  $s \in \hat{J}$ ,

$$E(\phi(s)) = E_* \quad \text{and} \quad \sum_j \mu_j \phi_j(s) = c_0 = \sum_j \mu_j u_j(t), \quad (3.4)$$

where  $E_* = \lim_{t \rightarrow \infty} E(\mathbf{u}(t))$ . By the result of the Lemma, then, we have

$$0 < \hat{c} \sqrt{E(\mathbf{u}(t)) - E_*} \leq \|Q\nabla E(\mathbf{u}(t))\| \quad (3.5)$$

whenever  $\mathbf{u}(t) \in \mathcal{N}$ , a small enough neighborhood of  $\hat{\mathbf{u}}$ . But then, by the classic argument of Łojasiewicz, and because  $Q = Q^2$  is self-adjoint,

$$-\frac{d}{dt} \sqrt{E(\mathbf{u}(t)) - E_*} = \frac{\langle \nabla E(\mathbf{u}(t)), Q^2 \nabla E(\mathbf{u}(t)) \rangle}{2\sqrt{E(\mathbf{u}(t)) - E_*}} = \frac{\|Q\nabla E(\mathbf{u}(t))\| \|\partial_t \mathbf{u}\|}{2\sqrt{E(\mathbf{u}(t)) - E_*}} \geq \frac{\hat{c}}{2} \|\partial_t \mathbf{u}\|. \quad (3.6)$$

On any interval  $[t_n, T]$  on which  $\mathbf{u}(t) \in \mathcal{N}$  it follows

$$\|\mathbf{u}(T) - \mathbf{u}(t_n)\| \leq \int_{t_n}^T \|\partial_t \mathbf{u}(\tau)\| d\tau \leq C\sqrt{E(\mathbf{u}(t_n)) - E_*}.$$

For large enough  $n$ , the right-hand side becomes arbitrarily small and it follows  $\mathbf{u}(t)$  remains inside  $\mathcal{N}$  for all  $t \geq t_n$ . This implies  $(\liminf \bar{f}, \limsup \bar{f}) \subset \hat{J}$ , a contradiction. Hence  $\lim_{t \rightarrow \infty} \bar{f}(t)$  exists.  $\square$

**Remark 1.** The inequality (3.5) can be interpreted as a Lojasiewicz inequality in the constrained-mean hypersurface  $\mathcal{M} = \{\mathbf{u} \in \mathbb{R}^N : \sum_j \mu_j u_j = c_0\}$ , since the gradient of  $E$  restricted to this surface can be interpreted as the projection  $Q\nabla E$  on the tangent space. Thus the inequality of Lemma 3 becomes a Lojasiewicz inequality near regular equilibria under the assumption that the identities in (3.4) hold.

This happens in the proof above where we assume  $\bar{f}(t)$  fails to converge. But it can also happen in certain degenerate situations when the Lojasiewicz inequality can be used to ensure a rate of convergence to equilibrium. A standard argument ensures a certain asymptotically exponential rate of convergence, *provided* the states  $\phi(s)$  with  $s = \bar{f}(t)$  can have the same average as  $\mathbf{u}(t)$  and the same limiting energy  $E_*$ . These provisos will not hold, however, in the scenarios to be considered in Sections 4 and 5, where arbitrarily slow rates of convergence are seen to be possible.  $\triangle$

The remainder of the proof of Theorem 1 goes as in [27], in principle. However, the proof in that paper appears to have a gap (in Lemma 3 in particular), so we provide a full corrected proof here for the convenience of the reader.

*Proof of Theorem 1.* Suppose for contradiction that some bounded solution  $\mathbf{u}$  of (3.1) fails to converge. Then  $a_j < b_j$  for some  $j$ , where

$$a_j = \liminf u_j(t), \quad b_j = \limsup u_j(t), \quad j = 1, \dots, N.$$

Due to Lemma 4, by adding a constant to  $f$  we may assume  $\bar{f}(t) \rightarrow 0$  as  $t \rightarrow \infty$ . By considering times  $t_{n,j} \rightarrow \infty$  such that  $u_j(t_{n,j})$  takes given limits inside  $(a_j, b_j)$ , we infer  $f(v) = 0$  for all  $v \in \bigcup_j [a_j, b_j]$ .

The idea of the remainder of the proof is that mass conservation  $\sum_k \mu_k u_k(t) = c_0$  must become violated, due to the synchrony implied by the equations  $\partial_t u_j = \bar{f}(t) = \partial_t u_k$  which must hold whenever  $u_j$  and  $u_k$  are respectively inside any nonempty open intervals  $(a_j, b_j)$ ,  $(a_k, b_k)$ .

Select a point  $\mathbf{v}$  in  $\omega(\mathbf{u})$  such that  $v_j \in (a_j, b_j)$  for  $j$  in some *maximal* set  $S$  of indices. With the notation  $B(x, r) = [x - r, x + r]$ , choose  $\varepsilon > 0$  so that  $B(v_j, 2\varepsilon) \subset (a_j, b_j)$  for all  $j \in S$ , and select  $t_n \rightarrow \infty$  such that  $\mathbf{u}(t_n) \rightarrow \mathbf{v}$  as  $n \rightarrow \infty$  and  $u_j(t_n) \in B(v_j, \varepsilon)$  for all  $j$  and  $n$ . Now fix some  $i \in S$  and define

$$T_n = \inf\{t > t_n : |u_i(t) - u_i(t_n)| > \varepsilon\}, \quad I_n = [t_n, T_n].$$

Then for all  $n$ ,  $t_n < T_n < \infty$  and  $u_i(t) \in B(v_i, 2\varepsilon)$  for all  $t \in I_n$ . Moreover, for any  $j \in S$ , by synchrony we have

$$u_j(t) - u_j(t_n) = u_i(t) - u_i(t_n) \in [-\varepsilon, \varepsilon] \quad \text{and} \quad u_j(t) \in B(v_j, 2\varepsilon)$$

for all  $t \in I_n$ . In particular, when  $t = T_n$  it follows there is a fixed sign  $\sigma \in \{-1, +1\}$  such that

$$u_j(T_n) - u_j(t_n) = u_i(T_n) - u_i(t_n) = \sigma\varepsilon. \tag{3.7}$$

By passing to a subsequence we may presume this holds for all  $n$  with  $\sigma$  independent of  $n$ .

We claim next that for all indices  $k \notin S$ ,

$$\operatorname{osc}_{I_n} u_k \rightarrow 0 \quad \text{as } n \rightarrow \infty, \tag{3.8}$$



where  $\text{osc}$  is the oscillation—supremum minus infimum on the indicated interval. Suppose not. Then for some  $k$ ,  $\text{osc}_{I_n} u_k \geq \hat{\varepsilon} > 0$  for infinitely many  $n$ . Hence  $b_k - a_k \geq \hat{\varepsilon}_n$ , and by continuity there exist  $\tau_n \in I_n$  such that  $u_k(\tau_n) = \hat{v}_k$  for some  $\hat{v}_k \in (a_k, b_k)$ . We may extract a suitable subsequence such that  $u_j(\tau_n)$  converges to some  $\hat{v}_j$  for all  $j$ . In particular we find  $\hat{v}_j \in (a_j, b_j)$  for all  $j \in S \cup \{k\}$ . This contradicts the maximality of  $S$ . Hence (3.8) holds.

From this it follows  $u_k(T_n) - u_k(t_n) \rightarrow 0$  for all  $k \notin S$ . Along the appropriate subsequence then, mass conservation together with (3.7) implies

$$\sum_j \mu_j v_j = \lim \sum_j \mu_j u_j(t_n) = \lim \sum_j \mu_j u_j(T_n) = \sigma \varepsilon \#S + \sum_j \mu_j v_j,$$

where  $\#S$  is the cardinality of  $S$ . This contradiction implies  $\mathbf{u}(t)$  tends to a limit.  $\square$

## 4 Non-convergence: the piecewise-linear case

In this section we describe solutions to (1.1) that do not converge as  $t \rightarrow \infty$ , for the case when  $f$  is piecewise linear with  $N$ -shaped graph, given by

$$f(z) = \begin{cases} z + 1 & z < -\frac{1}{2}, \\ -z & |z| < \frac{1}{2}, \\ z - 1 & z > \frac{1}{2}. \end{cases} \quad (4.1)$$

For  $|s| < \frac{1}{2}$ , the equation  $f(z) = s$  has the three solutions  $z_l(s) = -1 + s$ ,  $z_m(s) = -s$ , and  $z_r(s) = 1 + s$ . Since  $z_l + 2z_m + z_r \equiv 0$ , we see  $f$  fails to satisfy the nondegeneracy condition of Andrews and Ball [1]; this will be crucial in our construction. We presume the probability measure  $\nu$  is nonatomic. This implies that given any countable set  $(\mu_j)$  with  $\sum \mu_j = 1$ , there exists a measurable partition  $(\Omega_j)$  of  $\Omega$  such that  $\nu(\Omega_j) = \mu_j$  for all  $j$ . (This follows since  $\nu$  has the ‘‘Darboux property,’’ see [8, p. 28] and [15, p. 174(2)].)

### 4.1 Equilibria, and phase transition times

*Equilibria.* With  $f$  as in (4.1), equation (1.1) has a family of equilibria  $\hat{u}_s$  satisfying  $f(\hat{u}_s(x)) \equiv s$  for any constant  $s \in (-\frac{1}{2}, \frac{1}{2})$ , with  $\hat{u}_s(x) = z_j(s)$  on sets  $\hat{\Omega}_j$  of measure denoted  $\hat{\mu}_j$  for  $j = l, m, r$  to indicate the left, middle, and right phases, respectively. We fix the particular values

$$\hat{\mu}_l = \frac{1}{4}, \quad \hat{\mu}_m = \frac{1}{2}, \quad \hat{\mu}_r = \frac{1}{4}, \quad (4.2)$$

so that all these equilibria have mean zero, i.e.,

$$\int_{\Omega} \hat{u}_s(x) d\nu(x) = 0, \quad \text{independent of } s.$$

Our goal in this section is to describe a solution that has some nontrivial collection of these equilibria in its  $\omega$ -limit set (in the  $L^2$  topology).

*Phases and transition times.* In this section, we will only consider solutions taking values in the interval  $[a, b] = [-\frac{3}{2}, \frac{3}{2}]$ , which is positively invariant according to Lemma 2. For the remainder of this section we fix the values

$$a = -\frac{3}{2}, \quad \hat{b} = -\frac{1}{2}, \quad \hat{a} = \frac{1}{2}, \quad b = \frac{3}{2},$$

and define left, middle, and right phase intervals respectively by

$$\Phi_l = [a, \hat{b}], \quad \Phi_m = (\hat{b}, \hat{a}), \quad \Phi_r = [\hat{a}, b]. \quad (4.3)$$

For the solutions we consider, the left and right phase subintervals  $\Phi_l$  and  $\Phi_r$  are each pointwise stable. We define measures of sets corresponding to the left, middle, and right phases by

$$\nu_j(t) = \nu(A_j(t)), \quad A_j(t) = \{x \in \Omega : u(x, t) \in \Phi_j\}, \quad (4.4)$$

for each symbol  $j = l, m, r$  respectively. Then by pointwise stability, the left and right phases  $A_l(t)$  and  $A_r(t)$  and their measures are nondecreasing, while the middle phase  $A_m(t)$  and its measure  $\nu_m(t)$  are nonincreasing. Consequently a transition time (exit time) from the middle phase exists at each point, as follows.

**Lemma 5** (Phase transition times). *For each  $x$  with  $u(x, 0) \in \Phi_m$ , there exists  $\tau(x) \in (0, \infty]$  such that*

$$u(x, t) \in \begin{cases} \Phi_m, & 0 \leq t < \tau(x), \\ \Phi_l \cup \Phi_r, & t \geq \tau(x). \end{cases}$$

Moreover, as long as two points  $u(x, t)$  and  $u(y, t)$  remain in the middle phase  $\Phi_m$ , the difference grows exponentially, for we have

$$\partial_t(u(x, t) - u(y, t)) = u(x, t) - u(y, t).$$

**Corollary 1.** *If  $u(x, 0), u(y, 0) \in \Phi_m$ , then for  $0 \leq t < \tau(x) \wedge \tau(y)$  we have*

$$u(x, t) - u(y, t) = e^t(u(x, 0) - u(y, 0)).$$

## 4.2 Mean force and heuristics

*Evolution of mean force.* For the piecewise-linear nonlinearity in (4.1), it happens that  $\bar{f}(t)$  evolves in a strikingly simple way. Due to the fact that

$$f'(u) = \begin{cases} 1 & u < -\frac{1}{2}, \\ -1 & |u| < \frac{1}{2}, \\ 1 & u > \frac{1}{2}, \end{cases} \quad f'(u)f(u) = \begin{cases} u+1 & u < -\frac{1}{2}, \\ u+0 & |u| < \frac{1}{2}, \\ u-1 & u > \frac{1}{2}, \end{cases} \quad (4.5)$$

and  $\bar{f}(t)$  is Lipschitz, hence differentiable a.e., we find using (4.4) that with  $\bar{u} = \int_{\Omega} u \, d\nu$ , for a.e.  $t$ ,

$$\frac{d}{dt} \bar{f}(t) = \int_{\Omega} f'(u)(-f(u) + \bar{f}(t)) \, dx = -(\bar{u} + \nu_l - \nu_r) + (\nu_l - \nu_m + \nu_r) \bar{f}(t). \quad (4.6)$$

*Heuristics.* We can now explain the main idea behind our examples of non-convergence, by describing a simple calculation that shows how tiny perturbations from certain (always unstable) degenerate equilibria can produce slow, but eventually large, changes in  $\bar{f}(t)$ . We will consider solutions with mean  $\bar{u} = 0$ . Desiring some equilibrium  $\hat{u}$  as above to be in the  $\omega$ -limit set,  $\nu_m(t)$ , the measure of the middle phase, should approach  $\hat{\nu}_m = \frac{1}{2}$  from above. Thus we will perturb by moving small bits of the (stable) left and right phases to be in the (unstable) middle phase, close to but not exactly at the same value as  $\hat{u}$  takes.

