

Document Version

Final published version

Citation (APA)

Gennip, Y. V., Giga, Y., & Okamoto, J. (2026). Graph Gradient Flows: From Discrete to Continuum. In *Infosys Science Foundation Series in Mathematical Sciences* (pp. 163-250). (Infosys Science Foundation Series in Mathematical Sciences; Vol. Part F893). Springer. https://doi.org/10.1007/978-981-97-9812-4_2

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

**Green Open Access added to [TU Delft Institutional Repository](#)
as part of the Taverne amendment.**

More information about this copyright law amendment
can be found at <https://www.openaccess.nl>.

Otherwise as indicated in the copyright section:
the publisher is the copyright holder of this work and the
author uses the Dutch legislation to make this work public.

Chapter 2

Graph Gradient Flows: From Discrete to Continuum



Yves van Gennip, Yoshikazu Giga, and Jun Okamoto

Abstract This paper gives a framework to study a continuum limit of a gradient flow on a graph where the number of vertices increases in an appropriate way. As examples, we prove the convergence of a discrete total variation flow and a discrete Allen–Cahn flow on discretised tori to their respective continuum limits.

2.1 Introduction

We prove the convergence of two graph-based discrete gradient flows, the total variation flow, and the Allen–Cahn flow, to a continuum limiting problem when the graph approximates a Euclidean space. Although characterisation of continuum limits has been studied variationally at the energy level quite extensively in the past decade, mainly Γ -convergence results, the convergence problem for discrete flows has only been attracting more attention fairly recently.

If $\Phi : H \rightarrow \mathbb{R}$ is a Fréchet differentiable function on a Hilbert space $(H, \langle \cdot, \cdot \rangle)$, with gradient $\nabla_H \Phi$, then the gradient flow of Φ (w.r.t. the inner product on H) is a function $u : (0, \infty) \rightarrow H$ that satisfies $\dot{u} = -\nabla_H \Phi(u)$. If Φ is not differentiable, but convex, the differential equation is replaced by the differential inclusion $\dot{u} \in -\partial\Phi(u)$, where $\partial\Phi(u)$ denotes the subdifferential of Φ at u . The function Φ is often called the *energy*, even when there is no direct physical interpretation as

Y. van Gennip

Delft Institute of Applied Mathematics, Technische Universiteit Delft, Delft, The Netherlands

e-mail: y.vangennip@tudelft.nl

Y. Giga (✉)

Graduate School of Mathematical Sciences, The University of Tokyo, Meguro-ku, Tokyo, Japan

e-mail: labgiga@ms.u-tokyo.ac.jp

J. Okamoto

Faculty of Medicine Bldg. B, Institute for the Advanced Study of Human Biology, Kyoto University, Sakyo-ku, Kyoto, Japan

e-mail: okamoto.jun.8n@kyoto-u.ac.jp

such. An equivalent formulation can be given in terms of an *evolution variational inequality* (see Definition 2.1; recent in-depth overviews are given in [4, 52]). Such a formulation offers more flexibility when generalising the notion of gradient flow to metric spaces that lack a Hilbert space structure. Additionally, the variational setting can be exploited when comparing gradient flows that are formulated on different underlying spaces. We make crucial use of this advantage to prove convergence of gradient flows that are formulated on discrete graphs to a continuum gradient flow.

In the current work, we consider two families of gradient flows: total variation gradient flows with the energy Φ being the graph-based or continuum total variation functional (Sect. 2.4.1) and Allen–Cahn gradient flows with the energy Φ being the graph-based or continuum Allen–Cahn functional (also known as the Ginzburg–Landau functional or Modica–Mortola functional; Sect. 2.5.1).

To be able to prove convergence results, and thus to compare flows formulated on a graph with flows formulated on a continuum, we require a way to embed the discrete flows in continuum space. In this paper we restrict ourselves to graphs that are obtained as a regular (square or (hyper)cubic) discretisation of the flat torus. In that setting, we can associate functions on the graph with functions on the torus via piecewise-constant or piecewise-linear interpolation. We use the former for our convergence results for total variation flow and the latter for Allen–Cahn flow.

To avoid additional technical complexities, for the Allen–Cahn flows we only consider the one-dimensional case.

The two main results in this paper are Theorems 2.4 and 2.6. The former states that (under assumptions that will be given in detail later) the L^2 -difference between the (piecewise-constant interpolant of the) graph-based total variation flow and the continuum total variation flow is controlled by the L^2 -difference of their initial conditions. In particular, if both flows start from the same initial condition, they are equal. It has been already noted in [42] that if a function is constant on rectangles, it remains constant on those rectangles under the anisotropic ℓ^1 -total variation flow. The relation to the piecewise-constant interpolant of the graph-based total variation flow, however, was not mentioned there. Analysis of the space-discrete total variation flow by means of mode-decomposition methods is presented in the recent preprint [17].

The second main result, Theorem 2.6, states that (again under assumptions that will be specified in due time) on any finite time interval the difference between the (piecewise-linear interpolant of the) graph-based and continuum Allen–Cahn flows converges to zero, uniformly in time, as the mesh size of the discretisation of the one-dimensional torus goes to zero.

We start the paper with an abstract general framework based on evolution variational inequalities [5] for gradient flows. Within this framework we can derive Theorem 2.1 and Corollary 2.1 which, after carefully casting the total variation flow problem in the language of the abstract setting and confirming that the required assumptions are satisfied in Sects. 2.3 and 2.4, immediately lead to Theorem 2.4. Proving our result for the Allen–Cahn flows requires more work, because in this case it is not possible to work with piecewise-constant embeddings of the graph-based functions (since the continuum Allen–Cahn functional is not well-defined for functions with jumps) and the embedding via linear interpolation that we use instead

is not an isometry, which means the results from our abstract framework are not directly applicable. Moreover, inequality (2.6) between the graph-based functions Φ_h and continuum function Φ does not hold when Φ_h and Φ are the graph-based and continuum Allen–Cahn functionals, respectively. Since this inequality is one of the requirements for Theorem 2.1, we need to resort to using Theorem 2.2 instead, which requires a weaker condition (2.9) to be satisfied, but at the cost of leading to a weaker conclusion, namely only an asymptotic comparison (when the mesh size approaches zero), rather than an explicit estimate at positive mesh size.

In this paper we prove the convergence of the flows using the evolution variational inequality formulation of the flows directly. An alternative approach is possible based on [21, Theorem 4.11], which requires Γ -convergence of the discrete energies Φ_h to the continuum energy Φ , asymptotic local equicoercivity of the energies Φ_h [21, Definition 4.4], and well-preparedness of the initial conditions. The Γ -convergence results may be obtained in a way similar to the discrete-to-continuum Γ -convergence results in [61]. Alternatively, we see that condition (2.6) or condition (2.9) allows for the construction of recovery sequences in the proof of Γ -convergence, and lower semicontinuity of Φ together with (2.11) implies the limit inferior inequality that is required for Γ -convergence. The (asymptotically local) equicoercivity condition follows from condition (2.5) together with compactness properties of sublevel sets of Φ . We conclude that the a priori alternative approach from [21] corresponds to our approach using Theorem 2.2, or at the very least, can be employed after establishing the same conditions we use. We note that Theorem 2.1 does not require the limit inferior condition, but can only be expected to be useful in those special occasions where the discrete gradient flows are also continuum gradient flows (after embedding)—a situation which is not expected to be typical.

We note that the ideas from [21] (and relatedly [53]) have been generalised to functions that are not geodesically λ -convex (see Sect. 2.2 for a definition) by Sandier and Serfaty and others [8, 31, 52, 56–58]. A general framework provided by Mielke requires convergence of metrics as well as Γ -convergence of functionals to conclude convergence of gradient flows [46, Theorem 3.4.3]. It is assumed that the sets underlying the metric spaces in the sequence are independent of the limit parameter but the metrics can depend on it. We suspect that a suitable modification of such an idea can lead to similar convergence results as ours, but we do not pursue that course in the present paper. For other literature about this energy-dissipation-principle (EDP) convergence, which can be seen as an extension of the work by Sandier and Serfaty, we refer to [22, 47–50]. A key ingredient is the characterisation of gradient flows as minimisers (at value zero) of a De Giorgi functional. For example, in [26, Theorem 3.14] this formulation is used to establish the continuum limit of nonlocal-interaction equations on graphs as gradient flows in a probability space equipped with Wasserstein distance.

As examples of other recent literature that is interested in the convergence of graph-based gradient flows to continuum gradient flows, we mention [35, 36], in which the limit of the graph-based p -Laplacian evolution problem is established.

The paper is structured as follows. The general abstract setting is introduced in Sect. 2.2, and the setup of the graphs that we use, with embeddings and projec-

tions into and from the continuum space, is given in Sect. 2.3. The results from this section will be applied to obtain our two main convergence results, which are given in Sects. 2.4 (total variation flow) and 2.5 (Allen–Cahn flow). The main part of the paper ends with conclusions and suggestions for future work in Sect. 2.6, after which follow appendices that give detailed proofs for some properties of total variation for piecewise-constant functions (Appendix A2.7), properties of the operator Γ which is important in Sect. 2.5 (Appendix B2.8), and two Poincaré–Wirtinger-type inequalities that we need in the proof of Lemma 2.7 (Appendix C2.9).

2.2 An Abstract Framework

Let (M, d) be a complete metric space with a distance d . Let γ be a continuous function from an interval I to M , i.e., $\gamma \in C(I, M)$. Following [5] we say that γ is an absolutely continuous curve if there is an integrable function¹ f defined on I such that, for all $t, s \in I$,

$$d(\gamma(t), \gamma(s)) \leq \int_s^t f(r) dr.$$

For any pair of metric spaces M and \tilde{M} , by $C(M; \tilde{M})$ we will denote the space of continuous functions with domain M and codomain \tilde{M} . If additionally M and \tilde{M} are differentiable manifolds and k is a positive integer k , then $C^k(M; \tilde{M})$ will be the space of k times continuously differentiable functions with domain² M and codomain \tilde{M} . If $\tilde{M} = \mathbb{R}$, we write $C^k(M)$. Moreover, $C^\infty(M; \tilde{M}) = \bigcup_{k \in \mathbb{N}} C^k(M; \tilde{M})$.

A curve $\gamma : (0, \infty) \rightarrow M$ is locally absolutely continuous if it is absolutely continuous in any bounded interval $(a, b) \subset (0, \infty)$.

In this paper we use the notation Φ for a functional $\Phi : M \rightarrow \mathbb{R} \cup \{+\infty\}$. For $\lambda \in \mathbb{R}$, such a functional is called λ -convex on a curve $\gamma \in C([0, 1], M)$ if

$$\Phi(\gamma(t)) \leq (1 - t)\Phi(\gamma(0)) + t\Phi(\gamma(1)) - \frac{1}{2}\lambda t(1 - t)d^2(\gamma(0), \gamma(1))$$

for all $t \in [0, 1]$. We note that, if $\lambda_2 \leq \lambda_1$ and Φ is λ_1 -convex on γ , Φ is also λ_2 -convex on γ .

A curve $\gamma \in C([0, 1], M)$ is a constant-speed geodesic (from $\gamma(0)$ to $\gamma(1)$) if

$$d(\gamma(s), \gamma(t)) = d(\gamma(0), \gamma(1))(t - s) \quad \text{for all } s, t \in [0, 1] \quad \text{with } s \leq t.$$

¹ We note that if such an f exists, it necessarily is nonnegative almost everywhere on I .

² In a slight and common abuse of notation, we will sometimes write $u \in C^k(M; \tilde{M})$ if $u : \hat{M} \rightarrow \tilde{M}$ with $\hat{M} \supset M$ and $u|_M \in C^k(M; \tilde{M})$.

A functional Φ is said to be geodesically λ -convex if for any $z_0, z_1 \in \mathcal{D}(\Phi)$ there is a constant-speed geodesic $\gamma \in C([0, 1], M)$ such that $\gamma(0) = z_0, \gamma(1) = z_1$, and such that Φ is λ -convex on the curve γ . Here $\mathcal{D}(\Phi) := \{z \in M \mid \Phi(z) < \infty\}$ is the effective domain of Φ .

Remark 2.1 In a Hilbert space, geodesic λ -convexity of Φ is nothing but³ convexity of $\Phi(\cdot) - \frac{\lambda \|\cdot\|^2}{2}$. This follows from the fact that in a (real or complex) Hilbert space for each pair of distinct points a and b , there is a unique constant-speed geodesic from a to b , given by the straight line of the form $\gamma(t) = a + t(b - a)$. It can be checked directly that this is indeed a constant-speed geodesic. To prove uniqueness, let z be a point on a constant-speed geodesic $\tilde{\gamma}$ from $a := \tilde{\gamma}(0)$ to $b := \tilde{\gamma}(1)$, not equal to a or to b . Then, by the triangle inequality, $\|x + y\| \leq \|x\| + \|y\|$ for $x := z - a$, and $y := b - z$. By the definition of a constant-speed geodesic, if $t \in (0, 1)$ is such that $\tilde{\gamma}(t) = z$, we require

$$\begin{aligned} \|x\| &= \|z - a\| = \|b - a\|t = \|x + y\|t \quad \text{and} \\ \|y\| &= \|b - z\| = \|b - a\|(1 - t) = \|x + y\|(1 - t), \end{aligned} \tag{2.1}$$

so we require equality in the triangle inequality. In the case of a Hilbert space, this holds if and only if $(\|x\| + \|y\|)^2 - \|x + y\|^2 = 2(\|x\|\|y\| - \langle x, y \rangle) = 0$, where $\langle x, y \rangle$ denotes the inner product. Hence we require equality in the Cauchy–Schwarz inequality $\langle x, y \rangle \leq \|x\|\|y\|$, which is equivalent to x and y being linearly dependent, i.e., in geometric terms, x and y are parallel. Thus there exists a scalar c such that $y = cx$ and thus $b - z = c(z - a)$. We note that $c \neq -1$, since $a \neq b$. It follows that $z = \frac{1}{1+c}b + \frac{c}{1+c}a$. Moreover, by (2.1), $\|x + y\| = |1 + c|\|x\| = \frac{1}{t}\|x\|$ and thus $|\frac{1}{1+c}| = t \in (0, 1)$. Hence we can write $\frac{1}{1+c} = te^{i\vartheta}$ where $\vartheta := \text{Arg}(\frac{1}{1+c})$. We write i for the imaginary unit. Thus

$$\|\tilde{\gamma}(t) - b\| = \|z - b\| = |1 - te^{i\vartheta}| \|a - b\|.$$

By continuity of $\tilde{\gamma}$, if we take $t \uparrow 1$, we find $|1 - te^{i\vartheta}| \|a - b\| = 0$. Since $a \neq b$, this implies $e^{i\vartheta} = 1$. Hence $t = \frac{1}{1+c}$ and thus $\tilde{\gamma}(t) = z = tb + (1 - t)a = a + t(b - a)$.

We note that the uniqueness may not be true for a Banach space even if it has finite dimension. A simple example is \mathbb{R}^2 equipped with the ℓ^1 -norm,⁴ so that the unit ball is a square.

If $\lambda = 0$, then a geodesically λ -convex Φ is called geodesically convex; if $\lambda \leq 0$, then such Φ is often called geodesically semiconvex.⁵

³ Thus in particular, in a Hilbert space setting geodesic 0-convexity is equivalent to convexity in the sense of vector spaces.

⁴ See Sect. 2.3.1.

⁵ This terminology is not consistent in the literature: sometimes $\lambda < 0$ is required for semiconvexity, while some sources even allow any $\lambda \in \mathbb{R}$.

Definition 2.1 (*Evolution variational inequality formulation of gradient flow*)

Assume that Φ is geodesically λ -convex. A gradient flow of Φ is a locally absolutely continuous curve $u : (0, \infty) \rightarrow M$, which satisfies the evolution variational inequality

$$\frac{1}{2} \frac{d}{dt} d^2(u(t), v) + \frac{1}{2} \lambda d^2(u(t), v) \leq \Phi(v) - \Phi(u(t))$$

for almost every $t > 0$ and for all $v \in \mathcal{D}(\Phi)$. Formally we write $\dot{u} \in -\partial\Phi(u)$. If we wish to emphasise the space or metric in which the gradient flow takes place, we speak of an M -gradient flow.

Remark 2.2 Consider the setting of Definition 2.1. For a fixed $v \in \mathcal{D}(\Phi)$, the function $w \mapsto d^2(w, v)$ is Lipschitz continuous on any bounded subset of M . By continuity of u , if $0 < T_1 < T_2$, then $u([T_1, T_2]) \subset M$ is bounded. Since the composition of a Lipschitz-continuous and absolutely continuous function is again absolutely continuous, on any interval $[T_1, T_2]$, $t \mapsto d^2(u(t), v)$ is absolutely continuous and thus its derivative exists for almost every $t \in [T_1, T_2]$. As $T_1 > 0$ is arbitrary, the derivative $\frac{d}{dt} d^2(u(t), v)$ exists for almost every $t \in (0, \infty)$.

Lemma 2.1 *Let Φ be as in Definition 2.1. If $c \in M$ is such that, for all $v \in \mathcal{D}(\Phi)$, $\Phi(v) \geq \Phi(c)$, then the constant curve $u = c$ is a gradient flow of Φ .*

Proof Let $v \in \mathcal{D}(\Phi)$. If γ is a constant-speed geodesic from c to v , then by the geodesic λ -convexity of Φ we have for all $t \in [0, 1]$,

$$\Phi(\gamma(t)) \leq (1 - t)\Phi(c) + t\Phi(v) - \frac{1}{2} \lambda t(1 - t) d^2(c, v).$$

By assumption Φ has a global minimum at c , hence

$$t\Phi(c) \leq \Phi(\gamma(t)) - \Phi(c) + t\Phi(c) \leq t\Phi(v) - \frac{1}{2} \lambda t(1 - t) d^2(c, v).$$

Let $t \neq 0$ and divide by t to find $\Phi(c) \leq \Phi(v) - \frac{1}{2} \lambda (1 - t) d^2(c, v)$. Taking the limit $t \downarrow 0$, we conclude that $\Phi(c) \leq \Phi(v) - \frac{1}{2} \lambda d^2(c, v)$ and thus, since $\frac{d}{dt} d^2(c, v) = 0$, $\frac{1}{2} \frac{d}{dt} d^2(c, v) + \frac{1}{2} \lambda d^2(c, v) \leq \Phi(v) - \Phi(c)$. Therefore the constant curve $u = c$ satisfies the inequality from Definition 2.1 and hence is a gradient flow of Φ . \square

Lemma 2.2 *Let Φ be as in Definition 2.1 and additionally assume that Φ is lower semicontinuous. Let u be a gradient flow of Φ with $u((0, \infty)) \subset \mathcal{D}(\Phi)$ and continuous on $[0, \infty)$. Then $t \mapsto \Phi(u(t))$ is a non-increasing function on $[0, \infty)$.*

Moreover, if $I \subset [0, \infty)$ is a nondegenerate interval, Φ is as in Definition 2.1 (not necessarily lower semicontinuous) and u is as above in the first part of this lemma, then for almost every $t \in I$, $\Phi(u(t)) = \Phi(u(\inf I)) < +\infty$ if and only if, for all $t \in I$, $u(t) = u(\inf I)$.

