RELAXATION TO EQUILIBRIUM IN THE ONE-DIMENSIONAL CAHN–HILLIARD EQUATION∗

FELIX OTTO† AND MARIA G. WESTDICKENBERG‡

Abstract. We study the stability of a so-called kink profile for the one-dimensional Cahn–Hilliard problem on the real line. We derive optimal bounds on the decay to equilibrium under the assumption that the initial energy is less than three times the energy of a kink and that the initial $H^{-1}$ distance to a kink is bounded. Working with the $H^{-1}$ distance is natural, since the equation is a gradient flow with respect to this metric. Indeed, our method is to establish and exploit elementary algebraic and differential relationships among three natural quantities: the energy, the dissipation, and the $H^{-1}$ distance to a kink. Along the way it is necessary and possible to control the time-dependent shift of the center of the $L^2$ closest kink. Our result is different from earlier results because we do not assume smallness of the initial distance to a kink; we assume only boundedness.

Key words. energy–energy–dissipation, gradient flow, relaxation to equilibrium, stability

AMS subject classifications. 35K55, 49N99

DOI. 10.1137/130925517

1. Introduction. We introduce a nonlinear, energy-based method to study the stability of energy minimizers of the one-dimensional Cahn–Hilliard equation on the real line subject to $\pm 1$ boundary conditions at infinity. In particular, we are interested in the rate of convergence in time to a so-called kink or transition layer. We develop a nonlinear energy method that obtains optimal rates (see Remark 1 below) and is nonperturbative in the sense that we do not require our initial data to be close to the set of stable states; we require only that the distance be order-one (in the sense of Theorem 1.2).

The Cahn–Hilliard equation

\begin{equation}
    u_t = (G'(u) - u_{xx})_{xx} = 0
\end{equation}

is the $H^{-1}$ gradient flow of the energy

\[ E(u) := \int \frac{1}{2} u_x^2 + G(u) \, dx. \]

From our point of view, it is natural to exploit the gradient structure of the equation and to use the $H^{-1}$ distance as a tool in the analysis. For simplicity, we work with the canonical double-well potential $G(u) = (1 - u^2)^2/4$, but other symmetric double-well potentials with nondegenerate absolute minima are possible.

Our focus in this paper is the stability of so-called kink states. The centered kink $v$ satisfies

\begin{equation}
    v_{xx} - G'(v) = 0, \quad v \to \pm 1 \text{ as } x \to \pm \infty, \quad v(0) = 0.
\end{equation}

∗Received by the editors June 18, 2013; accepted for publication November 7, 2013; published electronically February 11, 2014.

†Max Planck Institute for Mathematics in the Sciences, Leipzig 04103, Germany (felix.otto@mis.mpg.de).

‡RWTH Aachen University, Aachen, Germany (maria@math1.rwth-aachen.de). This author was partially supported as an Alfred P. Sloan Research Fellow and by the National Science Foundation under grant DMS-0955051.
As is well known, the energy of the kink is

\[ E(v) = \int_{-1}^{1} \sqrt{2G(s)} \, ds =: c_0, \]

and \( v \) is a minimum energy state:

(1.3) \[ E(v) = \min \{ E(u) : u \in C^\infty \text{ and } u \to \pm 1 \text{ as } x \to \pm \infty \}. \]

An important feature of the energy minimization problem on the line is that it is degenerate in the sense that every shifted kink profile \( v(\cdot - c) \) also minimizes the energy. The energy minimizers are stable: A perturbation of a kink converges to a (shifted) kink profile under the gradient flow, and one is interested in analyzing and quantifying this convergence.

Before being more precise about our goal, we comment briefly on the similarity to and difference from the Allen–Cahn equation. The Cahn–Hilliard equation arises as a phenomenological model for the mixing of a binary alloy. It is related to the Allen–Cahn equation,

\[ u_t + (G'(u) - u_{xx}) = 0; \]

however, the analysis of the fourth-order Cahn–Hilliard equation is more complicated than that of the second-order Allen–Cahn equation. The main reason for this difference is that, in the setting of the Cahn–Hilliard equation on the real line, the linearized operator around a kink profile admits no spectral gap. On the other hand, the conservation of mass in the Cahn–Hilliard equation makes the equation more physical for some applications and helps in the analysis, as we now discuss.

In terms of the convergence to kink states described above, it is the mass conservation of the Cahn–Hilliard equation that allows one to determine \textit{to which kink state} \( v(\cdot - c_*) \) \textit{the solution converges}. This fact is important since the translation invariance of the energy means that energy decay alone does not give any information about the shift. Without loss of generality, we choose initial data \( u_0 \) such that the integral of \( u_0 - v \) vanishes, so that the limiting shift \( c_* = 0 \). It turns out that there are two different timescales for convergence: a faster timescale on which the solution \( u \) converges to the \( L^2 \) closest kink \( v(\cdot - c(t)) \) and a slower timescale on which \( u \) converges to the centered kink \( v \). In Theorem 1.2 below we capture these two timescales under natural assumptions on the initial data; see Remark 1.

Although we present and apply our method in the context of the Cahn–Hilliard equation, the result in Lemma 1.5 is general. In other words, the algebraic and differential relationships expressed in Lemma 1.5 describe sufficient conditions for a general gradient flow to satisfy the corresponding decay rates. Hence the method developed here may be of interest for other models, higher dimensional problems, or systems, if the required bounds can be established in that setting.

\textbf{1.1. Setup and notation.} We consider the solution \( u \) of the Cahn–Hilliard equation on the line

(1.4) \[
\begin{align*}
  u_t - (G'(u) - u_{xx})_{xx} &= 0 & \text{on } \mathbb{R} \times (0, \infty), \\
  u &\to \pm 1 & \text{for } x \to \pm \infty, \\
  u &= u_0 & \text{on } \mathbb{R} \times \{t = 0\}.
\end{align*}
\]
We define the shifted kink \( v_c(x) = v(x - c) \) as an \( L^2 \) projection of \( u \) onto the collection of shifted kink profiles. As an \( L^2 \) projection, \( v_c \) satisfies the Euler–Lagrange equation

\[
\int (u - v_c)v_{cx} \, dx = 0.
\]  

(1.5)

(Throughout the paper, the domain of integration is \( \mathbb{R} \) unless otherwise specified.) For order-one perturbations, the \( L^2 \) projection may not be unique and there is no reason to believe that \( c \) is continuous in time, but we will use neither uniqueness nor continuity (except for in heuristic arguments in subsection 1.3). We need only that (1.5) hold.

We will often express the difference between \( u \) and \( v \) or \( v_c \) as

\[
f := u - v \quad \text{and} \quad f_c := u - v_c,
\]

respectively. We note that

\[
v - v_c = f_c - f.
\]

(1.6)

We also point out for future reference that \( f \) and \( f_c \) satisfy

\[
f_t - \left( G'(v_c + f_c) - G'(v_c) - f_{cxx} \right)_{xx} = 0,
\]

(1.7)

\[
\int f_c v_{cx} \, dx = 0.
\]

(1.8)

In the next definition, we introduce the three quantities that we shall monitor. They are motivated by the gradient flow structure of the evolution.

**Definition 1.1 (energy gap, dissipation, squared distance).** For the Cahn–Hilliard equation in one space dimension, we define

- the energy gap

\[
\mathcal{E}(u) := E(u) - E(v) = E(u) - c_0 \geq 0,
\]

- the dissipation

\[
D(u) := \int \left( (u_{xx} - G'(u))_x \right)^2 \, dx,
\]

- and the squared induced distance to the fixed kink

\[
H(u) := \int F^2 \, dx \quad \text{for } F \in L^2 \text{ with } F_x = u - v,
\]

i.e., the squared \( H^{-1} \) norm of \( u - v \).

Our result obtains relaxation rates for the Cahn–Hilliard evolution (1.4) under the assumption that the initial \( H^{-1} \) distance is bounded and the initial energy is less than that of three transition layers; i.e., we will assume that the energy gap of the initial data satisfies

\[
\mathcal{E}(u) \leq 2c_0 - \epsilon \quad \text{for some } \epsilon > 0.
\]

(1.9)
Notice that (1.9) is obviously equivalent to $E(u) \leq 3c_0 - \epsilon$. Also notice that the assumption $H(u) < \infty$ implies in particular that $u - v$ has integral zero and leads to the convergence of $c$ to zero.

Throughout the paper, when we speak of $u$ and the value $c \in \mathbb{R}$, we understand that $c$ is the shift corresponding to an $L^2$ projection $v_c$ so that, in particular, $u$ and $v_c$ satisfy (1.8).

Throughout the paper, we will use the following notation.

**Notation 1.** We write $A \lesssim B$ if there exists a constant $C < \infty$—which is universal except for a possible dependence on the $\epsilon$ in (1.9)—such that

$$A \leq C B.$$  

We define $\gtrsim$ analogously and say $A \sim B$ if $B \lesssim A \lesssim B$.

We will also occasionally use the notation $\ll$ or $\gg$. For instance, $A(x) \ll B$ for $x \gg 1$ means that for every $\delta > 0$, there exists $M < \infty$ such that $x \geq M$ implies $A(x) \leq \delta B$. (Here also $M$ may depend on the $\epsilon$ in (1.9).)

Finally, we use $A \approx B$ or $A \gtrsim B$ as $L \uparrow \infty$ to indicate that for $L = L(\epsilon) \gg 1$, one has

$$|A - B| \ll 1 \quad \text{or} \quad B - A \ll 1,$$

respectively.

Our main result is the following.

**Theorem 1.2.** Consider initial data such that $H_0 := H(u_0)$ is finite and $\mathcal{E}_0 := \mathcal{E}(u_0) \leq 2c_0 - \epsilon$ for some $\epsilon > 0$. The solution of the Cahn–Hilliard equation (1.4) satisfies

\[
\begin{align*}
(1.10) & \quad \mathcal{E}(t) \leq \mathcal{E}_0, \\
(1.11) & \quad \mathcal{E}(t) \lesssim (H_0 + \mathcal{E}_0) t^{-1}, \\
(1.12) & \quad c^2(t) \lesssim (H_0 + \mathcal{E}_0)^{1/2} \mathcal{E}_0^{1/2}, \\
(1.13) & \quad c^2(t) \lesssim (H_0 + \mathcal{E}_0) t^{-1/2}, \\
(1.14) & \quad H(t) \lesssim H_0 + \mathcal{E}_0, \\
(1.15) & \quad D(t) \lesssim (H_0 + \mathcal{E}_0 + (H_0 + \mathcal{E}_0)^2) t^{-2},
\end{align*}
\]

where $\mathcal{E}(t) = \mathcal{E}(u(\cdot, t))$, $H(t) = H(u(\cdot, t))$, $D(t) = D(u(\cdot, t))$, and $c(t)$ is the shift associated to $u(\cdot, t)$ via the $L^2$ projection.

**Remark 1** (scale separation). In light of Lemma 1.3 below, (1.11) implies in particular that

\[
(1.16) \quad \|u(\cdot, t) - v_{c(t)}\|_{H^1} \lesssim (H_0 + \mathcal{E}_0)^{1/2} t^{-1/2}.
\]

The estimate (1.13), on the other hand, reveals the rate of relaxation to the centered kink $v$ via

\[
\begin{align*}
\|u(\cdot, t) - v\|_{H^1} & \lesssim \|u - v_c\|_{H^1} + \|v - v_c\|_{H^1} \\
& \lesssim (H_0 + \mathcal{E}_0)^{1/2} t^{-1/2} + \|v - v_c\|_{H^1} \\
& \lesssim (H_0 + \mathcal{E}_0)^{1/2} t^{-1/2} + |c| \\
& \lesssim (H_0 + \mathcal{E}_0)^{1/2} t^{-1/2},
\end{align*}
\]

where $c \ll 1$.
The estimates (1.16) and (1.17) describe a scale separation of the dynamics into the fast relaxation to some kink $v_{c(t)}$ with rate $t^{-\frac{1}{2}}$ and the slow relaxation to the kink $v$ (selected via the integral of the initial data) with rate $t^{-\frac{1}{4}}$. This scale separation is governed by a Stefan problem; see subsection 1.3, below. The role of the Stefan problem in Cahn–Hilliard dynamics has been pointed out by Pego [P]; see also the remark at the end of subsection 1.4.

Remark 2 (optimality and comparison to previous results). Below in subsection 1.3 we give a heuristic argument to show that, under our assumptions on the initial data, the rates expressed in (1.13) and (1.16) are optimal. In subsection 1.4, we compare these decay rates to those obtained previously for the one-dimensional problem by Bricmont, Kupiainen, and Taskinen [BKT], Carlen, Carvalho, and Orlandi [CCO01], and Howard [H07a].

1.2. Method. Our method exploits elementary algebraic and differential relationships among the quantities $E$, $D$, $H$, and $c$, which are presented in Lemmas 1.3 and 1.4, respectively. These relationships are combined in Lemma 1.5 to deduce the desired decay rates. The application of Lemma 1.5 to the Cahn–Hilliard problem yields Theorem 1.2 (see the proof at the end of section 2).

We first establish the essential “algebraic” relationships among $E$, $D$, $H$, and $c$.

LEMMA 1.3. Let the function $u$ be smooth and such that $H(u) < \infty$ and $E(u) \leq 2c_0 - \epsilon$ for some $\epsilon > 0$. Then we have

$$\begin{align*}
(1.18) & \quad E \sim \int f_c^2 + f_{cx}^2 \, dx, \\
(1.19) & \quad D \sim \int f_c^2 + f_{cxx}^2 + f_{cxxx}^2 \, dx, \\
(1.20) & \quad c^2 \lesssim (HE)^{\frac{1}{2}} + (|c| + 1)E, \\
(1.21) & \quad E \lesssim (HD)^{\frac{1}{2}} + (|c| + 1)^2D.
\end{align*}$$

We turn now to the essential differential relations among $E$, $D$, $H$, and $c$.