Imagine then that the initial data takes values near  $-1 + s$ ,  $-s$ ,  $1 + s$  on sets of measure

$$\nu_j = \hat{\nu}_j - \varepsilon_j, \quad (4.7)$$

for each symbol  $j = l, m, r$  respectively, with  $\varepsilon_l, \varepsilon_r > 0$  small and  $\varepsilon_m = -\varepsilon_l - \varepsilon_r$ . Suppose no phase changes occur over some interval of time during which the measures  $\nu_j$  do not change. Then during this time interval, (4.6) becomes

$$\frac{d}{dt} \bar{f}(t) = \varepsilon_l - \varepsilon_r - 2(\varepsilon_l + \varepsilon_r) \bar{f}(t). \quad (4.8)$$

Regardless of what the original value of  $s$  was,  $\bar{f}(t)$  is now forced to drift toward a particular equilibrium value determined by  $\varepsilon_l$  and  $\varepsilon_r$ , namely

$$\bar{f}^{\text{eq}} = \frac{1}{2} \frac{\varepsilon_l - \varepsilon_r}{\varepsilon_l + \varepsilon_r} \in \left(-\frac{1}{2}, \frac{1}{2}\right). \quad (4.9)$$

This value can be of order 1 no matter how small  $\varepsilon_l, \varepsilon_r$  are.

Now the idea to obtain persistent oscillations is to use the exponential growth rate of perturbations in the (unstable) middle phase to arrange that small bits of that phase will change alternately to the (stable) left and right phases. The time gaps between these changes should be large enough so that  $\bar{f}(t)$  is attracted near the prevailing value of  $\bar{f}^{\text{eq}}$ , and the pattern of changes should cause  $\varepsilon_l, \varepsilon_r$  to alternately decrease in a way that forces the value of  $\bar{f}^{\text{eq}}$  to alternately drift toward distinctly different values. We will show this can be done infinitely often, with the implication that  $\bar{f}(t)$  will fail to converge as  $t \rightarrow \infty$ , and the same for  $u(\cdot, t)$ .

### 4.3 Initial data and main result

With suitable initial data specified as follows, we can ensure that  $\bar{f}(t)$  fails to converge. We consider initial data taking infinitely many values, of the form  $u(x, 0) = v_0(x) - \bar{v}_0$  so that  $\bar{u} = 0$ , with

$$v_0(x) = \begin{cases} -1, 0, 1 & \text{in } \Omega_l, \Omega_m, \Omega_r \text{ respectively,} \\ (-1)^j \alpha_j & \text{in } \Omega_j, j = 0, 1, 2, \dots \end{cases} \quad (4.10)$$

We write  $\mu_j = \nu(\Omega_j)$  for  $j = l, m, r$  and  $0, 1, 2, \dots$ , and assume

$$\mu_l = \frac{1}{4} - \sum_{j \text{ odd}} \mu_j, \quad \mu_m = \frac{1}{2}, \quad \mu_r = \frac{1}{4} - \sum_{j \text{ even}} \mu_j. \quad (4.11)$$

**Theorem 2** (Counterexample to convergence). *Let  $f$  be given by (4.1) and consider initial data for (1.1) of the form  $u(x, 0) = v_0(x) - \bar{v}_0$  with  $v_0$  given as above. Let  $0 < \eta < 1$ , and assume  $0 < \mu_0 \leq \frac{1-\eta}{4}$  and*

$$\mu_j = \mu_0 \eta^j \quad (j = 0, 1, 2, \dots).$$

Assume  $0 < \alpha_0 < \frac{1}{4}$ , and that

$$0 < \alpha_{j+1} \leq \alpha_j \mu_j \quad \text{for } j = 0, 1, 2, \dots \quad (4.12)$$

Then:

- (i) The phase transition times  $\tau_j = \tau(\Omega_j)$  satisfy  $\tau_m = +\infty$  and  $\tau_{j+1} > \tau_j$  for all  $j \geq 0$ .
- (ii) If moreover for some positive sequence  $\beta_j$  decreasing to 0,

$$\alpha_{j+1} \leq \alpha_j \mu_j \beta_j^{1/\mu_j} \quad (4.13)$$

for all  $j$  sufficiently large, then

$$\limsup \bar{f}(t) - \liminf \bar{f}(t) = \frac{1-\eta}{1+\eta},$$

and as  $t \rightarrow \infty$  the solution  $u$  fails to converge in  $L^p$  for any  $p \in [0, \infty]$ . Its  $\omega$ -limit set consists of all the equilibria  $\hat{u}_s$  for  $|s| \leq \frac{1-\eta}{2(1+\eta)}$ .

Observe that

$$-\bar{v}_0 = \mu_l - \mu_r - \sum_{j \geq 0} (-1)^j \alpha_j \mu_j, \quad \mu_l - \mu_r = \frac{\mu_0}{1 + \eta}, \quad 0 < \sum_{j \geq 0} (-1)^j \alpha_j \mu_j < \alpha_0 \mu_0.$$

Thus the hypotheses imply  $0 < -\bar{v}_0 < \mu_0 \leq \frac{1}{4}$  and it follows that  $u(x, 0) \in \Phi_j$  for all  $x \in \Omega_j$   $j = l, m, r$ . Moreover  $\mu_l, \mu_r > 0$  and  $u(x, 0) \in \Phi_m$  for all  $x \in \Omega_j$  with  $j \geq 0$ , since  $|(-1)^j \alpha_j - \bar{v}_0| < \frac{1}{2}$ .

**Remark 2.** In case  $\Omega = [0, 1]$  and  $\nu$  is the Lebesgue measure on  $[0, 1]$ , we can ensure the initial data are monotonically increasing by an explicit choice of the  $\Omega_j$ , setting

$$\Omega_l = \left[0, \frac{1}{4} - \frac{\mu_0 \eta}{1 - \eta^2}\right), \quad \Omega_m = \left[\frac{1}{4}, \frac{3}{4}\right], \quad \Omega_r = \left(\frac{3}{4} + \frac{\mu_0}{1 - \eta^2}, 1\right],$$

and

$$\Omega_j = \begin{cases} \left(\frac{3}{4} + \frac{\mu_0 \eta^{j+2}}{1 - \eta^2}, \frac{3}{4} + \frac{\mu_0 \eta^j}{1 - \eta^2}\right) & \text{for even } j \geq 0, \\ \left[\frac{1}{4} - \frac{\mu_0 \eta^j}{1 - \eta^2}, \frac{1}{4} - \frac{\mu_0 \eta^{j+2}}{1 - \eta^2}\right) & \text{for odd } j \geq 1. \end{cases} \quad \triangle$$

#### 4.4 Ordering of phase transition times

In this subsection our goal is to prove part (i) of the theorem. The ideas for this part of the proof will also apply to the case of cubic nonlinearity with few changes, see Section 5 below.

To begin we set some notation. Let  $u_j(t)$  denote the value of  $u(x, t)$  for  $x \in \Omega_j$ ,  $j = l, m, r$  and  $0, 1, 2, \dots$ . Noting that  $u_j(0)$  lies in the middle phase  $\Phi_m$  for  $j = m$  and  $0, 1, 2, \dots$ , we let  $\tau_j = \tau(\Omega_j)$  denote the phase transition time for all  $x \in \Omega_j$ . For convenience we also write  $\tau_{-1} = 0$  and  $\alpha_{-1} = 1$ .

First, we claim  $\tau_m = \tau(\Omega_m) = +\infty$ . The proof is simple based on preservation of order, the invariance of the interval  $[a, b] = [-\frac{3}{2}, \frac{3}{2}]$ , and mass conservation. Preservation of order (Lemma 1) and the invariance of  $[a, b]$  implies that for all  $t \geq 0$ ,

$$a \leq u_j(t) < u_m(t) < u_k(t) \leq b \quad \text{for } j \text{ odd or } = l, \text{ and } k \text{ even or } = r. \quad (4.14)$$

Supposing  $\tau_m < \infty$ , we have either  $u_m(\tau_m) = \frac{1}{2}$  or  $-\frac{1}{2}$ . Consider the first case. Then by mass conservation and (4.14), at time  $t = \tau_m$ ,

$$0 = \bar{u} > a \left( \mu_l + \sum_{j \text{ odd}} \mu_j \right) + u(\tau_m) \left( \mu_m + \mu_r + \sum_{j \text{ even}} \mu_j \right) = -\frac{3}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4} = 0, \quad (4.15)$$

a contradiction. A similar contradiction obtains if  $u_m(\tau_m) = -\frac{1}{2}$ . This proves the claim.

Because now  $\min_{0 \leq t \leq T} |u_m(t) \pm \frac{1}{2}| > 0$  for all  $T$ , and  $\alpha_j \rightarrow 0$  as  $j \rightarrow \infty$ , by simple continuity with respect to initial data we can infer that  $\tau_j \rightarrow \infty$  as  $j \rightarrow \infty$ .

Next we claim the phase transition times  $\tau_j$  strictly increase with  $j$  for  $j \geq 0$ . The argument is based on a refinement of the estimates above. By preservation of order we have

$$u_j(t) < u_{j+2}(t) < u_m(t) < u_{k+2}(t) < u_k(t) \quad \text{for all } j \text{ odd and } k \text{ even.}$$

Hence  $\tau_{j+2} > \tau_j$  for all  $j \geq 0$ .

**Lemma 6.** For any  $t > 0$ , let

$$j_l(t) = \min\{j \text{ odd} : \tau_j \geq t\}, \quad j_r(t) = \min\{j \text{ even} : \tau_j \geq t\}.$$

Then we have the bounds

$$u_m(t) < \frac{1}{2} - \mu_{j_l(t)}, \quad u_m(t) > -\frac{1}{2} + \mu_{j_r(t)}.$$

*Proof.* The sets of points initially in the unstable phase that transition into the left and right stable phases at time  $t$  or later have measure respectively given by

$$\varepsilon_l(t) = \nu(\{x : \hat{b} \leq u(x, t) < u_m(t)\}), \quad \varepsilon_r(t) = \nu(\{x : u_m(t) < u(x, t) \leq \hat{a}\}), \quad (4.16)$$

which here satisfy

$$\varepsilon_l(t) = \hat{\mu}_l - \nu_l(t) = \sum_{\text{odd } j \geq j_l} \mu_j = \frac{\mu_{j_l}}{1 - \eta^2}, \quad \varepsilon_r(t) = \hat{\mu}_r - \nu_r(t) = \sum_{\text{even } k \geq j_r} \mu_k = \frac{\mu_{j_r}}{1 - \eta^2}.$$

Using these quantities we can obtain a bound on  $u_m(t)$  with inequalities similar to (4.15). Namely, preservation of order and invariance imply

$$0 = \bar{u} > a(\hat{\mu}_l - \varepsilon_l) - \hat{a}\varepsilon_l + u_m(t)(\hat{\mu}_m + \varepsilon_r) + \hat{a}(\hat{\mu}_r - \varepsilon_r), \quad (4.17)$$

$$0 = \bar{u} < \hat{b}(\hat{\mu}_l - \varepsilon_l) + u_m(t)(\varepsilon_l + \hat{\mu}_m) + \hat{a}\varepsilon_r + b(\hat{\mu}_r - \varepsilon_r). \quad (4.18)$$

Recalling  $a = -b = -\frac{3}{2}$  and  $\hat{b} = -\hat{a} = -\frac{1}{2}$  it follows

$$u_m(t) < \left(\frac{1}{2} + \varepsilon_r\right)^{-1} \left(\frac{1}{4} + \frac{1}{2}\varepsilon_r - \varepsilon_l\right) = \frac{1}{2} - \frac{2\varepsilon_l}{1 + 2\varepsilon_r} = \frac{1}{2} - \frac{2\mu_{j_l}}{1 - \eta^2 + 2\mu_{j_r}}, \quad (4.19)$$

$$u_m(t) > \left(\frac{1}{2} + \varepsilon_l\right)^{-1} \left(-\frac{1}{4} - \frac{1}{2}\varepsilon_l + \varepsilon_r\right) = -\frac{1}{2} + \frac{2\varepsilon_r}{1 + 2\varepsilon_l} = -\frac{1}{2} + \frac{2\mu_{j_r}}{1 - \eta^2 + 2\mu_{j_l}}. \quad (4.20)$$