Proof First we prove the following claim: If $v \in M$, then $t \mapsto d^2(u(t), v)$ is locally absolutely continuous on $(0, \infty)$. To prove this, let $(a, b) \subset (0, \infty)$, since u is locally absolutely continuous on $(0, \infty)$, if $a < b < \infty$, there exists an integrable function f on (a, b) such that, for all $s, t \in (a, b)$,

$$d(u(t), u(s)) \leq \int_s^t f(r) dr.$$

Since u is continuous on $[a, b]$, so is $t \mapsto d(u(t), v)$; hence this function is bounded on (a, b) . In particular, there is a $C > 0$ such that, for all $s, t \in (a, b)$, $|d(u(t), v) + d(u(s), v)| \leq C$; hence, for all $s, t \in (a, b)$,

$$\begin{aligned} |d^2(u(t), v) - d^2(u(s), v)| &= |d(u(t), v) + d(u(s), v)| |d(u(t), v) - d(u(s), v)| \\ &\leq C d(u(t), u(s)) \\ &\leq \int_s^t C f(r) dr, \end{aligned}$$

where we used the reverse triangle inequality for the first inequality. Hence $t \mapsto d^2(u(t), v)$ is absolutely continuous on (a, b) and thus locally absolutely continuous on $(0, \infty)$.

Next we prove that $\Phi \circ u$ is non-increasing. Let $t^* \in [0, \infty)$. If $\Phi(u(t^*)) = +\infty$, then it is trivially true that, for all $t \geq t^*$, $\Phi(u(t)) \leq \Phi(u(t^*))$. Now assume that $\Phi(u(t^*)) < +\infty$, so that $u(t^*) \in \mathcal{D}(\Phi)$. Define

$$S := \{t > t^* : \Phi(u(t)) > \Phi(u(t^*))\}.$$

Since Φ is lower semicontinuous and u is (absolutely) continuous, $\Phi \circ u$ is lower semicontinuous; hence S is an open set. For a proof by contradiction, assume that $S \neq \emptyset$, then S is an (at most) countable union of pairwise disjoint open intervals [54, Theorem 2.24]. Since S is bounded below, so is each of these intervals. Let $J \subset S$ be such an interval. Since S is open, $a := \inf S \in [t^*, \infty) \setminus S$, thus $\Phi(u(a)) \leq \Phi(u(t^*))$. In particular, $a \notin J$.

For notational convenience, define $g : J \rightarrow \mathbb{R}$ by $g(t) := d^2(u(t), u(a))$. Recall that g is locally absolutely continuous on J .

In the inequality in Definition 2.1, choose $v = u(a)$. Then, for almost every $t \in J$,

$$\frac{1}{2} \left[\frac{d}{dt} g(t) + \lambda g(t) \right] \leq \Phi(u(a)) - \Phi(u(t)) \leq \Phi(u(t^*)) - \Phi(u(t)) < 0.$$

We apply a Grönwall argument: We have, for almost every $t \in J$,

$$\frac{d}{dt} (e^{\lambda t} g(t)) = e^{\lambda t} \left[\frac{d}{dt} g(t) + \lambda g(t) \right] < 0.$$

Let $t \in J$ and integrate the inequality over (a, t) , using $g(a) = 0$, to find

$$e^{\lambda t} g(t) - e^{\lambda a} g(a) = e^{\lambda t} g(t) < 0$$

and hence

$$g(t) < 0,$$

which contradicts $g(t) = d^2(u(t), u(a)) \geq 0$. Therefore $S = \emptyset$ and thus, for all $t \geq t^*$, $\Phi(u(t)) \leq \Phi(u(t^*))$, which proves the first claim of the lemma.

To prove the second claim, let $I \subset [0, \infty)$ be a nondegenerate interval such that, for almost every $t \in I$, $\Phi(u(t)) = \Phi(u(b)) < +\infty$, where $b := \inf I$. The “if” statement follows directly since by assumption, if $t \in I \setminus \{0\}$, then $\Phi(u(t)) < +\infty$. For the “only if” claim, note that, by the inequality in Definition 2.1 with $v = u(b)$, we have for almost every $t \in I$, $h'(t) + \lambda h(t) \leq 0$, where $h(t) := d^2(u(t), u(b))$. By a similar argument as above for g , we find that $h(t) \leq 0$, hence $h(t) = 0$, and thus $u(t) = u(b)$ for all $t \in I$. \square

Remark 2.3 In the first part of Lemma 2.2 above, we assume that Φ is lower semicontinuous. In fact, in the proof we only use the fact that $\Phi \circ u$ is lower semicontinuous and thus the conclusions of the lemma remain true under this weaker assumption.

Remark 2.4 Definition 2.1 is a natural extension of a gradient flow in a Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \|$. By [5, Remark 1.1.3], in the Hilbert space setting the derivative $\dot{u}(t)$ of an absolutely continuous curve u exists for a.e. t . Assume t is such that $\dot{u}(t)$ exists. If Φ is convex ($\lambda = 0$), then by definition of the subdifferential of a convex function [24, Chap. I, Sect. 5.1] we see that $\dot{u}(t) \in -\partial\Phi(u(t))$ is equivalent to, for all $v \in \mathcal{D}(\Phi)$,

$$\langle -\dot{u}(t), v - u(t) \rangle \leq \Phi(v) - \Phi(u(t)).$$

The left-hand side equals $\frac{d}{dt} \frac{\|u(t) - v\|^2}{2}$ and thus u satisfies the evolution variational inequality (Definition 2.1) for $\lambda = 0$ if and only if $\dot{u}(t) \in -\partial\Phi(u(t))$.

For $\lambda \neq 0$ we interpret $\dot{u} \in -\partial\Phi(u)$ (for those t for which $\dot{u}(t)$ exists) as

$$\dot{u}(t) \in -\partial \left(\Phi - \lambda \frac{\| \cdot \|^2}{2} \right) (u(t)) - \lambda u(t), \tag{2.2}$$

which is equivalent to, for all $v \in \mathcal{D}(\Phi)$,

$$\langle -\dot{u}(t) - \lambda u(t), v - u(t) \rangle \leq \Phi(v) - \lambda \frac{\|v\|^2}{2} - \left(\Phi(u(t)) - \lambda \frac{\|u(t)\|^2}{2} \right).$$

Since

$$\langle \lambda u(t), u(t) - v \rangle + \lambda \frac{\|v\|^2}{2} - \lambda \frac{\|u(t)\|^2}{2} = \lambda \frac{\|u(t) - v\|^2}{2},$$

the function u satisfies the evolution variational inequality if and only if (2.2) is satisfied.

We recall a basic unique solvability result. The next result is a special case of [5, Theorem 4.0.4], where more general Φ are allowed.

Proposition 2.1 *Assume that $\Phi : M \rightarrow \mathbb{R} \cup \{+\infty\}$ is geodesically λ -convex and lower semicontinuous with $\Phi \not\equiv +\infty$. Assume furthermore that Φ is locally bounded below at some point $u_* \in \mathcal{D}(\Phi)$, i.e., there exists an $r_* > 0$ such that*

$$m_* := \inf \{ \Phi(v) \mid v \in M \text{ and } d(v, u_*) < r_* \} > -\infty.$$

Then for all $u_0 \in \overline{\mathcal{D}(\Phi)}$ (the closure of $\mathcal{D}(\Phi)$ in the topology of M), there exists a solution $u \in C([0, \infty), M)$ to $\dot{u} \in -\partial\Phi(u)$ satisfying $u(0) = u_0$. Furthermore, for all $t > 0$, $u(t) \in \mathcal{D}(\Phi)$. Moreover, such a solution is unique among all locally absolutely continuous curves u for which $u(t) \rightarrow u_0$ as $t \downarrow 0$.

Remark 2.5 Without the assumption of lower semicontinuity of Φ in Proposition 2.1 above, global-in-time existence with $u(t) \in \mathcal{D}(\Phi)$ for all $t > 0$ is not guaranteed, even though a local-in-time solution may exist. Consider for example the function $\Phi : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$, defined by

$$\Phi(v) := \begin{cases} -v, & \text{if } v < 0, \\ +\infty, & \text{if } v \geq 0, \end{cases}$$

which is convex and locally bounded (at all points in its effective domain), but not lower semicontinuous. From Definition 2.1—with $\lambda = 0$ and the Euclidean distance on \mathbb{R} for d —it follows that, for any $a > 0$, $u(t) = t - a$ satisfies the evolution variational inequality for $t < a$. However, at $t = a$ the curve leaves the effective domain of Φ , i.e., $u(a) \notin \mathcal{D}(\Phi)$. Moreover, any function \tilde{u} that satisfies the evolution variational inequality locally in time must be of the form $\tilde{u}(t) = t - b$, for some $b > 0$. Indeed, if $\Phi(\tilde{u}(t)) = +\infty$, the evolution variational inequality does not hold, thus $\tilde{u}(t) < 0$ has to hold for almost every $t > 0$ and thus, by continuity, for all $t > 0$. By Remark 2.4, the evolution variational inequality in this case is equivalent to the differential inclusion $\dot{\tilde{u}}(t) \in \partial\Phi(\tilde{u}(t)) = \{1\}$, where the equality follows from differentiability of Φ on $(-\infty, 0)$ with $\Phi'(v) = -1$ [24, Proposition 5.3]. Thus $\tilde{u}(t) = t - b$, as claimed and no global-in-time solutions exist. On the other hand, if we define $\Phi(0) := 0$ instead, then Φ is lower semicontinuous and indeed the function $u(t) = \min(t - a, 0)$ is a global solution.

Definition 2.2 If u is a gradient flow of Φ and $u(t) \rightarrow u_0 \in \overline{\mathcal{D}(\Phi)}$ as $t \downarrow 0$, then u is called a gradient flow of Φ starting from u_0 .

If the assumptions from Proposition 2.1 are satisfied, uniqueness allows us to call u the gradient flow of Φ starting from u_0 .

In the setting of Proposition 2.1, $\Phi(u(t)) < +\infty$ for all $t > 0$. Hence the evolution variational inequality from Definition 2.1, which was assumed to hold for all $v \in \mathcal{D}(\Phi)$, trivially also holds for all $v \in M \setminus \mathcal{D}(\Phi)$.

The results when M is a Hilbert space go back to the theory of maximal monotone operators [9] pioneered by Kōmura [41], where the first existence result has been established.

If M is a normed vector space, in particular if M is a Banach space, the assumption $m_* > -\infty$ for some u_* and r_* is automatically fulfilled. Indeed, by the Hahn–Banach theorem, any lower semicontinuous convex function that is not identically equal to $+\infty$ is bounded below by an affine continuous function [10, Proposition 1.10] and thus in particular locally bounded below. Since $\Phi - \lambda \|\cdot\|^2/2$ is convex for λ -convex functions Φ , such Φ are locally bounded from below if they are lower semicontinuous and not identically equal to $+\infty$. In specific instances, still we may note local boundedness below explicitly to emphasise this condition is satisfied.

Returning to the case of a general complete metric space M , to observe the uniqueness part of Proposition 2.1, let u and v be two solutions and let $s, t \in (0, \infty)$ be two “times” at which the evolution variational inequality holds, then

$$\frac{1}{2} \frac{d}{dt} d^2(v(t), u(s)) + \frac{1}{2} \lambda d^2(v(t), u(s)) \leq \Phi(u(s)) - \Phi(v(t))$$

and

$$\frac{1}{2} \frac{d}{dt} d^2(u(t), v(s)) + \frac{1}{2} \lambda d^2(u(t), v(s)) \leq \Phi(v(s)) - \Phi(u(t)).$$

Adding both sides and observing that

$$\frac{d}{dt} d^2(u(t), v(t)) = \frac{d}{dt} d^2(u(t), v(s)) \Big|_{s=t} + \frac{d}{dt} d^2(v(t), u(s)) \Big|_{s=t},$$

we take $s = t$ to get

$$\frac{1}{2} \frac{d}{dt} d^2(u(t), v(t)) + \lambda d^2(u(t), v(t)) \leq 0.$$

Via an integration, this implies a generalised contraction principle: for all $t \geq 0$,

$$d^2(u(t), v(t)) \leq e^{-2\lambda t} d^2(u(0), v(0)). \tag{2.3}$$

If the initial data $u(0)$ and $v(0)$ are the same, then we conclude that $u \equiv v$. This idea is useful to define our convergence result.

In what follows, we consider a family of complete metric spaces $(M_h, d_h)_{h \in \mathfrak{H}}$ approximating (M, d) as $h \rightarrow 0$ in some sense which is to be made precise. Here \mathfrak{H} is a nonempty subset of $(0, 1)$ of which 0 is a limit point. In particular, we allow \mathfrak{H} to be a proper subset of $(0, 1)$. In fact, from Sect. 2.3 onwards we only consider $h \in (0, 1)$ for which $\frac{1}{h}$ is an integer.

Let $i_h : M_h \rightarrow M$ be an embedding map which is an isometry, i.e., $d_h(v_1, v_2) = d(i_h v_1, i_h v_2)$, for all $v_1, v_2 \in M_h$. In particular, i_h is continuous. Let $p_h : M \rightarrow M_h$ be a mapping such that $p_h \circ i_h$ is the identity map on M_h . This is (in spirit⁶) a projection, in the sense that $p_h \circ i_h \circ p_h = p_h$ on M . We assume the following non-strict contraction property (i.e., 1-Lipschitz continuity) holds:

$$d_h(p_h w_1, p_h w_2) \leq d(w_1, w_2) \quad \text{for all } w_1, w_2 \in M. \quad (2.4)$$

In particular, p_h is continuous.

We want to consider functionals $\Phi : M \rightarrow \mathbb{R} \cup \{+\infty\}$ and, for all $h \in \mathfrak{H}$, $\Phi_h : M_h \rightarrow \mathbb{R} \cup \{+\infty\}$, for which gradient flows exist. Hence we will require that these functionals satisfy the conditions of Proposition 2.1, i.e., they are geodesically λ -convex, lower semicontinuous, not identically equal to $+\infty$, and locally bounded below at some point in their respective domains. If, for all $h \in \mathfrak{H}$, the functionals Φ_h and Φ satisfy these conditions with the same λ , we will say Φ_h and Φ satisfy condition (λ -AGS), in deference to [5, Theorem 4.0.4]. We note that we do not require the points in the effective domain at which local lower boundedness is satisfied to be the same for all functionals, nor the values of the lower bounds or the radii of the local neighbourhoods on which the bounds hold.

Additionally, we sometimes assume any or all of the following. When we do so, we will explicitly state this:

$$\Phi(i_h v) \leq \Phi_h(v) \quad \text{for all } h \in \mathfrak{H} \text{ and all } v \in M_h \quad (2.5)$$

$$\Phi_h(p_h w) \leq \Phi(w) \quad \text{for all } h \in \mathfrak{H} \text{ and all } w \in M. \quad (2.6)$$

$$d_h^2(v, p_h w) + d^2(i_h p_h w, w) = d^2(i_h v, w) \quad \text{for all } h \in \mathfrak{H}, \text{ all } v \in M_h, \text{ and all } w \in M. \quad (2.7)$$

The last assumption is similar to the Pythagorean theorem and thus says that as a projection p_h is “orthogonal”, in a sense. A typical example where this Pythagorean theorem holds is the case where M and all M_h are inner product spaces with inner products $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_h$, respectively, which satisfy $\langle i_h u, i_h v \rangle = \langle u, v \rangle_h$ and

$$\langle i_h p_h w - w, i_h u \rangle = 0, \quad (2.8)$$

for all $h \in \mathfrak{H}$, all $w \in M$, and all $u, v \in M_h$.

Inner product structures are not necessary for (2.7) to hold, as the following example shows. If (M, d) is a complete metric space such that d cannot be generated by an inner product and, for all $h \in \mathfrak{H}$, $(M_h, d_h) = (M, d)$, then we can define, for all $h \in \mathfrak{H}$, both i_h and p_h to be equal to the identity operator on M . In that case i_h

⁶ We will call p_h a projection, even though $p_h \circ p_h$ is not necessarily well-defined, let alone equal to p_h .

is an isometry, p_h satisfies (2.4), and $p_h \circ i_h$ is the identity operator. Moreover, (2.7) holds, since, for all $h \in \mathfrak{H}$ and for all $v, w \in M_h = M$,

$$d_h^2(v, p_h w) + d^2(i_h p_h w, w) = d^2(v, w) + d^2(w, w) = d^2(v, w) = d^2(i_h v, w).$$

Condition (2.5) implies that $i_h(\mathcal{D}(\Phi_h)) \subset \mathcal{D}(\Phi)$. Similarly, condition (2.6) implies that $p_h(\mathcal{D}(\Phi) \subset \mathcal{D}(\Phi_h)$.

Since $p_h \circ i_h$ is the identity map on M_h , we note that (2.5) and (2.6) together imply that, for all $v \in M_h$,

$$\Phi(i_h v) \leq \Phi_h(v) = \Phi_h(p_h i_h v) \leq \Phi(i_h v)$$

and thus $\Phi(i_h v) = \Phi_h(v)$.

On the other hand, if $\Phi(i_h v) = \Phi_h(v)$ holds for all $v \in M_h$, this is not sufficient to imply (2.6), as the latter condition has to hold for all⁷ $w \in M$ and while we can conclude that $\Phi_h(p_h i_h v) = \Phi(i_h v)$ for all $v \in M_h$, i_h is typically not surjective.

Theorem 2.1 *Assume there exists a $\lambda \leq 0$ such that Φ_h and Φ satisfy condition (λ -AGS). Assume that $\overline{\mathcal{D}(\Phi_h)} = M_h$. Let u^h be the gradient flow of Φ_h starting from $u_0^h \in M_h$. Assume moreover (2.5), (2.6), and (2.7). Then $i_h u^h$ is the gradient flow of Φ starting from $i_h u_0^h$.*

Corollary 2.1 *Assume there exists a $\lambda \leq 0$ such that Φ_h, M_h, Φ, M , and u^h satisfy all the hypotheses from Theorem 2.1. Let u be the gradient flow of Φ starting from $u_0 \in M = \overline{\mathcal{D}(\Phi)}$. Then*

$$d^2(i_h u^h(t), u(t)) \leq e^{-2\lambda t} d^2(i_h u_0^h, u_0), \quad \text{for all } t > 0.$$

If $i_h u_0^h \rightarrow u_0$ as $h \rightarrow 0$, this yields, for all $T > 0$, uniform convergence of $i_h u^h$ to u on $(0, T)$ as $h \rightarrow 0$. For example if $i_h p_h u_0 \rightarrow u_0$ in M as $h \rightarrow 0$ and $u_0^h = p_h u_0$, then $i_h u^h$ converges to u in M uniformly on $(0, T)$ as $h \rightarrow 0$, for any $T > 0$. Corollary 2.1 follows from Theorem 2.1 by the generalised contraction principle (2.3).

Proof (of Theorem 2.1) By definition our assumptions (2.5) and (2.6) yield

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} d_h^2(u^h, p_h w) + \frac{\lambda}{2} d_h^2(u^h, p_h w) &\leq \Phi_h(p_h w) - \Phi_h(u^h) \\ &\leq \Phi(w) - \Phi(i_h u^h) \quad \text{for all } w \in M. \end{aligned}$$

By (2.7) we see

$$\frac{d}{dt} d_h^2(u^h, p_h w) = \frac{d}{dt} d^2(i_h u^h, w),$$

⁷ This is related to the requirement in Definition 2.1 that the evolution variational inequality holds for all $v \in \mathcal{D}(\Phi)$.

since $d^2(i_h p_h w, w)$ does not depend on t . We thus conclude that

$$\frac{1}{2} \frac{d}{dt} d^2(i_h u^h, w) + \frac{\lambda}{2} d^2(i_h u^h, w) \leq \Phi(w) - \Phi(i_h u^h), \quad \text{for all } w \in M,$$

since $\lambda \leq 0$ and $d_h(u^h, p_h w) = d_h(p_h i_h u^h, p_h w) \leq d(i_h u^h, w)$ by the non-strict contraction property (2.4). This says that $i_h u^h$ is a gradient flow of Φ . Since $u^h(t) \rightarrow u_0^h$ as $t \downarrow 0$ and since i_h is continuous, we have $i_h u^h(t) \rightarrow i_h u_0^h$ as $t \downarrow 0$. Thus $i_h u^h$ is a gradient flow starting at u_0^h . \square

The next theorem gives convergence under a weaker assumption. Instead of (2.6) we consider the weaker assumption

$$\overline{\lim}_{h \rightarrow 0} \Phi_h(p_h w) \leq \Phi(w) \quad \text{for all } w \in M \quad (2.9)$$

and the bound in (2.10). Instead of (2.5) we assume (2.11).