LEMMA 1.4. Let the smooth function $u$ satisfy $H(u) < \infty$, $E(u) \leq 2c_0 - \epsilon$ for some $\epsilon > 0$ and the Cahn–Hilliard equation (1.1). It follows that

$$\begin{align*}
(1.22) & \quad \frac{dE}{dt} = -D, \\
(1.23) & \quad \frac{dH}{dt} \lesssim \left( (1 + |c|)c^2D \right)^{\frac{1}{2}} + \frac{E}{2}D^{\frac{1}{2}}, \\
(1.24) & \quad \frac{dD}{dt} \lesssim D^{\frac{3}{2}}.
\end{align*}$$

The last lemma contains an ODE argument that applies the algebraic and differential relations from Lemmas 1.3 and 1.4.

LEMMA 1.5. Suppose $E(t) \geq 0$, $D(t) \geq 0$, $H(t) \geq 0$, and $c^2(t) \geq 0$, $t \in [0, t_*]$, are related by the differential inequalities

$$\begin{align*}
(1.25) & \quad \frac{dE}{dt} = -D, \quad \frac{dH}{dt} \lesssim c^2 \left( (c^2D)^{\frac{1}{2}} + \frac{E}{2}D^{\frac{1}{2}} \right), \quad \frac{dD}{dt} \lesssim D^{\frac{3}{2}},
\end{align*}$$

and by the algebraic inequalities

$$\begin{align*}
(1.26) & \quad E \lesssim (HD)^{\frac{1}{2}} + c_s^2D \quad \text{and} \quad c^2 \lesssim (HE)^{\frac{1}{2}} + c_sE,
\end{align*}$$

where $c_s$ is some positive constant.
where \( c_* \) is a fixed number. Then we have

\[
\begin{align*}
(1.27) & \quad \mathcal{E}(t) \leq \mathcal{E}_0, \\
(1.28) & \quad \mathcal{E}(t) \lesssim (H_0 + c_*^2 \mathcal{E}_0) t^{-1}, \\
(1.29) & \quad c^2(t) \lesssim (H_0 + c_*^2 \mathcal{E}_0) \frac{t}{\mathcal{E}_0^2}, \\
(1.30) & \quad c^2(t) \leq (H_0 + c_*^2 \mathcal{E}_0) t^{-\frac{1}{2}}, \\
(1.31) & \quad H(t) \lesssim H_0 + c_*^2 \mathcal{E}_0, \\
(1.32) & \quad \mathcal{D}(t) \lesssim (H_0 + c_*^2 \mathcal{E}_0 + (H_0 + c_*^2 \mathcal{E}_0)^2) t^{-2},
\end{align*}
\]

where \( \mathcal{E}_0 = \mathcal{E}(0) \) and \( H_0 = H(0) \).

Remark 3 (boundary values). We remark for future reference that the assumption that \( H(u) < \infty \) and \( E(u) < \infty \) implies that \( u \) satisfies ±1 boundary conditions at ±\( \infty \). Indeed, it is not hard to deduce a bound on \( u - v \) in \( H^1 \). The boundary conditions then follow from

\[
\sup_{[M, \infty)} |f|^2 \leq \int_M^\infty f^2 + f_x^2 \, dx
\]

and the analogous estimate on \((-\infty, -M] \).

Remark 4 (pointwise dissipation bound). The pointwise bound (1.32) is obtained a posteriori from (1.28) and \( \frac{d\mathcal{D}}{dt} \lesssim \mathcal{D}^{\frac{3}{2}} \). The earlier estimates (1.27)–(1.31) are independent of the bound on \( \frac{d\mathcal{D}}{dt} \).

1.3. Discussion of optimality. In this section, we argue heuristically that—under only the assumption that the \( \dot{H}^{-1} \) norm of \( f = u - v \) is controlled—the rates of relaxation \( t^{-\frac{1}{2}} \) of \( f_c \) in \( L^2 \) and \( t^{-\frac{3}{4}} \) of \( c \) are optimal. We ignore the initial control of \( f_c \) in \( H^1 \) as well as the decay rate \( t^{-1} \) of \( f_{cx} \) in \( L^2 \) in this discussion, since the phenomenon of slow decay in \( t \) is a consequence of the problem’s being formulated on the whole line, where the lowest norm in terms of order of differentiation is the strongest norm in terms of capturing decay in \( x \) (and thus slowest to decay in \( t \)). As is standard and will become clear below, the \( t^{-\frac{3}{4}} \)-relaxation in \( L^2 \) is characteristic of the simple diffusion equation with initial data in \( \dot{H}^{-1} \). Our main goal here is to explain the \( t^{-\frac{1}{2}} \)-relaxation of \( c \).

For stronger decay properties on \( u - v \) such as compact support, we would expect the better rates for \( \|f_c\|_{L^2} \) and \( c(t) \) of \( t^{-\frac{3}{4}} \) and \( t^{-\frac{1}{2}} \), respectively. In the second part below, we give an additional argument for the convergence rate in the case of strong spatial decay.

For our heuristic discussion, we will simplify our problem. We begin by reformulating the problem in terms of \( (f_c, c) \) as

\[
\begin{align*}
(1.33) & \quad f_{ct} - \dot{c}v_{cx} - (-f_{cxx} + G'(v_c) + f_c - G'(v_c))_{xx} = 0, \\
(1.34) & \quad \int f_cv_{cx} \, dx = 0.
\end{align*}
\]

Indeed, (1.33) is just a combination of the three equations

\[
\begin{align*}
u_t - (-u_{xx} + G'(u))_{xx} &= 0, \\
-v_{cxx} + G'(v_c) &= 0, \\
v_{ct} &= -\dot{c}v_{cx},
\end{align*}
\]
and (1.34) is the Euler–Lagrange equation of the $L^2$-optimal projection onto the slow manifold.

The first heuristic simplification is to linearize in $f_c$, which seems justified since we are interested in the decay close to equilibrium:

$$f_{ct} - \dot{c} v_{cx} - (-f_{cx} + G''(v_c)f_{cx})_{xx} = 0,$$

$$\int f_c v_{cx} \, dx = 0. \tag{1.35}$$

The second heuristic simplification is to carry out a long wavelength approximation in $f_c$, which seems justified since it is the fact that our problem is formulated on the entire line that leads to an only algebraic decay in time in the first place. This allows us to make three simplifications:

- Thinking of testing (1.35) with functions of long wavelength, we may replace $-\dot{c} v_{cx}$ by $-\dot{c} \int v_{cx} \, dx \delta(x-c) = -2\dot{c} \delta(x-c)$, where $\delta(x-c)$ denotes the Dirac distribution in $x$ centered at $c$.
- Likewise, we may replace $G''(v_c)f_c$ by $G''(1)f_c$ and ignore $-f_{cx}$ in comparison in (1.35).
- Finally, we may replace $\int f_c v_{cx} \, dx$ by $f_c(c,t) \int v_{cx} \, dx = 2f_c(c,t)$.

These three simplifications lead to

$$f_{ct} - 2\dot{c} \delta(x-c) - G''(1)f_{cx} = 0,$$

$$f_c(c,t) = 0, \tag{1.36}$$

which can be classically reformulated as

$$f_{ct} - G''(1)f_{cx} = 0 \quad \text{for } x \neq c(t), \tag{1.37}$$

$$2\dot{c} + G''(1)[f_{cx}] = 0, \tag{1.38}$$

$$f_c = 0 \quad \text{for } x = c(t), \tag{1.39}$$

where $[f_{cx}]$ denotes the jump $f_{cx}$ experiences when crossing $x = c(t)$ (right value minus left value). These equations define the Stefan problem, an evolution for $(f_c,c)$ that amounts to a free boundary problem for the set $\{(x,t) \mid x \neq c(t)\}$ on which the diffusion equation (1.37) is to be solved together with the Dirichlet boundary conditions (1.39).

The third and last heuristic simplification is the linearization of (1.37)–(1.39). For this purpose we think of both $f_c$ and $c$ as being small (and thus their own linear perturbations), leading to

$$f_{ct} - G''(1)f_{cx} = 0 \quad \text{for } x \neq 0, \tag{1.40}$$

$$f_c = 0 \quad \text{for } x = 0, \tag{1.41}$$

$$2\dot{c} + G''(1)[f_{cx}] = 0. \tag{1.42}$$

Notice that this linear system is triangular: The evolution of $f_c$ decouples from that of $c$, which is slaved to that of $f_c$.

We now discuss the relaxation rate of (1.40)–(1.42). In order to do so, we have to heuristically translate our assumption of finite $H^{-1}$ norm of $u(\cdot,0) - v = f_c(\cdot,0) + (v_c(0) - v)$ into this setting. We first linearize in $c$, leading to $f_c(\cdot,0) + (v_c(0) - v) \approx f_c(\cdot,0) - c(0) v_c$. We then focus on large length scales, leading to $f_c(\cdot,0) - c(0) v_c \approx f_c(\cdot,0) - 2c(0)\delta$. Hence we translate our assumption into

$$\|f_c(\cdot,0) - 2c(0)\delta\|_{H^{-1}}^2 < \infty. \tag{1.43}$$
We thus have heuristically reduced the problem to the following question: What is the relaxation rate of $(1.40)$–$(1.42)$ under the assumption $(1.43)$? We can now give an explicit answer.

Note that $(1.43)$ can be reformulated as follows: There exists a function $F(x)$ such that

$$(1.44) \quad F_x = f_c(\cdot, 0) - 2c(0) \delta \quad \text{and} \quad \int F^2 \, dx < \infty.$$ 

This implies in particular that $\int f_c(x, 0) \, dx - 2c(0) = 0$. As a consequence of $(1.40)$–$(1.42)$, this condition is preserved in time; i.e., we have $\int f_c \, dx - 2c = 0$ for all $t \geq 0$. Hence the triangular structure of $(1.40)$–$(1.42)$ becomes even more explicit:

$$
\begin{align*}
  f_{ct} - G''(1)f_{cxx} &= 0 \quad \text{for } x \neq 0, \\
  f_c &= 0 \quad \text{for } x = 0, \\
  c &= \frac{1}{2} \int f_c \, dx.
\end{align*}
$$

We also note that $(1.44)$ can be reformulated in terms of $f_c(\cdot, 0)$ exclusively: There exists a function $F(x)$ such that

$$(0, \infty) \quad F_x = f_c(\cdot, 0) \quad \text{and} \quad \int F^2 \, dx < \infty.$$ 

By symmetry, it is enough to reformulate the problem on the positive half line. Suppose that there exists a function $F(x)$ such that

$$(0, \infty) \quad F_x = f_c(\cdot, 0) \quad \text{and} \quad \int F^2 \, dx < \infty,$$

and that $f_c$ solves the diffusion equation with Dirichlet boundary conditions:

$$f_{ct} - G''(1)f_{cxx} = 0 \quad \text{on } x > 0 \quad \text{and} \quad f_c = 0 \quad \text{at } x = 0.$$ 

What is the relaxation rate of $f_c$ in $L^2((0, \infty))$ and of $\int_{(0, \infty)} f_c \, dx$? By reflection (odd for $f_c$, even for $F$), we can reformulate the question as a question on the whole line. Suppose that there exists a function $F(x)$ such that

$$\int_{(0, \infty)} F^2 \, dx < \infty,$$

and that $f_c$ solves the diffusion equation

$$f_{ct} - G''(1)f_{cxx} = 0.$$ 

What is the relaxation rate of $f_c$ in $L^2$ and of $\int H f_c \, dx$ (where $H$ denotes the Heaviside function)? This be reformulated in terms of the Fourier transform in $x$: Suppose we have

$$\int \frac{1}{k^2} |\mathcal{F} f_c(k, 0)|^2 \, dk < \infty.$$ 

What is the decay rate of

$$\int |\exp(-G''(1)tk^2)\mathcal{F} f_c(k, 0)|^2 \, dk$$
and
\[
\int \frac{i}{k} \exp(-G''(1)tk^2)\mathcal{F}f_c(k,0)\,dk.
\]

The answer can be read off:
\[
\left| \int \exp(-G''(1)tk^2)\mathcal{F}f_c(k,0)\,dk \right|^2 \\
\leq \sup_k \left( k^2 \exp(-2G''(1)tk^2) \right) \int \frac{1}{k^2} |\mathcal{F}f_c(k,0)|^2 \,dk \\
\lesssim t^{-1} \int \frac{1}{k^2} |\mathcal{F}f_c(k,0)|^2 \,dk
\]

and
\[
\left| \int \frac{i}{k} \exp(-G''(1)tk^2)\mathcal{F}f_c(k,0)\,dk \right|^2 \\
\leq \int \exp(-2G''(1)tk^2)\,dk \int \frac{1}{k^2} |\mathcal{F}f_c(k,0)|^2 \,dk \\
\lesssim t^{-\frac{1}{2}} \int \frac{1}{k^2} |\mathcal{F}f_c(k,0)|^2 \,dk.
\]

This concludes the argument for the optimality of our result.