Since  $2 > 1 - \eta^2 + 2\mu_j$  for all  $j$ , this finishes the proof of the lemma.  $\square$

Now we finish the proof of part (i) of the theorem, considering even and odd cases separately. Let  $k \geq 0$  be even. Then by Corollary 1, at  $t = \tau_k$  we have

$$\alpha_k e^{\tau_k} = u_k(t) - u_m(t) = \frac{1}{2} - u_m(t) < 1.$$

We claim  $\tau_{k+1} > \tau_k$ . If not, then for  $t = \tau_{k+1}$  we have  $u_{k+1}(t) = -\frac{1}{2}$ ,  $j_l(t) = k + 1$  and  $j_r(t) \leq k$ , hence by the Lemma and Corollary 1 we have

$$\mu_k \leq \mu_{j_r(t)} < u_m(\tau_{k+1}) + \frac{1}{2} = \alpha_{k+1} e^{\tau_{k+1}} \leq \alpha_{k+1} e^{\tau_k} < \frac{\alpha_{k+1}}{\alpha_k}. \quad (4.21)$$

This contradicts (4.12), proving  $\tau_{k+1} > \tau_k$ . Similarly, for  $j$  odd, at  $t = \tau_j$  we have  $\alpha_j e^{\tau_j} = u_m(t) + \frac{1}{2} < 1$ , and if  $\tau_{j+1} \leq \tau_j$  then for  $t = \tau_{j+1}$  we have  $u_{j+1}(t) = \frac{1}{2}$ ,  $j_r(t) = j + 1$  and  $j_l(t) \leq j$ , hence

$$\mu_j < \frac{1}{2} - u_m(\tau_{j+1}) = \alpha_{j+1} e^{\tau_{j+1}} < \frac{\alpha_{j+1}}{\alpha_j}.$$

Thus we conclude  $\tau_{j+1} > \tau_j$  for all  $j = 0, 1, 2, \dots$ . This finishes the proof of part (i) of the Theorem.

## 4.5 Proof of non-convergence

For times  $t$  in any interval  $(\tau_{k-1}, \tau_k)$  between transition times ( $k \geq 0$ ),  $\bar{f}(t)$  evolves according to (4.6), which can be written using (4.16) as

$$\frac{d}{dt}\bar{f}(t) = -2\varepsilon_k(\bar{f}(t) - \bar{f}_k^{\text{eq}}), \quad \varepsilon_k = \varepsilon_l + \varepsilon_r = \frac{\mu_k}{1-\eta}, \quad \bar{f}_k^{\text{eq}} = \frac{(-1)^{k-1}}{2} \frac{1-\eta}{1+\eta}, \quad (4.22)$$

because for  $k$  even we have  $j_l(t) = k+1$ ,  $j_r(t) = k$ , and for  $k$  odd we have  $j_l(t) = k$ ,  $j_r(t) = k+1$ . Then because  $\varepsilon_k > \mu_k$ ,

$$|\bar{f}(\tau_k) - \bar{f}_k^{\text{eq}}| = |\bar{f}(\tau_{k-1}) - \bar{f}_k^{\text{eq}}| e^{-2\varepsilon_k(\tau_k - \tau_{k-1})} < e^{-\mu_k(\tau_k - \tau_{k-1})}. \quad (4.23)$$

Since  $\alpha_k e^{\tau_k} = \frac{1}{2} - u_m(\tau_k)$  for  $k$  even and  $\alpha_k e^{\tau_k} = u_m(\tau_k) + \frac{1}{2}$  for  $k$  odd, by Lemma 6 we infer  $\alpha_k e^{\tau_k} > \mu_{k+1}$  and  $\alpha_{k-1} e^{\tau_{k-1}} \leq 1$  in both cases. Hence for  $k$  sufficiently large,

$$e^{\tau_k - \tau_{k-1}} > \frac{\alpha_{k-1} \mu_{k+1}}{\alpha_k} > \eta^2 \beta_{k-1}^{-1/\mu_{k-1}},$$

due to the hypothesis (4.13), and it follows

$$\mu_k(\tau_k - \tau_{k-1}) > \mu_k \log \eta^2 - \eta \log \beta_{k-1} \rightarrow \infty \quad \text{as } k \rightarrow \infty.$$

Thus  $|\bar{f}(\tau_k) - \bar{f}_k^{\text{eq}}| \rightarrow 0$  as  $k \rightarrow \infty$ , and this entails the result in part (ii) of the Theorem.

## 4.6 A gradient inequality, insufficient for convergence

It is curious to note that for the piecewise-linear nonlinearity in (4.1), a gradient inequality holds that is very similar to the one from Lemma 3 that holds generally in the case of finite range.

**Lemma 7.** *Suppose  $u \in B(\Omega)$  takes values  $u(x) \in \Phi_j$  for  $x \in \Omega_j$  for  $j = l, m, r$ , where  $\Omega_l \cup \Omega_m \cup \Omega_r = \Omega$ . Let  $s = \overline{f(u)} = \int_{\Omega} f(u) d\nu$  and let  $\phi(x) = -1 + s, -s, 1 + s$  in  $\Omega_l, \Omega_m, \Omega_r$  respectively. Then*

$$\left| E(u) - E(\phi) - s \int_{\Omega} (u - \phi) d\nu \right| \leq \frac{1}{2} \int_{\Omega} |f(u) - \overline{f(u)}|^2 d\nu.$$

*Proof.* Since  $u(x)$  and  $\phi(x)$  belong to the same phase for all  $x$ , Taylor expansion of the piecewise-quadratic primitive  $F$  of  $f$ , and the facts that  $f' = \pm 1$  and  $f(\phi(x)) = s = \overline{f(u)}$  everywhere, yield

$$F(u) - F(\phi) - s(u - \phi) = (f(\phi) - s)(u - \phi) + \frac{1}{2} f'(\phi)(u - \phi)^2 = \pm \frac{1}{2} (u - \phi)^2$$

and

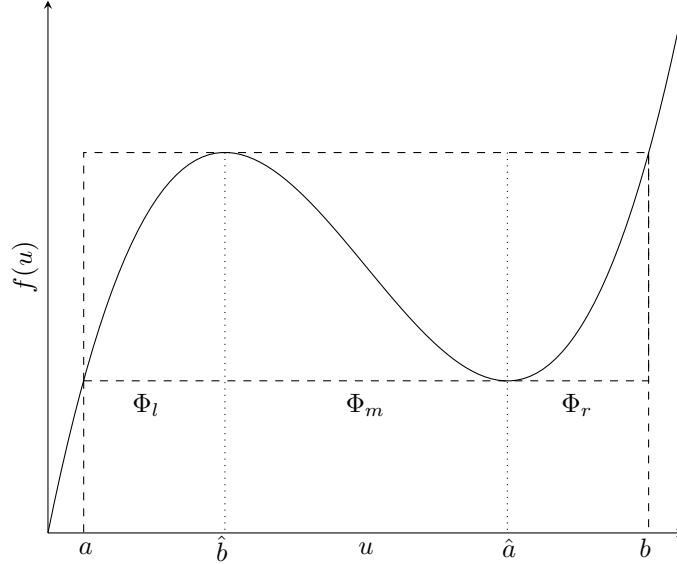
$$|f(u) - \overline{f(u)}|^2 = |f(u) - f(\phi)|^2 = |u - \phi|^2.$$

Upon integration, the Lemma follows.  $\square$

The difference with the finite range case is that the equilibrium states  $\phi$  in this Lemma are chosen with values  $\phi(x)$  in the same phase as  $u(x)$  at each point, but here this means  $\phi$  may not have the same average as  $u$ , and may never be in the  $L^2$   $\omega$ -limit set of the solution.

## 5 Non-convergence for a cubic nonlinearity

In order to demonstrate that the possibility of non-convergence of solutions of (1.1) is not due to any lack of analyticity of the nonlinear function  $f$ , we extend our analysis from the previous section to deal with the

Figure 1: Cubic  $f$  and phases.

case that  $f$  is cubic and nonmonotone, fixing

$$f(u) = u^3 - u. \quad (5.1)$$

It will be evident that our analysis can extend to other nonlinearities with  $N$ -shaped graph that admit a linear relation between distinct roots of  $f(z) = s$ , but we fix  $f$  in the form (5.1) for simplicity.

Moreover, to show that non-convergence is not restricted to solutions having countable range or limited regularity, we allow initial data of a more general type. When  $\Omega$  is the interval  $[0, 1]$  or a bounded domain in  $\mathbb{R}^d$ , for example, our assumptions will permit initial data and solutions to be  $C^\infty$  smooth.

### 5.1 Phases, equilibria, and transition times

Our solutions will take values in the phase intervals given by

$$\Phi_l = [a, \hat{b}], \quad \Phi_m = (\hat{b}, \hat{a}), \quad \Phi_r = [\hat{a}, b], \quad (5.2)$$

with

$$a = -\frac{2}{\sqrt{3}}, \quad \hat{b} = -\frac{1}{\sqrt{3}}, \quad \hat{a} = \frac{1}{\sqrt{3}}, \quad b = \frac{2}{\sqrt{3}}.$$

See Fig.1.

According to Lemma 2, the interval  $[a, b]$  is invariant and the phase intervals  $\Phi_l, \Phi_r$  are pointwise stable. Because of this, Lemma 5 holds in this context *mutatis mutandi*, and transition times  $\tau(x) \in (0, \infty]$  are well defined for states  $u(x, t)$  initially in  $\Phi_m$  to exit into either  $\Phi_l$  or  $\Phi_r$ .

For each  $s$  in the interval  $\hat{J} := (f(a), f(b))$ , the equation  $f(z) = s$  has a solution  $z_j(s) \in \Phi_j$ ,  $j = l, m, r$ . These three roots of the cubic equation  $f(z) = s$  for  $s \in \hat{J}$  satisfy the trace relation  $z_l(s) + z_m(s) + z_r(s) = 0$ . Our solutions will have asymptotic limits among the degenerate family of equilibria  $\hat{u}_s$  taking the values  $z_j(s)$  on sets  $\hat{\Omega}_j$ ,  $j = l, m, r$  of measure  $\hat{\mu}_j$ , with

$$\hat{\mu}_l = \hat{\mu}_m = \hat{\mu}_r = \frac{1}{3}. \quad (5.3)$$

By consequence of the trace relation, the equilibria  $\hat{u}_s$  all have mean  $\int_{\Omega} \hat{u}_s = 0$  independent of  $s$ .

## 5.2 Heuristics: solutions with three values

The main idea for non-convergence with the cubic nonlinearity is similar to that for the piecewise linear case: Perturbing the equilibria  $\hat{u}_s$  by moving a tiny amount of mass from the stable phases into the unstable phase can cause a slow drift by a large amount. The mean force does not appear to evolve in such a simple way as before, so we provide a different motivation.

Consider a solution taking three values  $u_j(t) \in \Phi_j$ ,  $j = l, m, r$ , on sets  $\Omega_j$  respectively having measures

$$\mu_l = \frac{1}{3} - \varepsilon_l, \quad \mu_m = \frac{1}{3} + \varepsilon_l + \varepsilon_r, \quad \mu_r = \frac{1}{3} - \varepsilon_r, \quad (5.4)$$

for small positive constants  $\varepsilon_l, \varepsilon_r$ , and assume that  $0 = \bar{u} = \mu_l u_l + \mu_m u_m + \mu_r u_r$ , which entails

$$\frac{1}{3}(u_l + u_m + u_r) = \varepsilon_r(u_r - u_m) - \varepsilon_l(u_m - u_l). \quad (5.5)$$

Note that  $u_r - u_m$  and  $u_m - u_l$  are positive, and that  $z_r(s) - z_m(s)$  and  $z_m(s) - z_l(s)$  change in opposite directions as  $s$  increases. Thus we are motivated to examine the dynamics of the ‘‘phase ratio’’

$$R = \frac{u_r - u_m}{u_m - u_l} \quad (5.6)$$

as a proxy for the level of  $\bar{f}(t)$ . We find that

$$\begin{aligned} \partial_t R &= - \left( \frac{f(u_r) - f(u_m)}{u_r - u_m} - \frac{f(u_m) - f(u_l)}{u_m - u_l} \right) R \\ &= -(u_l + u_r + u_m)(u_r - u_l)R, \end{aligned} \quad (5.7)$$

since  $u^3 - v^3 = (u^2 + uv + v^2)(u - v)$ . By (5.5) this becomes

$$\begin{aligned} \partial_t R &= 3 \left( \varepsilon_l(u_m - u_l) - \varepsilon_r(u_r - u_m) \right) (u_r - u_l)R \\ &= 3\varepsilon_r \left( \frac{\varepsilon_l}{\varepsilon_r} - R \right) (u_r - u_l)(u_r - u_m). \end{aligned} \quad (5.8)$$

Thus the ratio  $R$  is driven to approach  $\varepsilon_l/\varepsilon_r$  at a slow exponential rate. Similar to the piecewise linear case, the key to obtain non-convergence will be to ensure that solutions behave like these three-value solutions over long time intervals, with the ratio  $\varepsilon_l/\varepsilon_r$  effectively held close to constant, but forced to change substantially infinitely many times.