Theorem 2.2 *Assume there exists a $\lambda \leq 0$ such that Φ_h and Φ satisfy condition (λ -AGS) and let conditions (2.7) and (2.9) be satisfied. Let u^h be the gradient flow of Φ_h starting from $u_0^h \in M_h = \overline{\mathcal{D}(\Phi_h)}$ and let u be the gradient flow of Φ starting from $u_0 \in M = \overline{\mathcal{D}(\Phi)}$. Assume further that there exists a $T \in (0, \infty]$, a $\delta > 0$, and a nonnegative function $\Psi : M \rightarrow \mathbb{R} \cup \{+\infty\}$ such that the function $t \mapsto e^{2\lambda t} \Psi(u(t))$ is in $L^1(0, T)$ ⁸ and, for all $w \in M$ and for all $h \in (0, \delta) \cap \mathfrak{H}$,*

$$\Phi_h(p_h w) \leq \Psi(w). \quad (2.10)$$

Moreover, let there be a function $\varepsilon : (0, 1) \rightarrow (0, \infty)$ such that $\varepsilon(h) \rightarrow 0$ as $h \rightarrow 0$ and such that, for all $t \in [0, T]$ and for all $h \in (0, \delta) \cap \mathfrak{H}$,

$$\Phi(i_h u^h(t)) \leq \Phi_h(u^h(t)) + \varepsilon(h). \quad (2.11)$$

If $i_h u_0^h \rightarrow u_0$ and

- $T \in (0, \infty)$ or
- $T = \infty$, $\lambda < 0$, and $\liminf_{t \rightarrow \infty} \Phi(u(t)) > -\infty$,

then

$$\lim_{h \rightarrow 0} \sup_{t \in [0, T]} e^{2\lambda t} d(i_h u^h(t), u(t)) = 0.$$

If $T = \infty$, we interpret $[0, T]$ as $[0, \infty)$.

⁸ By $L^1(0, T)$ we denote the set of Lebesgue integrable functions on $(0, T)$.

Proof As in the proof of uniqueness, leading up to (2.3), we observe that

$$\frac{1}{2} \frac{d}{dt} d_h^2(u^h(t), p_h u(s)) + \frac{\lambda}{2} d_h^2(u^h(t), p_h u(s)) \leq \Phi_h(p_h u(s)) - \Phi_h(u^h(t)), \tag{2.12}$$

$$\frac{1}{2} \frac{d}{dt} d^2(u(t), i_h u^h(s)) + \frac{\lambda}{2} d^2(u(t), i_h u^h(s)) \leq \Phi(i_h u^h(s)) - \Phi(u(t)), \tag{2.13}$$

for all s and t at which the evolution variational inequality is satisfied. By the chain rule and (2.7), we see that

$$\begin{aligned} \frac{d}{dt} d^2(u(t), i_h u^h(t)) &= \frac{d}{dt} d^2(u(t), i_h u^h(s)) \Big|_{s=t} + \frac{d}{dt} d^2(i_h u^h(t), u(s)) \Big|_{s=t} \\ &= \frac{d}{dt} d^2(u(t), i_h u^h(s)) \Big|_{s=t} + \frac{d}{dt} d_h^2(u^h(t), p_h u(s)) \Big|_{s=t}. \end{aligned}$$

Adding (2.12) and (2.13) we now observe that

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} d^2(u(t), i_h u^h(t)) + \lambda d^2(u(t), i_h u^h(t)) \\ &\leq \frac{1}{2} \frac{d}{dt} d^2(u(t), i_h u^h(t)) + \frac{\lambda}{2} d^2(u(t), i_h u^h(t)) + \frac{\lambda}{2} d_h^2(u^h(t), p_h u(t)) \\ &\leq \Phi(i_h u^h(t)) - \Phi_h(u^h(t)) + \Phi_h(p_h u(t)) - \Phi(u(t)) \\ &\leq \varepsilon(h) + \Phi_h(p_h u(t)) - \Phi(u(t)), \end{aligned}$$

where for the first inequality we used that, by the non-strict contraction property (2.4), $d_h^2(u^h(t), p_h u(t)) = d_h^2(p_h i_h u^h(t), p_h u(t)) \leq d^2(i_h u^h(t), p_h u(t))$ and that $\lambda \leq 0$. The final inequality follows from (2.11).

For any real-valued function f that is differentiable at t , we have

$$e^{-2\lambda t} \frac{d}{dt} (e^{2\lambda t} f(t)) = \frac{df}{dt}(t) + 2\lambda f(t).$$

Using this identity for $f(t) := \frac{d}{dt} d^2(u(t), i_h u^h(t))$, we multiply the inequality above by $2e^{2\lambda t}$ and integrate both sides of $[0, t]$, for a fixed $t \in [0, T]$, to get

$$\begin{aligned} e^{2\lambda t} d^2(u(t), i_h u^h(t)) &\leq d^2(u_0, i_h u_0^h) \\ &\quad + 2 \int_0^t e^{2\lambda \tau} [\Phi_h(p_h u(\tau)) - \Phi(u(\tau)) + \varepsilon(h)] d\tau. \end{aligned} \tag{2.14}$$

Next we claim that if a family of measurable functions f_h on $[0, T]$ satisfies, for almost every $t \in (0, T)$, $\overline{\lim}_{h \rightarrow 0} f_h(t) \leq 0$ and there exists a $g \in L^1(0, T)$ with $g \geq 0$

such that, for all h , $f_h \leq g$ on $(0, T)$, then

$$\overline{\lim}_{h \rightarrow 0} \sup_{0 < t < T} \int_0^t f_h(\tau) d\tau \leq 0. \quad (2.15)$$

We emphasise that the case $T = \infty$ is allowed. To prove (2.15), first we note that, for all $\eta > 0$, there exists a $t_{h,\eta} \in (0, T)$ such that

$$\sup_{0 < t < T} \int_0^t f_h(\tau) d\tau \leq \int_0^{t_{h,\eta}} f_h(\tau) d\tau + \eta.$$

For all h and for all $\eta > 0$, we set

$$\bar{f}_h(\tau) := \begin{cases} f_h(\tau) & \text{if } 0 \leq \tau \leq t_{h,\eta}, \\ 0 & \text{if } t_{h,\eta} < \tau \leq T. \end{cases}$$

Since $\bar{f}_h \leq g$, applying Fatou's lemma [28, Sect. 1.3, Theorem 1] for $g - \bar{f}_h \geq 0$ yields

$$\int_0^T \overline{\lim}_{h \rightarrow 0} (g - \bar{f}_h)(\tau) d\tau \leq \overline{\lim}_{h \rightarrow 0} \int_0^T (g - \bar{f}_h)(\tau) d\tau.$$

Since $g \in L^1(0, T)$, this implies

$$\int_0^T \overline{\lim}_{h \rightarrow 0} (-\bar{f}_h(\tau)) d\tau \leq \overline{\lim}_{h \rightarrow 0} \left(- \int_0^T \bar{f}_h(\tau) d\tau \right)$$

and thus

$$\overline{\lim}_{h \rightarrow 0} \int_0^T \bar{f}_h(\tau) d\tau \leq \int_0^T \overline{\lim}_{h \rightarrow 0} \bar{f}_h(\tau) d\tau.$$

We observe that, for almost all $\tau \in (0, T)$, $\overline{\lim}_{h \rightarrow 0} \bar{f}_h(\tau) \leq 0$ and hence

$$\overline{\lim}_{h \rightarrow 0} \int_0^{t_{h,\eta}} f_h(\tau) d\tau = \overline{\lim}_{h \rightarrow 0} \int_0^T \bar{f}_h(\tau) d\tau \leq 0.$$

Thus we obtain that

$$\overline{\lim}_{h \rightarrow 0} \sup_{0 < t < T} \int_0^t f_h(\tau) d\tau \leq 0 + \eta.$$

Sending η to zero, we establish our claim in (2.15).

Writing

$$f_h(\tau) := e^{2\lambda\tau} [\Phi_h(p_h u(\tau)) - \Phi(u(\tau)) + \varepsilon(h)]$$

and taking the supremum over $(0, T)$ on both sides of inequality (2.14), we get

$$\sup_{0 < t < T} e^{2\lambda t} d^2(u(t), i_h u^h(t)) \leq d^2(u_0, i_h u_0^h) + 2 \sup_{0 < t < T} \int_0^t f_h(\tau) d\tau. \quad (2.16)$$

We define

$$\tilde{g}(\tau) := e^{2\lambda\tau} [\Psi(u(\tau)) - \Phi(u(T)) + \varepsilon(h)] \quad \text{and} \quad g(\tau) := (\tilde{g}(\tau))^+,$$

where, for $x \in \mathbb{R}$, $x^+ := \max(0, x)$ denotes the positive part of x . In the case that $T = \infty$, we set $\Phi(u(\infty)) := \liminf_{t \rightarrow \infty} \Phi(u(t)) > -\infty$.

By construction $g \geq 0$. By (2.10) we have that, for all $h \in (0, \delta) \cap \mathfrak{H}$ and all $\tau \in [0, T]$, $\Phi_h(p_h u(\tau)) \leq \Psi(u(\tau))$. Furthermore, by Lemma 2.2, the function $t \mapsto \Phi(u(t))$ is non-increasing on $[0, \infty)$. Therefore the functions f_h satisfy, for all $h \in (0, \delta) \cap \mathfrak{H}$, $f_h \leq \tilde{g} \leq g$ on $(0, T)$.

Moreover, by assumption the function $t \mapsto e^{2\lambda t} \Psi(u(t))$ is in $L^1(0, T)$ and hence, if $T < \infty$, then $\tilde{g} \in L^1(0, T)$ and thus $g \in L^1(0, T)$. In the case where $T = \infty$, we also have assumed that $\lambda < 0$ and $\Phi(u(\infty)) > -\infty$. Hence it follows that $e^{2\lambda t} [-\Phi(u(\infty)) + \varepsilon(h)]^+$ is integrable on $(0, \infty)$. Since $g(\tau) \leq e^{2\lambda\tau} \{[\Psi(u(\tau))]^+ + [-\Phi(u(\infty)) + \varepsilon(h)]^+\}$, we again deduce that $g \in L^1(0, T)$.

By Proposition 2.1 we know that, for all $t > 0$, $u(t) \in M$ and therefore, by assumption (2.9) and $\lim_{h \rightarrow 0} \varepsilon(h) = 0$, we have that, for all $t \in (0, T)$, $\overline{\lim}_{h \rightarrow 0} f_h(t) \leq 0$. Thus all the required assumptions for (2.15) are satisfied. Applying (2.15) to the inequality in (2.16), we obtain

$$\overline{\lim}_{h \rightarrow 0} \sup_{0 < t < T} e^{2\lambda t} d^2(u(t), i_h u^h(t)) \leq \overline{\lim}_{h \rightarrow 0} d^2(u_0, i_h u_0^h) + 0.$$

Since $\underline{\lim}_{h \rightarrow 0} \sup_{0 < t < T} d^2(u(t), i_h u^h(t)) \geq 0$, the result follows. □

Remark 2.6 If $T < \infty$ in Theorem 2.2, we may replace the assumption of integrability of the function $t \mapsto e^{2\lambda t} \Psi(u(t))$ on $(0, T)$ by an assumption of integrability of $t \mapsto \Psi(u(t))$ on $(0, T)$, since the latter implies the former. Moreover, in that case the conclusion of the theorem,

$$\lim_{h \rightarrow 0} \sup_{t \in [0, T]} e^{2\lambda t} d(i_h u^h(t), u(t)) = 0,$$

implies that

$$\lim_{h \rightarrow 0} \sup_{t \in [0, T]} d(i_h u^h(t), u(t)) = 0,$$

since, by nonnegativity of λ ,

$$\sup_{t \in [0, T]} e^{2\lambda t} d(i_h u^h(t), u(t)) \leq \sup_{t \in [0, T]} d(i_h u^h(t), u(t))$$

$$\leq e^{-2\lambda T} \sup_{t \in [0, T]} e^{2\lambda t} d(i_h u^h(t), u(t)).$$

We use Theorem 2.1 and Corollary 2.1 to prove convergence of discrete total variation flows to a continuum total variation flow in Theorem 2.4. These results are not applicable in our proof of convergence of discrete Allen–Cahn flows to a continuum Allen–Cahn flow in Theorem 2.6, as explained in detail in Sect. 2.5.4. Instead, we first approximate the continuum flow by different discrete flows and use Theorem 2.2 to prove convergence of those new flows to the continuum flow in Theorem 2.7, after which Theorem 2.8 establishes that the new flows approximate the discrete flows, resulting in Theorem 2.6.

2.3 Setting of the Problem

We consider an anisotropic total variation flow and an Allen–Cahn flow on “cubic” graphs and their convergence to their continuum versions when the mesh size of the graphs tends to zero. In this section we define the graphs, the corresponding metric spaces M_h and their continuum counterpart M , the embeddings i_h and “projections” p_h , and the functionals which generate the gradient flows.

2.3.1 The Graphs

We construct a graph G_h by discretising a flat torus $\mathbb{T}^n := (\mathbb{R}/\mathbb{Z})^n$ with mesh size $h \in (0, 1)$. In the remainder of this paper, we assume that h is such that $1/h$ is an integer, i.e.,

$$\mathfrak{H} := \left\{ h \in (0, 1) \mid \exists k \in \mathbb{N} \ h = \frac{1}{k} \right\}.$$

For any $n \in \mathbb{N}$, we denote the Euclidean norm on \mathbb{R}^n by $|\cdot|$. We use the same notation for the induced Euclidean norms on \mathbb{T}^n .

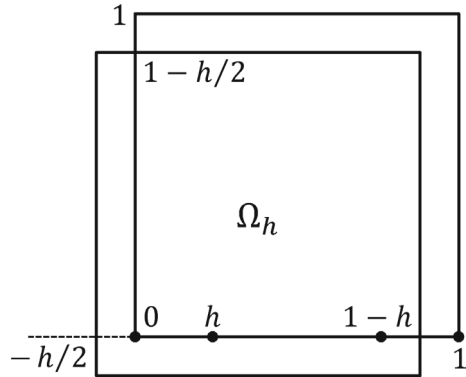
Define the graphs $G_h = (V_h, E_h)$ via their node sets V_h and edge sets E_h :

$$V_h := (h(\mathbb{Z}/h^{-1}\mathbb{Z}))^n = \{0, h, 2h, \dots, 1 - h\}^n,$$

$$E_h := \left\{ (z, \tilde{z}) \in V_h \times V_h \mid |z - \tilde{z}|_{\ell^1} = h \right\} \\ \cup \left\{ (z, \tilde{z}) \in V_h \times V_h \mid \exists j \ |z_j - \tilde{z}_j| = 1 - h \text{ and } \forall i \neq j \ z_i = \tilde{z}_i \right\}.$$

Here $|x|_{\ell^1} := \sum_{i=1}^n |x_i|$. We parametrise the torus \mathbb{T}^n by the hypercube $\Omega_h := [-h/2, 1 - h/2]^n$ with periodic boundary conditions (see Fig. 2.1 for the case with $n = 2$). Then the first set in E_h contains the edges within Ω_h and the second set the

Fig. 2.1 The square $\Omega_h := [-h/2, 1 - h/2]^2$ with periodic boundary conditions parametrises the torus \mathbb{T}^2 . The black circles indicate some of the nodes in the corresponding node set V_h , where the node at $(0, 1)$ is identified with the one at $(0, 0)$



edges that “go over” the boundary of Ω_h and capture the periodicity on \mathbb{T}^n . We note that $(z, \tilde{z}) \in E_h$ if and only if $(\tilde{z}, z) \in E_h$. We identify these edges with each other so that the graphs are undirected. When it is clear from the context what h is, we may also write $z \sim \tilde{z}$ to indicate $(z, \tilde{z}) \in E_h$.

We note that elsewhere in the literature (e.g., in [61] and some of the references therein) the graph G_h is viewed as an edge-weighted graph, in which a weight h^{n-1} is assigned to each edge (and to each pair of unconnected nodes a weight of zero). In this paper, we have chosen to directly incorporate the dependence on the mesh size into the definitions of our inner products and functionals which we give below, e.g., the factors h^{n-1} and h^n in $\|\nabla u\|_1$ and $\langle \cdot, \cdot \rangle_h$, respectively (see Remark 2.9).

Using the notation $[n] := \{1, \dots, n\}$, for $z \in V_h$ we define

$$Q_z^h := \left\{ hy + z \in \mathbb{R}^n \mid \forall i \in [n] \ -1/2 \leq y_i < 1/2 \right\},$$

which is a “half-open” n -cube centred at the vertex z with edges of length h . With

$$I_{z_i}^h := [z_i - h/2, z_i + h/2), \tag{2.17}$$

we have $Q_z^h = \prod_{i=1}^n I_{z_i}^h := I_{z_1}^h \times \dots \times I_{z_n}^h$. We note that, if f is integrable on \mathbb{T}^n , then

$$\int_{\mathbb{T}^n} f(x) dx = \sum_{z \in V_h} \int_{Q_z^h} f(x) dx.$$

2.3.2 Function Spaces, Embedding, and Projection

Let $h \in \mathfrak{H}$. We define the L^2 inner product of functions $v_1, v_2 : V_h \rightarrow \mathbb{R}$ as

$$\langle v_1, v_2 \rangle_h := \sum_{z \in V_h} v_1(z)v_2(z)h^n.$$

We denote the inner product space of all functions $v : V_h \rightarrow \mathbb{R}$ equipped with $\langle \cdot, \cdot \rangle_h$ by L_h^2 . If we want to consider the set of real-valued vertex functions $v : V_h \rightarrow \mathbb{R}$ without assuming a priori it is equipped with a specific inner product, we denote this by \mathcal{V}_h . For functions $v \in \mathcal{V}_h$ we also define the norm $\|v\|_\infty := \max_{z \in V} |v(z)|$.

Since L_h^2 is a finite (n -)dimensional inner product space, it is complete (and thus a Hilbert space) and the corresponding norm $\|v\|_h := \sqrt{\langle v, v \rangle_h}$ is finite for all $v \in L_h^2$. In Sect. 2.4, L_h^2 will play the role that M_h played in Sect. 2.2. In Sect. 2.5 we will introduce a different inner product on \mathcal{V}_h and the resulting Hilbert space \overline{L}_h^2 will play the role of M_h in that section instead (see Sect. 2.5.4).

The role of M will be played by the Lebesgue space $L^2(\mathbb{T}^n)$ consisting of all functions $w : \mathbb{T}^n \rightarrow \mathbb{R}$ which have finite norm $\|w\|_{L^2(\mathbb{T}^n)} := \sqrt{\langle w, w \rangle_{L^2(\mathbb{T}^n)}}$, where

$$\langle w_1, w_2 \rangle_{L^2(\mathbb{T}^n)} := \int_{\mathbb{T}^n} w_1(x)w_2(x) dx = \int_{\Omega_h} w_1(x)w_2(x) dx,$$

for functions $w_1, w_2 \in L^2(\mathbb{T}^n)$. We emphasise that the value of the integral on the right does not depend on the choice of h . For later use, for $w : \mathbb{T}^n \rightarrow \mathbb{R}$ and $p \in \mathbb{N}$ we also define the norms $\|w\|_{L^p(\mathbb{T}^n)} := \left(\int_{\mathbb{T}^n} |w(x)|^p dx \right)^{1/p}$ and $\|w\|_{L^\infty(\mathbb{T}^n)} := \text{esssup}_{x \in \mathbb{T}^n} |w(x)|$.