We now give the heuristic argument that for a \textit{compactly supported} small initial perturbation, the shift \(c\) will in fact decay at the faster rate \(t^{-\frac{1}{2}}\)—and also several norms (including the \(L^2\) norm) of \(f_c\) decay faster than when the initial perturbation is just finite in \(H^{-1}\). We will proceed in a more formal fashion than the back-of-the-envelope discussion above. The starting point is the proper linearization of the Cahn–Hilliard equation around the \textit{fixed} stationary solution \(v\), expressed in terms of the (infinitesimal) perturbation \(f = u - v\):

\[
(1.45) \quad f_t - (-f_{xx} + G''(v)f)_{xx} = 0.
\]

We split \(f\) into its projection onto \(v_x\) (the basis of the kernel of \(-\partial_x^2 + G''(v))\) and the remainder,

\[
(1.46) \quad f = f_c + cv_x \quad \text{with} \quad c := \frac{\int f v_x \,dx}{\int v_x^2 \,dx}.
\]

and note that the pair \((f_c, c)\) is the \textit{linearized} version of the above-considered pair coming from the (nonlinear) \(L^2\)-projection onto the slow manifold. We use (1.46) to unfold (1.45) into a set of equations similar to (1.33) and (1.34):

\[
(1.47) \quad f_c t + \lambda v_x - (-f_{cxx} + G''(v)f_c)_{xx} = 0 \quad \text{and} \quad \int f_c v_x \,dx = 0,
\]

where \(\lambda := \dot{c}\) is the Lagrange multiplier corresponding to the constraint \(\int f_c v_x \,dx = 0\).

We now carry out an abridged version of matched asymptotic expansion for (1.47). For the two outer solutions (one on the left half line, the other on the right half line), we make the following ansatz:

\[
(1.48) \quad x = t^\frac{3}{2} y, \quad t = \exp(\tau), \quad f_c = t^{-1} f_c^\pm \quad \text{for} \pm x \gg 1.
\]
The rescaling of $x$ and $t$ is motivated by diffusive scaling, the rescaling of $f_c$ itself by the (effective) Dirichlet boundary condition at $y = 0$. Indeed, because of the exponential decay of $v_x$ (in the $x$ variable), $\int f_c v_x \, dx = 0$ implies that to leading order
\begin{equation}
(1.49) \quad f_c^\pm = 0 \quad \text{for } y = 0.
\end{equation}

Plugging in the ansatz (1.48) into (1.47), we first notice that because of the exponential decay of $\lambda v_x$, the $\lambda v_x$-term is irrelevant for $y \neq 0$. Likewise, $G''(v)$ can be replaced by $G''(1)$. Hence we obtain, in each interval $y \in (\pm \infty, 0)$ and $y \in (0, +\infty)$ separately,
\begin{equation}
0 = t^{-2} \left( -f_c^\pm - \frac{1}{2} y f_{cy}^\pm + f_{ct}^\pm - G''(1) f_{cyy}^\pm \right) + t^{-3} f_{cyyy}^\pm,
\end{equation}
so that to leading order in the time asymptotics, one has
\begin{equation}
(1.50) \quad -f_c^\pm - \frac{1}{2} y f_{cy}^\pm + f_{ct}^\pm - G''(1) f_{cyy}^\pm = 0 \quad \text{for } y \neq 0.
\end{equation}

In fact, it is more telling to rewrite the Dirichlet problem (1.49) and (1.50) in terms of the two antiderivatives $F_c^\pm$ (i.e., $F_c^\pm = f_c^\pm$) that we normalize in such a way that they decay for $\pm y \rightarrow \infty$ (note that therefore, in general, $F_c^+(0, t) \neq F_c^-(0, t)$). Indeed, in terms of $F_c^\pm$, the Dirichlet problem turns into the simpler Neumann problem
\begin{equation}
F_c^\pm - \left( \frac{1}{2} y f_c^\pm \right)_y - G''(1) F_{cy}^\pm = 0 \quad \text{for } y \neq 0 \quad \text{and} \quad F_{cy}^\pm = 0 \quad \text{for } y = 0.
\end{equation}

All solutions of this problem with compact support initially converge for $\tau \uparrow \infty$ exponentially to Gaussians (different ones on each half line):
\begin{align*}
F_c^\pm \xrightarrow{\tau \uparrow \infty} \text{const}^+ \exp \left( -\frac{y^2}{4G''(1)} \right).
\end{align*}

On the level of $f_c^\pm$, this translates into
\begin{equation}
(1.51) \quad f_c^\pm \xrightarrow{\tau \uparrow \infty} A^\pm y \exp \left( -\frac{y^2}{4G''(1)} \right).
\end{equation}

The two constants $A^\pm$ are determined by the initial data.

We now turn to the inner solution and make the ansatz:
\begin{equation*}
t = \exp(\tau), \quad f_c = t^{-\frac{3}{2}} f_c^0, \quad \text{and} \quad \lambda = t^{-\frac{3}{2}} \lambda^0.
\end{equation*}

For this ansatz, (1.47) turns into
\begin{equation*}
0 = t^{-\frac{3}{2}} \left( \lambda^0 v_x - \left( -f_{cxx}^0 + G''(v) f_c^0 \right)_{xx} \right) + t^{-5/2} \left( -\frac{3}{2} f_c^0 + f_{ct}^0 \right),
\end{equation*}
so that to leading order, we obtain the quasi-static equation
\begin{equation}
(1.52) \quad \lambda^0 v_x - \left( -f_{cxx}^0 + G''(v) f_c^0 \right)_{xx} = 0,
\end{equation}

together with $\int f_c^0 v_x \, dx = 0$. It is convenient to introduce the antiderivative $V$ of the kink $v$ (i.e., $V_x = v$) normalized by $\int V v_x \, dx = 0$. Then (1.52) can be rewritten as
\begin{equation}
(1.53) \quad -\lambda^0 V - \mu^0 x \text{ const} - f_{cxx}^0 + G''(v) f_c^0 = 0.
\end{equation}
Testing (1.53) with \( v_x \) (the basis of the kernel of the \( L^2 \)-symmetric \(-\partial^2_x + G''(v) \) and thus orthogonal to its image) yields the “solvability condition” \( \text{const} = 0 \) (by normalization of \( V \) and since \( v_x \) is even). Hence (1.53) simplifies to

\[
- f^{0}_{xx} + G''(v)f^{0}_c = \lambda^0 V + \mu^0 x.
\]

The two time-dependent constants \( \lambda^0 \) and \( \mu^0 \) are determined by “matching” the outer and inner solutions. Indeed, we can compare them on the level of the first spatial derivatives, since the scaling is \( t^{-\frac{3}{2}} \) both from coming outside and from inside:

\[
\partial_x f \approx t^{-\frac{3}{2}} f^{\pm}_{cy} \quad \text{and} \quad \partial_x f \approx t^{-\frac{3}{2}} f^{0}_{cx}.
\]

From (1.54) we read off

\[
G''(1)f^{0}_{cx} \approx \left\{ \begin{array}{ll}
\lambda^0 + \mu^0 & \text{for } x \gg 1 \\
-\lambda^0 + \mu^0 & \text{for } -x \gg 1
\end{array} \right.,
\]

whereas from (1.51), we obtain for \( \tau \uparrow \infty \)

\[
f^{\pm}_{cy} \approx A^\pm \quad \text{for } y \approx 0.
\]

Hence the inner coefficients \( \lambda^0 \) and \( \mu^0 \) are determined via a linear set of equations by the two outer (time asymptotic) coefficients \( A^\pm \), which in turn are determined by the initial conditions and generically do not vanish. Hence \( \lambda^0 \) generically does not vanish, which means \( \dot{c} \sim t^{-\frac{3}{2}} \) and thus as claimed \( c \sim t^{-\frac{1}{2}} \).

1.4. Discussion of previous results in the literature. The convergence to the stationary profile has been studied in [BKT, CCO01, H07a] under slightly different assumptions and with different methods. Here we discuss these previous results (in the notation of our paper) and point out similarities and differences compared to our work.

We remark that in this literature review, we will not be careful about checking whether the obtained constants are universal or whether and how they depend on the initial data. We will denote these constants by \( C \). Our focus here is purely on the exponent in the convergence rates.

Bricmont, Kupiainen, and Taskinen [BKT] prove stability of the kink in a weighted \( L^\infty \) norm using a renormalization group approach. It is a perturbative result: They assume that the initial datum is given by

\[
u(x,0) = v(x) + f(x) \quad \text{with} \quad \sup_{\mathbb{R}} \left( (1 + |x|^p)f(x) \right) \leq \delta
\]

for some \( \delta > 0 \) sufficiently small and any \( p > 3 \). Under this condition, they conclude [BKT, Theorem 1.1] that

\[
\sup_x |u(t,x) - v(x)| \to 0 \quad \text{for } t \uparrow \infty.
\]

In fact, they obtain the convergence rate to the fixed kink

\[
\sup_x |u(x,t) - v(x)| \leq Ct^{-\frac{1}{2}};
\]

cf. [BKT, Proposition 2.1]. According to the heuristic explanation in subsection 1.3, this rate is optimal given the assumption of strong decay at infinity.
In terms of the method, their result relies on a very careful and fairly explicit analysis of the semigroup generated by the linearization (1.45) [BKT, section 3]. By passing to the antiderivative variable $F$ (i.e., $F_x = f$), they rewrite the generator in the $L^2$-symmetric fashion (as opposed to the $H^{-1}$-symmetric original version):

$$-\partial_x(-\partial_x^2 + G''(v))\partial_x.$$  

This operator is positive semidefinite, with continuous spectrum $[0, \infty)$; moreover, $v$ is in its kernel, but this does not yield a point spectrum since $v$ is not in $L^2$. The main ingredient is [BKT, Proposition 3.2], which shows that, for large times, the parabolic Green’s function behaves like that of simple diffusion with Dirichlet boundary conditions at $x = 0$ plus a low-rank part that decays like $t^{-\frac{1}{2}}$ coming from $v$. Korvola, Kupiainen, and Taskinen [KKT] extend the method to derive convergence rates for the Cahn–Hilliard equation in dimensions $d \geq 3$.

A second paper studying relaxation to the kink is the paper by Carlen, Carvalho, and Orlandi [CCO01], which proves convergence in terms of the energy gap and $L^1$ norm. The method—developed by the authors in previous work—is to exploit a system of ODEs for the decay of the energy gap and the (slower) growth of a quantity related to the second moments. Specifically, fix any $\varepsilon > 0$. They assume spatial decay for the initial data in the sense that

$$\int x^2 (u_0 - v_c)^2 \, dx \lesssim 1,$$

and that the initial data is close in $L^2$ to a shifted kink:

$$(1.57) \quad \int (u_0 - v_c)^2 \, dx \leq \delta$$

for some $\delta$ sufficiently small. They then show [CCO01, Theorem 1.1] that the energy gap satisfies

$$\mathcal{E}(u) \leq C(1 + t)^{-(\frac{3}{4} - \varepsilon)}$$

and that the distance to the centered kink $v$ satisfies

$$\int |u(x,t) - v(x)| \, dx \leq C(1 + t)^{-(\frac{3}{4} - \varepsilon)} .$$

The method applied in [CCO01] is a technique developed by the authors for a related nonlocal equation, which they first treated in one space dimension, and which Carlen and Orlandi later also treated in dimensions two and three; see [CCO00, CO] and the references therein.

Like our method, the method of [CCO01] is energy-based and nonlinear. Also like our result, their result explicitly expresses the multiscale nature of the relaxation phenomenon, capturing a faster rate of relaxation to $v_c$ and a slower rate of relaxation to $v$.

There are at least two important differences between the result of [CCO01] and ours, however. First, like that in [BKT], their result is perturbative in the sense that they require closeness rather than just boundedness in $L^2$ (cf. (1.57)). Second, their rates fall short of optimal because they rely on second moments rather than the $H^{-1}$ norm. At first it may seem surprising that this choice makes a significant difference: Second moments are stronger than the $H^{-1}$ norm but equivalent on the
level of scaling. However, even on the level of the linear problem (cf. subsection 1.3 above), one can see that the $\dot{H}^{-1}$ norm leads to a better rate, as we now explain.

Essentially, the method of [CCO01] applied to the linear problem described in subsection 1.3 amounts to deriving the system

$$\frac{d}{dt} \frac{1}{2} \int f_c^2 \, dx = - \int f_{cx}^2 \, dx,$$

$$\frac{d}{dt} \int x^2 f_c^2 \, dx \leq 2 \int f_c^2 \, dx.$$ 

Actually, one can slightly improve the second relation using

$$\int x^2 f_{cx}^2 \, dx \geq \frac{1}{4} \int f_c^2 \, dx,$$

so that the system of ODEs improves to

$$\frac{d}{dt} \frac{1}{2} \int f_c^2 \, dx = - \int f_{cx}^2 \, dx,$$

$$\frac{d}{dt} \int x^2 f_c^2 \, dx \leq \frac{3}{2} \int f_c^2 \, dx.$$ 

The argument of [CCO01] then employs the interpolation inequality

$$\int x^2 f_{cx}^2 \, dx \int f_{cx}^2 \, dx \geq \frac{9}{4} \left( \int f_c^2 \, dx \right)^2$$

(which is at the root of Heisenberg’s uncertainty principle) to link the three quantities appearing in (1.58). Setting $A := \int x^2 f_{cx}^2 \, dx$, $B := \int f_c^2 \, dx$, and $C := \int f_{cx}^2 \, dx$ and inserting this bound into (1.58) yields the system of ODEs

$$\dot{A} \leq \frac{3}{2} B, \quad \dot{B} \leq -\frac{9}{2} \frac{B^2}{A}.$$ 

As explained in [CCO01], such a differential system implies time decay of $B = \int f_c^2 \, dx$ with rate $t^{-\frac{7}{4}}$, which, however, falls short of the optimal $t^{-1}$ rate for the squared $L^2$ norm of $f_c$ under the assumption of weak spatial decay (see the optimality discussion in subsection 1.3 above).