## 5.3 Initial data and main result

We will consider initial data structured in a way roughly similar to the piecewise linear case, but will now allow for small transition zones  $\tilde{\Omega}_j \subset \Omega_j$ . In case  $\Omega = [0, 1]$  or a bounded domain in  $\mathbb{R}^d$ , say, these transition zones permit the initial data to be chosen to smoothly interpolate between locally constant values in the rest of  $\Omega$ . The resulting solution is then a smooth function of  $x$  and  $t$ .

Similar to before, we write  $\mu_j = \nu(\Omega_j)$  for  $j = l, m, r$  and all  $j \geq 0$ , and suppose that

$$\mu_l = \frac{1}{3} - \sum_{j \text{ odd}} \mu_j, \quad \mu_m = \frac{1}{3}, \quad \mu_r = \frac{1}{3} - \sum_{j \text{ even}} \mu_j, \quad \mu_j = \mu_0 \eta^j, \quad j = 0, 1, 2, \dots, \quad (5.9)$$

where  $\eta > 0$  is sufficiently small; it will suffice to suppose  $\eta \leq \frac{1}{8}$ . Further, we take  $\tilde{\Omega}_j \subset \Omega_j$  to satisfy

$$\tilde{\mu}_j = \nu(\tilde{\Omega}_j) = \theta \mu_j, \quad \text{with } \theta \in [0, \eta^2]. \quad (5.10)$$



Like before, we will assume the initial data satisfy  $u(x, 0) = v_0(x) - \bar{v}_0$ , but where now

$$v_0(x) = \begin{cases} -1, 0, 1 & \text{in } \Omega_l, \Omega_m, \Omega_r \text{ respectively,} \\ (-1)^j \alpha_j & \text{in } \Omega_j \setminus \tilde{\Omega}_j, \quad j = 0, 1, 2, \dots \end{cases} \quad (5.11)$$

Furthermore, setting  $\alpha_{-2} = \alpha_{-1} = 1$  we require that

$$0 < \alpha_j \leq (-1)^j v_0(x) \leq \alpha_{j-2} \quad \text{in } \tilde{\Omega}_j, \quad j = 0, 1, 2, \dots \quad (5.12)$$

This means  $v_0(x)$  is between  $(-1)^j \alpha_j$  and  $(-1)^j \alpha_{j-2}$  whenever  $x \in \tilde{\Omega}_j$ , for all  $j \geq 0$ . Note that we recover piecewise constant initial data by taking either  $\theta = 0$  or  $u(x, 0) \equiv \alpha_j$  for all  $x \in \Omega_j$ . The positive constants  $\alpha_j$  must be small and decrease to zero sufficiently rapidly as described below.

Under the mild smallness conditions

$$\eta \leq \frac{1}{8}, \quad \mu_0 \leq \frac{1}{10}, \quad \alpha_0 \leq \frac{1}{2}, \quad \theta \leq \eta^2, \quad (5.13)$$

we can ensure that the initial values are in the correct phases, with  $u(x, 0) \in \Phi_j$  whenever  $x \in \Omega_j$  for  $j = l, m, r$ , and  $(-1)^j \alpha_j - \bar{v}_0 \in \Phi_m$  for all  $j \geq 0$ : Observe that

$$\bar{v}_0 = \mu_r - \mu_l + \sum_{j \geq 0} \left( (-1)^j \alpha_j \mu_j + \int_{\tilde{\Omega}_j} (v_0(x) - (-1)^j \alpha_j) d\nu(x) \right). \quad (5.14)$$

We have that  $\mu_r - \mu_l = -\frac{\mu_0}{1+\eta}$ , that  $0 < \sum_{j \geq 0} (-1)^j \alpha_j \mu_j < \alpha_0 \mu_0$ , and

$$\sum_{j \geq 0} \left| \int_{\tilde{\Omega}_j} (v_0(x) - (-1)^j \alpha_j) d\nu(x) \right| \leq \sum_{j \geq 0} \theta \mu_j \alpha_{j-2} \leq \frac{\eta^2 \mu_0}{1-\eta}.$$

Then (5.13) implies  $\alpha_0 < \frac{1}{1+\eta}$ , hence  $|\bar{v}_0| \leq \mu_0 \left( \frac{1}{1+\eta} + \frac{\eta^2}{1-\eta} \right) \leq \mu_0$ . Noting  $b-1 = \frac{2\sqrt{3}-3}{3} > \frac{1}{10} \geq \mu_0$ , it follows that  $-1 + \bar{v}_0 \in \Phi_l$ , that  $1 + \bar{v}_0 \in \Phi_r$ , and that  $\alpha_0 + |\bar{v}_0| \leq \frac{6}{10} < \hat{a}$ . This will ensure all the initial values are in the correct phases as stated.

Our main result in this section may now be stated as follows.

**Theorem 3** (Non-convergence with cubic  $f$ ). *Let the initial values  $u(x, 0) = v_0(x) - \bar{v}_0$  as described above. Assume (5.13) and assume  $(\alpha_j)_{j \geq 0}$  is a positive decreasing sequence satisfying*

$$\alpha_j \leq \mu_j \left( \frac{\alpha_{j-1}}{2} \right)^{1/\mu_j} \quad \text{for all } j \geq 1. \quad (5.15)$$

*Then: (i) The phase transition times  $\tau = \tau(\Omega_j \setminus \tilde{\Omega}_j)$  satisfy  $\tau_m = +\infty$  and  $\tau_{j+1} > \tau_j$  for all  $j \geq 0$ , with*

$$e^{\tau_j} \leq \left( \frac{\alpha_j}{2} \right)^{-1/\mu_{j+1}}. \quad (5.16)$$

*(ii) If moreover for sufficiently large  $j$  we have*

$$\alpha_j \leq \frac{1}{24} \left( \frac{\alpha_{j-1}}{2} \right)^{1/\mu_j} e^{-2\kappa_j}, \quad \text{where } \kappa_j := \frac{3}{\mu_{j+1}} \log \left( \frac{18}{\mu_{j+1}} \right), \quad (5.17)$$

*then  $u(\cdot, t)$  does not converge as  $t \rightarrow \infty$  (in any  $L^p$ ,  $1 \leq p < \infty$ ).*

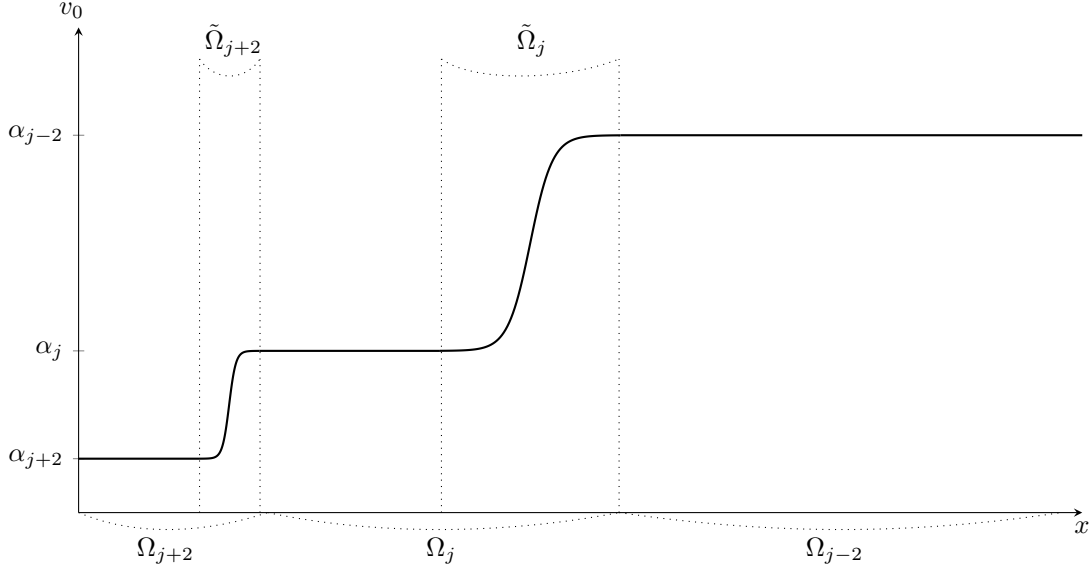


Figure 2: Schematic illustration of smooth initial data near  $\Omega_j$  for  $j$  even, with transition zones  $\tilde{\Omega}_j$  and  $\tilde{\Omega}_{j+2}$ .

**Remark 3.** Note  $\kappa_j \geq 1$  for all  $j \geq 0$ , as  $\mu_{j+1} \leq \mu_0 \leq \frac{1}{10}$ . Condition (5.17) is much stronger than (5.15) as

$$e^{-2\kappa_j} = \left(\frac{\mu_{j+1}}{18}\right)^{6/\mu_{j+1}} \ll \mu_j. \quad \triangle$$

**Remark 4** (Smooth initial data). To construct smooth initial data satisfying the assumptions of Theorem 3, first consider the case  $\Omega = [0, 1]$  with intervals  $\Omega_j$  of length  $\mu_j$  defined as in Remark 2, but with the numbers  $\frac{1}{4}$  and  $\frac{3}{4}$  replaced by  $\frac{1}{3}$  and  $\frac{2}{3}$  respectively. Fix a smooth, nondecreasing “ramp” function  $\Theta : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\Theta(x) = 0$  for  $x \leq 1 - \theta$  and  $\Theta(x) = 1$  for  $x \geq 1$ . Then set

$$v_0(x) = \begin{cases} -1, 0, 1 & \text{in } \Omega_l, \Omega_m, \Omega_r \text{ respectively,} \\ \tilde{v}_j(x) & \text{in } \Omega_j, \quad j = 0, 1, 2, \dots, \end{cases}$$

where  $\tilde{v}_j$  is a smooth function interpolating between  $(-1)^j \alpha_j$  and  $(-1)^j \alpha_{j-2}$  on  $\Omega_j$ , defined by

$$\tilde{v}_j(x) = \begin{cases} \alpha_j + (\alpha_{j-2} - \alpha_j) \Theta\left(\frac{x - c_j}{\mu_j}\right) & \text{for } j \text{ even, with } c_j = \inf \Omega_j, \\ -\alpha_j - (\alpha_{j-2} - \alpha_j) \Theta\left(\frac{c_j - x}{\mu_j}\right) & \text{for } j \text{ odd, with } c_j = \sup \Omega_j. \end{cases}$$

Then  $\tilde{v}_j(x) = (-1)^j \alpha_j$  on  $\Omega_j \setminus \tilde{\Omega}_j$ , where  $\tilde{\Omega}_j \subset \Omega_j$  is the closed interval of length  $\theta \mu_j$  at the right end of  $\Omega_j$  for  $j$  even (resp. at the left end for  $j$  odd). Clearly  $v_0$  is smooth everywhere in  $[0, 1]$  except possibly at the endpoints of  $\Omega_m = [\frac{1}{3}, \frac{2}{3}]$ . However,  $v_0$  is smooth at these endpoints also, as a consequence of the fact that in  $\tilde{\Omega}_j$  we have  $\partial_x^k v_0 = O(\mu_j^{-k} \alpha_{j-2})$  which approaches zero as  $j \rightarrow \infty$  for each fixed  $k \geq 1$ .

We can make a similar construction of smooth initial data on a domain  $\Omega = B(0, r) \subset \mathbb{R}^d$  for suitable  $r > 0$  using a radial construction. Further, such radial initial data  $u$  can be composed with any smooth volume-preserving diffeomorphism from  $\mathbb{R}^d$  to  $\mathbb{R}^d$  to produce more general smooth initial data in  $\mathbb{R}^d$  with the same distribution of values.  $\triangle$

## 5.4 Ordering of transition times

We prove part (i) of Theorem 3 in this subsection. Henceforth, for  $j = l, m, r$  we let  $u_j(t)$  denote the value of  $u(x, t)$  in  $\Omega_j$ . For  $j = 0, 1, 2, \dots$ , we let  $u_j(t)$  denote the value of  $u(x, t)$  in  $\Omega_j \setminus \tilde{\Omega}_j$ , and we define  $\tau_j = \tau(\Omega_j \setminus \tilde{\Omega}_j)$  be the corresponding phase transition time.