We define the embedding operator $i_h : L_h^2 \rightarrow L^2(\mathbb{T}^n)$. If $v \in L_h^2$ and $x \in \mathbb{T}^n$, we set

$$i_h v(x) := v(z),$$

where $z \in V_h$ is such that $x \in Q_z^h$. We note that, by our identification of \mathbb{T}^n with Ω_h (with periodic boundary conditions) for each x, z is uniquely determined.

Next we define the projection operator $p_h : L^2(\mathbb{T}^n) \rightarrow L_h^2$. If $w \in L^2(\mathbb{T}^n)$ and $z \in V_h$, then

$$p_h w(z) := \frac{1}{|Q_z^h|} \int_{Q_z^h} w(x) dx = h^{-n} \int_{Q_z^h} w(x) dx,$$

where $|\cdot|$ denotes the Lebesgue measure and the second equality follows since $|Q_z^h| = h^n$, for all $z \in V_h$.

We end this subsection with some useful properties of i_h and p_h . First we introduce some new notation which we will require in the proof of Theorem 2.3 (viii) and at later points in this paper. For $i \in [n]$, we define

$$\tilde{Q}_{z,i}^h := \prod_{j \in [n] \setminus \{i\}} I_{z_j}^h, \tag{2.18}$$

where again \prod denotes the Cartesian product of sets. We refer to (2.17) for the definition of $I_{z_j}^h$.

Theorem 2.3 *Let $h \in \mathfrak{H}$.*

- (i) *If $v_1, v_2 \in L^2_h$, then $\langle i_h v_1, i_h v_2 \rangle_{L^2(\mathbb{T}^n)} = \langle v_1, v_2 \rangle_h$. In particular, it follows that $i_h(L^2_h) \subset L^2(\mathbb{T}^n)$ as required, and that i_h is an isometry.*
- (ii) *The composition $p_h \circ i_h$ is the identity on L^2_h .*
- (iii) *If $w \in L^2(\mathbb{T}^n)$ and $x \in Q^h_z$, we have*

$$i_h p_h w(x) = h^{-n} \int_{Q^h_z} w(y) dy.$$

- (iv) *If $w_1, w_2 \in L^2(\mathbb{T}^n)$, then*

$$\langle i_h p_h w_1, w_2 \rangle_{L^2(\mathbb{T}^n)} = \langle p_h w_1, p_h w_2 \rangle_h = \langle w_1, i_h p_h w_2 \rangle_{L^2(\mathbb{T}^n)}.$$

In particular the operator $i_h \circ p_h$ is self-adjoint on $L^2(\mathbb{T}^n)$.

- (v) *Condition (2.8) is satisfied and thus the Pythagorean equality (2.7) holds.*
- (vi) *If $f \in C(\mathbb{T}^n)$, then $i_h p_h f \rightarrow f$ uniformly, as $h \rightarrow 0$.*
- (vii) *For all $w \in L^2(\mathbb{T}^n)$, $i_h p_h w \rightarrow w$ in $L^2(\mathbb{T}^n)$ as $h \rightarrow 0$.*
- (viii) *For all $w \in L^2(\mathbb{T}^n)$ and for all $z \in V_h$,*

$$\begin{aligned} \int_{Q^h_z} (i_h p_h w)(x) dx &= \int_{Q^h_z} w(x) dx \quad \text{and} \\ \int_{Q^h_z} |i_h p_h w|(x) dx &\leq \int_{Q^h_z} |w|(x) dx. \end{aligned}$$

Proof (i) A direct computation shows

$$\langle i_h v_1, i_h v_2 \rangle_{L^2(\mathbb{T}^n)} = \sum_{z \in V_h} \int_{Q^h_z} i_h v_1(x) i_h v_2(x) dx = \sum_{z \in V_h} v_1(z) v_2(z) h^n = \langle v_1, v_2 \rangle_h.$$

- (ii) Let $v \in L^2_h$ and $z \in V_h$, then

$$p_h i_h v(z) = h^{-n} \int_{Q^h_z} i_h v(x) dx = h^{-n} \int_{Q^h_z} v(z) dx = v(z).$$

- (iii) This follows directly from $i_h p_h w(x) = p_h w(z)$ and the definition of p_h .

- (iv) We compute

$$\begin{aligned} \langle i_h p_h w_1, w_2 \rangle_{L^2(\mathbb{T}^n)} &= \sum_{z \in V_h} \int_{Q^h_z} (i_h p_h w_1)(x) w_2(x) dx \\ &= \sum_{z \in V_h} \int_{Q^h_z} \left(h^{-n} \int_{Q^h_z} w_1(y) dy \right) w_2(x) dx \end{aligned}$$

$$\begin{aligned}
&= \sum_{z \in V_h} \left(h^{-n} \int_{Q_z^h} w_1(y) dy \right) \left(h^{-n} \int_{Q_z^h} w_2(x) dx \right) h^n \\
&= \langle p_h w_1, p_h w_2 \rangle_h,
\end{aligned}$$

which proves the first equality. Continuing the computation to prove the second equality, we find

$$\begin{aligned}
\sum_{z \in V_h} \int_{Q_z^h} \left(h^{-n} \int_{Q_z^h} w_1(y) dy \right) w_2(x) dx &= \sum_{z \in V_h} \int_{Q_z^h} w_1(y) \left(h^{-n} \int_{Q_z^h} w_2(x) dx \right) dy \\
&= \sum_{z \in V_h} \int_{Q_z^h} w_1(y) (i_h p_h w_2)(y) dy = \langle w_1, i_h p_h w_2 \rangle_{L^2(\mathbb{T}^n)}.
\end{aligned}$$

Since $i_h \circ p_h$ is a bounded operator, as follows from part (iii), the symmetry just proven implies self-adjointness.

- (v) Since $i_h \circ p_h$ is self-adjoint on $L^2(\mathbb{T}^n)$ by part (iv) and $p_h \circ i_h$ is the identity on L_h^2 , we have, for $u \in L_h^2$ and $w \in L^2(\mathbb{T}^n)$,

$$\langle i_h p_h w - w, i_h u \rangle_{L^2(\mathbb{T}^n)} = \langle w, i_h p_h i_h u - i_h u \rangle_{L^2(\mathbb{T}^n)} = 0.$$

- (vi) By compactness of \mathbb{T}^n , f is uniformly continuous. Let $\varepsilon > 0$, then there exists a $\delta > 0$ such that, if $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. Let $h < n^{-1/2}\delta$, $x \in \mathbb{T}^n$, and let $z \in V_h$ be such that $x \in Q_z^h$. If $y \in Q_z^h$, then $|x - y| < \sqrt{n}h = \delta$. Thus

$$\begin{aligned}
|i_h p_h f(x) - f(x)| &= \left| h^{-n} \int_{Q_z^h} f(y) dy - f(x) \right| = \left| h^{-n} \int_{Q_z^h} (f(y) - f(x)) dy \right| \\
&\leq h^{-n} \int_{Q_z^h} |f(y) - f(x)| dy < h^{-n} \int_{Q_z^h} \varepsilon dy = \varepsilon.
\end{aligned}$$

- (vii) Let $f \in C(\mathbb{T}^n)$. By the previous point, $i_h p_h f \rightarrow f$ uniformly as $h \rightarrow 0$. Since $i_h \circ p_h$ is self-adjoint, we have, for $w \in L^2(\mathbb{T}^n)$,

$$\langle i_h p_h w, f \rangle_{L^2} = \langle w, i_h p_h f \rangle_{L^2} \rightarrow \langle w, f \rangle_{L^2} \quad \text{as } h \rightarrow 0.$$

By the Cauchy–Schwarz inequality, applied on a single cube Q_z^h , we get, for all $\tilde{w} \in L^2(\mathbb{T}^n)$,

$$\left(\int_{Q_z^h} \tilde{w}(y) dy \right)^2 \leq \left(\int_{Q_z^h} \tilde{w}(y) dy \right) \left(\int_{Q_z^h} 1^2 dy \right) = \left(\int_{Q_z^h} \tilde{w}(y) dy \right) h^n.$$

If $\|\tilde{w}\|_{L^2(\mathbb{T}^n)} = 1$, then

$$\begin{aligned} \|i_h p_h \tilde{w}\|_{L^2(\mathbb{T}^n)}^2 &= \sum_{z \in V_h} \int_{Q_z^h} \left(h^{-n} \int_{Q_z^h} \tilde{w}(y) dy \right)^2 dx = \sum_{z \in V_h} h^{-n} \left(\int_{Q_z^h} \tilde{w}(y) dy \right)^2 \\ &\leq \sum_{z \in V_h} \int_{Q_z^h} \tilde{w}(y) dy = \int_{\mathbb{T}^n} \tilde{w}(y) dy = 1. \end{aligned}$$

Hence the operator norm of $i_h \circ p_h$ satisfies

$$\|i_h \circ p_h\|_{op} := \sup \left\{ \|i_h p_h \tilde{w}\|_{L^2(\mathbb{T}^n)} \mid \tilde{w} \in L^2(\mathbb{T}^n) \text{ and } \|\tilde{w}\|_{L^2(\mathbb{T}^n)} = 1 \right\} \leq 1.$$

Let $\tilde{w} \in L^2(\mathbb{T}^n)$. Since $C(\mathbb{T}^n)$ is dense in $L^2(\mathbb{T}^n)$, there exists a sequence (f_m) in $C(\mathbb{T}^n)$ which converges to \tilde{w} in $L^2(\mathbb{T}^n)$ as $m \rightarrow \infty$. Combining the Cauchy-Schwarz inequality with the bound on the operator norm, we find

$$\begin{aligned} \langle i_h p_h w, \tilde{w} \rangle_{L^2(\mathbb{T}^n)} &= \langle i_h p_h w, f_m \rangle_{L^2(\mathbb{T}^n)} + \langle i_h p_h w, \tilde{w} - f_m \rangle_{L^2(\mathbb{T}^n)} \\ &\leq \langle i_h p_h w, f_m \rangle_{L^2(\mathbb{T}^n)} + \|i_h p_h w\|_{L^2(\mathbb{T}^n)} \|\tilde{w} - f_m\|_{L^2(\mathbb{T}^n)} \\ &\leq \langle i_h p_h w, f_m \rangle_{L^2(\mathbb{T}^n)} + \|w\|_{L^2(\mathbb{T}^n)} \|\tilde{w} - f_m\|_{L^2(\mathbb{T}^n)}. \end{aligned}$$

Taking first $h \rightarrow 0$ and then $m \rightarrow \infty$, the right-hand side vanishes and thus $i_h p_h w \rightharpoonup w$ (i.e., weakly) in $L^2(\mathbb{T}^n)$. By lower semicontinuity of the $L^2(\mathbb{T}^n)$ norm with respect to weak $L^2(\mathbb{T}^n)$ convergence and again the estimate on the operator norm, we have that

$$\|w\|_{L^2(\mathbb{T}^n)} \leq \liminf_{h \rightarrow 0} \|i_h p_h w\|_{L^2(\mathbb{T}^n)} \leq \liminf_{h \rightarrow 0} \|w\|_{L^2(\mathbb{T}^n)} = \|w\|_{L^2(\mathbb{T}^n)}.$$

Therefore $\|i_h p_h w\|_{L^2(\mathbb{T}^n)} \rightarrow \|w\|_{L^2(\mathbb{T}^n)}$. Together with the weak convergence $p_h w \rightharpoonup w$ in $L^2(\mathbb{T}^n)$, this implies that $i_h p_h w \rightarrow w$ (i.e., strongly) in $L^2(\mathbb{T}^n)$.

(viii) For all $j \in [n]$, define the operator $q_h^j : L^2(\mathbb{T}^n) \rightarrow L^2(\mathbb{T}^n)$ by, for all $w \in L^2(\mathbb{T}^n)$ and all $x \in Q_z^h$,

$$q_h^j w(x) := h^{-1} \int_{I_{z_j}^h} w(\tilde{x}_j, x_j) dx_j = h^{-1} \int_{I_{z_j}^h} w(x) dx_j.$$

Assume $\{\alpha_1, \dots, \alpha_n\} = [n]$. Let $i, j \in [n]$ with $i \neq j$ and assume that $\alpha_{n-1} = i$ and $\alpha_n = j$. Let $x \in Q_z^h$, then

$$q_h^{\alpha_{n-1}} q_h^{\alpha_n} w(x) = q_h^i q_h^j w(x) = h^{-1} \int_{I_{z_i}^h} q_h^j w(x) dx_i = h^{-2} \int_{I_{z_i}^h} \int_{I_{z_j}^h} w(x) dx_i dx_j.$$

Repeating this argument for $\alpha_1, \dots, \alpha_{n-2}$, we find that

$$q_h^{\alpha_1} \dots q_h^{\alpha_n} w = i_h p_h w. \tag{2.19}$$

Furthermore, we compute for all $j \in [n]$ and all $z \in V_h$,

$$\begin{aligned} \int_{Q_z^h} |q_h^j w(x)| dx &= \int_{Q_z^h} \left| h^{-1} \int_{I_{z_j}^h} w(\tilde{x}_j, y) dy \right| dx \\ &= \int_{I_{z_j}^h} \int_{\tilde{Q}_{z,j}^h} \left| h^{-1} \int_{I_{z_j}^h} w(\tilde{x}_j, y) dy \right| dx_j d\tilde{x}_j \\ &= \int_{\tilde{Q}_{z,j}^h} \left| \int_{I_{z_j}^h} w(\tilde{x}_j, y) dy \right| d\tilde{x}_j \leq \int_{\tilde{Q}_{z,j}^h} \int_{I_{z_j}^h} |w(\tilde{x}_j, y)| dy d\tilde{x}_j \\ &= \int_{Q_z^h} |w(x)| dx. \end{aligned}$$

In the absence of the absolute value function in the integrals above, the inequality becomes an equality. Repeatedly applying this (in)equality according to (2.19) gives the desired result. □

As mentioned before, in Sect. 2.5.4 we introduce a different inner product on \mathcal{V}_h which gives rise to the Hilbert space \overline{L}_h^2 . We postpone the introductions of a corresponding new embedding operator I_h (Sect. 2.5.4) and projection operator P_h (Sect. 2.5.7) to Sect. 2.5, where they will be needed for our proofs for the Allen–Cahn gradient flows.

2.4 Continuum Limit of a Discrete Total Variation Flow

In this section we consider the L_h^2 -gradient flow of the discrete total variation φ_{TV}^h and want to claim convergence to anisotropic total variation flow, i.e., the $L^2(\mathbb{T}^n)$ -gradient flow of the continuum anisotropic total variation φ_{TV} . First we define these functionals in Sect. 2.4.1.

2.4.1 Discrete and Continuum Total Variation Functionals

Let $h \in \mathfrak{H}$. The discrete (anisotropic) graph total variation functional $\varphi_{TV}^h : L_h^2 \rightarrow \mathbb{R}$ is defined by

$$\varphi_{TV}^h(u) := \frac{1}{2} \sum_{z, \tilde{z} \sim z} h^{n-1} |u(z) - u(\tilde{z})|,$$

where we have used the shorthand notation $\sum_{z, \tilde{z} \sim z} := \sum_{z \in V_h} \sum_{\tilde{z} \in V_h: \tilde{z} \sim z}$. Since L_h^2 is a Hilbert space and thus closed, we note that $\mathcal{D}(\varphi_{TV}^h) = L_h^2$.

The anisotropic total variation functional for functions $u : \mathbb{T}^n \rightarrow \mathbb{R}$ is

$$\int_{\mathbb{T}^n} |Du|_{l^1} := \sup \left\{ \int_{\mathbb{T}^n} u(x) \operatorname{div} g(x) dx \mid g \in C^1(\mathbb{T}^n; \mathbb{R}^n), \forall x \in \mathbb{T}^n |g(x)|_\infty \leq 1 \right\},$$

where $|g(x)|_\infty := \max_{1 \leq i \leq n} |g_i(x)|$. If u is regular enough to admit a weak partial derivative in each direction, the notation $\int_{\mathbb{T}^n} |Du(x)|_{l^1} dx$ is compatible with the earlier definition of $|\cdot|_{l^1}$.

If $g = (g_1, \dots, g_n)$, the condition $|g(x)|_\infty \leq 1$ is equivalent to, for all $i \in [n]$, $|g_i(x)| \leq 1$, hence

$$\sup_g \int_{\mathbb{T}^n} u(x) \operatorname{div} g(x) dx = \sup_g \sum_{i=1}^n \int_{\mathbb{T}^n} u(x) \frac{\partial g_i}{\partial x_i}(x) dx = \sum_{i=1}^n \sup_{g_i} \int_{\mathbb{T}^n} u(x) \frac{\partial g_i}{\partial x_i}(x) dx, \tag{2.20}$$

where the supremum on the right-hand side is taken over all $g_i \in C^1(\mathbb{T}^n; \mathbb{R})$ that satisfy, for all $x \in \mathbb{T}^n$, $|g_i(x)| \leq 1$. (We allow ourselves the slight misuse of notation, to stress that the set of admissible functions for each supremum depends on i .)

We denote the space of functions of bounded variation, i.e., the space of all $u \in L^1(\mathbb{T}^n)$ (i.e., Lebesgue integrable functions on \mathbb{T}^n) for which $\int_{\mathbb{T}^n} |Du|_{l^1} < +\infty$, by $BV(\mathbb{T}^n)$. It can be shown that $\int_{\mathbb{T}^n} |Du|_{l^1}$ and the isotropic total variation (as in, for example, [34]) are equivalent seminorms and thus $BV(\mathbb{T}^n)$ as defined above agrees with the standard definition based on the isotropic total variation. It is worth noting that $BV(\mathbb{T}^n)$ is not the same as $BV(\operatorname{int} \Omega_h)$ (where int denotes the topological interior), since the anisotropic total variation on \mathbb{T}^n also measures variations on the boundary of Ω_h (taking into account the periodic boundary conditions), whereas the anisotropic total variation on $\operatorname{int} \Omega_h$ does not.

An important property of the anisotropic total variation is that, for $u \in L^2_h$,

$$\int_{\mathbb{T}^n} |Di_h u|_{l^1} = \phi_{TV}^h(u). \tag{2.21}$$

We provide a proof of (2.21) in Appendix A2.7.

We are now ready to define the continuum counterpart of the discrete anisotropic total variation functional, $\varphi_{TV} : L^2(\mathbb{T}^n) \rightarrow \mathbb{R} \cup \{+\infty\}$, by

$$\varphi_{TV}(u) := \begin{cases} \int_{\mathbb{T}^n} |Du|_{l^1}, & \text{if } u \in BV(\mathbb{T}^n) \cap L^2(\mathbb{T}^n), \\ +\infty, & \text{if } u \in L^2(\mathbb{T}^n) \setminus BV(\mathbb{T}^n). \end{cases} \tag{2.22}$$

Because $BV(\mathbb{T}^n) \cap L^2(\mathbb{T}^n)$ is dense in $L^2(\mathbb{T}^n)$ (as follows from the fact that the space of infinitely often continuously differentiable real-valued functions on \mathbb{T}^n , $C(\mathbb{T}^n)$, is dense in $L^2(\mathbb{T}^n)$ [1, Corollary 2.30]), we have $\overline{\mathcal{D}(\varphi_{TV})} = L^2(\mathbb{T}^n)$.