If one considers the $\dot{H}^{-1}$ norm instead of second moments, one obtains a better rate. Indeed, let $F$ satisfy $F_x = f_c$ and define $A := \int F^2 \, dx$, $B := \int f_c^2 \, dx$, and $C := \int f_{cx}^2 \, dx$. We obtain the system of inequalities

$$\dot{A} = -2B, \quad \dot{B} = -2C.$$ 

Then, using the simple interpolation inequality $B^2 \leq AC$, the system becomes

$$\dot{A} = -2B, \quad \dot{B} \leq -\frac{B^2}{A}.$$ 

Since $\dot{A} \leq 0$, we read off $A \leq A_0$ and obtain from $\dot{B} \leq -2B^2/A_0$ immediately that $B$ decays with the optimal $t^{-1}$ rate.

Finally, we consider the results derived by Howard in [H07a]. Assuming spatial decay of the initial perturbation, he derives convergence rates both for the $L^\infty$ norm
of $f_c$ and for the shift. Howard is particularly interested in obtaining sharp rates of convergence in time under low rates of spatial decay of the initial perturbation in $L^\infty$. For instance, Howard shows [H07a, Theorem 1.2] that for

\begin{equation}
|u_0(x) - v(x)| \leq \delta (1 + |x|)^{-\frac{3}{2}},
\end{equation}

for some $\delta$ sufficiently small, one obtains convergence rates

\begin{equation}
||f_c||_\infty \leq C(1 + \sqrt{t})^{-\frac{3}{4}}, \quad |c(t)| \leq C(1 + t)^{-\frac{1}{4}}.
\end{equation}

Formally, this result is optimal and is similar to ours in the following sense. The weak spatial decay rate $(1 + |x|)^{-\frac{3}{2}}$ in (1.59) is equivalent in terms of scaling to our assumption of bounded $\dot{H}^{-1}$ distance to $v$, and the $t^{-\frac{1}{4}}$ decay of $f_c$ in (1.60) is the same that one obtains from

\begin{equation}
||f_c||_\infty \lesssim \left( ||f_c||_{L^2}, ||f_{cx}||_{L^2} \right)^{\left( \frac{1}{1.18}, \frac{1}{1.19} \right)} \lesssim (ED)^{\frac{1}{4}}
\end{equation}

if one uses our bounds (1.11) and (1.15).

Under the stronger spatial decay assumption

\begin{equation}
|u_0(x) - v(x)| \leq \delta (1 + |x|)^{-2},
\end{equation}

Howard’s method also captures the optimal rates for strong decay

\begin{equation}
||f_c||_\infty \leq C(1 + t)^{-1}, \quad |c(t)| \leq C(1 + t)^{-\frac{1}{2}}.
\end{equation}

See the remark around equation (1.28) in [H07a]. As Howard points out, the spatial decay rate $(1 + |x|)^2$ is the critical rate of spatial decay, after which the result saturates.

In terms of method, Howard’s analysis is based on a careful analysis of the spectrum of the linear operator via Evans function techniques. He exploits the spectral information to derive pointwise estimates on the Green’s function of the linear equation. Howard has extended his method to higher dimensional problems and, together with Kwon, to systems; see [H07b, HK] and the references therein.

In distinction to Howard, we work with the weaker $\dot{H}^{-1}$ norm, which is quite natural given the $\dot{H}^{-1}$ gradient flow structure of the equation. Under this weaker assumption on the initial data, our method returns optimal results via a shorter and less technical method. In particular, our method may be easier to adapt to problems for which explicit estimates on the Green’s function of the linear problem are difficult or inaccessible. It is important to point out, however, that our method saturates with the $\dot{H}^{-1}$ assumption. In other words, our method is oblivious to the faster rates of decay—captured by [BKT, H07a]—that one obtains under more stringent decay assumptions on the initial data.

Perhaps the most salient feature of our result in distinction to the three other results described above is that [BKT, CCO01, H07a] are all perturbative in the sense that $\delta$ in (1.55), (1.57), and (1.59) must be assumed to be sufficiently small. While these three previous works hence assume closeness to a kink state, our result assumes only that the distance is bounded. Handling order-one perturbations of a kink state is necessary, for instance, if one wants to consider the coarsening problem of multiple kinks that collide with each other.

In some sense, our method is an attempt to generalize to the mildly nonconvex case the relationships that hold for gradient flows with respect to a convex energy
functional. The dynamic bounds on energy, dissipation, and distance in the convex setting were observed already by Brezis [B]. Indeed, when $\mathcal{E}$ is convex, one obtains

$$\frac{d\mathcal{E}}{dt} = -D, \quad \frac{dH}{dt} \leq 0, \quad \mathcal{E} \leq (HD)^{\frac{1}{2}}, \quad \frac{dD}{dt} \leq 0,$$

from which it is easy to deduce

$$\mathcal{E} \leq \mathcal{E}_0, \quad \mathcal{E} \leq \frac{H_0}{l}, \quad H \leq H_0, \quad D \leq \frac{4H_0}{l^2}.$$  

Naturally the first equation in (1.61) and the first estimate in (1.62) hold independently of convexity. Our result expresses that in the mildly nonconvex case, the second to fourth estimates in (1.61) hold up to “error terms” as expressed in (1.25) and (1.26). The generalization of the second to fourth estimates from (1.62) is then expressed in (1.28), (1.31), and (1.32).

There is a similarity between our method and the use of negative norms to derive convergence rates as described in the book of Lemarie-Rieusset [L] (see also the cited references) and exploited recently by Guo and Tice [GT], Guo and Wang [GW], and Sohinger and Strain [SS]. Lemarie-Rieusset [L] derives an optimal decay rate for the Navier–Stokes equation starting from a negative Besov space. More recently in [GT], Guo and Tice use negative norms in their derivation of convergence rates for the equations of viscous surface waves. In [GW], Guo and Wang apply this idea to obtain optimal rates of convergence for dissipative equations, including the Navier–Stokes equations and the Boltzmann equation. Sohinger and Strain derive improved results for the Boltzmann equation in [SS]. Although our method is different, there is a similarity in the observation that negative norms can play an important role in optimal convergence rates.

As mentioned above, previous results have been derived for convergence to stationarity in higher space dimensions, both for the Cahn–Hilliard equation and similar equations; see [KKT, CO, H07b]. Unsurprisingly, the dynamics in $d > 1$ is even richer than in $d = 1$. As described by Pego in [P], the Cahn–Hilliard equation in higher dimensions is described in the short term by the Stefan problem and on a longer timescale by the Mullins–Sekerka model. It would be interesting to see whether our method can capture this behavior; this is the subject of current and future investigations. The method developed in this paper is also used by Esselborn in [E] to obtain convergence rates to equilibrium in the thin film equation.

1.5. Organization. In section 2, we prove Lemmas 1.3–1.5 with the exception of the energy–energy–dissipation estimates (equations (1.18) and (1.19) from Lemma 1.3). At the end of the section, we prove Theorem 1.2, which is easy to do with Lemmas 1.3–1.5 in hand. Then in section 3 we give the proof of estimates (1.18) and (1.19).

2. Proofs other than energy–energy–dissipation. The proof of (1.18) and (1.19) from Lemma 1.3 is lengthy and is deferred to section 3. Estimates (1.18) and (1.19) are similar to the bounds on energy and energy–dissipation established in previous work [OR]. We now proceed with the proof of equations (1.20) and (1.21) of Lemma 1.3.

Proof of (1.20) and (1.21) of Lemma 1.3. We first address (1.20). We will use the identity (1.6) and the fact that

(i) the integral of $f$ vanishes because of $f = F_x$ with $H = \int P^2 dx < \infty$;
(ii) \( \int f^2 dx \lesssim \mathcal{E} \) by (1.18).

In order to make the first item more quantitative, we select a smooth cut-off function \( \eta(\hat{x}) \), i.e., a function such that \( \eta = 1 \) for \( \hat{x} \in [-1, 1] \) and \( \eta = 0 \) for \( \hat{x} \not\in [-2, 2] \). We set \( \eta_L(x) = \eta(\frac{x}{L}) \) and claim that

\[
||| \int \eta_L f dx ||| \lesssim (L^{-1} H)^{\frac{1}{2}}.
\]

Indeed, we have

\[
||| \int \eta_L f dx ||| = ||| \int \eta_{Lx} F dx ||| \quad \text{because } F_x = f
\]

\[
\leq \left( \int \eta_{Lx}^2 dx \int F^2 dx \right)^{\frac{1}{2}}
\]

\[
= \left( L^{-1} \int \eta_{x}^2 d\hat{x} \int F^2 dx \right)^{\frac{1}{2}}.
\]

We use (2.1) in the form of

\[
||| \int \eta_L f dx ||| \ll |c| \quad \text{for } L \gg c^{-2} H.
\]

We also note that since \( v - v_c \) decays exponentially on scale one away from \([-|c|, |c|]\), we have

\[
\int \eta_L (v - v_c) dx \approx \int v - v_c dx = 2c \quad \text{for } L \gg |c| + 1.
\]

From (1.6), (2.2), and (2.3), we obtain

\[
2c \approx \int \eta_L f_c dx \quad \text{for } L \gg \max\{c^{-2} H, |c| + 1\}.
\]

We now observe that

\[
||| \int \eta_L f_c dx ||| \leq \left( \int \eta_{Lx}^2 dx \int f_c^2 dx \right)^{\frac{1}{2}}
\]

\[
= \left( L \int \eta^2 d\hat{x} \int f_c^2 dx \right)^{\frac{1}{2}}
\]

\[
\overset{(1.18)}{\lesssim} (L\mathcal{E})^{\frac{1}{2}}.
\]

Estimates (2.4) and (2.5) combine to

\[
|c| \lesssim \left( \max\{c^{-2} H, |c| + 1\} \mathcal{E} \right)^{\frac{1}{2}},
\]

which we rewrite as

\[
c^2 \lesssim \max\{c^{-2} H \mathcal{E}, (|c| + 1) \mathcal{E}\} \quad \text{or} \quad c^2 \lesssim \max\{(H \mathcal{E})^{\frac{1}{2}}, (|c| + 1) \mathcal{E}\}.
\]

This establishes the estimate (1.20) of Lemma 1.3.
We turn now to (1.21). We first claim that it is enough to show

\[
E \lesssim (HD)^{\frac{1}{2}} + ((|c| + 1)c^2 D)^{\frac{1}{2}} + D.
\]

Indeed, inserting (1.20) into (2.6) yields

\[
E \lesssim (HD)^{\frac{1}{2}} + \left((|c| + 1) \left(HE^{\frac{1}{2}} + (|c| + 1)E \right) D\right)^{\frac{1}{2}} + D
\]

\[
\lesssim (HD)^{\frac{1}{2}} + \mathcal{E}^{\frac{1}{2}} (HD)^{\frac{1}{2}} ((|c| + 1)^2 D)^{\frac{1}{2}} + \mathcal{E}^{\frac{1}{2}} ((|c| + 1)^2 D)^{\frac{1}{2}} + D,
\]

so that (1.21) follows from a twofold application of Young’s inequality. We now turn to (2.6). In view of (1.18) and (1.19), it is enough to show that

\[
\int f^2_c \, dx \lesssim \left( \int F^2 \, dx \int f^2_{cx} \, dx \right)^{\frac{1}{2}} + \left( (|c| + 1)c^2 \int f^2_{cx} \, dx \right)^{\frac{1}{2}}.
\]

The first (and leading-order) right-hand side term can be quickly established:

\[
\int f^2_c \, dx = \int ff_c \, dx + \int (v - v_c)f_c \, dx \quad \text{by (1.6)}
\]

\[
= - \int \int Ff_{cx} \, dx + \int (v - v_c)f_c \, dx \quad \text{by } F_x = f
\]

\[
\leq \left( \int F^2 \, dx \int f^2_{cx} \, dx \right)^{\frac{1}{2}} + \int |v - v_c|f_c \, dx.
\]

Hence, it is enough to show that

\[
\int |v - v_c|f_c \, dx \lesssim \left( (|c| + 1)c^2 \int f^2_{cx} \, dx \right)^{\frac{1}{2}}.
\]

To this end, we observe by the Cauchy–Schwarz inequality that

\[
\int |v - v_c|f_c \, dx \lesssim \left( (x - c)^2 + 1 \right) \left( v_c - v \right)^2 dx \int \frac{1}{(x - c)^2 + 1} f^2_c \, dx \right)^{\frac{1}{2}}.
\]

We turn to the first factor in (2.9). Since the integrand depends in a continuous way on $c$, it is enough to bound the behavior for very small and very large $c$. For $|c| \ll 1$ we have

\[
\int ((x - c)^2 + 1)(v_c - v)^2 dx \approx \int (x^2 + 1)(v_c)^2 dx \sim c^2.
\]

For $c \gg 1$, on the other hand, we note that $v - v_c$ is, up to a factor of 2 and exponential tails of scale 1, close to the characteristic function of the interval $(0, c)$, so that we have

\[
\int ((x - c)^2 + 1)(v - v_c)^2 dx \approx 2 \int_0^c ((x - c)^2 + 1) dx \sim c^3.
\]
Now (2.10) and (2.11) combine to

\[
\int ((x - c)^2 + 1)(v - v_c)^2 \, dx \lesssim \left( |c| + 1 \right) c^2.
\]

The proof of (2.8) is completed by addressing the second factor in (2.9). Because we will need the following Hardy-type estimate several times, we formulate it as a lemma and give the proof immediately below.

**Lemma 2.1.** For a smooth function \( f_c \) satisfying (1.8), we have

\[
\int \frac{1}{(x - c)^2 + 1} f_c^2 \, dx \lesssim \int f_{cx}^2 \, dx.
\]

Inserting (2.12) and (2.13) into (2.9) yields (2.8).