**Lemma 8.** *The transition time  $\tau_m = \tau(\Omega_m) = +\infty$ . That is,  $u_m(t) \in \Phi_m$  for all  $t \geq 0$ .*

*Proof.* By preservation of order we have

$$\nu(\{x : u(x, t) < u_m(t)\}) = \hat{\mu}_l = \frac{1}{3}, \quad \nu(\{x : u_m(t) < u(x, t)\}) = \hat{\mu}_r = \frac{1}{3}.$$

By invariance of  $[a, b]$ , if  $u_m(t)$  escapes  $\Phi_m$  on the right at some finite time  $t_*$ , then  $u_m(t_*) = \hat{a} = -\frac{1}{2}a$  and

$$0 = \bar{u} > a\hat{\mu}_l + u_m(t_*)(\hat{\mu}_m + \hat{\mu}_r) = 0,$$

a contradiction. Similarly, if  $u_m(t_*) = \hat{b} = -\frac{1}{2}b$ ,

$$0 = \bar{u} < u_m(t_*)(\hat{\mu}_l + \hat{\mu}_m) + b\hat{\mu}_r = 0.$$

Hence  $u_m(t) \in \Phi_m$  for all  $t \geq 0$ . □

Note now that by (5.12) and preservation of order, for any  $x \in \Omega_j$ ,  $u(x, t)$  can exit  $\Phi_m$  only at  $\hat{b}$  if  $j$  is odd, and only at  $\hat{a}$  if  $j$  is even. Since  $(-1)^j(u(x, t) - u_j(t)) \geq 0$ , the transition time

$$\tau(x) \leq \tau_j \quad \text{for any } x \in \tilde{\Omega}_j, \quad j = 0, 1, 2, \dots \quad (5.18)$$

Let  $\varepsilon_l(t), \varepsilon_r(t)$  be as defined in (4.16). Equivalently we have

$$\begin{aligned} \varepsilon_l(t) &= \nu(\{x : \hat{b} \leq u(x, 0) < u_m(0) \text{ and } \tau(x) \geq t\}), \\ \varepsilon_r(t) &= \nu(\{x : \hat{a} \geq u(x, 0) > u_m(0) \text{ and } \tau(x) \geq t\}). \end{aligned} \quad (5.19)$$

These functions are left continuous in  $t$ . Because

$$\nu(\Omega_j \setminus \tilde{\Omega}_j) = (1 - \theta)\mu_j \quad \text{and} \quad \sum_{k=0}^{\infty} \mu_{j+2k} = \frac{\mu_j}{1 - \eta^2},$$

by the assumption  $\theta \leq \eta^2$  from (5.13) we have that whenever  $t \leq \tau_j$  (so  $u_j(t) \in [\hat{b}, \hat{a}] = \bar{\Phi}_m$ ),

$$(1 - \eta^2)\mu_j \leq (1 - \theta)\mu_j \leq \begin{cases} \varepsilon_l(t) & \text{for } j \text{ odd} \\ \varepsilon_r(t) & \text{for } j \text{ even} \end{cases} \leq \frac{\mu_j}{1 - \eta^2}. \quad (5.20)$$

**Lemma 9** (Bounds on  $u_m$ ). *Let  $j_l(t)$  and  $j_r(t)$  be defined as in Lemma 6. Then*

$$\hat{a} - u_m(t) > \frac{(\hat{b} - a)\varepsilon_l}{\hat{\mu}_m + \varepsilon_r} > \mu_{j_l(t)}, \quad u_m(t) - \hat{b} > \frac{(b - \hat{a})\varepsilon_r}{\hat{\mu}_m + \varepsilon_l} > \mu_{j_r(t)}. \quad (5.21)$$

*Proof.* We will proceed as in the proof of Lemma 6. Preservation of order and invariance imply that

$$\begin{aligned} 0 &= \bar{u} > a(\hat{\mu}_l - \varepsilon_l) + \hat{b}\varepsilon_l + u_m(t)(\hat{\mu}_m + \varepsilon_r) + \hat{a}(\hat{\mu}_r - \varepsilon_r), \\ 0 &= \bar{u} < \hat{b}(\hat{\mu}_l - \varepsilon_l) + u_m(t)(\varepsilon_l + \hat{\mu}_m) + \hat{a}\varepsilon_r + b(\hat{\mu}_r - \varepsilon_r). \end{aligned}$$

The first inequality implies

$$\hat{a} - u_m(t) > \frac{\hat{a}(\hat{\mu}_m + \varepsilon_r) + a(\hat{\mu}_l - \varepsilon_l) + \hat{b}\varepsilon_l + \hat{a}(\hat{\mu}_r - \varepsilon_r)}{\hat{\mu}_m + \varepsilon_r} = \frac{(\hat{b} - a)\varepsilon_l}{\hat{\mu}_m + \varepsilon_r},$$

where we used  $a\hat{\mu}_l + \hat{a}\hat{\mu}_m + \hat{a}\hat{\mu}_r = 0$ . Similarly, using the second inequality and  $\hat{b}\hat{\mu}_l + \hat{b}\hat{\mu}_m + b\hat{\mu}_r = 0$ , we obtain

$$u_m(t) - \hat{b} > \frac{-\hat{b}(\hat{\mu}_l - \varepsilon_l) - \hat{a}\varepsilon_r - b(\hat{\mu}_r - \varepsilon_r) - \hat{b}(\hat{\mu}_m + \varepsilon_l)}{\hat{\mu}_m + \varepsilon_l} = \frac{(b - \hat{a})\varepsilon_r}{\hat{\mu}_m + \varepsilon_l}.$$

Finally, the remaining bounds in (5.21) follow by applying the first inequality in (5.20) in the numerators and the bounds  $\varepsilon_l, \varepsilon_r \leq \mu_0/(1 - \eta^2) < \frac{1}{9}$  in the denominators.  $\square$

In order to obtain the proper ordering of transition times, we need to control the expansion rate of  $|u_m(t) - u_j(t)|$  inside the unstable phase  $\Phi_m$ . For this purpose, note that

$$|f(u) - f(v)| \geq h|u - v| \quad \text{whenever} \quad \begin{cases} u, v \in \Phi_m \text{ with } (\hat{a} - v) \wedge (v - \hat{b}) \geq h, \text{ or} \\ u, v \in \Phi_r \text{ with } v - \hat{a} \geq h, \text{ or} \\ u, v \in \Phi_l \text{ with } \hat{b} - v \geq h. \end{cases} \quad (5.22)$$

To see this, suppose  $u, v \in \Phi_m$  and  $h \leq \hat{a} - v \leq v - \hat{b}$ . Then necessarily  $h \leq \hat{a}$ , as  $(\hat{a} - v) \wedge (v - \hat{b}) \leq \hat{a}$ . Explicitly computing, since  $uv$  and  $v^2$  are each less than  $\hat{a}(\hat{a} - h)$ , and  $3\hat{a}^2 = 1$ , we have

$$\left| \frac{f(u) - f(v)}{u - v} \right| = 1 - u^2 - uv - v^2 \geq 1 - \hat{a}^2 - 2\hat{a}(\hat{a} - h) = 2\hat{a}h \geq h.$$

By symmetry, we can deduce the same inequality when  $(v - \hat{b}) \leq (\hat{a} - v)$ . Similarly, when  $u, v \in \Phi_r$  and  $v - \hat{a} \geq h$ ,

$$\frac{f(u) - f(v)}{u - v} = u^2 + uv + v^2 - 1 \geq (\hat{a} + h)^2 + \hat{a}(\hat{a} + h) + \hat{a}^2 - 1 = \sqrt{3}h + h^2 \geq h$$

and the case when  $u, v \in \Phi_l$  can be verified by analogous calculations.

*Proof of Theorem 3 part (i).* By preservation of order we have  $\tau_{j+2} > \tau_j$  for all  $j \geq 0$ . Supposing that  $\tau_{j+1} \leq \tau_j$  for some  $j$ , we may take  $j$  minimal. Then for  $0 \leq t \leq \tau_{j+1} \leq \tau_j$ , both  $j_l(t), j_r(t) \leq j + 1$ , so the bounds in Lemma 9 apply to yield

$$(\hat{a} - u_m) \wedge (u_m - \hat{b}) > \mu_{j+1}.$$

In case  $j$  is odd, we infer that for all  $t \leq \tau_{j+1}$ ,

$$\partial_t(u_m - u_j) = f(u_j) - f(u_m) \geq \mu_{j+1}(u_m - u_j),$$

whence at  $t = \tau_{j+1}$ ,

$$\hat{a} - \hat{b} > |u_m - u_j| \geq \alpha_j e^{\mu_{j+1}\tau_{j+1}}. \quad (5.23)$$

In case  $j$  is even, the same inequality follows in similar fashion by computing  $\partial_t(u_j - u_m)$ . Now in either case, since  $|f'| \leq 1$  in  $\Phi_m$  and  $u_{j+1}(\tau_{j+1}) = \hat{a}$  or  $\hat{b}$ , use of Gronwall's inequality yields

$$\mu_{j+1} \leq |u_m(\tau_{j+1}) - u_{j+1}(\tau_{j+1})| \leq \alpha_{j+1} e^{\tau_{j+1}}. \quad (5.24)$$

The inequalities (5.23)–(5.24) imply

$$\alpha_{j+1} \geq \mu_{j+1} e^{-\tau_{j+1}} > \mu_{j+1} \left( \frac{\alpha_j}{\hat{a} - \hat{b}} \right)^{1/\mu_{j+1}},$$

which contradicts the assumption (5.15) since  $\hat{a} - \hat{b} < 2$ . Hence  $\tau_{j+1} > \tau_j$  for all  $j$ . The bound (5.16) follows because in (5.23) we can now replace  $\tau_{j+1}$  by  $\tau_j$ .  $\square$

## 5.5 Analysis of non-convergence

By the result of part (i) of Theorem 3, we have  $\tau_{j+1} > \tau_j$  for all  $j$ . Then it follows that  $|j_l(t) - j_r(t)| = 1$  for all  $t$ , and whenever  $t \leq \tau_j$ , necessarily both  $j_l(t), j_r(t) \leq j + 1$ . Thus by Lemma 9,

$$\hat{b} + \mu_{j+1} < u_m(t) < \hat{a} - \mu_{j+1} \quad \text{whenever } t \leq \tau_j. \quad (5.25)$$

In this section our goal is to prove part (ii) of Theorem 3. The proof is more involved than in the piecewise-linear case. We proceed by examining the evolution of the phase ratio, then establish estimates involving exponential contraction in the stable phases, and finish by an argument by contradiction.

### 5.5.1 Evolution of the phase ratio

Our strategy to obtain non-convergence is to study the evolution of the phase ratio  $R$  defined exactly as in subsection 5.2, by

$$R = \frac{u_r - u_m}{u_m - u_l}. \quad (5.26)$$

The evolution equation (5.7) continues to hold in the present context. In order to obtain an analog of (5.8), we need to express the sum  $u_l + u_m + u_r$  differently using conservation of mass. For this purpose we alter the definition of  $\varepsilon_l(t), \varepsilon_r(t)$  to always include whole pieces, as follows:

For any  $t \geq 0$ , let  $j_m(t)$  indicate the index of the *next* value  $u_j$  to change phase (by leaving  $\bar{\Phi}_m = [\hat{b}, \hat{a}]$ ), so

$$j_m(t) = \min\{j : \tau_j \geq t\} = j_l(t) \wedge j_r(t).$$

Then  $t \in (\tau_{j_m-1}, \tau_{j_m}]$ . This means that if  $j = j_m$  or  $j_m + 1$ , then  $u_j(t) \in \bar{\Phi}_m$  but  $u_{j+2}(t) \notin \bar{\Phi}_m$ , so  $u(\bar{\Omega}_j, t)$  may be split between phases. For all *other*  $j$ , the sets  $u(\Omega_j, t)$  are entirely in one phase—the unstable phase  $\bar{\Phi}_m$  if  $t < \tau_j$ , and one of the stable phases  $\Phi_l$  or  $\Phi_r$  if  $t > \tau_j$ , for  $j$  odd or even respectively. Accounting only for those  $j$  for which  $u_j$  lies in (the closure of) the unstable phase, define

$$\hat{\varepsilon}_l(t) = \nu \left( \bigcup \left\{ \Omega_j : \hat{b} \leq u_j(t) < u_m(t) \right\} \right) = \sum_{\text{odd } j \geq j_m(t)} \mu_j, \quad (5.27)$$

$$\hat{\varepsilon}_r(t) = \nu \left( \bigcup \left\{ \Omega_j : u_m(t) < u_j(t) \leq \hat{a} \right\} \right) = \sum_{\text{even } j \geq j_m(t)} \mu_j. \quad (5.28)$$

**Remark 5.** The relation with  $\varepsilon_l, \varepsilon_r$  is as follows. Let us denote the part of  $\Omega_j$  outside  $\bar{\Phi}_m$  by

$$\nu_j(t) = \nu(\{x \in \Omega_j : u(x, t) < \hat{b} \text{ or } u(x, t) > \hat{a}\}),$$

and note  $0 \leq \nu_j(t) \leq \theta \mu_j$  for  $j \geq j_m(t)$  because  $u_j \in [\hat{b}, \hat{a}]$ . Then

$$\begin{aligned} j = j_m(t) \text{ odd} &\implies \varepsilon_l = \hat{\varepsilon}_l - \nu_{j_m}, & \varepsilon_r = \hat{\varepsilon}_r - \nu_{j_m+1}, \\ j = j_m(t) \text{ even} &\implies \varepsilon_r = \hat{\varepsilon}_r - \nu_{j_m}, & \varepsilon_l = \hat{\varepsilon}_l - \nu_{j_m+1}. \end{aligned} \quad (5.29)$$