We note that, for all $h \in \mathfrak{H}$, ϕ_{TV}^h is convex (geodesically 0-convex), as is φ_{TV} . These functionals are nonnegative and thus bounded below. Furthermore, ϕ_{TV}^h is

continuous. The lower semicontinuity of φ_{TV} is less obvious but it is standard [34], [28, Sect. 5.2, Theorem 1], [15, Appendix A] (note that these results survive the move from the $L^1(\mathbb{T}^n)$ topology to the $L^2(\mathbb{T}^n)$ topology). Since these functionals are also not identically equal to $+\infty$, φ_{TV}^h and φ_{TV} satisfy condition (0-AGS). Moreover, equation (2.21), which can be written as $\varphi_{TV}(i_h u) = \varphi_{TV}^h(u)$, implies that condition (2.5) is satisfied with $\Phi_h = \varphi_{TV}^h$ and $\Phi = \varphi_{TV}$. By Theorem 2.3(v) we know that (2.7) is satisfied.

Formally (ignoring issues when $\frac{\partial u}{\partial x_j} = 0$), the $L^2(\mathbb{T}^n)$ -gradient flow of φ_{TV} is a solution to

$$u_t = \operatorname{div} \left(\sum_{j=1}^n \frac{\partial}{\partial x_j} / \left| \frac{\partial u}{\partial x_j} \right| \right).$$

2.4.2 Convergence of Flows

We will establish the following result.

Theorem 2.4 *Let $h \in \mathfrak{H}$ and let u^h be the L_h^2 -gradient flow of φ_{TV}^h with initial data $u_0^h \in L_h^2$. Let u be the $L^2(\mathbb{T}^n)$ -gradient flow of φ_{TV} with initial data $u_0 \in L^2(\mathbb{T}^n)$. Then, for all $t \geq 0$,*

$$\|i_h u^h(t) - u(t)\|_{L^2(\mathbb{T}^n)} \leq \|i_h u_0^h - u_0\|_{L^2(\mathbb{T}^n)}.$$

We will apply our abstract results (Theorem 2.1 and Corollary 2.1) to $\Phi = \varphi_{TV}$ on $M = L^2(\mathbb{T}^n)$ and $\Phi_h = \varphi_{TV}^h$ on $M_h = L_h^2$. As established in Sect. 2.4.1, these functionals satisfy condition (0-AGS). Moreover, $L_h^2 = \mathcal{D}(\varphi_{TV}^h)$, $L^2(\mathbb{T}^n) = \overline{\mathcal{D}(\varphi_{TV})}$, and conditions (2.7) and (2.5) hold. In order to apply Theorem 2.1 and Corollary 2.1, it remains to show that condition (2.6) is satisfied as well. This is the content of the following lemma.

Lemma 2.3 *Let $h \in \mathfrak{H}$. For all $w \in L^2(\mathbb{T}^n)$,*

$$\varphi_{TV}^h(p_h w) \leq \varphi_{TV}(w).$$

Proof If $w \in L^2(\mathbb{T}^n) \setminus BV(\mathbb{T}^n)$, the inequality is trivially true. Now assume that $w \in BV(\mathbb{T}^n)$.

By (2.21), (2.60), and Corollary 2.2, we have that

$$\begin{aligned} \varphi_{TV}^h(p_h w) &= \int_{\mathbb{T}^n} |Di_h p_h w|_{l^1} = \sum_{i=1}^n \int_{\mathbb{T}^n} |D_{x_i} i_h p_h w| \\ &= \sum_{i=1}^n \sup_{g \in H_{h,i}} \int_{\mathbb{T}^n} i_h p_h w(x) \frac{\partial g}{\partial x_i}(x) dx, \end{aligned}$$

where we recall the definition of the sets $H_{h,i}$ from (2.66). In particular, if $g \in H_{h,i}$, then $\frac{\partial g}{\partial x_i}$ is constant a.e. on each set $\text{int } Q_z^h$. Thus, if we denote its value on $\text{int } Q_z^h$ by $c_z \in \mathbb{R}$, then

$$\begin{aligned} \int_{\mathbb{T}^n} i_h p_h w(x) \frac{\partial g}{\partial x_i}(x) dx &= \sum_{z \in V_h} \int_{Q_z^h} i_h p_h w(x) \frac{\partial g}{\partial x_i}(x) dx = \sum_{z \in V_h} c_z \int_{Q_z^h} i_h p_h w(x) dx \\ &= \sum_{z \in V_h} c_z \int_{Q_z^h} w(x) dx = \sum_{z \in V_h} \int_{Q_z^h} w(x) \frac{\partial g}{\partial x_i}(x) dx \\ &= \int_{\mathbb{T}^n} w(x) \frac{\partial g}{\partial x_i}(x) dx, \end{aligned}$$

where we used Theorem 2.3 (viii) to establish the third equality above.

We deduce that

$$\sup_{g \in H_{h,i}} \int_{\mathbb{T}^n} i_h p_h w(x) \frac{\partial g}{\partial x_i}(x) dx = \sup_{g \in H_{h,i}} \int_{\mathbb{T}^n} w(x) \frac{\partial g}{\partial x_i}(x) dx \leq \int_{\mathbb{T}^n} |D_{x_i} w|,$$

where the inequality follows from the fact that each set $H_{h,i}$ is a subset of the corresponding set of admissible functions in the supremum in (2.61). Applying (2.60) once more, we conclude from (2.22) that

$$\varphi_{TV}^h(p_h w) \leq \sum_{i=1}^n \int_{\mathbb{T}^n} |D_{x_i} w| = \int_{\mathbb{T}^n} |Dw|_{l^1} = \varphi_{TV}(w).$$

□

Proof (of Theorem 2.4) By Eq. (2.21), Lemma 2.3, and Theorem 2.3, we know that $\Phi = \varphi_{TV}$ and $\Phi_h = \varphi_{TV}^h$ satisfy conditions (2.5), (2.6), and (2.7). Moreover, by our discussion in Sect. 2.4.1 they also satisfy condition (0-AGS). We are able to apply Theorem 2.1 and Corollary 2.1 to get the desired result. □

Remark 2.7 In [42] it is proved that u^h is the gradient flow of φ_{TV} in $L^2(\mathbb{T}^n)$ by calculating the minimal section of the subdifferential of φ_{TV} , i.e., the element of the subdifferential (with respect to the $L^2(\mathbb{T}^n)$ inner product) $\partial \varphi_{TV}(u^h(t))$ that is smallest in $L^2(\mathbb{T}^n)$ norm, at each time t . Our result gives a direct proof that u^h is indeed the gradient flow of φ_{TV} in $L^2(\mathbb{T}^n)$.

2.4.3 Convergence of Energy

We are interested in the convergence of the “energy” $\varphi_{TV}^h(u^h)$ to $\varphi_{TV}(u)$.

Theorem 2.5 *Let $h \in \mathfrak{S}$. Let u^h and u be as in Theorem 2.4 and let $T > 0$. Then*

$$\left| \int_0^T (\varphi_{TV}^h(u^h(s)) - \varphi_{TV}(u(s))) ds \right| \leq \left(\|i_h u_0^h\|_{L^2(\mathbb{T}^n)} + \|u_0\|_{L^2(\mathbb{T}^n)} \right) \|i_h u_0^h - u_0\|_{L^2(\mathbb{T}^n)}.$$

Proof By definition $\dot{u}^h \in -\partial\varphi_{TV}^h(u^h)$ and $\dot{u} \in -\partial\varphi_{TV}(u)$. Since φ_{TV}^h and φ_{TV} are positively one-homogeneous, we see that,⁹ for all $v \in \partial\varphi_{TV}^h(u^h)$ and for all $w \in \partial\varphi_{TV}(u)$ we have the Euler identity (see, for example, [6, Lemma 1.7])

$$\langle u^h, v \rangle_h = \varphi_{TV}^h(u^h), \quad \langle u, w \rangle_{L^2(\mathbb{T}^n)} = \varphi_{TV}(u).$$

Choosing $v = \dot{u}^h$ and $w = \dot{u}$ yields

$$\varphi_{TV}^h(u^h) = \frac{1}{2} \frac{d}{dt} \|u^h\|_h^2, \quad \varphi_{TV}(u) = \frac{1}{2} \frac{d}{dt} \|u\|_{L^2(\mathbb{T}^n)}^2.$$

These equalities hold almost everywhere on $(0, \infty)$. From this observation it follows that

$$\begin{aligned} \int_0^T (\varphi_{TV}^h(u^h(s)) - \varphi_{TV}(u(s))) ds &= \frac{1}{2} \left(\|i_h u^h(T)\|_{L^2(\mathbb{T}^n)}^2 - \|u(T)\|_{L^2(\mathbb{T}^n)}^2 \right) \\ &\quad - \frac{1}{2} \left(\|i_h u_0^h\|_{L^2(\mathbb{T}^n)}^2 - \|u_0\|_{L^2(\mathbb{T}^n)}^2 \right), \end{aligned} \quad (2.23)$$

where we have also used the isometry property of i_h (Theorem 2.3).

By the contraction property from Theorem 2.4 we see that

$$\begin{aligned} \|i_h u^h(T) - 0\|_{L^2(\mathbb{T}^n)} &\leq \|i_h u_0^h - 0\|_{L^2(\mathbb{T}^n)}, \quad \|u(T) - 0\|_{L^2(\mathbb{T}^n)} \leq \|u_0 - 0\|_{L^2(\mathbb{T}^n)} \text{ and} \\ \|i_h u^h(T) - u(T)\|_{L^2(\mathbb{T}^n)} &\leq \|i_h u_0^h - u_0\|_{L^2(\mathbb{T}^n)}. \end{aligned}$$

For the first two inequalities we used that the constant solutions $u^h = 0$ and $u = 0$ are gradient flows of φ_{TV}^h starting from 0 and of φ_{TV} starting from 0, respectively (Lemma 2.1). Since, for all $a, b \in \mathbb{R}$, $a^2 - b^2 = (a + b)(a - b)$, we find, using the inequalities above and the reverse triangle inequality, that

$$\begin{aligned} &\left| \|i_h u^h(T)\|_{L^2(\mathbb{T}^n)}^2 - \|u(T)\|_{L^2(\mathbb{T}^n)}^2 \right| \\ &= \left| \|i_h u^h(T)\|_{L^2(\mathbb{T}^n)} + \|u(T)\|_{L^2(\mathbb{T}^n)} \right| \left| \|i_h u^h(T)\|_{L^2(\mathbb{T}^n)} - \|u(T)\|_{L^2(\mathbb{T}^n)} \right| \\ &\leq \left| \|i_h u^h(T)\|_{L^2(\mathbb{T}^n)} + \|u(T)\|_{L^2(\mathbb{T}^n)} \right| \|i_h u^h(T) - u(T)\|_{L^2(\mathbb{T}^n)} \\ &\leq \left| \|i_h u_0^h\|_{L^2(\mathbb{T}^n)} + \|u_0\|_{L^2(\mathbb{T}^n)} \right| \|i_h u_0^h - u_0\|_{L^2(\mathbb{T}^n)} \end{aligned}$$

⁹ Assume φ is a convex function on a (real or complex) vector space V . This function is positively one-homogeneous if, for all $u \in V$ and for all $r > 0$, $\varphi(ru) = r\varphi(u)$. If $v \in \partial\varphi(u)$, then, for all $w \in V$, $\varphi(w) - \varphi(u) \geq \langle w - u, v \rangle$. Choosing $w = 0$ and $w = 2u$ leads to $\varphi(u) = \langle u, v \rangle$.

and

$$\begin{aligned} & \left| \|i_h u_0^h\|_{L^2(\mathbb{T}^n)}^2 - \|u_0\|_{L^2(\mathbb{T}^n)}^2 \right| \\ &= \left| \|i_h u_0^h\|_{L^2(\mathbb{T}^n)} + \|u_0\|_{L^2(\mathbb{T}^n)} \right| \left| \|i_h u_0^h\|_{L^2(\mathbb{T}^n)} - \|u_0\|_{L^2(\mathbb{T}^n)} \right| \\ &\leq \left(\|i_h u_0^h\|_{L^2(\mathbb{T}^n)} + \|u_0\|_{L^2(\mathbb{T}^n)} \right) \|i_h u_0^h - u_0\|_{L^2(\mathbb{T}^n)}. \end{aligned}$$

By (2.23) we get the desired estimate. \square

2.5 Continuum Limit of a Discrete Allen–Cahn Flow

In this section we will study the convergence of discrete Allen–Cahn gradient flows to a continuum Allen–Cahn gradient flow. These flows are determined, in the sense of Definition 2.1, by the discrete and continuum Allen–Cahn functionals, φ_{AC}^h and φ_{AC} , respectively. We start by defining these functionals in Sect. 2.5.1. From Sect. 2.5.3 onwards we restrict ourselves to the case $n = 1$.

2.5.1 Discrete and Continuum Allen–Cahn Functionals

Let $h \in \mathfrak{H}$. The discrete Allen–Cahn functional consists of two terms, the Dirichlet energy $\varphi_D^h : \mathcal{V}_h \rightarrow \mathbb{R}$ and the double-well potential energy $\varphi_W^h : \mathcal{V}_h \rightarrow \mathbb{R}$, which we define separately:

$$\begin{aligned} \varphi_D^h(u) &:= \frac{1}{4} \sum_{z: \bar{z} \sim z} h^{n-2} (u(z) - u(\bar{z}))^2, \\ \varphi_W^h(u) &:= \sum_{z \in V_h} h^n W(u(z)). \end{aligned}$$

Here W is a double-well potential with wells of equal depth; in this paper, we shall fix $W(x) = \alpha(x^2 - 1)^2/4$ for a given $\alpha > 0$. Now we define the discrete Allen–Cahn functional $\varphi_{AC}^h : \mathcal{V}_h \rightarrow \mathbb{R}$ as

$$\varphi_{AC}^h := \varphi_D^h + \varphi_W^h.$$

We note here that we have chosen \mathcal{V}_h as the domain of φ_D^h , φ_W^h , and φ_{AC}^h . At various times we will want to emphasise a specific inner product structure (and induced topology) on the domain, in which case we will write L_h^2 or \bar{L}_h^2 for the domain. The latter of these two Hilbert spaces will be defined in Sect. 2.5.4.

The continuum versions of the Dirichlet energy, $\varphi_D : L^2(\mathbb{T}^n) \rightarrow \mathbb{R} \cup \{+\infty\}$, and the double-well energy, $\varphi_W : L^2(\mathbb{T}^n) \rightarrow \mathbb{R} \cup \{+\infty\}$, are given by

$$\varphi_D(u) := \begin{cases} \frac{1}{2} \int_{\mathbb{T}^n} |\nabla u(x)|^2 dx, & \text{if } u \in H^1(\mathbb{T}^n), \\ +\infty, & \text{otherwise,} \end{cases} \quad \text{and}$$

$$\varphi_W(u) := \begin{cases} \int_{\mathbb{T}^n} W(u(x)) dx, & \text{if } W \circ u \in L^1(\mathbb{T}^n), \\ +\infty, & \text{otherwise.} \end{cases}$$

Here $|\nabla u|$ denotes the Euclidean norm of the (weak) gradient of u . We remind ourselves that $H^m(\mathbb{T}^n)$ is¹⁰ the space of functions in $L^2(\mathbb{T}^n)$ that are $m \in \mathbb{N}$ times weakly differentiable with each of their weak derivatives also being in $L^2(\mathbb{T}^n)$. It is equipped with the (squared) norm $\|w\|_{H^m(\mathbb{T}^n)}^2 := \|w\|_{L^2(\mathbb{T}^n)}^2 + \sum_{l=1}^m \|w^{(l)}\|_{L^2(\mathbb{T}^n)}^2$, where $w^{(l)}$ denotes the l th derivative of w .

We note that φ_W is $L^2(\mathbb{T}^n)$ -coercive since, by Jensen's inequality [27, Appendix B Theorem 2],

$$\int_{\mathbb{T}^n} W(u) dx = \int_{\mathbb{T}^n} (u^2(x) - 1)^2 dx \geq \left(\int_{\mathbb{T}^n} (u^2(x) - 1) dx \right)^2 = \left(\|u\|_{L^2(\mathbb{T}^n)}^2 - 1 \right)^2.$$

The continuum Allen–Cahn functional $\varphi_{AC} : L^2(\mathbb{T}^n) \rightarrow \mathbb{R} \cup \{+\infty\}$ is then defined to be

$$\varphi_{AC}(u) := \varphi_D(u) + \varphi_W(u).$$

Neither φ_{AC}^h nor φ_{AC} are convex, but they are geodesically $(-\alpha)$ -convex, uniformly in h (for φ_{AC}^h). This follows since φ_D^h and φ_D are convex and since $\left(W(x) - \frac{(-\alpha)}{2}x^2 \right)'' = \alpha(3x^2 - 1) + \alpha \geq 0$, so that W is geodesically $(-\alpha)$ -convex, and thus so are φ_W^h and φ_W . We emphasise that α does not depend on h .

The functional φ_{AC}^h is continuous, if \mathcal{V}_h is equipped with any norm-induced topology (as those topologies are all equivalent since \mathcal{V}_h is finite-dimensional). Moreover, since the integrand of $\varphi_{AC}(u)$, $\frac{1}{2}|\nabla u|^2 + W(u)$, is smooth as a function of $(\nabla u, u)$ and convex in ∇u , φ_{AC} is lower semicontinuous with respect to weak $H^1(\mathbb{T}^n)$ convergence [27, §8.2 Theorem 1] and thus, by the compact Rellich–Kondrachov embedding [1, Theorem 6.3] also with respect to (strong) $L^2(\mathbb{T}^n)$ convergence. Since φ_{AC}^h and φ_{AC} are not identically equal to $+\infty$ and nonnegative (and thus bounded below), they satisfy the $(-\alpha)$ -AGS condition.

The Allen–Cahn functional (discrete or continuum) also goes by various other names in the literature, such as Ginzburg–Landau functional [7] or Modica–Mortola functional [51].

¹⁰ And similarly if \mathbb{T}^n is replaced by an open subset of itself or of \mathbb{R}^n . Recall that $w \in L^2(\mathbb{T}^n)$ is the weak derivative of $u \in L^2(\mathbb{T}^n)$ with respect to \mathbb{T}^n -coordinate x_i , if, for all $\varphi \in C^\infty(\mathbb{T}^n)$, $\langle w, \varphi \rangle_{L^2(\mathbb{T}^n)} = \langle u, \frac{\partial}{\partial x_i} \varphi \rangle_{L^2(\mathbb{T}^n)}$.

2.5.2 Additional Setup

We are interested in discrete Allen–Cahn flows, i.e., solutions of $\dot{u} \in -\partial\varphi_{AC}^h(u)$, and in particular in their convergence to a continuum Allen–Cahn flow, i.e., a solution of $\dot{u} \in -\partial\varphi_{AC}(u)$ (Theorem 2.6). An explicit form of the equation for the Allen–Cahn flow is

$$\dot{u} = \Delta u - W'(u) \quad \text{in } \mathbb{T}^n \times (0, \infty), \tag{2.24}$$

which is often called the Allen–Cahn equation [3]. As usual Δ denotes the (continuum) Laplacian $\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$ where x_i are the coordinates on \mathbb{T}^n . This explicit form of the gradient flow equation can be computed directly since the subdifferential in (2.2) is a singleton. Similarly an explicit form of the equation for the discrete Allen–Cahn flow can be computed:

$$\dot{u} = \Delta_h u - W'(u) \quad \text{in } V_h \times (0, \infty), \tag{2.25}$$

where

$$(\Delta_h u)(z) := h^{-2} \sum_{\tilde{z} \in V_h: \tilde{z} \sim z} (u(\tilde{z}) - u(z)) \tag{2.26}$$

defines the graph Laplacian [16, 63] for $u \in L_h^2$ and $z \in V_h$. This discrete flow equation has been studied in more detail in [11–13, 62]. Details of the derivation in the one-dimensional ($n = 1$) case follow in Lemma 2.6.