**Proof of Lemma 2.1.** On the one hand, by Hardy’s inequality we have

\[
\int \frac{1}{(x - c)^2 + 1} (f_c - f_c(c))^2 \, dx \lesssim \int f_{cx}^2 \, dx.
\]

On the other hand, we note that \( v_{cx} = v_x (\cdot - c) \) is nonnegative, of integral 2, and exponentially decaying on scale 1. Hence (1.8) implies

\[
f_c(c) = -\frac{1}{2} \int (f_c - f_c(c)) v_{cx} \, dx,
\]

and thus

\[
f_c^2(c) \lesssim \int f_{cx}^2 \, dx,
\]

which we rewrite as

\[
\int \frac{1}{(x - c)^2 + 1} f_c^2(c) \, dx \lesssim \int f_{cx}^2 \, dx.
\]

Now (2.14) and (2.15) combine by the triangle inequality to give (2.13).

**Proof of Lemma 1.4.** Equality (1.22) is classical.

We now turn to (1.23). We note that (1.7), on the level of \( F \), translates into

\[
F_t - (G'(v_c + f_c) - G'(v_c) - f_{cx x})_x = 0.
\]

We thus obtain

\[
\begin{aligned}
\frac{d}{dt} \int F^2 \, dx &= \int F F_t \, dx \\
&= \int F (G'(v_c + f_c) - G'(v_c) - f_{cx x})_x \, dx \quad \text{by (2.16)} \\
&= -\int f (G'(v_c + f_c) - G'(v_c) - f_{cx x}) \, dx \quad \text{by } F_x = f \\
&= -\int (f_c + v_c - v)(G'(v_c + f_c) - G'(v_c) - f_{cx x}) \, dx \quad \text{by (1.6)} \\
&= -\int f_{cx}^2 + G''(v_c) f_c^2 + \int (G'(v_c + f_c) - G'(v_c) - G''(v_c) f_c) f_c \, dx \\
&- \int (v_c - v)(G'(v_c + f_c) - G'(v_c)) \, dx - \int (v_c - v)_x f_{cx} \, dx.
\end{aligned}
\]
Recalling the positivity of the linearized energy (cf. (3.2)), we obtain
\[
\frac{d}{dt} \frac{1}{2} \int F^2 \, dx \\
\leq - \int (G'(v_c + f_c) - G'(v_c) - G''(v_c)f_c) \, f_c \, dx \\
+ \int (v - v_c)(G'(v_c + f_c) - G'(v_c)) \, dx + \int (v - v_c)f_{cx} \, dx \\
=: I + II + III.
\]
(2.17)

Hence it remains to show that
\[
I + II + III \lesssim ((|c| + 1)c^2 D)^{\frac{1}{2}} + \mathcal{E}^\frac{3}{4} D^\frac{1}{4}.
\]

We remark for reference below that we will make use of the elementary inequality
\[
\sup |f_c| \lesssim \left( \int f_c^2 \, dx \int f_{cx}^2 \, dx \right)^{\frac{1}{4}}
\]
in two ways. By estimating both terms on the right-hand side using (1.18) from Lemma 1 or estimating one term by (1.18) and the other by (1.19), we obtain
\begin{align}
(2.18) \quad \sup |f_c| & \lesssim \mathcal{E}^{\frac{3}{4}} \lesssim 1 \quad \text{or} \quad \sup |f_c| \lesssim (\mathcal{E}D)^{\frac{1}{4}},
\end{align}
respectively.

An application of the mean value theorem and the $L^\infty$ bound from the first part of (2.18) establish
\[
II = \int (v - v_c)(G'(v_c + f_c) - G'(v_c)) \, dx \\
\lesssim \int |v - v_c||f_c| \, dx \overset{(2.8),(1.19)}{\lesssim} ((|c| + 1)c^2 D)^{\frac{1}{2}}.
\]
(2.19)

Next we consider term $I$ from (2.17), which in view of the first bound in (2.18) satisfies
\[
I = - \int (G'(v_c + f_c) - G'(v_c) - G''(v_c)f_c) \, f_c \, dx \lesssim \int |f_c|^3 \, dx \\
\leq \sup |f_c| \int f_c^2 \, dx.
\]
Applying the second part of (2.18) and (1.18) from Lemma 1.3 now returns
\begin{align}
(2.20) \quad I \lesssim \mathcal{E}^\frac{3}{2} D^{\frac{1}{2}} \lesssim \mathcal{E}^\frac{3}{2} D^{\frac{1}{2}},
\end{align}
where in the second equality we have recalled the assumed bound on the energy gap.

For term $III$ from (2.17), we have by the Cauchy–Schwarz inequality that
\[
III = \int (v - v_c)f_{cx} \, dx \leq \left( \int (v_x - v_{cx})^2 \, dx \int f_{cx}^2 \, dx \right)^{\frac{1}{2}}.
\]
We note that both
\[
\left( \int (v_x - v_{cx})^2 \, dx \right)^{\frac{1}{2}} \leq 2 \left( \int v_x^2 \, dx \right)^{\frac{1}{2}} \lesssim 1
\]
and
\[
\left( \int (v_x - v_{cx})^2 \, dx \right)^{\frac{1}{2}} \leq |c| \left( \int v_{xx}^2 \, dx \right)^{\frac{1}{2}} \lesssim |c|,
\]
from which it follows that
\[
\int (v_x - v_{cx})^2 \, dx \lesssim \min \{ c^2, 1 \} \lesssim (|c| + 1)c^2.
\]
This estimate together with (1.19) yields
\[
(2.21) \quad III \lesssim (|c| + 1)c^2 D^{\frac{1}{2}}.
\]
The combination of (2.19), (2.20), and (2.21) completes the proof of (1.23).

We turn now to the proof of (1.24). It will be convenient to abbreviate
\[
g := u_{xx} - G'(u),
\]
so that \( D = \int g_x^2 \, dx \). Using the Cahn–Hilliard equation (1.4) in the form \( u_t = -g_{xx} \), we compute directly
\[
\frac{d}{dt} \frac{1}{2} \int \left( (u_{xx} + G'(u))_x \right)^2 \, dx = -\int g_{xx}(-\partial_x^2 + G''(u))u_t \, dx
\]
(2.22)
\[
= -\int h_x^2 + G''(v_c)h^2 \, dx + \int (G''(u) - G''(v_c))h^2 \, dx,
\]
where \( v_c \) is as always the \( L^2 \)-projection of \( u \) and \( h := (-u_{xx} + G'(u))_{xx} = g_{xx} \). The idea is to use the lower bound on the linearized energy from Lemma 3.1 below for functions that are orthogonal to \( v_{cx} \). Therefore we decompose \( h = h_0 + \alpha v_{cx} \), where
\[
\alpha := \frac{\int h v_{cx} \, dx}{\int v_{cx}^2 \, dx}
\]
and consequently \( \int h_0 v_{cx} \, dx = 0 \). Notice that an integration by parts and the exponential decay of \( v_{cx}, v_{cxx} \) give easily
(2.23)
\[
\alpha^2 \lesssim D.
\]
We substitute the decomposition of \( h \) into the first term in (2.22) to deduce
\[
\int h_x^2 + G''(v_c)h^2 \, dx = \int h_0^2 + G''(v_c)h_0^2 \, dx
\]
(2.24)
\[
\gtrsim \int h_0^2 \, dx.
\]
In the first equality, we use \(-v_{cxx} + G'(v_c) = 0\) to deduce that both \( \alpha \)-dependent terms vanish. In the second line, we apply the linearized energy gap estimate from Lemma 3.1 below to \( h_0 \).
It now suffices to establish

\[(2.25) \quad \int (G''(u) - G''(v_c)) h^2 \, dx \lesssim D^\frac{3}{4} \left( \int h_{0x}^2 + h_0^2 \, dx \right)^{\frac{1}{2}} + D^{\frac{2}{3}}. \]

Indeed, combining (2.22) and (2.24), the estimate (2.25) and an application of Young’s inequality return the bound (1.24).

In order to show (2.25), we begin with the estimate

\[
\int (G''(u) - G''(v_c)) h^2 \, dx \\
\leq \sup_{|\tau| \leq 1 + ||u - v_c||_{\infty}} |G'''(\tau)| \int |u - v_c| h^2 \, dx \\
(2.18) \lesssim \int |u - v_c| h_0^2 \, dx + \alpha^2 \int |u - v_c| v_{cx}^2 \, dx,
\]

where in the last line we have substituted \( h = h_0 + \alpha v_{cx} \) and applied Young’s inequality. On the one hand, for the 0th order in \( \alpha \) term, we use

\[
\int |u - v_c| h_0^2 \, dx \\
\leq ||u - v_c||_{\infty} \int h_0^2 \, dx \\
(2.27) \lesssim (\mathcal{E} D)^{\frac{1}{2}} \int h_0^2 \, dx \lesssim D^{\frac{1}{2}} \int h_0^2 \, dx,
\]

where we have used the bound from (2.18) to control \( ||u - v_c||_{\infty} \) and we have recalled the uniform bound on \( \mathcal{E} \). We now estimate the \( L^2 \) norm of \( h_0 \) via

\[
\int h_0^2 \, dx = \int h_0 (h - \alpha v_{cx}) \, dx \\
\leq \left| \int h_0 g_{xx} \, dx \right| + |\alpha| \left( \int h_0^2 \, dx \int v_{cx}^2 \, dx \right)^{\frac{1}{2}} \\
\leq \left( \int h_{0x}^2 \, dx \int g_{xx}^2 \, dx \right)^{\frac{1}{2}} + |\alpha| \left( \int h_0^2 \, dx \right)^{\frac{1}{2}} \\
(2.28) \lesssim D^{\frac{1}{2}} \left( \int h_{0x}^2 + h_0^2 \, dx \right)^{\frac{1}{2}}.
\]

For the second order in \( \alpha \) term, we write

\[
\alpha^2 \int |u - v_c| v_{cx}^2 \, dx \\
\leq \alpha^2 \left( \int \frac{(u - v_c)^2}{1 + (x - c)^2} \, dx \right)^{\frac{1}{2}} \\
(2.23) \lesssim D \left( \int \frac{(u - v_c)^2}{1 + (x - c)^2} \, dx \right)^{\frac{1}{2}} \\
(2.29) \lesssim D^{\frac{3}{2}},
\]
where in the first bound we have used the exponential decay of $v_{ex}$, and in the last line we have applied Lemma 2.1 and (1.19).

The combination of (2.26), (2.27), (2.28), and (2.29) establishes (2.25) and hence completes the proof of (1.24).

**Proof of Lemma 1.5.** Throughout the proof, the reader may think of $c_*=1$: The way $c_*$ intervenes is dictated by scaling. Inequality (1.27) is immediate from the first item in (1.25). We first address (1.31). To that purpose, we insert the second item in (1.25) into the second item in (1.25) and derive

$$
\frac{dH}{dt} \leq c_{*}^{\frac{1}{2}} \left[ (H^{\frac{1}{2}}E^{\frac{1}{2}} + (c_{*}E)^{\frac{1}{2}})D^{\frac{1}{2}} + E^{\frac{1}{2}}D^{\frac{1}{2}} \right].
$$

We combine this with the change of variables $]0, \infty[ \ni t \leftrightarrow \mathcal{E} \in (0, \mathcal{E}_0]$; according to the first item in (1.25) we have $-\frac{d}{d\mathcal{E}} = D^{-1} \frac{d}{dt}$. Hence (2.30) turns into

$$
\frac{dH}{d\mathcal{E}} \leq c_{*}^{\frac{1}{2}} \left[ (H^{\frac{1}{2}}E^{\frac{1}{2}} + (c_{*}E)^{\frac{1}{2}})D^{\frac{1}{2}} + \mathcal{E}^{\frac{1}{2}}D^{\frac{1}{2}} \right].
$$

We now insert the first item in (1.26), which we rewrite as $D^{-1} \leq HE^{-2} + c_{*}^{2}E^{-1}$, into the preceding differential inequality:

$$
\begin{align*}
-\frac{dH}{d\mathcal{E}} & \leq c_{*}^{\frac{1}{2}} \left[ (H^{\frac{1}{2}}E^{\frac{1}{2}} + (c_{*}E)^{\frac{1}{2}})(H^{\frac{1}{2}}E^{-1} + c_{*}E^{-\frac{1}{2}}) \\
& \quad + \mathcal{E}^{\frac{1}{2}}(H^{\frac{1}{2}}E^{-\frac{1}{2}} + c_{*}^{\frac{1}{2}}E^{-\frac{1}{2}}) \right] \\
& \leq c_{*}^{\frac{1}{2}} \left( H^{\frac{1}{2}}E^{-\frac{1}{2}} + c_{*}H^{\frac{1}{2}}E^{-\frac{1}{2}} + c_{*}^{\frac{1}{2}}H^{\frac{1}{2}}E^{-\frac{1}{2}} + c_{*}^{\frac{1}{2}} \right) \\
& \leq C_{0}c_{*}^{\frac{3}{2}}(H^{\frac{1}{2}}E^{-\frac{1}{2}} + c_{*}^{\frac{1}{2}}) \quad \text{by Young’s inequality},
\end{align*}
$$

where $C_{0} < \infty$ denotes some universal constant. This last differential inequality can be reformulated as

$$
-\frac{d(H + C_{0}c_{*}^{2}\mathcal{E})}{d\mathcal{E}} \leq C_{0}c_{*}^{\frac{1}{2}}H^{\frac{1}{2}}E^{-\frac{1}{2}} \leq C_{0}c_{*}^{\frac{1}{2}}(H + C_{0}c_{*}^{2}\mathcal{E})^{\frac{1}{2}}E^{-\frac{1}{2}},
$$

which can in turn be used to deduce

$$
-\frac{d(H + C_{0}c_{*}^{2}\mathcal{E})^{\frac{1}{2}}}{d\mathcal{E}} \leq C_{0} \frac{d}{d\mathcal{E}}(c_{*}^{2}\mathcal{E})^{\frac{1}{2}}.
$$

Integrating this differential inequality over $[\mathcal{E}, \mathcal{E}_0]$, we obtain

$$
(H + C_{0}c_{*}^{2}\mathcal{E})^{\frac{1}{2}} - (H_{0} + C_{0}c_{*}^{2}\mathcal{E}_0)^{\frac{1}{2}} \leq C_{0}(c_{*}^{2}\mathcal{E}_0)^{\frac{1}{4}},
$$

which yields (1.31).