$\triangle$

Recall that

$$\bar{u} = \mu_l u_l + \mu_m u_m + \mu_r u_r + \sum_{j \geq 0} \left( (1 - \theta) \mu_j u_j + \int_{\tilde{\Omega}_j} u \right). \quad (5.30)$$

Since  $\nu(\tilde{\Omega}_j) = \theta \mu_j$ , for any constant  $v \in \mathbb{R}$  we can write

$$(1 - \theta) \mu_j u_j + \int_{\tilde{\Omega}_j} u = (1 - \theta) \mu_j (u_j - v) + \int_{\tilde{\Omega}_j} (u - v) d\nu + \mu_j v.$$

In view of (5.9) and (5.27)–(5.28) then, we find that

$$\bar{u} = u_l(\hat{\mu}_l - \hat{\varepsilon}_l) + u_m(\hat{\mu}_m + \hat{\varepsilon}_l + \hat{\varepsilon}_r) + u_r(\hat{\mu}_r - \hat{\varepsilon}_r) + H(t), \quad (5.31)$$

where  $H(t) = H_l(t) + H_m(t) + H_r(t)$ , with

$$H_l(t) = \sum_{\text{odd } j < j_m(t)} \left( (1 - \theta) \mu_j (u_j - u_l) + \int_{\tilde{\Omega}_j} (u(x, t) - u_l) d\nu(x) \right), \quad (5.32)$$

$$H_m(t) = \sum_{j \geq j_m(t)} \left( (1 - \theta) \mu_j (u_j - u_m) + \int_{\tilde{\Omega}_j} (u(x, t) - u_m) d\nu(x) \right), \quad (5.33)$$

$$H_r(t) = \sum_{\text{even } j < j_m(t)} \left( (1 - \theta) \mu_j (u_j - u_r) + \int_{\tilde{\Omega}_j} (u(x, t) - u_r) d\nu(x) \right). \quad (5.34)$$

Since  $\bar{u} = 0$  and by (5.3) we obtain our desired relation,

$$\frac{1}{3}(u_l + u_m + u_r) = \hat{\varepsilon}_r(u_r - u_m) - \hat{\varepsilon}_l(u_m - u_l) - H(t). \quad (5.35)$$

Now, by using (5.35) in the evolution equation (5.7) for the phase ratio  $R$ , we infer that

$$\partial_t R = 3 \left( \hat{\varepsilon}_l(u_m - u_l) - \hat{\varepsilon}_r(u_r - u_m) + H(t) \right) (u_r - u_l) R, \quad (5.36)$$

which can be compared to equation (5.8) for solutions with three values. As this comparison suggests, our aim is show that  $H(t)$  is tiny enough over large enough time intervals that non-convergence follows.

### 5.5.2 Estimates in the stable phases

**Lemma 10** (Estimates on  $\bar{f}, u_l, u_r$ ). *Let  $h_j = \frac{1}{3}\mu_{j+1}$  for all  $j$ . Then whenever  $t \leq \tau_j$  we have*

$$u_l(t) < \hat{b} - h_j, \quad u_r(t) > \hat{a} + h_j, \quad f(\hat{a} + h_j) < \bar{f}(t) < f(\hat{b} - h_j). \quad (5.37)$$

*Proof.* We prove the bounds on  $\bar{f}(t)$  first. Observe

$$\begin{aligned} \bar{f}(t) &> f(a)(\hat{\mu}_l - \varepsilon_l) + f(u_m)(\hat{\mu}_l + \varepsilon_l) + f(\hat{a})\hat{\mu}_r \\ &= f(\hat{a}) + (f(u_m) - f(\hat{a}))(\hat{\mu}_l + \varepsilon_l) \\ &> f(\hat{a}) + (f(\hat{a} - \mu_{j+1}) - f(\hat{a}))(\hat{\mu}_r + \mu_{j+1}), \end{aligned}$$

$$\begin{aligned}
\bar{f}(t) &< f(\hat{b})\hat{\mu}_l + f(u_m)(\hat{\mu}_m + \varepsilon_r) + f(b)(\hat{\mu}_r - \varepsilon_r) \\
&= f(\hat{b}) + (f(u_m) - f(\hat{b}))(\hat{\mu}_m + \varepsilon_r) \\
&\leq f(\hat{b}) + (f(\hat{b} + \mu_{j+1}) - f(\hat{b}))(\hat{\mu}_m + \mu_{j+1}).
\end{aligned}$$

Since  $0 = f'(\hat{b}) = f'(\hat{a})$ , Taylor expansion gives, for  $0 < h < 1/\sqrt{3} = \hat{a} = -\hat{b}$ ,

$$\begin{aligned}
f(\hat{a} - h) - f(\hat{a}) &= 3\hat{a}h^2 - h^3 > 2\hat{a}h^2 > h^2, \\
f(\hat{b} + h) - f(\hat{b}) &= 3\hat{b}h^2 + h^3 < -2\hat{a}h^2 < -h^2,
\end{aligned}$$

Hence

$$\bar{f}(t) > f(\hat{a}) + \mu_{j+1}^2(\frac{1}{3} + \mu_{j+1}) > f(\hat{a} + \frac{1}{3}\mu_{j+1}), \quad (5.38)$$

$$\bar{f}(t) < f(\hat{b}) - \mu_{j+1}^2(\frac{1}{3} + \mu_{j+1}) < f(\hat{b} - \frac{1}{3}\mu_{j+1}). \quad (5.39)$$

This proves the claimed bounds on  $\bar{f}$ .

Note that initially  $u_l(0) = -1 - \bar{v}_0 < \hat{b} - h_0$  and  $u_r(0) = 1 - \hat{v}_0 > \hat{a} + h_0$ , since  $|\bar{v}_0| \leq \mu_0 \leq \frac{1}{10}$  and  $h_0 \leq \frac{1}{240}$ . Then the claimed bounds on  $u_l$  and  $u_r$  follow from the bounds on  $\bar{f}$ , the evolution equation (1.1), and the monotonicity of  $f$  on the invariant intervals  $[a, \hat{b}]$  and  $[\hat{a}, b]$ .  $\square$

**Lemma 11.** *Whenever  $\tau_j < t < \tau_k$  we have:*

$$u_r - u_j \leq e^{-h_k(t-\tau_j)} \quad \text{for } j, k \text{ even}, \quad u_j - u_l \leq e^{-h_k(t-\tau_j)} \quad \text{for } j, k \text{ odd}.$$

*Proof.* Suppose  $\tau_j < t < \tau_k$ . In case  $j, k$  are both even, we know  $u_r \geq \hat{a} + h_k$  by Lemma 10, so

$$\frac{f(u_r) - f(u_j)}{u_r - u_j} \geq \frac{f(\hat{a} + h_k) - f(\hat{a})}{h_k} = 3\hat{a}h_k + h_k^2 > h_k.$$

It follows

$$\partial_t(u_r - u_j) = -(u_r - u_j) \frac{f(u_r) - f(u_j)}{u_r - u_j} < -h_k(u_r - u_j),$$

hence

$$u_r(t) - u_j(t) \leq (u_r(\tau_j) - u_j(\tau_j))e^{-h_k(t-\tau_j)} \leq (b - \hat{a})e^{-h_k(t-\tau_j)} < e^{-h_k(t-\tau_j)}. \quad (5.40)$$

In case  $j, k$  are odd, we know  $u_l \leq \hat{b} - h_k$ , hence

$$\frac{f(u_j) - f(u_l)}{u_j - u_l} \geq \frac{f(\hat{b}) - f(\hat{b} - h_k)}{h_k} = -3\hat{b}h_k + h_k^2 > h_k.$$

It follows

$$\partial_t(u_j - u_l) = -(u_j - u_l) \frac{f(u_j) - f(u_l)}{u_j - u_l} < -h_k(u_j - u_l),$$

thus

$$u_j(t) - u_l(t) \leq (u_j(\tau_j) - u_l(\tau_j))e^{-h_k(t-\tau_j)} \leq (\hat{b} - a)e^{-h_k(t-\tau_j)} < e^{-h_k(t-\tau_j)}. \quad (5.41)$$

$\square$

**Lemma 12** (Bounds for  $H(t)$ ). *For  $0 \leq \tau_{k-1} < t < \tau_k$  we have*

$$|H_m(t)| \leq 2\mu_k(\alpha_k e^t + \theta), \quad |H_l(t) + H_r(t)| \leq 2\mu_0 e^{-h_k(t-\tau_{k-1})}.$$

*Proof.* For the given range of  $t$  we have  $j_m(t) = k$ . To prove the bound on  $H_m(t)$  defined by (5.33), we use

the bound  $|f'| \leq 1$  in  $\Phi_m$  to infer  $|u_m - u_j| \leq \alpha_j e^t$  for all  $j \geq k$ , and the bound  $|u - u_m(t)| < b - \hat{b} = \sqrt{3}$  in  $\tilde{\Omega}_j$ . Then since  $\eta \leq \frac{1}{8}$  we infer

$$|H_m(t)| \leq \sum_{j \geq k} ((1 - \theta)\mu_j \alpha_j e^t + \theta \mu_j (b - \hat{b})) \leq \frac{\mu_k}{1 - \eta} (\alpha_k e^t + \sqrt{3}\theta) \leq 2\mu_k (\alpha_k e^t + \theta). \quad (5.42)$$

By Lemma 11 we find

$$|H_l(t) + H_r(t)| \leq \sum_{j < k} \mu_j e^{-h_k(t - \tau_j)} \leq 2\mu_0 e^{-h_k(t - \tau_{k-1})}. \quad \square$$

Note that  $H_m(t)$  can be kept small for any specified time by forcing the  $\alpha_k$  to decay faster, whereas the exponential contraction in the stable phase will force  $H_l(t), H_r(t)$  to be small for  $t - \tau_{k-1}$  large enough. We will see that smallness of  $H(t)$  implies lower bounds on the drift of the phase ratio  $R$  in (5.36), leading to non-convergence.

### 5.5.3 Proof of non-convergence

In this subsection we complete the proof of part (ii) of Theorem 3. For use below, recall  $h_k = \frac{1}{3}\mu_{k+1}$ , and note that  $\kappa_k$  satisfies

$$\kappa_k = \frac{1}{h_k} \log \left( \frac{6}{h_k} \right), \quad h_k e^{\mu_k \kappa_k} = 6, \quad 2\mu_0 e^{-h_k \kappa_k} \leq \frac{1}{9} \mu_{k+1}. \quad (5.43)$$

*Proof of Theorem 3 part (ii).* 1. We argue by contradiction. Supposing that  $\lim_{t \rightarrow \infty} u(\cdot, t)$  exists, there is some  $\hat{s} \in [f(a), f(b)]$  such that as  $t \rightarrow \infty$ ,

$$\bar{f}(t) \rightarrow \hat{s} \quad \text{and} \quad u_j(t) \rightarrow z_j(\hat{s}) \quad \text{for } j = l, m, r.$$

We will consider the cases  $\hat{s} \leq 0$  and  $\hat{s} > 0$  separately. First consider the case  $\hat{s} \leq 0$ . Then necessarily  $z_l(\hat{s}) \leq -1$  and  $0 \leq z_m(\hat{s}) \leq z_r(\hat{s}) \leq 1$ , and as  $t \rightarrow \infty$  we have

$$R(t) \rightarrow \hat{R} := \frac{z_r(\hat{s}) - z_m(\hat{s})}{z_m(\hat{s}) - z_l(\hat{s})} \in [0, 1]. \quad (5.44)$$

In particular, if  $T$  is large enough, then for all  $t > T$  we have

$$u_l(t) \leq -1 + \frac{1}{2}\eta^2, \quad u_m(t) \geq -\frac{1}{2}\eta^2, \quad \text{and} \quad R(t) < 2. \quad (5.45)$$

We will contradict the last conclusion by showing that for any sufficiently large odd  $k$ , necessarily  $R(t_k) \geq 12$  for some  $t_k \in (\tau_{k-1}, \tau_k)$ .

2. We claim that for any sufficiently large odd  $k$  with  $\tau_{k-1} > T$ ,

$$\tau_k > \tau_{k-1} + 2\kappa_k. \quad (5.46)$$

Indeed, since  $k$  is odd and  $\eta \leq \frac{1}{8}$ ,

$$\alpha_k e^{\tau_k} = u_m(\tau_k) - u_k(\tau_k) = u_m(\tau_k) - \hat{b} > -\frac{1}{2}\eta^2 + \frac{1}{\sqrt{3}} > \frac{1}{4}.$$

But in light of (5.16) and the condition (5.17), we get that for all  $t \leq \tau_{k-1} + 2\kappa_k$ ,

$$\alpha_k e^t \leq \alpha_k \left( \frac{2}{\alpha_{k-1}} \right)^{1/\mu_k} e^{2\kappa_k} \leq \frac{1}{24}. \quad (5.47)$$



Thus (5.46) holds.