It will also be useful to define a graph gradient $\nabla_h : V_h \times V_h \rightarrow \mathbb{R}$; for $u \in L_h^2$ and $z, \tilde{z} \in V_h$,

$$\nabla_h u(z, \tilde{z}) = \begin{cases} h^{-1}(u(\tilde{z}) - u(z)), & \text{if } \tilde{z} \sim z, \\ 0, & \text{otherwise.} \end{cases}$$

We note that $\nabla_h u$ is defined on $V_h \times V_h$. For such functions $\chi, \phi : V_h \times V_h \rightarrow \mathbb{R}$ we introduce the inner product

$$\langle \chi, \phi \rangle_{h,h} := \frac{1}{2} \sum_{z; \tilde{z} \sim z} h^n \chi(z, \tilde{z}) \phi(z, \tilde{z})$$

and corresponding norm $\|\chi\|_{h,h} := \sqrt{\langle \chi, \chi \rangle_{h,h}}$. Then we have, for $u, v \in \mathcal{V}_h$,

$$\begin{aligned} \langle -\Delta_h u, v \rangle_h &= \sum_{z; \tilde{z} \sim z} h^{n-2} (u(z) - u(\tilde{z})) v(z) = \frac{1}{2} \sum_{z; \tilde{z} \sim z} h^{n-2} (u(z) - u(\tilde{z})) (v(z) - v(\tilde{z})) \\ &= \langle \nabla_h u, \nabla_h v \rangle_{h,h} = \langle u, -\Delta_h v \rangle_h. \end{aligned} \tag{2.27}$$

and $\varphi_D^h(u) = \frac{1}{2} \|\nabla_h u\|_{h,h}^2$.

Remark 2.8 Two different sign conventions for the (graph) Laplacian appear in the literature. Our choice in (2.26) is in line with the typical definition in the partial differential equations' literature which has the Laplacian as negative semidefinite. This differs by an overall minus sign from the common definition in the spectral graph literature, which has the graph Laplacian as positive semidefinite.

Remark 2.9 It is also interesting to compare how the choices we have made in this paper regarding the scaling with h correspond to the scaling choices in [62, Sect. 2] which are common in the literature. In [62] the scaling of the inner products, operators, and summands in the functionals depends on the edge weights ω , the node degrees $d \sim \omega$, and two parameters, r and q , as follows:

$$\begin{aligned} \langle \cdot, \cdot \rangle_h &\sim d^r \sim \omega^r, & \langle \cdot, \cdot \rangle_{h,h} &\sim \omega^{2q-1}, \\ \nabla_h &\sim \omega^{1-q}, & \Delta_h &\sim \omega^{1-r}, \\ \text{summands in } \phi_D^h &\sim \omega, & \text{summands in } \phi_{TV}^h &\sim \omega^q. \end{aligned}$$

We note that for the functionals ϕ_D^h and ϕ_{TV}^h , we consider the scaling of the summands only, i.e., we do not consider the fact that the number of summands is $|V_h| = h^{-n}$. According to our choices in this paper, we require the summands of ϕ_D^h to be proportional to h^{n-2} , which suggests the choice $\omega = h^{n-2}$ for the edge weights between neighbouring nodes. Comparisons for the other objects lead to

$$\omega^r = h^n, \quad \omega^{2q-1} = h^n, \quad \omega^{1-q} = h^{-1}, \quad \omega^{1-r} = h^{-2}, \quad \text{and} \quad \omega^q = h^{n-1}.$$

It follows that, when $n \neq 2$,

$$q = \frac{n-1}{n-2} \quad \text{and} \quad r = 2q - 1 = \frac{n}{n-2}.$$

This implies that $q > 1$ and $r > 1$, which are not choices that are typically considered. When $n = 2$, we have $\omega = 1$, which cannot be made compatible with the other requirements.

2.5.3 Restriction to One-Dimensional Case ($n = 1$)

To avoid additional technical complexity, we consider the convergence problem only for the one-dimensional case, i.e., $n = 1$. In this case the vertex set is $V_h = \{0, h, 2h, \dots, 1-h\}$. We remind ourselves that $h \in \mathfrak{H}$ and thus $|V_h| = h^{-1}$ is an integer. It will be useful to write $[\frac{1}{h} - 1]_0 := \{0\} \cup [\frac{1}{h} - 1] = \{0, 1, \dots, \frac{1}{h} - 1\}$, so that to each $k \in [\frac{1}{h} - 1]_0$ corresponds one and only one $kh \in V_h$.

Remember that \mathcal{V}_h denotes the set of real-valued vertex functions $v : V_h \rightarrow \mathbb{R}$, without assuming a specific inner product structure.

To simplify notation, if $v \in \mathcal{V}_h$, we write $v_k := v(kh)$. Similarly, if $\chi : V_h \times V_h \rightarrow \mathbb{R}$, then we write $\chi_{k,l} := \chi(kh, lh)$. By periodicity $v_{\frac{1}{h}} = v_0$, $v_{\frac{1}{h}+1} = v_1$, $v_{\frac{1}{h}-1} = v_{-1}$, etc. Periodicity imposes similar relationships on $\chi_{k,l}$.

For future reference it is convenient to explicitly state to what some of the relevant definitions reduce in this one-dimensional case:

$$\begin{aligned} (\Delta_h u)_k &= h^{-2} (u_{k+1} + u_{k-1} - 2u_k), & (\nabla_h u)_{k,k\pm 1} &= h^{-1} (u_{k\pm 1} - u_k), \\ \langle u, v \rangle_h &= \sum_{k=0}^{\frac{1}{h}-1} h u_k v_k, & \langle \chi, \phi \rangle_{h,h} &= \frac{1}{2} \sum_{k=0}^{\frac{1}{h}-1} h (\chi_{k,k+1} \phi_{k,k+1} + \chi_{k,k-1} \phi_{k,k-1}), \\ \varphi_D^h(u) &= \frac{1}{2} \sum_{k=0}^{\frac{1}{h}-1} h^{-1} (u_{k+1} - u_k)^2. \end{aligned}$$

We note that if χ and ϕ are both symmetric (i.e., $\chi_{k,l} = \chi_{l,k}$ and $\phi_{k,l} = \phi_{l,k}$) or both skew-symmetric (i.e., $\chi_{k,l} = -\chi_{l,k}$ and $\phi_{k,l} = -\phi_{l,k}$), then, by periodicity,

$$\langle \chi, \phi \rangle_{h,h} = \sum_{k=0}^{\frac{1}{h}-1} h \chi_{k,k+1} \phi_{k,k+1}. \tag{2.28}$$

In particular, this holds if χ and ϕ are both graph gradients of a vertex function (and thus skew-symmetric).

We also observe that $\Delta_h u$ is equal to the second order central difference approximation of the second derivative of a function $u : \mathbb{T} \rightarrow \mathbb{R}$.

Finally, we note that by the Sobolev embedding [1, Theorem 4.12] and the finite measure of \mathbb{T} , $H^1(\mathbb{T}) \subset C(\mathbb{T}) \subset L^4(\mathbb{T})$. Hence, if $u \in H^1(\mathbb{T})$, then $W \circ u \in L^1(\mathbb{T})$ and thus $\varphi_W(u) < +\infty$. It follows that $H^1(\mathbb{T}) \subset \mathcal{D}(\varphi_{AC})$. Hence, by density of $H^1(\mathbb{T})$ in $L^2(\mathbb{T})$ (as follows from the density of $C^\infty(\mathbb{T})$ in $L^2(\mathbb{T})$), we have $\overline{\mathcal{D}(\varphi_{AC})} = L^2(\mathbb{T})$.

2.5.4 Extension Operator and Induced Inner Product

In our study of the total variation flow in Sect. 2.4 we used an embedding operator $i_h : L^2_h \rightarrow L^2(\mathbb{T}^n)$, which preserved the total variation in the sense of formula (2.21). This, in turn, ensured that condition (2.5) was satisfied as an equality. In the present case, the situation is more complicated as we cannot find an embedding operator that preserves the Allen–Cahn functional in a similar way. We define a new embedding operator.

Let $h \in \mathfrak{H}$. We define $I_h : \mathcal{V}_h \rightarrow L^2(\mathbb{T})$ via piecewise-linear interpolation

$$I_h u(x) := \frac{u_{k+1} - u_k}{h} (x - kh) + u_k,$$

where $k \in [\frac{1}{h} - 1]_0$ is such that $kh \leq x < (k+1)h$, i.e., $x \in I_{kh}^h = Q_{kh}^h$ in the notation of Sect. 2.3.1.

A direct computation shows that, for all $u \in \mathcal{V}_h$,

$$\begin{aligned} \varphi_D(I_h u) &= \frac{1}{2} \int_{\mathbb{T}} |(I_h u)'(x)|^2 dx = \frac{1}{2} \sum_{k=0}^{\frac{1}{h}-1} \int_{kh}^{(k+1)h} |(I_h u)'(x)|^2 dx \\ &= \frac{1}{2} \sum_{k=0}^{\frac{1}{h}-1} h \left(\frac{u_{k+1} - u_k}{h} \right)^2 dx = \frac{1}{2} \sum_{k=0}^{\frac{1}{h}-1} h^{-1} (u_{k+1} - u_k)^2 dx \\ &= \varphi_D^h(u), \end{aligned} \tag{2.29}$$

where we used that $u_{\frac{1}{h}} - u_{\frac{1}{h}-1} = u_0 - u_{\frac{1}{h}-1}$ by periodicity. We note that this does not suffice to conclude that (2.5) holds for the full functionals φ_{AC} and φ_{AC}^h . In fact, when we prove Theorem 2.7 below, we will use Theorem 2.2 rather than Theorem 2.1, to avoid needing requirement (2.5). The identity in (2.29) above, will be very useful in the proof of Theorem 2.7.

We note that $I_h : \mathcal{V}_h \rightarrow L^2(\mathbb{T})$ is a linear operator, but unlike i_h , when the domain is equipped with our usual inner product, $I_h : L_h^2 \rightarrow L^2(\mathbb{T})$ is not an isometry. This is problematic, as we required our embedding to be an isometry in Sect. 2.2. Therefore we introduce the inner product $(\cdot, \cdot)_h$ that is induced by I_h , i.e., for $v_1, v_2 \in \mathcal{V}_h$,

$$(v_1, v_2)_h := \langle I_h v_1, I_h v_2 \rangle_{L^2(\mathbb{T})}.$$

To differentiate the Hilbert space of vertex functions equipped with this new inner product from L_h^2 , we denote the set \mathcal{V}_h equipped with $(\cdot, \cdot)_h$ by \overline{L}_h^2 . We note that, by definition, this ensures that $I_h : \overline{L}_h^2 \rightarrow L^2(\mathbb{T})$ is an isometry. We denote the norm for $v \in \overline{L}_h^2$ by $\|v\|_{\overline{L}_h^2} := \sqrt{(v, v)_h}$. We emphasise that, as sets, \mathcal{V}_h , L_h^2 , and \overline{L}_h^2 are equal, i.e., $v \in \mathcal{V}_h$ if and only if $v \in L_h^2$, which in turn is true if and only if $v \in \overline{L}_h^2$.

Lemma 2.4 *The operator I_h satisfies, for all $u, v \in \mathcal{V}_h$,*

$$\langle I_h u, I_h v \rangle_{L^2(\mathbb{T})} = \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} \left[2u_k v_k + \frac{1}{2}(u_k v_{k+1} + u_{k+1} v_k) \right].$$

Moreover, $\frac{1}{3} \|u\|_h^2 = \frac{1}{3} \|i_h u\|_{L^2(\mathbb{T})}^2 \leq \|I_h u\|_{L^2(\mathbb{T})}^2 = \|u\|_{\overline{L}_h^2}^2 \leq \|i_h u\|_{L^2(\mathbb{T})}^2 = \|u\|_h^2$.

Proof For $a, b \in \mathbb{R}$, a direct computation shows that

$$\int_0^h \left| \frac{b-a}{h} x + a \right|^2 dx = \frac{h}{3} (a^2 + b^2 + ab).$$

This gives for $u \in \mathcal{V}_h$,

$$\begin{aligned}
 \|I_h u\|_{L^2(\mathbb{T})}^2 &= \sum_{k=0}^{\frac{1}{h}-1} \int_{kh}^{(k+1)h} \left| \frac{u_{k+1} - u_k}{h} (x - kh) + u_k \right|^2 dx \\
 &= \sum_{k=0}^{\frac{1}{h}-1} \int_0^h \left| \frac{u_{k+1} - u_k}{h} x + u_k \right|^2 dx \\
 &= \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} [u_k^2 + u_{k+1}^2 + u_k u_{k+1}]. \tag{2.30}
 \end{aligned}$$

Let $u, v \in \mathcal{V}_h$. By the polarisation identity, we now get

$$\begin{aligned}
 4\langle I_h u, I_h v \rangle_{L^2(\mathbb{T})} &= \|I_h(u+v)\|_{L^2(\mathbb{T})}^2 - \|I_h(u-v)\|_{L^2(\mathbb{T})}^2 \\
 &= \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} [(u_k + v_k)^2 + (u_{k+1} + v_{k+1})^2 + (u_k + v_k)(u_{k+1} + v_{k+1}) \\
 &\quad - (u_k - v_k)^2 - (u_{k+1} - v_{k+1})^2 - (u_k - v_k)(u_{k+1} - v_{k+1})] \\
 &= \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} [4u_k v_k + 4u_{k+1} v_{k+1} + 2u_k v_{k+1} + 2u_{k+1} v_k],
 \end{aligned}$$

from which the first desired identity follows, by periodicity.

Using periodicity in (2.30), we also find that

$$\|I_h u\|_{L^2(\mathbb{T})}^2 = \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} (2u_k^2 + u_k u_{k+1}).$$

Using Young's inequality $|u_k u_{k+1}| \leq (u_k^2 + u_{k+1}^2)/2$ and periodicity again yields

$$\begin{aligned}
 \frac{1}{3} \|u\|_h^2 &= \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} u_k^2 = \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} \left[2u_k^2 - \frac{u_k^2 + u_{k+1}^2}{2} \right] \leq \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} (2u_k^2 + u_k u_{k+1}) \\
 &= \|I_h u\|_{L^2(\mathbb{T})}^2 \leq \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} \left[2u_k^2 + \frac{u_k^2 + u_{k+1}^2}{2} \right] \\
 &= \|u\|_h^2.
 \end{aligned}$$

By the isometry property of Theorem 2.3 the required inequalities now follow. \square

Remark 2.10 Let $a, b \in \mathbb{R}$ and $m \in \mathbb{N}$. A direct computation shows

$$\int_0^h \left| \frac{b-a}{h}x + a \right|^m dx = \begin{cases} \frac{h}{(b-a)(m+1)} [b^{m+1} - a^{m+1}], & \text{if } m \text{ is even,} \\ \frac{h}{(b-a)(m+1)} [(\operatorname{sgn} b)b^{m+1} - (\operatorname{sgn} a)a^{m+1}], & \text{if } m \text{ is odd.} \end{cases}$$

A proof by induction shows that $b^{m+1} - a^{m+1} = (b-a) \sum_{l=0}^m a^l b^{m-l}$, hence if m is even we have

$$\int_0^h \left| \frac{b-a}{h}x + a \right|^m dx = \frac{h}{m+1} \sum_{\ell=0}^m a^\ell b^{m-\ell}.$$

This implies in particular that, for m even and for $u \in \mathcal{V}_h$,

$$\begin{aligned} \|I_h u\|_{L^m(\mathbb{T})}^m &:= \int_{\mathbb{T}} |I_h u|^m dx = \sum_{k=0}^{\frac{1}{h}-1} \frac{h}{m+1} \sum_{l=0}^m u_k^l u_{k+1}^{m-l} \\ &\leq \sum_{k=0}^{\frac{1}{h}-1} \frac{h}{m+1} \sum_{l=0}^m \left(l \frac{|u_k|^m}{m} + (m-l) \frac{|u_k|^m}{m} \right) \\ &= \sum_{k=0}^{1/h-1} h |u_k|^m := \|u\|_{L_h^m}^m. \end{aligned} \tag{2.31}$$

The inequality follows from Young's inequality, $|u_k^l u_{k+1}^{m-l}| \leq \frac{l}{m} |u_k|^m + \frac{m-l}{m} |u_{k+1}|^m$, and periodicity in k .

On the other hand, when m is even, l is odd, and $1 \leq l \leq m-1$, Young's inequality $|u_k u_{k+1}| \leq (u_k^2 + u_{k+1}^2)/2$ tells us that

$$\begin{aligned} u_k^l u_{k+1}^{m-l} &= u_k^{l-1} u_{k+1}^{m-l-1} (u_k u_{k+1}) \geq u_k^{l-1} u_{k+1}^{m-l-1} \left(-\frac{1}{2} u_k^2 - \frac{1}{2} u_{k+1}^2 \right) \\ &= -\frac{1}{2} u_k^{l+1} u_{k+1}^{m-l-1} - \frac{1}{2} u_k^{l-1} u_{k+1}^{m-l+1}. \end{aligned}$$

Using periodicity in k , it follows that

$$\begin{aligned} \sum_{l=0}^m u_k^l u_{k+1}^{m-l} &= \frac{1}{2} u_k^m + \frac{1}{2} u_{k+1}^m + \sum_{\substack{l=1 \\ l \text{ is odd}}}^{m-1} \left(u_k^l u_{k+1}^{m-l} + \frac{1}{2} u_k^{l+1} u_{k+1}^{m-l-1} + \frac{1}{2} u_k^{l-1} u_{k+1}^{m-l+1} \right) \\ &\geq \frac{1}{2} u_k^m + \frac{1}{2} u_{k+1}^m + \sum_{\substack{l=1 \\ l \text{ is odd}}}^{m-1} \left(-\frac{1}{2} u_k^{l+1} u_{k+1}^{m-l-1} - \frac{1}{2} u_k^{l-1} u_{k+1}^{m-l+1} \right) \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2}u_k^{l+1}u_{k+1}^{m-l-1} + \frac{1}{2}u_k^{l-1}u_{k+1}^{m-l+1} \Big) \\
& = \frac{1}{2}u_k^m + \frac{1}{2}u_{k+1}^m.
\end{aligned}$$

Summing over k and using periodicity again, we thus find that, for m even,

$$\begin{aligned}
\|I_h u\|_{L^m(\mathbb{T})}^m &= \frac{h}{m+1} \sum_{k=0}^{\frac{1}{h}-1} \sum_{l=0}^m u_k^l u_{k+1}^{m-l} \geq \frac{h}{m+1} \sum_{k=0}^{\frac{1}{h}-1} \left(\frac{1}{2}|u_k|^m + \frac{1}{2}|u_{k+1}|^m \right) \\
&= \frac{h}{m+1} \sum_{k=0}^{\frac{1}{h}-1} |u_k|^m \\
&= \frac{1}{m+1} \|u\|_{L_h^m}^m. \tag{2.32}
\end{aligned}$$

We note that $\|\cdot\|_h = \|\cdot\|_{L_h^2}$, so that the inequalities from Lemma 2.4 correspond to the case $m = 2$.