We now turn to (1.28). To this purpose, we note that, thanks to (1.31), it follows from the first item in (1.26) that

$$
D \geq \min \{ H^{-1}\mathcal{E}^{2}, c_{*}^{-2}\mathcal{E} \} \geq \min \{ (H_{0} + c_{*}^{2}\mathcal{E}_0)^{-1}\mathcal{E}^{2}, c_{*}^{-2}\mathcal{E} \} \\
= (H_{0} + c_{*}^{2}\mathcal{E}_0)^{-1}\mathcal{E}^{2} \quad \text{because } \mathcal{E} \leq \mathcal{E}_0.
$$

We insert this into the first item in (1.25) to obtain the differential inequality

$$
\frac{d\mathcal{E}}{dt} \leq -(H_{0} + c_{*}^{2}\mathcal{E}_0)^{-1}\mathcal{E}^{2} \quad \text{and thus } \quad \frac{d}{dt}\mathcal{E}^{-1} \geq C_{0}^{-1}(H_{0} + c_{*}^{2}\mathcal{E}_0)^{-1},
$$
where $C_0 < \infty$ denotes some universal constant. Integrating over $[0,t]$ yields
\[ E^{-1} \geq C_0^{-1}(H_0 + c_0^2 E_0)^{-1} t \] and thus \[ E \leq C_0 (H_0 + c_0^2 E_0) t^{-1}, \]
as desired. This establishes (1.28).

We now turn to (1.29) and (1.30). Using $E \leq E_0$ (cf. (1.27)), we rewrite the second item of (1.26) as
\[ c_0^2 \lesssim (H_0 + c_0^2 E_0)^{1/2} E_0^{1/2}. \]

On the one hand, (1.29) follows from inserting (1.31) and (1.27) into this estimate. On the other hand, (1.30) follows from inserting (1.31) and (1.28) into this estimate.

Finally, as mentioned in Remark 4, the bound (1.32) on the dissipation follows from (1.28) and $dD/dt \lesssim D^{3/2}$; cf. Lemma 1.4. Indeed, from $dE/dt = -D$ and (1.28), it follows for any $0 < t < T$ that
\[ \int_t^T D(s) \, ds \lesssim \frac{H_0 + c_0^2 E_0}{t}. \]

On the other hand, for any $0 < s < T$, integrating the bound $dD/dt \lesssim D^{3/2}$ (i.e., $-dD/dt - 1/2 \lesssim 1$) from $s$ to $T$ yields
\[ D(s) \gtrsim D(T) \left( \frac{1}{1 + D^{3/2}(T)(T-s)} \right)^{1/2}. \]

Combining (2.31) and (2.32) leads to
\[ \frac{H_0 + c_0^2 E_0}{t} \gtrsim D(T) \int_t^T \frac{1}{\left(1 + D^{3/2}(T)(T-s)\right)} \, ds = D^{3/2}(T) \int_0^{D^{3/2}(T)(T-t)} \frac{1}{(1+\sigma)^2} \, d\sigma \geq \min \left\{ D^{3/2}(T), D(T)(T-t) \right\}. \]

Plugging $t = T/2$ into the relation above implies (1.32).

**Proof of Theorem 1.2.** According to Lemmas 1.3 and 1.4, we may apply Lemma 1.5 with $c_* := \sup_{t \in [0,t_*]} |c(t)| + 1$. In particular, we obtain by (1.29) that
\[ c_0^2 \lesssim (H_0 + c_0^2 E_0)^{1/2} E_0^{1/2} + 1, \]
and thus by Young’s inequality
\[ c_0^2 \lesssim (H_0 E_0)^{1/2} + E_0^2 + 1 \quad \text{and} \quad H_0 + c_0^2 E_0 \lesssim H_0 + E_0 + E_0^2, \]
independently of the time horizon $t_*$. In view of these bounds, the six estimates (1.27), (1.28), (1.29), (1.30), (1.31), and (1.32) translate one-to-one into the six estimates (1.10), (1.11), (1.12), (1.13), (1.14), and (1.15). 

3. **Proof of energy–energy–dissipation relationships.** In this section we establish the scaling of the energy gap and dissipation. The estimates and methods in this section are similar to those employed in [OR], although the context here (infinite domain, energy assumption $\mathcal{E} \leq 2c_0 - \epsilon$) is different from the context of that paper (bounded domain, positivity assumption $\mathcal{E} \leq 2c_0 - \epsilon$).
3.1. **Linear estimates.** Let \( v_c = v(\cdot - c) \) for some \( c \in \mathbb{R} \), with \( v \) satisfying (1.2).

In this subsection we establish energy–energy–dissipation estimates for the linearized quantities

- the linearized energy gap \( \int f_x^2 + G''(v_c)f^2 \, dx \), and
- the linearized dissipation \( \int (-f_{xx} + G'(v_c)f)_x^2 \, dx \).

**Lemma 3.1 (linearized energy gap estimate).** Suppose that \( f \in H^1 \) satisfies

\[
\int f_{xx} \, dx = 0.
\]

Then

\[
\int f^2 \, dx \lesssim \int f_x^2 + G''(v_c)f^2 \, dx.
\]

**Lemma 3.2 (linearized dissipation estimate).** For any \( f \) with \( f_x \in L^2 \) and

\[
\int f_{xx} \, dx = 0,
\]

we have

\[
\int f_x^2 \, dx \lesssim \int (-f_{xx} + G'(v_c)f)_x^2 \, dx.
\]

The first lemma is part of Proposition 3.2 of [OR]. For completeness, we give a proof below. Lemma 3.2 is new and is also proved below. In the proofs, we use the following results.

**Lemma 3.3.** For any \( f \in H^1 \), we have

\[
\int f_x^2 + G''(v_c)f^2 \, dx \geq 0.
\]

**Lemma 3.4.** Suppose that \( f \in C^3 \) satisfies \( f_x \in L^2 \) and \( f_{xx} \in L^2 \) as well as

\[
(-f_{xx} + G'(v_c)f)_x = 0.
\]

Then \( f = \alpha v_{cx} \) for some \( \alpha \in \mathbb{R} \).

**Lemma 3.5.** For all \( u \) satisfying \( \lim_{x \to \pm \infty} u(x) = \pm 1 \), we have

\[
||u||_{\infty} \lesssim 1 + E(u).
\]

Lemma 3.3, the positivity of the linearized energy, follows from the fact that \( v_c \) is an absolute energy minimizer of the nonlinear energy. The proof of Lemma 3.4 is given below. We omit the proof of Lemma 3.5, which is well known. The result follows from the classical observation of Modica and Mortola that for any points \( x_1 \leq x_2 \in \mathbb{R} \), the energy on \( (x_1, x_2) \) satisfies the lower bound

\[
E_{(x_1, x_2)}(u) \geq \int_{x_1}^{x_2} |u_x| \sqrt{2G(u(x))} \, dx \geq |\Phi(u(x_2)) - \Phi(u(x_1))|,
\]

where \( \Phi \) is the antiderivative

\[
\Phi(u) := \int_1^u \sqrt{2G(s)} \, ds.
\]
Sending $x_2 \to \infty$ and choosing $x_1$ to maximize $u$ over $\mathbb{R}$ produces the bound.

Proof of Lemma 3.1. Without loss of generality (because of translation invariance), we may assume that $c = 0$. The proof of the proposition is by contradiction. Assume that there exists an $H^1$-sequence $\{f_n\}_{n=1}^\infty$ such that

(3.5) \[ \int f_n v_x \, dx = 0, \]
(3.6) \[ \int f_n^2 \, dx = 1, \]
(3.7) \[ \int f_{nx}^2 + G''(v)f_n^2 \, dx \to 0 \quad \text{as } n \to \infty. \]

It follows that $f_n$ is uniformly bounded in $H^1$, so that we can extract a subsequence and a limit function $f \in H^1$ such that

(3.8) \[ f_n \rightharpoonup f \quad \text{weakly in } H^1, \]
(3.9) \[ f_n \to f \quad \text{in } C^0_{\text{loc}}, \]
(3.10) \[ \int f v_x \, dx = 0, \]
(3.11) \[ \int f_x^2 + G''(v)f^2 \, dx \leq 0, \]

where (3.9) follows from the embedding $H^1 \subset C^0_{\text{loc}}$ together with the Arzelà–Ascoli theorem, (3.10) follows from (3.5) and weak convergence, and (3.11) follows from (3.9) and weak lower semicontinuity.

According to Lemma 3.3, (3.11) implies that $f$ minimizes the linearized energy gap functional and hence satisfies the Euler–Lagrange equation

\[ -f_{xx} + G''(v)f = 0. \]

Given (3.9), we deduce from this equation that $f \in C^3$ and $f_{xx} \in L^2$. Lemma 3.4 then yields $f = \alpha v_x$, which together with (3.10) forces $\alpha = 0$ and hence $f \equiv 0$.

We will now show that

\[ f_n \to 0 \quad \text{in } L^2, \]

contradicting the fact that $\|f_n\|_{L^2} = 1$ for all $n$. According to (3.7), it suffices to show that

\[ \int |G'(1) - G''(v)|f_n^2 \, dx \to 0 \quad \text{as } n \to \infty. \]

On the one hand, for any fixed $X < \infty$, (3.9) implies that

\[ \int_{-X}^X |G'(1) - G''(v)|f_n^2 \, dx \to 0 \quad \text{as } n \to \infty. \]

On the other hand, the bound $\|f_n\|_{L^2} = 1$ and the exponential convergence of $v$ to 1 for $|x|$ large give

\[ \int_{\mathbb{R}\backslash(-X,X)} |G'(1) - G''(v)|f_n^2 \, dx \leq \sup_{|x| \geq X} |G'(1) - G''(v(x))| \to 0 \quad \text{as } X \to \infty \]

uniformly in $n$.  \[ \Box \]
**Proof of Lemma 3.2.** Once again, we may without loss of generality assume that $c = 0$. As we did for the energy bound, we will use an indirect argument. Suppose to the contrary that there exists a sequence of functions $f_n$ such that

\[(3.12) \quad \int f_n v_x \, dx = 0,\]
\[(3.13) \quad \int f_{nx}^2 \, dx = 1,\]
\[(3.14) \quad \int \left( \left( - f_{nxx} + G''(v)f_n \right)_x \right)^2 \, dx \to 0.\]

We claim that (3.13) improves to

\[(3.15) \quad \int \frac{1}{1 + x^2} f_n^2 + f_{nx}^2 + f_{nxx}^2 + f_{nxxx}^2 \, dx \lesssim 1.\]

The first term follows from $\int f v_x \, dx = 0$ and Lemma 2.1. By (3.13) and interpolation, the bound on $f_{nxx}$ will follow once we have established the bound on the third derivatives. To this end, we remark that for any function $f$ with $\int f v_x \, dx = 0$, we have

\[\int f_{xxx}^2 \, dx \lesssim \int (f_{xx} - G'(v)f_x)^2 \, dx + \int f_x^2 \, dx.\]

Indeed, we observe that on the one hand, we have

\[\int f_{xxx}^2 \, dx \lesssim \int (f_{xxx} - (G''(v)f)_x)^2 \, dx + \int \left( \left( G''(v)f \right)_x \right)^2 \, dx,\]

while on the other hand, we estimate

\[\int \left( \left( G''(v)f \right)_x \right)^2 \, dx \lesssim \int (G'''(v)v_x f)^2 \, dx + \int (G''(v)f_x)^2 \, dx\]

\[\lesssim \int f_x^2 \, dx,\]

where the second inequality follows from the properties of $v$ and a second application of Lemma 2.1.

It follows from the estimates (3.12) through (3.15) that we can extract a limit function $f$ such that

\[(3.16) \quad f_n \to f \quad \text{weakly in } H^3_{loc},\]
\[(3.17) \quad \int f_x^2 + f_{xx}^2 + f_{xxx}^2 \, dx \leq 1,\]
\[(3.18) \quad f_n \to f \quad \text{uniformly in } C^2_{loc},\]
\[(3.19) \quad \int f v_x \, dx = 0 \quad \text{by (3.12) and weak convergence,}\]
\[(3.20) \quad \int \left( \left( - f_{xx} + G''(v)f \right)_x \right)^2 \, dx \leq 0,\]

where the last line follows from (3.16) and weak lower semicontinuity. Naturally, one concludes from (3.20) that

\[- f_{xx} + G''(v)f \right)_x = 0\]
almost everywhere, which together with (3.18) implies in particular that \( f \in C^3 \).
According to (3.17) and (3.19), we can invoke Lemma 3.4 to conclude that \( f \equiv 0 \).