3. For  $k$  odd and  $T < \tau_{k-1} < t < \tau_k$ , we have

$$j_m(t) = k, \quad \hat{\varepsilon}_l(t) = \frac{\mu_k}{1 - \eta^2}, \quad \hat{\varepsilon}_r(t) = \eta \hat{\varepsilon}_l,$$

and equation (5.36) takes the form

$$\partial_t R = 3\mu_k \left( (1 - \eta R) \frac{u_m - u_l}{1 - \eta^2} + \frac{H(t)}{\mu_k} \right) (u_r - u_l) R. \quad (5.48)$$

Now we can deduce from Lemma 12, (5.13), (5.47) and (5.43) that for  $\tau_{k-1} + \kappa_k < t < \tau_{k-1} + 2\kappa_k$ ,

$$\frac{|H_m(t)|}{\mu_k} \leq 2(\alpha_k e^t + \theta) \leq \frac{1}{6}, \quad \frac{|H_l(t) + H_r(t)|}{\mu_k} \leq \frac{2\mu_0 e^{-h_k \kappa_k}}{\mu_k} \leq \frac{1}{9}. \quad (5.49)$$

It follows from (5.45), (5.48), the fact  $u_r - u_l > \hat{a} - \hat{b} > 1$  and  $\eta R \leq \frac{1}{4}$  that for all  $t$  in this range,

$$\partial_t R \geq \mu_k R. \quad (5.50)$$

Using (5.25) and Lemma 10 we can ensure  $u_r - u_m \geq 4h_k$ , hence for  $t = \tau_{k-1} + \kappa_k$ ,

$$R(\tau_{k-1} + \kappa_k) = \frac{u_r - u_m}{u_m - u_l} \geq \frac{4h_k}{\hat{a} - a} \geq 2h_k.$$

Using (5.43), we infer that at time  $t_k := \tau_{k-1} + 2\kappa_k$ ,

$$R(t_k) \geq 2h_k e^{\mu_k \kappa_k} \geq 12. \quad (5.51)$$

This contradicts  $R(t) < 2$  for all  $t > T$ , and concludes the analysis in the case  $\hat{s} \leq 0$ .

4. The treatment in the case  $\hat{s} > 0$  is broadly similar. In this case, we can say that

$$R(t)^{-1} \rightarrow \tilde{R} := \frac{z_m(\hat{s}) - z_l(\hat{s})}{z_r(\hat{s}) - z_m(\hat{s})} \in [0, 1], \quad (5.52)$$

and find  $T$  large enough so that for all  $t > T$ ,

$$u_r(t) > 1 - \frac{1}{2}\eta^2, \quad u_m(t) \leq \frac{1}{2}\eta^2, \quad \text{and} \quad R(t)^{-1} < 2. \quad (5.53)$$

Now taking  $k$  even and sufficiently large, such that  $\tau_{k-1} > T$ ,

$$\alpha_k e_k^\tau = u_k(\tau_k) - u_m(\tau_k) > \hat{a} - \frac{1}{2}\eta^2 > \frac{1}{4},$$

while (5.47), and hence (5.46), follow as before. For  $k$  even and  $T < \tau_{k-1} < t < \tau_k$ ,

$$j_m(t) = k, \quad \hat{\varepsilon}_l(t) = \eta \hat{\varepsilon}_r, \quad \hat{\varepsilon}_r(t) = \frac{\mu_k}{1 - \eta^2},$$

and we find (5.36) equivalent to

$$\partial_t R^{-1} = 3\mu_k \left( (1 - \eta R^{-1}) \frac{u_r - u_m}{1 - \eta^2} - \frac{H(t)}{\mu_k} \right) (u_r - u_l) R^{-1} \quad (5.54)$$

As before, for  $\tau_{k-1} + \kappa_k < t < \tau_{k-1} + 2\kappa_k$ , the bounds (5.49) hold, and we can infer

$$\partial_t R^{-1} \geq \mu_k R^{-1} \quad (5.55)$$

for all  $t$  in this interval. At the time  $t = \tau_{k-1} + \kappa_k$  we have

$$R(\tau_{k-1} + \kappa_k)^{-1} = \frac{u_m - u_l}{u_r - u_m} \geq \frac{4h_k}{b - \hat{b}} \geq 2h_k, \quad (5.56)$$

and infer  $R(\tau_{k-1} + 2\kappa_k)^{-1} \geq 12$  like before, obtaining a contradiction. This finishes the proof.  $\square$

## 5.6 Unstable nature of non-convergence

Now we present a proof of Proposition 1.

*Proof of Proposition 1.* 1. Suppose  $u$  does not converge in  $L^2(\Omega)$  as  $t \rightarrow \infty$ . As  $f = u^3 - u$  is not constant on any open interval, we deduce  $\bar{f}(t)$  does not converge either; see [24, Lemma 3.4]. Hence we may choose an open interval  $J$  such that

$$\liminf \bar{f}(t) < \inf J < \sup J < \limsup \bar{f}(t), \quad (5.57)$$

and such that  $\bar{J}$  contains only regular values of  $f$ , omitting both critical values  $f(a)$  and  $f(b)$ . Moreover, due to the fact from (1.8) that  $\|\partial_t u\|_{L^2} \rightarrow 0$  as  $t \rightarrow \infty$ , we infer by differentiating (1.3) that the Lipschitz function  $\bar{f}(t)$  has derivative  $\partial_t \bar{f}(t) \rightarrow 0$  as  $t \rightarrow \infty$  in its set of differentiability. If we let  $J_m$  denote the “middle third” of  $J$ , then it follows there exist sequences  $T_k \rightarrow \infty$  and  $\tau_k \rightarrow \infty$  such that

$$\bar{f}(t) \in J_m \quad \text{for all } t \in I_k := [T_k, T_k + \tau_k] \text{ and all } k \in \mathbb{N}. \quad (5.58)$$

Letting  $\delta = |J_m|$  denote the length of  $J_m$ , we have  $|s - \hat{s}| \geq \delta$  whenever  $s \in J_m$  and  $\hat{s} \notin J$ .

2. We first dispose of the possibility that  $\sup J < f(a)$  or  $f(b) < \inf J$ . In this case  $f$  has a unique local inverse  $z(s)$  defined for  $s \in J$  satisfying  $f(z(s)) = s$ , and  $f$  is strictly monotone increasing on  $z(J)$ . Thus the interval  $z(J)$  is pointwise stable during each interval  $I_k$  (cf. Lemma 2), for when  $\bar{f}(t) \in J_m$ , we have

$$-f(v) + \bar{f}(t) \leq -\delta < 0 \quad \text{if } v \geq \sup z(J), \quad \text{and} \quad -f(v) + \bar{f}(t) \geq \delta > 0 \quad \text{if } v \leq \inf z(J). \quad (5.59)$$

Moreover, whenever  $k$  is so large that  $\tau_k = |I_k| \geq \hat{\tau} := 2M/\delta$  where  $M := \sup |u| + 1$ , then

$$u(x, T_k + \hat{\tau}) \in z(J) \quad \text{for all } x \in \Omega. \quad (5.60)$$

The reason is that  $|u(x, T_k)| \leq M$ , and the quantity  $u(x, t)$ , if not initially in  $z(J)$ , must monotonically move toward it with speed exceeding  $\delta$ , by (5.59). Since  $\hat{\tau}\delta = 2M$ ,  $u(x, t)$  must enter  $z(J)$  before time  $T_k + \hat{\tau}$ , and cannot escape as long as  $t \in I_k$ .

But now, since (5.60) holds, Lemma 2 implies the interval  $z(J)$  becomes positively invariant and therefore  $u(x, t) \in z(J)$  for all large  $t$ . This forces  $\bar{f}(t) \in J$  ever after, contradicting the choice of  $J$  in (5.57). By consequence we must have

$$J \subset \hat{J} = (f(a), f(b)).$$

In particular,  $a < f^{-1}(J) < b$ , i.e.,  $a < v < b$  whenever  $f(v) \in J$ .

3. By the invariance arguments of Lemma 2, the phase intervals  $[-M, \hat{b}]$  and  $[\hat{a}, M]$  are pointwise stable during the intervals  $I_k$  when  $\bar{f}(t) \in J_m$ . Supposing  $k$  is so large that  $\tau_k > \hat{\tau}$ , for a similar reason as in step 2 it follows that if  $u(x, T_k) \in [-M, \hat{b}]$  then  $u(x, T_k + \hat{\tau}) \in z_l(J)$ , and if  $u(x, T_k) \in [\hat{a}, M]$  then  $u(x, T_k + \hat{\tau}) \in z_r(J)$ . In particular this implies that there exists some  $T_* = T_k + \hat{\tau}$  such that  $u(x, T_*) \in [a, b]$  for all  $x$ .

For  $t > T_*$ ,  $[a, b]$  is positively invariant and the phase intervals  $\Phi_l$  and  $\Phi_r$  are pointwise stable.

Then the sets defined for  $t > T_*$  by

$$\Omega_j(t) = \{x : u(x, t) \in \Phi_j\} \quad \text{for } j = l, m, r,$$

are monotonic for  $t > T_*$ . Indeed, the set  $\Omega_m(t)$  decreases in time whereas the sets  $\Omega_l(t)$  and  $\Omega_r(t)$  increase in time. Thus, for each  $j = l, m, r$ , the quantities

$$\Omega_j^\infty = \lim_{t \uparrow \infty} \Omega_j(t) \quad \text{and} \quad \mu_j^\infty = \nu(\Omega_j^\infty)$$

exist. Let  $\check{\Omega}(t) = \Omega_m(t) \setminus \Omega_m^\infty$  denote the “bad set” where  $u(x, t)$  is not in the phase it eventually enters.

4. We next claim that

$$\sum_j \mu_j^\infty z_j(s) = \bar{u} \quad \text{for all } s \in J. \quad (5.61)$$

Fix  $s \in J$ , and define  $\phi(x) = z_j(s)$  for  $x \in \Omega_j^\infty$ ,  $j = l, m, r$ , so that  $f(\phi(x)) = s$  for all  $x$  and  $\bar{\phi} = \sum_j \mu_j^\infty z_j(s)$ . Note that for some  $\beta > 0$  we have

$$|v - z_j(s)| \leq \beta |f(v) - s| \quad \text{for all } v \in \Phi_j, j = l, m, r,$$

hence for  $t > T_*$ ,

$$|u(x, t) - \phi(x)| \leq \beta |f(u(x, t)) - s| \quad \text{for all } x \notin \check{\Omega}(t).$$

Taking  $t$  along any sequence  $t_k \rightarrow \infty$  such that  $\bar{f}(t_k) = s$  and  $t_k > T_*$ , we deduce that

$$|\bar{u} - \bar{\phi}|^2 \leq \int_{\Omega} |u(x, t) - \hat{u}(x)|^2 d\nu \leq \beta^2 \int_{\Omega \setminus \check{\Omega}(t)} |f(u(x, t)) - \bar{f}(t_k)|^2 d\nu + 2M\nu(\check{\Omega}(t)) \rightarrow 0$$

as  $k \rightarrow \infty$ , since  $\int_{\Omega} |f(u) - \bar{f}|^2 d\nu = \|\partial_t u\|_{L^2}^2 \rightarrow 0$ . Hence  $\bar{u} = \bar{\phi}$ , and this proves (5.61).

Property (5.61) implies that  $\mu_j^\infty = \frac{1}{3}$  for each  $j = l, m, r$ , by Proposition 12 of [3], which concerns relations between roots of cubic-like analytic functions. Then it follows  $\bar{u} = 0$ , since the  $z_j$  are the three roots of the cubic  $u^3 - u - s$ .

5. For the remainder of the proof, fix some  $\hat{x} \in \Omega_m^\infty$  and let  $c = u(\hat{x}, 0)$ . We claim that

$$\Omega_m^\infty = \hat{E} \quad \text{where} \quad \hat{E} := \{x \in \Omega : u(x, 0) = c\}. \quad (5.62)$$

Here  $\hat{E}$  denotes the level set where  $u(x, 0) = c = u(\hat{x}, 0)$ . We can then infer that  $\nu(\hat{E}) = \mu_m^\infty = \frac{1}{3}$ , and this will almost finish the proof.

By (1.1),  $u(x, t) = u(\hat{x}, t)$  for all  $x \in \hat{E}$  and all  $t \geq 0$ , so  $\hat{E} \subset \Omega_m^\infty$ . Suppose then that some  $x$  exists in  $\Omega_m^\infty \setminus \hat{E}$ . Then  $u(x, 0) \neq u(\hat{x}, 0)$ , yet both  $u(x, t)$  and  $u(\hat{x}, t)$  lie in  $\Phi_m$  for all  $t \geq T_*$ . It remains to show this leads to a contradiction.