Lemma 2.5 *Let $u \in \mathcal{V}_h$, then*

$$\|I_h u - i_h u\|_{L^2(\mathbb{T})}^2 = \frac{h^2}{2} \|\nabla_h u\|_{h,h}^2 = h^2 \varphi_D^h(u) \leq h^2 \varphi_{AC}^h(u).$$

Proof We compute

$$\begin{aligned}
\|I_h u - i_h u\|_{L^2(\mathbb{T})}^2 &= \sum_{k=0}^{\frac{1}{h}-1} \int_{kh}^{(k+1)h} \left| \frac{u_{k+1} - u_k}{h} (x - kh) + u_k - u_k \right|^2 dx \\
&= \sum_{k=0}^{\frac{1}{h}-1} \int_0^h \left(\frac{u_{k+1} - u_k}{h} \right)^2 x^2 dx = \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} (u_{k+1} - u_k)^2 \\
&= \frac{h^2}{2} \|\nabla_h u\|_{h,h}^2 = h^2 \varphi_D^h(u) \leq h^2 \varphi_{AC}^h(u).
\end{aligned}$$

□

As a consequence of Lemma 2.4 we have, for $u, v \in \mathcal{V}_h$,

$$\begin{aligned}
(u, v)_h &= \frac{h}{3} \sum_{k=0}^{1/h-1} \left[2u_k v_k + \frac{1}{2}(u_k v_{k+1} + u_{k+1} v_k) \right] \\
&= \frac{h}{3} \sum_{k=0}^{1/h-1} \left[2u_k + \frac{u_{k-1} + u_{k+1}}{2} \right] v_k = \langle \Gamma u, v \rangle_h, \tag{2.33}
\end{aligned}$$

where we used periodicity and defined the operator Γ by

$$(\Gamma u)_k := \frac{1}{3} \left(2u_k + \frac{u_{k-1} + u_{k+1}}{2} \right). \quad (2.34)$$

We can consider Γ as an operator from L_h^2 to L_h^2 or from \bar{L}_h^2 to \bar{L}_h^2 . We note that $\langle \Gamma u, v \rangle_h = (u, v)_h = \langle I_h u, I_h v \rangle_{L^2(\mathbb{T})}$ and thus $\langle \Gamma u, u \rangle_h = \|u\|_h^2 = \|I_h u\|_{L^2(\mathbb{T})}^2$. Hence, by Lemma 2.4,

$$\frac{1}{3} \|u\|_h^2 \leq \langle \Gamma u, u \rangle_h \leq \|u\|_h^2. \quad (2.35)$$

The operator Γ can be identified with a circulant $1/h$ by $1/h$ -matrix, if one regards u as a $1/h$ -dimensional vector. This identification is helpful in deriving some interesting properties of Γ and its matrix exponential in Appendix B2.8.

Defining the operators τ_+ and τ_- by

$$(\tau_{\pm} u)_k := u_{k\pm 1}, \quad (2.36)$$

we have $\Gamma = \frac{2}{3} \text{Id} + \frac{1}{6} \tau_+ + \frac{1}{6} \tau_-$.

By direct computation we find that τ_{\pm} and Δ_h commute: $\tau_{\pm} \Delta_h u = \Delta_h \tau_{\pm} u$. It follows that Γ and Δ_h also commute. It is also useful to observe that

$$(\nabla_h \tau_{\pm} u)_{k,k+1} = h^{-1} ((\tau_{\pm} u)_{k+1} - (\tau_{\pm} u)_k) = h^{-1} (u_{k+1\pm 1} - u_{k\pm 1}) = (\nabla_h u)_{k\pm 1, k+1\pm 1},$$

and thus by (2.28), periodicity, and Young's inequality,

$$\begin{aligned} \langle \nabla_h u, \nabla_h \Gamma u \rangle_{h,h} &= h \sum_{k=0}^{\frac{1}{h}-1} \left[(\nabla_h u)_{k,k+1} \left(\nabla_h \left(\frac{2}{3} u + \frac{1}{6} \tau_- u + \frac{1}{6} \tau_+ u \right) \right)_{k,k+1} \right] \\ &= h \sum_{k=0}^{\frac{1}{h}-1} \left[\frac{2}{3} (\nabla_h u)_{k,k+1}^2 + \frac{1}{6} (\nabla_h u)_{k,k+1} (\nabla_h u)_{k-1,k} \right. \\ &\quad \left. + \frac{1}{6} (\nabla_h u)_{k,k+1} (\nabla_h u)_{k+1,k+2} \right] \\ &= h \sum_{k=0}^{\frac{1}{h}-1} \left[\frac{2}{3} (\nabla_h u)_{k,k+1}^2 + \frac{1}{3} (\nabla_h u)_{k,k+1} (\nabla_h u)_{k-1,k} \right] \\ &\geq h \sum_{k=0}^{\frac{1}{h}-1} \left[\frac{2}{3} (\nabla_h u)_{k,k+1}^2 - \frac{1}{6} (\nabla_h u)_{k,k+1}^2 - \frac{1}{6} (\nabla_h u)_{k-1,k}^2 \right] \\ &= h \sum_{k=0}^{\frac{1}{h}-1} \left[\frac{2}{3} (\nabla_h u)_{k,k+1}^2 - \frac{1}{3} (\nabla_h u)_{k,k+1}^2 \right] = h \sum_{k=0}^{\frac{1}{h}-1} \frac{1}{3} (\nabla_h u)_{k,k+1}^2 \\ &= \frac{1}{3} \|\nabla_h u\|_{h,h}^2. \end{aligned} \quad (2.37)$$

We note that by periodicity we also have

$$\|u - \tau_+ u\|_h^2 = \|u - \tau_- u\|_h^2 = h^2 \|\nabla_h u\|_{h,h}^2. \tag{2.38}$$

Again by periodicity, we have that $\langle \tau_{\pm} u, v \rangle_h = \langle u, \tau_{\mp} v \rangle_h$ and thus Γ is a self-adjoint operator on L_h^2 . It follows that $\Gamma - \frac{1}{3}\text{Id}$ is also self-adjoint; moreover, by Lemma 2.4 it is positive semidefinite on L_h^2 :

$$\left\langle \left(\Gamma - \frac{1}{3}\text{Id} \right) u, u \right\rangle_h = \|I_h u\|_{L^2(\mathbb{T})}^2 - \frac{1}{3} \|u\|_h^2 \geq 0.$$

As an aside, a direct computation also shows that $\langle (\Gamma - \text{Id}/3) u, u \rangle_h$ is proportional to the *signless* graph Dirichlet functional¹¹ on \mathcal{V}_h (see, for example, [40, Sect. 2.1]):

$$\begin{aligned} \left\langle \left(\Gamma - \frac{1}{3}\text{Id} \right) u, u \right\rangle_h &= \sum_{k=0}^{\frac{1}{h}-1} h \left(\frac{1}{3} u_k^2 + \frac{1}{6} u_{k-1} u_k + \frac{1}{6} u_{k+1} u_k \right) = \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} (u_{k+1} u_k + u_k^2) \\ &= \frac{h}{3} \sum_{k=0}^{\frac{1}{h}-1} \left(u_{k+1} u_k + \frac{1}{2} u_k^2 + \frac{1}{2} u_{k+1}^2 \right) = \frac{h}{6} \sum_{k=0}^{\frac{1}{h}-1} (u_k + u_{k+1})^2 \geq 0, \end{aligned} \tag{2.39}$$

where we used periodicity twice. We point out that the computation above also proves that Γ itself is positive definite on L_h^2 , since $\langle \Gamma u, u \rangle_h \geq \frac{1}{3} \langle u, u \rangle_h$. In particular, since L_h^2 has finite dimension, the (linear) inverse operator $\Gamma^{-1} : L_h^2 \rightarrow L_h^2$ exists.

By Lemma 2.4 it also follows that $\text{Id} - \Gamma$ is self-adjoint. It is also positive semidefinite on L_h^2 , since, by (2.33), $\langle (\text{Id} - \Gamma) u, u \rangle_h = \|u\|_h^2 - \|I_h u\|_{L^2(\mathbb{T})}^2 \geq 0$. It will be useful to have an explicit expression for this difference of norms:

$$\begin{aligned} \|u\|_h^2 - \|I_h u\|_{L^2(\mathbb{T})}^2 &= \sum_{k=0}^{\frac{1}{h}-1} h \left[\frac{1}{2} u_k^2 + \frac{1}{2} u_{k+1}^2 - \frac{1}{3} u_k^2 - \frac{1}{3} u_{k+1}^2 - \frac{1}{3} u_k u_{k+1} \right] \\ &= \frac{1}{6} \sum_{k=0}^{\frac{1}{h}-1} h (u_{k+1} - u_k)^2 \\ &= \frac{h^2}{6} \|\nabla_h u\|_{h,h}^2, \end{aligned} \tag{2.40}$$

where we used (2.30) and periodicity.

¹¹ The plus sign in $u_k + u_{k+1}$ on the penultimate right-hand side of (2.39), rather than a minus sign, makes it (proportional to) the *signless* graph Dirichlet functional.

Since Γ is a positive definite operator on L_h^2 , the square root $\Gamma^{\frac{1}{2}} : L_h^2 \rightarrow L_h^2$ exists and is positive definite. Moreover, since Γ is symmetric, so¹² is $\Gamma^{\frac{1}{2}}$. By (2.35), it follows that

$$\begin{aligned} (\Gamma u, u)_h &= \langle \Gamma^2 u, u \rangle_h = \langle \Gamma^{\frac{1}{2}} \Gamma u, \Gamma^{\frac{1}{2}} u \rangle_h = \langle \Gamma \Gamma^{\frac{1}{2}} u, \Gamma^{\frac{1}{2}} u \rangle_h \geq \frac{1}{3} \langle \Gamma^{\frac{1}{2}} u, \Gamma^{\frac{1}{2}} u \rangle_h \\ &= \frac{1}{3} \langle \Gamma u, u \rangle_h = \frac{1}{3} (u, u)_h, \end{aligned}$$

thus Γ is also positive definite as operator on \overline{L}_h^2 . Hence, since \overline{L}_h^2 has finite dimension, the (linear) inverse operator $\Gamma^{-1} : \overline{L}_h^2 \rightarrow \overline{L}_h^2$ exists. Furthermore, since $(\Gamma u, u)_h = \langle \Gamma^2 u, u \rangle_h = (u, \Gamma u)_h$, Γ is also symmetric as operator on \overline{L}_h^2 .

We end this subsection with the computation of the subdifferential of φ_{AC}^h with respect to the L_h^2 and \overline{L}_h^2 inner products. For the definition of the subdifferential of a convex function, we refer to [24, Chap. I, Sect. 5.1]. To be able to distinguish the subdifferential with respect to $\langle \cdot, \cdot \rangle_h$ (which is the one we have been using until now) from the subdifferential with respect to $(\cdot, \cdot)_h$, from now on we will write $\partial_{L_h^2}$ for the former and $\partial_{\overline{L}_h^2}$ for the latter.

We recall from (2.2) that, since φ_{AC}^h is geodesically $(-\alpha)$ -convex, we are interested in the shifted subdifferentials $-\partial_{L_h^2} (\varphi_{AC}^h + \frac{\alpha}{2} \|\cdot\|_h^2)(u) + \alpha u$ and

$$-\partial_{\overline{L}_h^2} \left(\varphi_{AC}^h + \frac{\alpha}{2} \|\cdot\|_h^2 \right) (u) + \alpha u.$$

Lemma 2.6 *Let $u \in \mathcal{V}_h$. The subdifferential of φ_D^h at u with respect to the L_h^2 inner product is the singleton $\partial_{L_h^2} \varphi_D^h(u) = \{-\Delta_h u\} \subset L_h^2$. Moreover,*

$$-\partial_{L_h^2} \left(\varphi_{AC}^h + \frac{\alpha}{2} \|\cdot\|_h^2 \right) (u) + \alpha u = \{\Delta_h u - W' \circ u\} \subset L_h^2.$$

¹² Since Γ is self-adjoint, there is a spectral decomposition $\Gamma = \sum_j \lambda_j P_j$ where the P_j are the orthogonal projections onto the eigenspaces corresponding to the eigenvalues λ_j , which all are real. The summation is over finitely many j , since L_h^2 has finite dimension. Since Γ is positive definite, we have, for all j , $\lambda_j > 0$. Then $\Gamma^{1/2}$ is defined as

$$\Gamma^{1/2} := \sum_j \lambda_j^{1/2} P_j.$$

Since P_j is self-adjoint, this $\Gamma^{1/2}$ must be self-adjoint. This argument can be generalised to self-adjoint operators on a Hilbert space. For any real-valued continuous function f and self-adjoint operator A on a Hilbert space, $f(A)$ is defined by the spectral decomposition [64, Sect. XI.6, Theorem 1] and $f(A)$ is self-adjoint [64, Sect. XI.12, Theorem 3]. We also refer to [55, Chap. 10] for further details about square roots of positive-semidefinite self-adjoint operators.

The subdifferential of φ_D^h at u with respect to the \bar{L}_h^2 inner product is the singleton $\partial_{\bar{L}_h^2} \varphi_D^h(u) = \{-\Gamma^{-1} \Delta_h u\} \subset \bar{L}_h^2$. Furthermore,

$$-\partial_{\bar{L}_h^2} \left(\varphi_{AC}^h + \frac{\alpha}{2} \|\cdot\|_h^2 \right) (u) + \alpha u = \{\Gamma^{-1} \Delta_h u - \Gamma^{-1} (W' \circ u)\} \subset L_h^2.$$

Proof First we note that, for all $v \in \mathcal{V}_h$,

$$\begin{aligned} \varphi_D^h(u+v) - \varphi_D^h(u) &= h^{-1} \sum_{k=0}^{\frac{1}{h}-1} [(u_{k+1} - u_k)(v_{k+1} - v_k) + \frac{1}{2}(v_{k+1} - v_k)^2] \\ &= \langle \nabla_h u, \nabla_h v \rangle_{h,h} + \frac{1}{2} \|\nabla_h v\|_{h,h}^2 \geq \langle -\Delta_h u, v \rangle_h, \end{aligned}$$

where we used (2.27) and (2.28). Since φ_D^h is convex, it follows from the definition of subdifferential that $-\Delta_h u \in \partial_{L_h^2} \varphi_D^h(u)$. Because the domain V_h of φ_D^h has finite dimension and since the functional is smooth, the subdifferential is in fact a singleton (and agrees with the gradient of the functional), thus $\partial_{L_h^2} \varphi_D^h(u) = \{-\Delta_h u\}$.

Moreover, by (2.33) we have $\langle -\Delta_h u, v \rangle_h = \langle -\Gamma \Delta_h u, v \rangle_h$ and thus

$$\varphi_D^h(u+v) - \varphi_D^h(u) \geq \langle -\Gamma^{-1} \Delta_h u, v \rangle_h.$$

As above, it follows that $-\Gamma^{-1} \Delta_h u \in \partial_{\bar{L}_h^2} \varphi_D^h(u)$ and, since the subdifferential is a singleton, in fact $\partial_{\bar{L}_h^2} \varphi_D^h(u) = \{-\Gamma^{-1} \Delta_h u\}$.

Since $W(\cdot) + \frac{\alpha}{2} \|\cdot\|_h^2$ is convex and differentiable, we have, for all $v \in \mathcal{V}_h$,

$$W(u+v) + \frac{\alpha}{2} \|u+v\|_h^2 - W(u) - \frac{\alpha}{2} \|u\|_h^2 \geq \langle W' \circ u + \alpha u, v \rangle_h.$$

Hence, for all $v \in \mathcal{V}_h$,

$$\varphi_{AC}^h(u+v) + \frac{\alpha}{2} \|u+v\|_h^2 - \varphi_{AC}^h(u) - \frac{\alpha}{2} \|u\|_h^2 \geq \langle -\Delta_h u + W' \circ u + \alpha u, v \rangle_h$$

and similarly

$$\begin{aligned} \varphi_{AC}^h(u+v) + \frac{\alpha}{2} \|u+v\|_h^2 - \varphi_{AC}^h(u) - \frac{\alpha}{2} \|u\|_h^2 &\geq \langle -\Delta_h u + W' \circ u + \alpha \Gamma u, v \rangle_h \\ &= \langle -(\Gamma^{-1} \Delta_h u - \Gamma^{-1} (W' \circ u) - \alpha u), v \rangle_h, \end{aligned}$$

where the equality follows from (2.33).

The remaining results now follow, using again the fact that the subdifferentials are singletons. \square

2.5.5 Convergence Result

Our main result will be the convergence of discrete Allen–Cahn gradient flows to a continuum Allen–Cahn gradient flow, which we formalise in Theorem 2.6.

We remind ourselves that a discrete Allen–Cahn gradient flow is a solution of $\dot{u} \in -\partial_{L^2_h} \varphi_{AC}^h(u)$. It satisfies equation (2.25), which (by Lemma 2.6) for $n = 1$ is explicitly given by

$$\dot{u}_k = (\Delta_h u)_k - W'(u_k) \tag{2.41}$$

for $k \in [\frac{1}{h} - 1]_0$. A continuum Allen–Cahn flow is an $L^2(\mathbb{T})$ -gradient flow of φ_{AC} in $L^2(\mathbb{T})$ and is given by (2.24) (with $n = 1$). Given initial data, the unique existence of such a flow is guaranteed by Proposition 2.1, since φ_{AC} is $(-\alpha)$ -convex. The same proposition, or the Picard–Lindelöf theorem for ordinary differential equations (ODE) [37, Theorem I.3.1], also guarantees the unique existence of a discrete Allen–Cahn gradient flow, i.e., a solution of (2.41), given initial data.

Theorem 2.6 *Let $u_0 \in L^2(\mathbb{T})$ and, for all¹³ $h \in \mathfrak{H}$, $u_0^h \in L^2_h$ (for $n = 1$). Assume that there exists an $h_* > 0$, such that $h \mapsto \|u_0^h\|_\infty$ is bounded on $(0, h_*) \cap \mathfrak{H}$. Furthermore, assume that $h\|\nabla u_0^h\|_{h,h} \rightarrow 0$ and $I_h u_0^h \rightarrow u_0$ in $L^2(\mathbb{T})$ as $h \rightarrow 0$ in \mathfrak{H} . Let u^h be the solution of (2.41) starting from u_0^h . Let u be the solution of (2.24) starting from u_0 . Then, for all $T \in (0, \infty)$,*

$$\lim_{h \rightarrow 0} \sup_{t \in [0, T]} \|I_h u^h(t) - u(t)\|_{L^2(\mathbb{T})} = 0.$$

Remark 2.11 By Lemma 2.4, $\|\cdot\|_{L^2(\mathbb{T})} \circ i_h$ and $\|\cdot\|_{L^2(\mathbb{T})} \circ I_h$ are equivalent norms on \mathcal{V}_h . Hence the result of Theorem 2.6 also holds if the operator I_h is replaced by i_h in the assumptions of Theorem 2.6. Moreover, using Lemma 2.5 we have

$$\begin{aligned} \|i_h u^h(t) - u(t)\|_{L^2(\mathbb{T})} &\leq \|i_h u^h(t) - I_h u(t)\|_{L^2(\mathbb{T})} + \|I_h u^h(t) - u(t)\|_{L^2(\mathbb{T})} \\ &\leq h (\varphi_{AC}^h(u(t)))^{1/2} + \|I_h u^h(t) - u(t)\|_{L^2(\mathbb{T})}. \end{aligned}$$

By Lemma 2.2

$$\varphi_{AC}^h(u(t)) \leq \varphi_{AC}^h(u_0^h) \leq \frac{1}{2} \|\nabla_h u_0^h\|_{h,h}^2 + C_h,$$

with $C_h := \sup \{W(s) \mid |s| \leq \|u_0^h\|_\infty\} = \max(W(0), W(\|u_0^h\|_\infty))$. (The last equality follows from the specific definition of W that we are using.) Hence, for $h \in (0, h_*) \cap \mathfrak{H}$,

$$h (\varphi_{AC}^h(u(t)))^{1/2} \leq \frac{h}{\sqrt{2}} \|\nabla_h u_0^h\|_{h,h} + C_h,$$

where $C := \max_{h \in (0, h_*) \cap \mathfrak{H}} C_h$. Thus, by the assumptions on the initial condition u_0^h ,

¹³ We can also restrict ourselves to a proper subset of \mathfrak{H} , as long as it contains a sequence converging to zero.

$$\lim_{h \rightarrow 0} h (\varphi_{AC}^h(u(t)))^{1/2} = 0.$$

Hence we can also replace I_h by i_h in the conclusion of Theorem 2.6 and the result still holds.