We will now show that
\[
(3.21) \quad f_{nx} \to 0 \quad \text{in } L^2,
\]
which will contradict (3.13). We claim that it suffices to establish
\[
(3.22) \quad \int \left( (G''(v) - G''(1)) f_n \right)^2 dx \to 0 \quad \text{as } n \to \infty.
\]
Indeed, (3.14), (3.22), and the triangle inequality would imply
\[
\int \left( (f_{nx} + G''(1) f_n) \right)^2 dx \to 0 \quad \text{as } n \to \infty,
\]
which, in light of
\[
\int \left( (f_{nx} + G''(1) f_n) \right)^2 dx
= \int f_{nxxx}^2 - 2 f_{nxxx} G''(1) f_{nx} + (G''(1) f_{nx})^2 dx
= \int f_{nxxx}^2 + 2 G''(1) (f_{nx})^2 + ((G''(1) f_{nx})^2 dx
\geq \int (G''(1) f_{nx})^2 dx,
\]
implies (3.21). In order to establish (3.22), we argue as at the end of the previous proof that
\[
\int \left( (G''(v) v f_n) \right)^2 dx \to 0,
\]
\[
\int \left( (G''(v) - G''(1)) f_{nx} \right)^2 dx \to 0,
\]
using on the one hand the uniform convergence of \( f_n \) and \( f_{nx} \) to zero on compact sets and on the other hand the uniform bound (3.15) together with the exponential decay of \( v_x \) and \( G''(v) - G''(1) \) for \( |x| \) large.

Proof of Lemma 3.4. As usual, we may assume without loss of generality that \( c = 0 \). By assumption, there exists \( \lambda \in \mathbb{R} \) such that
\[
(3.23) \quad f_{xx} - G''(v) f = -\lambda.
\]
From a direct calculation, it follows that
\[
(f_x v_x + \lambda v - f v_{xx})_x = f_{xx} v_x + \lambda v_x - f v_{xxx}
\overset{(3.23)}{=} (G''(v) v_x - v_{xx}) f
\overset{(1.2)}{=} 0.
\]
Hence, there exists \( \mu \in \mathbb{R} \) such that
\[
f_x v_x + \lambda v - f v_{xx} = \mu,
\]
which we can write as
\[(3.24) \quad f_x v_x - f v_{xx} = \mu - \lambda v.\]

Since \(f_x\) and \(f_{xx}\) are in \(L^2\), we have
\[
\sup_{x \in \mathbb{R}} \left( (1 + |x|)^{-\frac{1}{2}} |f(x)| \right) < \infty, \quad \sup_{x \in \mathbb{R}} \left( (1 + |x|)^{-\frac{1}{2}} |f_x(x)| \right) < \infty.
\]

Since \(v_x\) and \(v_{xx}\) decay exponentially for \(|x| \to \infty\), we conclude that \(f_x v_x - f v_{xx} \to 0\) as \(|x| \to \infty\). On the other hand, the right-hand side of (3.24) has limits \(\mu \mp \lambda\) as \(x \to \pm \infty\), so that
\[0 = \mu \mp \lambda.\]

Inserting \(\lambda = \mu = 0\) into (3.24) implies
\[
\left( \frac{f}{v_x} \right)_x = 0,
\]
and we conclude that \(f = \alpha v_x\) for some \(\alpha \in \mathbb{R}\).

**3.2. Proof of nonlinear energy estimates.** Here we prove (1.18) of Lemma 1.3, that is,
\[(3.25) \quad \mathcal{E} \sim \int f_c^2 + f_c^{2x} dx.\]

Throughout this subsection, the function \(v_c\) corresponding to a given function \(u\) is an \(L^2\)-projection satisfying (1.5). As usual,
\[f_c := u - v_c.\]

Recall also that according to Remark 3, \(u\) satisfies \(\pm 1\) boundary conditions at \(\pm \infty\). We will use this fact below.

We break the proof into a series of five steps.

**Step 1:** If \(\mathcal{E} \ll 1\), then \(||f_c||_\infty \ll 1\).

**Step 2:** If \(||f_c||_\infty \ll 1\), then \(\int f_c^2 dx \lesssim \mathcal{E}\).

**Step 3:** If \(\mathcal{E} \gtrsim 1\) and \(\mathcal{E} \leq 2c_0 - \epsilon\), then \(\int f_c^2 dx \leq \mathcal{E}\).

**Step 4:** \(\int f_c^2 dx \lesssim \mathcal{E}\) and \(\mathcal{E} \lesssim 1\) imply \(\int f_c^2 + f_c^{2x} dx \lesssim \mathcal{E}\).

**Step 5:** \(\mathcal{E} \lesssim 1\) implies \(\mathcal{E} \lesssim \int f_c^2 + f_c^{2x} dx\).

It is in Step 3 above that the constants first acquire a dependence on the \(\epsilon\) from (1.9).

For reference below, we recall for the reader the energy identity
\[(3.26) \quad E(u) - E(v_c) = \int \frac{1}{2} \left( u_x^2 - v_c^{2x} \right) + G(u) - G(v_c) \, dx \]
\[= \int \frac{1}{2} f_c^2 + G(u) - G(v_c) - G'(v_c) f_c \, dx,
\]
which in turn follows from the identity
\[0 = \int (- v_{xxx} + G'(v_c)) f_c \, dx = \int v_{cx} f_{cx} + G'(v_c) f_c \, dx.\]
**Step 1.** First we show that a small energy gap implies $L^\infty$ closeness to *some* kink, and then we will show that this improves to closeness to *the particular* kink $v_c$, which as always is the $L^2$-projection of $u$.

The boundary conditions for $u$, $v_c$ at infinity imply

$$
\int u_x \sqrt{2G(u)} \, dx = \int v_{cx} \sqrt{2G(v_c)} \, dx = E(v_c),
$$

which leads to the identity

$$
E(u) - E(v_c) = \frac{1}{2} \int (u_x - \sqrt{2G(u)})^2 \, dx.
$$

From this identity and the assumption on the energy gap, we observe that we can express

$$
u_x = \sqrt{2G(u)} + r
$$

for a function $r$ that is small in $L^2$. According to the boundary conditions, $u$ has a zero at some location $\tilde{c} \in \mathbb{R}$. The function $v_{\tilde{c}} = v(x - \tilde{c})$ satisfies

$$
v_{\tilde{c}x} = \sqrt{2G(v_{\tilde{c}})},
$$

$$
v_{\tilde{c}}(\tilde{c}) = 0.
$$

Therefore, according to ODE theory, for any $\varepsilon > 0$ and $X < \infty$, there exists $\delta > 0$ such that

$$
\int r^2 \, dx \leq \delta \quad \Rightarrow \quad \sup_{(\tilde{c}-X, \tilde{c}+X)} |u - v_{\tilde{c}}| \leq \varepsilon.
$$

We now argue that $u$ is also close to $v_{\tilde{c}}$ on the interval $(\tilde{c} + X, \infty)$ and $(-\infty, \tilde{c} - X)$. Notice that because of the behavior of $v_{\tilde{c}}$ and by choosing $X \gg 1$, this amounts to arguing that

$$(3.27) \quad \sup_{(\tilde{c}+X, \infty)} |u - 1| \ll 1 \quad \text{and} \quad \sup_{(-\infty, \tilde{c}-X)} |u - (-1)| \ll 1,
$$

where we emphasize that the smallness is uniform with respect to $\tilde{c}$.

According to our previous argument, uniformly with respect to $\tilde{c}$ we have the estimates

$$(3.28) \quad \begin{cases}
u(\tilde{c} + X) \geq -\varepsilon + v_{\tilde{c}}(\tilde{c} + X) \geq 1 - 2\varepsilon & \text{for } X \text{ large}, \\
u(\tilde{c} - X) \leq \varepsilon + v_{\tilde{c}}(\tilde{c} - X) \leq 1 + 2\varepsilon & \text{for } X \text{ large}.
\end{cases}
$$

We conclude by the usual method of Modica and Mortola (cf. (3.4)) that the energy of $u$ satisfies

$$
E(\tilde{c} - X, \tilde{c} + X)(u) \lesssim c_0.
$$

The assumption on the energy gap then implies

$$(3.29) \quad E(\tilde{c} + X, \infty)(u) \ll 1, \quad E(-\infty, \tilde{c} - X)(u) \ll 1.
$$
The method of Modica and Mortola together with (3.29) and the boundary conditions at ±∞ yields (3.27), so that we have now deduced uniform $L^\infty$ closeness to some kink:

$$||u - v_{\varepsilon}||_\infty \ll 1.$$  

We now improve from closeness to an arbitrary kink $v_{\varepsilon}$ to closeness to the $L^2$-projection $v_c$. For this we use the orthogonality condition, which we write as

$$0 = \int f_c v_{cx} \, dx = \int \left( (u - v_{\varepsilon}) + (v_{\varepsilon} - v_c) \right)v_{cx} \, dx.$$  

From the properties of $v_c$ and the equality

$$\left| \int (v_{\varepsilon} - v_c)v_{cx} \, dx \right| = \left| \int (u - v_c)v_{cx} \, dx \right|$$  

$$\leq ||u - v_c||_\infty \int v_{cx} \, dx \ll 1,$$

we conclude $|\tilde{c} - c| \ll 1$ and hence also

$$||u - v_{\varepsilon}||_\infty \ll 1.$$  

**Step 2.** The smallness of $||f_c||_\infty$ from Step 1 allows us to pass from the linear to the nonlinear estimate of the energy gap when $E \ll 1$. We add and subtract a quadratic term in (3.26) to express the energy gap in the form

$$E(u) - E(v_c) = \int \left[ \frac{1}{2}(f_{cx}^2 + G''(v_c)f_c^2) \right] \, dx$$  

$$+ \int \left( G(u) - G(v_c) - G'(v_c)f_c - \frac{1}{2}G''(v_c)f_c^2 \right) \, dx$$  

$$\geq \int \left[ \frac{1}{2}(f_{cx}^2 + G''(v_c)f_c^2) \right] \, dx - C||f_c||_\infty \int f_c^2 \, dx,$$

where we have used Taylor’s formula and $||f_c||_\infty \lesssim 1$ to write

$$\left| G(u) - G(v_c) - G'(v_c)f_c - \frac{1}{2}G''(v_c)f_c^2 \right| \lesssim f_c^3.$$  

Applying the linear bound from Lemma 3.1 to $f_c$, and invoking the smallness of $||f_c||_\infty$, the estimate in (3.30) improves to $\int f_c^2 \, dx \lesssim E$, as desired.

**Step 3.** According to the assumption, we have $E(u) \leq 3c_0 - \varepsilon$. We say that $u$ has a $\delta$-transition layer on $(x_-, x_+)$ if $u(x_-) = -1 + \delta$ and $u(x_+) = 1 - \delta$, or vice versa, and if $(x_-, x_+)$ is minimal in the sense that the same is not true for any proper subset of $(x_-, x_+)$. Choosing $\delta = \delta(\varepsilon)$ sufficiently small with respect to $\varepsilon$, the Modica–Mortola trick (cf. (3.4)) and the energy bound imply that $u$ can have at most two $\delta$-transitions. But then by the boundary conditions, $u$ can have at most one and moreover must have exactly one such $\delta$-transition.

If we define $\tilde{c} := \inf\{x: u(x) = 1 - \delta\}$, then we have

$$u < 1 - \delta \quad \text{for} \quad x < \tilde{c}, \quad u > 1 + \delta \quad \text{for} \quad x > \tilde{c},$$

so that, in particular, there exists $C(\delta) < \infty$ such that

$$(u - (-1))^2 \leq C(\delta)G(u) \quad \text{for} \quad x < \tilde{c},$$

$$\quad \left| u - 1 \right|^2 \leq C(\delta)G(u) \quad \text{for} \quad x > \tilde{c}.$$  

(3.31)
Let \( \chi_\varepsilon \) denote the characteristic function
\[
\chi_\varepsilon(x) = \begin{cases} 
-1, & x < \tilde{c}, \\
1, & x \geq \tilde{c}.
\end{cases}
\]

From (3.31) we have
\[
\int |u - \chi_\varepsilon|^2 \, dx \leq C(\delta) \int G(u) \, dx \lesssim C(\delta),
\]
while, on the other hand, the properties of \( v_\varepsilon \) imply that
\[
\int (v_\varepsilon - \chi_\varepsilon)^2 \, dx \lesssim 1.
\]
Hence, the triangle inequality returns
\[
\int |u - v_\varepsilon|^2 \, dx \lesssim C(\delta).
\]
Finally, we observe that
\[
\int |u - v_\varepsilon|^2 \, dx \leq \int |u - v_{\varepsilon}|^2 \, dx \lesssim C(\delta) \lesssim C(\delta) \mathcal{E},
\]
where the first inequality follows because \( v_\varepsilon \) is the \( L^2 \)-projection of \( u \) and the last inequality comes from the assumed lower bound on the energy gap.

**Step 4.** Given the assumption on \( \int f_\varepsilon^2 \, dx \), it is enough to show
\[
\int f_\varepsilon^2 \, dx \lesssim \mathcal{E} + \int f_{\varepsilon}^2 \, dx.
\]
This estimate is an immediate consequence of the formula (3.26), Taylor’s formula, and the uniform bound on \( u \) from Lemma 3.5.

**Step 5.** This step also follows immediately from the identity (3.26), Taylor’s formula, and the uniform bound on \( u \) from Lemma 3.5.

### 3.3. Proof of nonlinear dissipation estimates

Here we prove estimate (1.19) from Lemma 1.3, that is,
\[
D \sim \int f_{\varepsilon}^2 + f_{\varepsilon}^2 + f_{\varepsilon}^2 \, dx.
\]
As in the preceding subsection, the function \( v_\varepsilon \) corresponding to a given function \( u \) is the \( L^2 \)-projection satisfying (1.5) and, as usual, \( f_\varepsilon := u - v_\varepsilon \). Recall also that according to Remark 3, \( u \) satisfies \( \pm 1 \) boundary conditions at \( \pm \infty \). We will use this fact below.

The proof consists of five steps.