Because  $f$  is decreasing on  $\Phi_m$  and

$$\partial_t(u(x, t) - u(\hat{x}, t)) = -f(u(x, t)) + f(u(\hat{x}, t)),$$

the difference  $h(t) = |u(x, t) - u(\hat{x}, t)|$  is increasing for all  $t > T_*$ . Moreover,  $\partial_t h(t) \geq \eta$  for some  $\eta > 0$ , such that  $f(v) - f(w) \geq \eta$  whenever  $v, w \in \bar{\Phi}_m$  with  $v + h(T_*) < w$ . This forces  $h(t) > \hat{a} - \hat{b}$  after time  $T_* + (\hat{a} - \hat{b})/\eta$ , which contradicts that both  $u(x, t)$  and  $u(\hat{x}, t)$  lie in  $\Phi_m = (\hat{b}, \hat{a})$ .

Hence  $\Omega_m^\infty = \hat{E}$ . For each point  $x \in E_+ := \{x \in \Omega : u(x, 0) > c\}$ , preservation of order and the argument just made imply that  $u(x, t) \in \Phi_r$  for  $t$  large enough, and for each point  $x \in E_- := \{x \in \Omega : u(x, 0) < c\}$ , necessarily  $u(x, t) \in \Phi_l$  for  $t$  large enough. Then it follows  $E_+ = \Omega_r^\infty$  and  $E_- = \Omega_l^\infty$ , whence  $\nu(E_+) = \mu_r^\infty = \frac{1}{3}$  and  $\nu(E_-) = \mu_l^\infty = \frac{1}{3}$ . This completes the proof.  $\square$

## 6 Sensitivity of convergence rates

In this section we comment on the possibility of curiously high sensitivity of convergence rates of solutions of the system (1.1) to perturbations of parameters involving degenerate equilibria. Consider the three-value case for piecewise-linear  $f$ , recalling from Section 4.2 that, upon fixing  $\nu_j = \hat{\mu}_j - \varepsilon_j$  for  $j = l, m, r$  with  $\varepsilon_m = -\varepsilon_l - \varepsilon_r$ , (4.6) implies

$$\frac{d}{dt}\bar{f}(t) = -2(\varepsilon_r + \varepsilon_l)\bar{f}(t) + (\varepsilon_l - \varepsilon_r - \bar{u}).$$

When  $\varepsilon_r + \varepsilon_l > 0$  and  $-2\varepsilon_r < \bar{u} < 2\varepsilon_l$ ,

$$\frac{d}{dt}\bar{f}(t) = -2(\varepsilon_r + \varepsilon_l)(\bar{f}(t) - \bar{f}^{\text{eq}}), \quad \bar{f}^{\text{eq}} = \frac{\varepsilon_l - \varepsilon_r - \bar{u}}{2(\varepsilon_r + \varepsilon_l)},$$

and  $\bar{f}(t)$  contracts towards  $\bar{f}^{\text{eq}} \in (-\frac{1}{2}, \frac{1}{2})$  at an  $O(\varepsilon_r + \varepsilon_l)$ -exponential rate.

On the other hand, setting  $\varepsilon_l = \varepsilon_r = 0$ , we see

$$\bar{f}(t) = \bar{f}(0) \text{ for all } t \geq 0.$$

Thus, when  $u_m(0) = \bar{f}(0)$ ,  $u_m$  is stationary and we observe an  $O(1)$  exponential convergence rate of the solution, as

$$\partial_t u_j(t) = -(u_j(t) - \bar{f}(0)) \quad \text{for } j = l, r.$$

In fact, this can also be seen via the gradient inequality in Lemma 7; letting  $\phi(x) = -1 + s, -s, 1 + s$  in  $\Omega_l, \Omega_m, \Omega_r$  respectively for  $s \in (-\frac{1}{2}, \frac{1}{2})$ , we see  $\int_{\Omega} \phi(s) d\nu = 0 = \bar{u}$  and

$$\frac{d}{dt}E(\phi(s)) = \sum_{j=l,m,r} \hat{\mu}_j f(\phi(s)) \phi'_j(s) = s(\hat{\mu}_l + \hat{\mu}_r - \hat{\mu}_m) = 0.$$

Thus the gradient inequality in Lemma 7 becomes a Łojasiewicz inequality with  $O(1)$  constant, implying exponential convergence at a rate that is  $O(1)$  — see also Remark 1.

Similar sensitivity can be observed for the cubic nonlinearity. Considering again the three-valued case and setting  $R = \frac{u_r - u_m}{u_m - u_l}$ , recall from (5.7) that

$$\partial_t R = -(u_l + u_r + u_m)(u_r - u_l)R.$$

As

$$\bar{u} = \sum_{j=l,m,r} \mu_j u_j = \frac{1}{3}(u_l + u_r + u_m) + \varepsilon_l(u_m - u_l) - \varepsilon_r(u_r - u_m),$$

we have

$$\begin{aligned} \partial_t R &= -3(\varepsilon_r(u_r - u_m) - \varepsilon_l(u_m - u_l) + \bar{u})(u_r - u_l)R, \\ \partial_t R^{-1} &= -3(\varepsilon_l(u_m - u_l) - \varepsilon_r(u_r - u_m) - \bar{u})(u_r - u_l)R^{-1}. \end{aligned}$$

For  $\bar{u} = 0$  and for small  $\varepsilon_r, \varepsilon_l > 0$ , the ratio  $R$  evolves toward the equilibrium  $\varepsilon_l/\varepsilon_r$  at a slow exponential rate that is  $O(\varepsilon_r)$ . And for  $\varepsilon_r = \varepsilon_l = 0$ , when  $\bar{u} > 0$  (resp.  $-\bar{u} > 0$ ), the ratio  $R$  (resp.  $R^{-1}$ ) contracts exponentially toward zero at a rate that is  $O(\bar{u})$ .

In case  $\bar{u} = \varepsilon_l = \varepsilon_r = 0$ , however, the ratio  $R$  is invariant in time, and  $O(1)$  exponential convergence can be observed. For instance, if

$$\bar{f}(0) = u_m(0), \quad u_l(0) = -u_r(0),$$

we see  $u_m$  remains constant at 0 and  $u_l = -u_r$ , as  $R(t) = R(0) = 1$  implies  $3\bar{u} = u_r + u_l + u_m = 3u_m = 0$ .

As  $f$  is symmetric about 0, this means  $\bar{f}(t) = 0$  for all  $t \geq 0$ , and thus

$$\partial_t u_j = -u_j(u_j + 1)(u_j - 1) \quad \text{for } j = l, r.$$

Then  $u_r$  and  $u_l$  converge exponentially towards 1 and  $-1$  respectively with  $O(1)$  rate.

In summary, even for low-dimensional dynamics where convergence to equilibrium always occurs, the exponential *rate* of convergence for the gradient system (1.1) can suddenly jump from  $O(1)$  to arbitrarily small values upon perturbation of parameters. Whether this phenomenon can occur more broadly in other kinds of gradient systems remains to be seen.

## Acknowledgements

This work has been partially supported by the National Science Foundation under grants DMS 2106534, DMS 1814991 and DMS 2206069. The authors are grateful to Sir John Ball for remarks which helped clarify and correct several arguments.

## References

- [1] G. ANDREWS AND J. BALL, *Asymptotic behaviour and changes of phase in one-dimensional nonlinear viscoelasticity*, Journal of Differential Equations, 44 (1982), pp. 306–341.
- [2] H. ATTOUCH, J. BOLTE, AND B. F. SVAITER, *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*, Math. Program., 137 (2013), pp. 91–129.
- [3] J. M. BALL AND Y. ŞENGÜL, *Quasistatic nonlinear viscoelasticity and gradient flows*, J. Dynam. Differential Equations, 27 (2015), pp. 405–442.
- [4] J. M. BALL, P. J. HOLMES, R. D. JAMES, R. L. PEGO, AND P. J. SWART, *On the dynamics of fine structure*, J. Nonlinear Sci., 1 (1991), pp. 17–70.
- [5] J. BOLTE, A. DANIILIDIS, AND A. LEWIS, *The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems*, SIAM J. Optim., 17 (2007), pp. 1205–1223.
- [6] L. BOTTOU, F. E. CURTIS, AND J. NOCEDAL, *Optimization methods for large-scale machine learning*, SIAM Review, 60 (2018), pp. 223–311.
- [7] Z. DENKOWSKA AND M. P. DENKOWSKI, *A long and winding road to definable sets*, J. Singul., 13 (2015), pp. 57–86.
- [8] N. DINULEANU, *Vector measures*, International Series of Monographs in Pure and Applied Mathematics, Vol. 95, Pergamon Press, Oxford-New York-Toronto; VEB Deutscher Verlag der Wissenschaften, Berlin, 1967.
- [9] M. FORTI, P. NISTRI, AND M. QUINCAMPOIX, *Convergence of neural networks for programming problems via a nonsmooth Lojasiewicz inequality*, IEEE Transactions on Neural Networks, 17 (2006), pp. 1471–1486.
- [10] G. FRIESECKE AND J. B. MCLEOD, *Dynamics as a mechanism preventing the formation of finer and finer microstructure*, Arch. Rational Mech. Anal., 133 (1996), pp. 199–247.
- [11] ———, *Dynamic stability of non-minimizing phase mixtures*, Proc. Roy. Soc. London Ser. A, 453 (1997), pp. 2427–2436.

- [12] N. GARCÍA TRILLOS, B. HOSSEINIE, AND D. SANZ-ALONSO, *From optimization to sampling through gradient flows*, Notices Amer. Math. Soc., 70 (2023), pp. 905–917.
- [13] J. K. HALE AND P. MASSATT, *Asymptotic behavior of gradient-like systems*, in Dynamical systems, II (Gainesville, Fla., 1981), Academic Press, New York, 1982, pp. 85–101.
- [14] J. K. HALE AND G. RAUGEL, *Convergence in gradient-like systems with applications to PDE*, Z. Angew. Math. Phys., 43 (1992), pp. 63–124.
- [15] P. R. HALMOS, *Measure Theory*, D. Van Nostrand Co., Inc., New York, N. Y., 1950.
- [16] A. HARAUX AND M. A. JENDOUBI, *The convergence problem for dissipative autonomous systems*, SpringerBriefs in Mathematics, Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2015. Classical methods and recent advances, BCAM SpringerBriefs.
- [17] D. HAUER AND J. M. MAZÓN, *Kurdyka-Lojasiewicz-Simon inequality for gradient flows in metric spaces*, Trans. Amer. Math. Soc., 372 (2019), pp. 4917–4976.
- [18] D. HILHORST, H. MATANO, T. N. NGUYEN, AND H. WEBER, *On the large time behavior of the solutions of a nonlocal ordinary differential equation with mass conservation*, J. Dynam. Differential Equations, 28 (2016), pp. 707–731.
- [19] ———, *Generation of interface for solutions of the mass conserved allen-cahn equation*, SIAM Journal on Mathematical Analysis, 52 (2020), pp. 2624–2654.
- [20] A. JENTZEN AND A. RIEKERT, *Convergence analysis for gradient flows in the training of artificial neural networks with ReLU activation*, Journal of Mathematical Analysis and Applications, 517 (2023), p. 126601.
- [21] S. LOJASIEWICZ, *Ensembles semi-analytiques*, Institut des Hautes Etudes Scientifiques, 1965.
- [22] J. A. NOHEL AND R. L. PEGO, *Nonlinear stability and asymptotic behavior of shearing motions of a non-Newtonian fluid*, SIAM J. Math. Anal., 24 (1993), pp. 911–942.
- [23] J. A. NOHEL, R. L. PEGO, AND A. E. TZAVARAS, *Stability of discontinuous steady states in shearing motions of a non-Newtonian fluid*, Proc. Roy. Soc. Edinburgh Sect. A, 115 (1990), pp. 39–59.
- [24] A. NOVICK-COHEN AND R. L. PEGO, *Stable patterns in a viscous diffusion equation*, Trans. Amer. Math. Soc., 324 (1991), pp. 331–351.
- [25] J. PALIS, JR. AND W. DE MELO, *Geometric theory of dynamical systems*, Springer-Verlag, New York-Berlin, 1982. An introduction, Translated from the Portuguese by A. K. Manning.
- [26] R. L. PEGO, *Phase transitions in one-dimensional nonlinear viscoelasticity: admissibility and stability*, Arch. Rational Mech. Anal., 97 (1987), pp. 353–394.
- [27] ———, *Stabilization in a gradient system with a conservation law*, Proc. Amer. Math. Soc., 114 (1992), pp. 1017–1024.
- [28] H. ROBBINS AND S. MONRO, *A Stochastic Approximation Method*, The Annals of Mathematical Statistics, 22 (1951), pp. 400 – 407.
- [29] J. RUBINSTEIN AND P. STERNBERG, *Nonlocal reaction—diffusion equations and nucleation*, IMA Journal of Applied Mathematics, 48 (1992), pp. 249–264.
- [30] Y. ŞENGÜL, *Nonlinear viscoelasticity of strain rate type: an overview*, Proceedings of the Royal Society A, 477 (2021), p. 20200715.

- [31] L. SIMON, *Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems*, Ann. of Math. (2), 118 (1983), pp. 525–571.