**Step 1:** \( \mathcal{E} \lesssim 2c_0 - \epsilon \) and \( D \ll 1 \) imply \( \|f_\varepsilon\|_\infty \ll 1. \)

**Step 2:** \( \|f_\varepsilon\|_\infty \ll 1 \) implies \( \int f_\varepsilon^2 \, dx \lesssim D. \)

**Step 3:** \( D \gtrsim 1 \) and \( \mathcal{E} \lesssim 2c_0 - \epsilon \) imply \( \int f_\varepsilon^2 \, dx \lesssim D. \)

**Step 4:** \( \int f_\varepsilon^2 \, dx \lesssim D \) and \( \mathcal{E} \lesssim 1 \) imply \( \int f_{\varepsilon}^2 + f_{\varepsilon}^2 + f_{\varepsilon}^2 \, dx \lesssim D. \)

**Step 5:** \( \mathcal{E} \lesssim 1 \) implies \( D \lesssim \int (f_{\varepsilon}^2 + f_{\varepsilon}^2) \, dx. \)
Here the relevant identity is
\[-u_{xx} + G'(u) = -f_{cxx} + G'(u) - G'(v_c)\]
\[= (-f_{cxx} + G''(v_c)f_c) + (G'(u) - G'(v_c) - G''(v_c)f_c).\]  
(3.33)

In both Steps 1 and 3 above, the constants depend on the \(\epsilon\) from (1.9).

**Step 1.** As in Step 2 for the energy bound, the point is to use the smallness of \(||f_c||_\infty\) from Step 1 to pass from a linear to a nonlinear dissipation estimate. To this end, we invoke identity (3.33) and the triangle inequality to write
\[
\int (\left((u_{xx} - G'(u))_x\right)^2 \, dx \geq \frac{1}{2} \int (\left(-f_{cxx} + G''(v_c)f_c\right)_x^2 \, dx
- \int (\left(G'(u) - G'(v_c) - G''(v_c)f_c\right)_x^2 \, dx.
\]

According to the linear dissipation bound from Lemma 3.2 and the Hardy-type inequality (2.13), it now suffices to argue
\[
\int (G'(u) - G'(v_c) - G''(v_c)f_c)_x^2 \, dx \lesssim ||f_c||_\infty \int \frac{1}{(x-c)^2 + 1} f_c^2 + f_{cxx}^2 \, dx.
\]
(3.34)

To show (3.34), we use the Taylor representation
\[
G'(u) - G'(v_c) - G''(v_c)(u - v_c)
= (u - v_c)^2 \int_0^1 G'''(v_c + \theta (u - v_c))(1 - \theta) \, d\theta,
\]
from which we deduce
\[
\left| (G'(u) - G'(v_c) - G''(v_c)f_c)_x \right|
= 2f_c f_{cxx} \int_0^1 G'''(v_c + \theta f_c)(1 - \theta) \, d\theta
+ f_c^2 \int_0^1 (v_{cxx} + \theta f_{cxx})G^{(4)}(v_c + \theta f_c)(1 - \theta) \, d\theta
\lesssim |f_c f_{cxx}| + |v_{cxx}| f_c^2 + f_{cxx} f_c^2,
\]
where we have used \(||f_c||_\infty \lesssim 1\) to bound the factors involving the third and fourth derivatives of \(G\). Recalling \(|f_c| \ll 1\) and the exponential decay of \(v_c\) at infinity, this leads immediately to (3.34).

**Step 3.** Here we use the lower bound on the energy gap from (3.25) to deduce
\[
\int f_{cxx}^2 \, dx \overset{(3.25)}{\lesssim} E \lesssim 1 \lesssim D \quad (\text{since } E \lesssim 1, \ D \gtrsim 1).
\]

**Step 4.** Given the interpolation inequality
\[
\int f_{cxxx}^2 \, dx \lesssim \int f_{cxx}^2 + f_{cxxx}^2 \, dx,
\]
it is enough to show that
\[ \int f_{xxx}^2 \, dx \lesssim D + \int f_{c}^2 \, dx. \]

We will use (3.33) in the form
\[ (3.35) \quad \int \left( -f_{xxx} + (G'(u) - G'(v_c)) \right)^2 \, dx = D. \]

Indeed, from (3.35) and the triangle inequality, it suffices to show that
\[ (3.36) \quad \int \left( (G'(u) - G'(v_c)) \right)^2 \, dx \lesssim \int f_{c}^2 \, dx. \]

To this end, we invoke the Taylor formula
\[ G'(u) - G'(v_c) = (u - v_c) \int_0^1 G''(v_c + \theta(u - v_c))(1 - \theta) \, d\theta \]
to write
\[ (G'(u) - G'(v_c))_x = f_{c} \int_0^1 G''(v_c + \theta(u - v_c))(1 - \theta) \, d\theta \]
\[ + f_{c} \int_0^1 G'''(v_c + \theta f_{c})(1 - \theta)(v_{cx} + \theta f_{cx}) \, d\theta. \]

Recalling the uniform bound on \( u \) (and hence \( f_{c} \)) inherited from the energy bound (see Lemma 3.5), we deduce
\[ \int \left( (G'(u) - G'(v_c)) \right)_x^2 \, dx \lesssim \int f_{c}^2 \, dx + \int v_{cx}^2 f_{c}^2 \, dx. \]

Lemma 2.1 and the exponential decay of \( v_{cx} \) then imply (3.36).

**Step 5.** The combination of (3.35), the triangle inequality, and (3.36) immediately yields the upper bound on the dissipation.

**Step 1.** Finally, we turn to the task of showing that a small dissipation implies \( L^\infty \) closeness of \( u \) to the kink \( v_c \). The argument breaks into four steps: (a) a uniform bound on \( |u_{xx} - G'(u)| \), (b) a uniform bound on the so-called discrepancy \( u_{xx}^2 / 2 - G(u) \), (c) an ODE argument that deduces from this information the uniform closeness to some kink state \( v_c \), and (d) refinement to uniform closeness to the particular kink state \( v_c \).

For convenience, let us define
\[ g := u_{xx} - G'(u). \]

We begin by showing that an average of \( g \) is small on large intervals. Consider an interval \( I_* \subset \mathbb{R} \) of length \( L \gg 1 \) to be fixed below. Let \( I \) be the interval extended by \( L \) to the left and right. Define a positive cut-off function \( \eta : I \to [0, 1] \) such that
\[ \eta = 1, \quad x \in I_*, \]
\[ \eta = \eta_x = 0, \quad x \text{ on the boundary of } I, \]
\[ |\eta_x| \lesssim \frac{1}{L} \quad \text{and} \quad |\eta_{xx}| \lesssim \frac{1}{L^2}. \]
We define the \( \eta \)-average of \( g \) as
\[
\langle g \rangle_\eta := \frac{\int g \eta \, dx}{\int \eta \, dx}.
\]
Recall the energy bound and, by Lemma 3.5, the uniform bound on \( u \). It follows that
\[
|\langle g \rangle_\eta| \lesssim \frac{1}{L} \left| \int_I u \eta \, dx \right| = \frac{1}{L} \left| \int_I \left| \eta \right| \, dx \right|
\]
\[
\lesssim \frac{1}{L^2} \left| \int_I \left| u \right| \, dx + \frac{1}{L} \left( \int_I |G'(u)| \, dx \right)^{\frac{1}{2}} \right|
\]
\[
\lesssim \frac{1}{L^2} \int_I 1 + G(u) \, dx + \frac{1}{L^2} \left( \int_I G(u) \, dx \right)^{\frac{1}{2}}
\]
(3.37)

On the one hand, we use this fact together with the Poincaré inequality
(3.38)
\[
\int_I g - \langle g \rangle_\eta \, dx \lesssim L^2 \int_I g_x^2 \, dx
\]
to conclude that
\[
\int_I g^2 \, dx \lesssim L |\langle g \rangle_\eta|^2 + L^2 \int_I g_x^2 \, dx
\]
(3.39)
\[
\lesssim 1 + L^2 \int_I g_x^2 \leq 1 + L^2 D.
\]
On the other hand, we estimate
\[
\sup_I |\eta g| \lesssim |\langle g \rangle_\eta| + \int_I |(\eta g)_x| \, dx
\]
\[
\leq |\langle g \rangle_\eta| + \int_I |\eta_x g| + |\eta g_x| \, dx
\]
\[
\lesssim \frac{1}{L^2} + \frac{1}{L^2} \left( \int_I g^2 \, dx \right)^{\frac{1}{2}} + \left( \int_I g_x^2 \, dx \right)^{\frac{1}{2}}
\]
(3.37)
\[
\lesssim \frac{1}{L^2} + (LD)^{\frac{1}{2}} \lesssim D^{\frac{1}{2}},
\]
where for the last inequality we have optimized in \( L \). According to the definition of \( \eta \), we recover
\[
\sup_I |g| = \sup_I |\eta g| \leq \sup_I |\eta g| \lesssim D^{\frac{1}{2}}.
\]
By translation invariance, this improves to the estimate on the full line:
(3.40)
\[
||g||_\infty \lesssim D^{\frac{1}{2}}.
\]
Our next task is to turn this bound into a bound on the so-called discrepancy
\[ \xi := \frac{1}{2} u_x^2 - G(u). \]

Let \( L \gg 1 \) be a large number to be fixed below. Because of the bound on the energy, we have for any interval \( I \subset \mathbb{R} \) of length \( L \) that
\begin{equation}
\frac{1}{L} \int_I \xi \, dx \leq \frac{1}{L} \int_I \left( \frac{1}{2} u_x^2 + G(u) \right) \, dx \lesssim \frac{1}{L}.
\end{equation}

On the other hand, we have the derivative bound
\begin{equation}
\int_I |\xi_x| \, dx = \int_I |u_x g| \, dx \\
\leq ||g||_{L^\infty} \left( \int_I u_x^2 \, dx \right)^{1/2} \\
\lesssim D^{1/2} L^{1/2},
\end{equation}
where in the last line we have recalled the energy bound and the estimate (3.40). Combining (3.41) and (3.42) yields
\begin{equation}
\sup_I |\xi| \leq \frac{1}{L} \int_I \xi \, dx + \int_I |\xi_x| \, dx \\
\lesssim \frac{1}{L} + D^{1/2} L^{1/2}.
\end{equation}

Optimizing in \( L \) gives \( L \sim D^{-1/6} \) and \( \sup_I |\xi| \lesssim D^{1/6} \). Once again taking into account translation invariance, we obtain
\[ ||\xi||_{L^\infty} \lesssim D^{1/6}. \]

We now turn our attention to an ODE argument for the distance from \( u \) to a shifted kink. We observe that, according to the boundary conditions, \( u \) has at least one zero. Consider any zero \( \tilde{c} \) of \( u \). Then \( u \) solves
\begin{equation}
\begin{cases}
 u_x = \pm \sqrt{2(G(u) + \xi)}, \\
 u(\tilde{c}) = 0,
\end{cases}
\end{equation}
with \( \sup |\xi| \ll 1 \). Notice that \( G(0) > 0 \) implies that a unique sign is taken in the differential equation on a neighborhood of \( \tilde{c} \). Assume for the moment that the equation is satisfied with a positive sign so that \( u \) is increasing to the right of \( \tilde{c} \). Let \( x_m > \tilde{c} \) denote the first maximum point of \( u \) greater than \( \tilde{c} \) if it exists and \( \infty \) otherwise. Now consider the kink \( v_{\tilde{c}} \) that satisfies
\[ v_x = \sqrt{2G(v)}, \]
\[ v(\tilde{c}) = 0. \]

Then by the smallness of the discrepancy and standard ODE theory, we have
\begin{equation}
\sup_{(\tilde{c}, \min\{x_m, \tilde{c}+X\})} |u - v_{\tilde{c}}| \ll 1 \quad \text{for } X \gg 1.
\end{equation}
If $x_m$ is a maximum of $u$, then $u_x(x_m) = 0$ and $|\xi| \ll 1$ implies $G(u(x_m)) \ll 1$. But then, since $u(x_m) > 0$, we conclude that $u(x_m) \approx 1$. By applying (3.44), we deduce $v(x_m) \approx 1$, which in turn yields $x_m \gg 1$. Obviously the same statement applies if $x_m = \infty$.

Hence (3.44) improves to

$$\sup_{(\tilde{c}, \tilde{c}+X)} |u - v_{\tilde{c}}| \ll 1 \quad \text{for} \quad X \gg 1.$$ 

The analogous argument to the left of $\tilde{c}$ returns

$$\sup_{(\tilde{c}-X, \tilde{c}+X)} |u - v_{\tilde{c}}| \ll 1 \quad \text{for} \quad X \gg 1.$$ 

The argument above and the method of Modica and Mortola (cf. (3.4)) imply in particular that zeros are well separated and that there is a contribution to the energy of close to $c_0$ in the neighborhood of any zero. Because of the energy bound $E(u) \leq 3c_0 - \epsilon$ and the boundary conditions at $\pm \infty$, $u$ therefore must have one and only one zero $\tilde{c}$.

Now consider what happens to the right of $\tilde{c} + X$. According to the preceding argument, $u$ is positive on $(\tilde{c} + X, \infty)$. Because $|\xi| \ll 1$, any local extremum $u_*$ must satisfy $G(u_*) \approx 0$; hence, by the positivity of $u$, any such point $u_*$ must satisfy $u_* \approx 1$. Applying the analogous argument on $(-\infty, \tilde{c} - X)$, we conclude

$$\begin{cases} 
    u \approx +1 & \text{on} \ (\tilde{c} + X, \infty), \\
    u \approx -1 & \text{on} \ (-\infty, \tilde{c} - X).
\end{cases}$$

Combining this fact with the behavior of $v_{\tilde{c}}$ on the corresponding intervals, we recover global $L^\infty$ closeness:

$$||u - v_{\tilde{c}}|| \ll 1.$$ 

The final step is to move from closeness to an arbitrary kink $v_{\tilde{c}}$ to closeness to the $L^2$-projection $v_{\tilde{c}}$. For this we use the orthogonality condition as in Step 1 of the proof of (1.18).

REFERENCES