Unfortunately, we cannot apply our abstract Theorem 2.2 immediately since the embedding mapping I_h is not an isometry from L_h^2 to $L^2(\mathbb{T})$. To circumvent this difficulty, we proceed in two steps.

First we consider the gradient flow of φ_{AC}^h with respect to the \overline{L}_h^2 inner product. By Lemma 2.6 its explicit form is

$$\dot{U}_k = (\Gamma^{-1}(\Delta_h U - W' \circ U))_k, \tag{2.42}$$

for $k \in [\frac{1}{h} - 1]_0$. As for (2.41), standard ODE techniques guarantee that, given initial data, a unique solution to (2.42) exists. Since $I_h : \overline{L}_h^2 \rightarrow L^2(\mathbb{T})$ is an isometry, we can apply Theorem 2.2 to compare a solution of (2.42) with a solution of (2.24). The following theorem makes this precise.

Theorem 2.7 *Assume the same hypotheses concerning u_0 , u_0^h , and u as in Theorem 2.6. Let U^h be the solution of (2.42) starting from u_0^h . Then, for all $T \in (0, \infty)$,*

$$\lim_{h \rightarrow 0} \sup_{t \in [0, T]} \|I_h U^h(t) - u(t)\|_{L^2(\mathbb{T})} = 0.$$

We shall prove Theorem 2.7 as an application of Theorem 2.2 in Sect. 2.5.8 with $M_h = \overline{L}_h^2$ and $M = L^2(\mathbb{T})$.

The second step in our proof of Theorem 2.6 is a comparison between u^h , the solution of (2.41), and U^h , the solution of (2.42).

Theorem 2.8 *Assume the same hypotheses concerning u_0^h and u^h as in Theorem 2.6 and the same hypothesis concerning U^h as in Theorem 2.7. Then, for all $T \in (0, \infty)$,*

$$\lim_{h \rightarrow 0} \sup_{t \in [0, T]} \|I_h u^h(t) - I_h U^h(t)\|_{L^2(\mathbb{T})} = 0.$$

We defer the proof of Theorem 2.8 to Sect. 2.5.6.

Proof (Proof of Theorem 2.6) Our main convergence result Theorem 2.6 follows from Theorems 2.7 and 2.8. □

Remark 2.12 Since we allow the choice $T = \infty$ in Theorem 2.2, one may wonder why we do not allow this choice in Theorems 2.7 and 2.8 and consequently also not in Theorem 2.6.

This choice would introduce nontrivial difficulties, such as “infinite” constants N and C_N in (2.49) and (2.50), respectively, in the proof of Theorem 2.9 (which in turn is needed in the proof of Theorem 2.8) and the unboundedness of $t \mapsto e^{\alpha t}$ in (2.57) in the proof of Theorem 2.7.

The remaining parts of Sect. 2.5 are organised as follows. We actually prove our “second step” (Theorem 2.8) first, in Sect. 2.5.6. The reason for this reordering of our steps is that one of the results from that section, Lemma 2.42, is also required in our proof of Theorem 2.7. In Sect. 2.5.6 we prove Theorem 2.9 and in Sect. 2.5.8 Theorem 2.7, which also completes the proof of Theorem 2.6. This proof is an application of Theorem 2.2. We thus require a projection P_h as a counterpart to the embedding I_h . We introduce P_h in Sect. 2.5.7 and prove some of its properties.

2.5.6 Comparison of Two Discrete Flows (proof of Theorem 2.8)

In this subsection, we prove Theorem 2.8. In fact we consider equations more general than (2.42) and (2.41). Theorem 2.8 will then follow as a special case of Theorem 2.9.

As usual, if there is no further specification, we assume that $h \in \mathfrak{H}$. Up to the proof of Theorem 2.8 at the end of this section, h is fixed; therefore we have suppressed the dependence on h in the notation u and U for solutions of (2.43) and (2.44), respectively.

Let $F : \mathbb{R} \rightarrow \mathbb{R}$ be a continuously differentiable function with $F(0) = 0$ and, for all $x \in \mathbb{R}$, $F(x)x \geq 0$. Let $\mu \in \mathbb{R}$. We consider, for all $k \in [\frac{1}{h} - 1]_0$,

$$\dot{u}_k = (\Delta_h u)_k + \mu u_k - F(u_k), \quad (2.43)$$

as a generalisation of (2.41). We recover (2.41) if $F(x) = W'(x) + \mu x$ and $\mu > \alpha$. We note that $W'(x) = \alpha x(x^2 - 1)$ and $W''(x) = \alpha(3x^2 - 1)$, so that in this case indeed $F(0) = W'(0) = 0$ and $F'(x) = W''(x) + \mu > 3\alpha x^2 \geq 0$.

As a generalisation of (2.42), we consider, for all $k \in [\frac{1}{h} - 1]_0$,

$$\Gamma \dot{U}_k = (\Delta_h U + \mu U - F \circ U)_k. \quad (2.44)$$

By the Picard–Lindelöf theorem [37, Theorem I.3.1], the initial value problems corresponding to (2.43) and (2.44) are guaranteed to have unique continuously differentiable solutions locally in time, i.e., at least up to some time $T^* > 0$. By $C^1([0, T^*]; \mathcal{V}_h)$ we denote the set of functions $u : [0, T^*] \rightarrow \mathcal{V}_h$ that are continuously differentiable (considering one-sided derivatives at the endpoints of the interval). Any operator (such as Δ_h or Γ) which is defined on \mathcal{V}_h has a natural action on $C^1([0, T^*]; \mathcal{V}_h)$, defined by $(\Delta_h u)(t) = \Delta_h u(t)$, $(\Gamma u)(t) = \Gamma u(t)$, etc.

We start by proving a minimum principle for supersolutions and a comparison principle for solutions of (2.43). If U solves equation (2.44), then it satisfies the required assumptions of part (i) of Proposition 2.2 below, since, for all $k \in [\frac{1}{h} - 1]_0$,

$$b_k = -\mu + \int_0^1 F'(\theta U_k) d\theta$$

is bounded below if U_k is bounded. This is the case for continuous solutions U_k on the compact interval $[0, T]$.

Proposition 2.2

(i) (Minimum principle). Let $T > 0$ and assume that there exist a function $b : [0, T] \times V_h \rightarrow \mathbb{R}$ that is bounded below and a function $v \in C^1([0, T]; \mathcal{V}_h)$ such that, for all $k \in [\frac{1}{h} - 1]_0$,

$$\dot{v}_k \geq (\Delta_h v)_k - b_k v_k.$$

If, for all $k \in [\frac{1}{h} - 1]_0$, $v_k(0) \geq 0$, then, for all $k \in [\frac{1}{h} - 1]_0$ and for all $t \in [0, T]$, $v_k(t) \geq 0$.

(ii) (Growth estimate). Let $T > 0$ and assume that $u \in C^1([0, T]; \mathcal{V}_h)$ is a solution of (2.43) starting from $u_0 \in \mathcal{V}_h$. Then, for all $t \in [0, T]$, $\|u(t)\|_\infty \leq \|u_0\|_\infty e^{\mu t}$.

Proof (i) Without loss of generality we may assume that $b < 0$, for if it is not, let $\tilde{\mu} \in \mathbb{R}$ be such that

$$\tilde{\mu} > - \inf_{(k,t) \in [\frac{1}{h}-1]_0 \times [0,T]} b_k(t).$$

Since b is bounded below, this is well-defined. Then we consider the following inequality for $\tilde{v}(t) := e^{-\tilde{\mu}t} v(t)$:

$$\begin{aligned} \dot{\tilde{v}}(t) &= -\tilde{\mu} e^{-\tilde{\mu}t} v(t) + e^{-\tilde{\mu}t} \dot{v}(t) \geq -\tilde{\mu} \tilde{v}(t) + e^{-\tilde{\mu}t} \Delta_h v(t) - e^{-\tilde{\mu}t} b(t)v(t) \\ &= \Delta_h \tilde{v}(t) - (b(t) + \tilde{\mu}) \tilde{v}(t). \end{aligned}$$

By definition of $\tilde{\mu}$, we have $b + \tilde{\mu} > 0$. Moreover, $\tilde{v}_k(t) \geq 0$ if and only if $v_k(t) \geq 0$.

Suppose that the conclusion of part (i) were false. Since v is continuous in t , the minimum of $v_k(t)$ over all $(k, T) \in [\frac{1}{h} - 1]_0 \times [0, T]$ exists and is negative (i.e., < 0).

Take a minimiser $(k_0, t_0) \in [\frac{1}{h} - 1]_0 \times (0, T]$. Then $\dot{v}_{k_0}(t_0) \leq 0$ and

$$(\Delta_h v)_{k_0}(t_0) = \frac{(v_{k_0+1}(t_0) - v_{k_0}(t_0)) + (v_{k_0-1}(t_0) - v_{k_0}(t_0))}{h^2} \geq 0,$$

and thus $0 \geq -b_{k_0}(t_0)v_{k_0}(t_0)$. This contradicts $v_{k_0}(t_0) < 0$.

(ii) We will prove that, for all $t \in [0, T]$,

$$\max_k u_k(t) \leq \|u_0\|_\infty e^{\mu t}.$$

A symmetric argument, of which we will not give the details, yields the estimate from below by $-\|u_0\|_\infty e^{\mu t}$.

Define $\bar{u} \in C^1(\mathbb{R}; \mathcal{V}_h)$ by $\bar{u}(t) := \|u_0\|_\infty e^{\mu t}$. We observe that, for all $t \in \mathbb{R}$, $\bar{u}(t)$ is constant on V_h and thus $\Delta_h \bar{u} = 0$. Moreover, for all $t \in \mathbb{R}$ and all $k \in [\frac{1}{h} - 1]_0$, $(F \circ \bar{u})_k(t) = F(\|u_0\|_\infty e^{\mu t}) \geq 0$, since $F(0) = 0$ and $F(x) \geq 0$ for strictly positive¹⁴ x . Thus \bar{u} is a supersolution of (2.43), i.e., for all $t \in \mathbb{R}$,

$$\dot{\bar{u}}(t) = \mu \bar{u}(t) = \Delta_h \bar{u}(t) + \mu \bar{u}(t) \geq \Delta_h \bar{u}(t) + \mu \bar{u}(t) - F \circ \bar{u}(t).$$

Thus $v := \bar{u} - u$ solves, for all $t \in [0, T]$,

$$\dot{v} \geq \Delta_h v + \mu v - F \circ \bar{u} + F \circ u = \Delta_h v - b v,$$

where we defined, for all $k \in [\frac{1}{h} - 1]_0$ and for all $t \in [0, T]$,

$$b_k(t) := -\mu + \int_0^1 F'(\theta \bar{u}_k(t) + (1 - \theta)u_k(t)) d\theta.$$

Since u and \bar{u} are continuous, the sets $u([0, T])$ and $\bar{u}([0, T])$ are compact and thus, by Tychonoff's theorem so is $[0, 1] \times u([0, T]) \times \bar{u}([0, T])$. Since the function $\mathbb{R}^3 \rightarrow \mathbb{R}$, $(\theta, x, y) \mapsto \theta x + (1 - \theta)y$ is continuous, the set $\Theta := \left\{ \theta x + (1 - \theta)y \in \mathbb{R} \mid (\theta, x, y) \in [0, 1] \times u([0, T]) \times \bar{u}([0, T]) \right\}$ is compact. Because F' is continuous, the set $F'(\Theta)$ is compact and in particular bounded. This implies that the function b is bounded and thus we can apply part (i) of the current proposition. Doing so yields, for all $t \in [0, T]$ and for all $k \in [\frac{1}{h} - 1]_0$, $v_k(t) \geq 0$. The proof is now complete. \square

Remark 2.13 As a corollary to Proposition 2.2, the results from that proposition can be extended to include the case $T = \infty$, if we then interpret $t \in [0, T]$ to mean $t \in [0, \infty)$. For a given $t \in [0, \infty)$, if we let $\tilde{T} > 0$ be large enough such that $t \in [0, \tilde{T}]$, then the results of both parts of Proposition 2.2 hold for that t .

Remark 2.14 We note that the growth estimate from Proposition 2.2 (ii) allows us to extend our earlier local-in-time existence results for the initial value problems associated with (2.43) to global-in-time existence results.¹⁵ Given $u_0^h \in \mathcal{V}_h$, by the

¹⁴ To avoid ambiguities regarding zero, we call numbers from $(0, \infty)$ and $(-\infty, 0)$ strictly positive and strictly negative, respectively.

¹⁵ We remark that Proposition 2.1 also gives us global-in-time existence—based on an energy principle—but without guarantee of differentiability.

local-in-time result there exists $T^* > 0$ such that a solution $u^h \in C^1([0, T^*]; \mathcal{V}_h)$ exists of (2.43) with initial condition $u^h(0) = u_0^h$. Since the right-hand side of (2.43) is defined for all $u \in \mathcal{V}_h$ and continuous with respect to u , by standard ODE results (e.g., [37, Theorem I.2.1]), the maximal right-interval of existence of the solution is $[0, \infty)$ unless there exists a time T^{**} such that the solution can be extended to $[0, T^{**})$ and $\lim_{t \rightarrow T^{**}} u(t) = \pm\infty$. By the growth estimate it follows that such a T^{**} does not exist and thus the solution exists for all $t > 0$.

In the following proposition we compare a solution u of (2.43) with Γu .

Proposition 2.3 *Let $h \in \mathfrak{H}$ and let u be a solution of (2.43) with initial data $u_0 \in \mathcal{V}_h$. Then, for all t in the domain of u ,*

$$\begin{aligned} e^{-2\mu t} \|u(t) - \Gamma u(t)\|_h^2 + 2 \int_0^t e^{-2\mu s} \|\nabla_h(u(s) - \Gamma u(s))\|_{h,h}^2 ds \\ \leq \frac{h^2}{9} \|\nabla_h u_0\|_{h,h}^2. \end{aligned}$$

Proof To simplify notation, we suppress the explicit dependence on t (which is assumed to be in the domain of u) in the first half of this proof.

Remembering the definition of τ_{\pm} from (2.36), we have

$$3(u - \Gamma u) = \frac{1}{2} [(u - \tau_+ u) + (u - \tau_- u)]. \tag{2.45}$$

Hence

$$\begin{aligned} \|u - \Gamma u\|_h^2 &= \frac{1}{6^2} \|(u - \tau_+ u) + (u - \tau_- u)\|_h^2 \leq \frac{1}{18} (\|u - \tau_+ u\|_h^2 + \|u - \tau_- u\|_h^2) \\ &= \frac{1}{18} (\|v_+\|_h^2 + \|v_-\|_h^2), \end{aligned}$$

where we used the notation $v_{\pm} := u - \tau_{\pm} u$. We also used that $\|a + b\|^2 \leq (\|a\| + \|b\|)^2 = \|a\|^2 + \|b\|^2 + 2\|a\|\|b\| \leq 2\|a\|^2 + 2\|b\|^2$, which holds for any norm. Because ∇_h is a linear operator, we thus also have

$$\|\nabla_h(u - \Gamma u)\|_{h,h}^2 = \frac{1}{6^2} \|\nabla_h v_+ + \nabla_h v_-\|_{h,h}^2 \leq \frac{1}{18} (\|\nabla_h v_+\|_{h,h}^2 + \|\nabla_h v_-\|_{h,h}^2).$$

Since $\tau_{\pm} u$ solves (2.43), calculating the inner product $\langle u - \tau_{\pm} u, \dot{u} - \tau_{\pm} \dot{u} \rangle_h = \frac{1}{2} \frac{d}{dt} \|v_{\pm}\|_h^2$ yields

$$\frac{1}{2} \frac{d}{dt} \|v_{\pm}\|_h^2 - \mu \|v_{\pm} u\|_h^2 + \|\nabla_h v_{\pm}\|_{h,h}^2 = -\langle u - \tau_{\pm} u, F \circ u - F \circ \tau_{\pm} u \rangle_h.$$

Here we used (2.27). The last term is nonpositive since F is nondecreasing. This implies

$$\begin{aligned} & \frac{d}{dt} \left(e^{-2\mu t} \|v_{\pm}(t)\|_h^2 \right) + 2e^{-2\mu t} \|\nabla_h v_{\pm}(t)\|_{h,h}^2 \\ &= e^{-2\mu t} \left(\frac{d}{dt} \|v_{\pm}(t)\|_h^2 - 2\mu \|v_{\pm}(t)\|_h^2 + 2 \|\nabla_h v_{\pm}(t)\|_{h,h}^2 \right) \leq 0. \end{aligned}$$

Integrating this inequality over $(0, t)$ and using (2.38), we find

$$\begin{aligned} 0 &\geq e^{-2\mu t} \|v_{\pm}(t)\|_h^2 - \|v_{\pm}(0)\|_h^2 + 2 \int_0^t e^{-2\mu s} \|\nabla_h v_{\pm}(s)\|_{h,h}^2 ds \\ &= e^{-2\mu t} \|v_{\pm}(t)\|_h^2 - h^2 \|\nabla_h u_0\|_{h,h}^2 + 2 \int_0^t e^{-2\mu s} \|\nabla_h v_{\pm}(s)\|_{h,h}^2 ds. \end{aligned} \quad (2.46)$$

Recalling the estimates for $\|u - \Gamma u\|_h^2$ and $\|\nabla_h(u - \Gamma u)\|_{h,h}^2$ earlier in this proof, the desired inequality follows:

$$\begin{aligned} & e^{-2\mu t} \|u(t) - \Gamma u(t)\|_h^2 + 2 \int_0^t e^{-2\mu s} \|\nabla_h(u(s) - \Gamma u(s))\|_{h,h}^2 ds \\ &\leq \frac{1}{18} \left(e^{-2\mu t} \|v_+(t)\|_h^2 + 2 \int_0^t e^{-2\mu s} \|\nabla_h v_+(s)\|_{h,h}^2 ds \right. \\ &\quad \left. + e^{-2\mu t} \|v_-(t)\|_h^2 + 2 \int_0^t e^{-2\mu s} \|\nabla_h v_-(s)\|_{h,h}^2 ds \right) \\ &\leq \frac{h^2}{9} \|\nabla_h u_0\|_{h,h}^2. \end{aligned}$$

□

Theorem 2.9 *Let $h \in \mathfrak{H}$, $\mu \geq 0$, and $T \in (0, \infty)$. Assume $u \in C^1([0, T]; \mathcal{V}_h)$ solves (2.43) starting from $u_0 \in \mathcal{V}_h$ and $U \in C^1([0, T]; \mathcal{V}_h)$ solves (2.44) starting from $U_0 \in \mathcal{V}_h$. Then there exists a constant C_N , depending only on μ , T , and $\|u_0\|_{\infty}$ (details in (2.49) and (2.50)), such that, for all $t \in [0, T]$ and for all $\delta > 0$,*

$$\|\Gamma^{1/2}(u(t) - U(t))\|_h^2 \leq \left[h^2 e^{2\mu T} \left(\frac{1}{18} + \frac{C_N}{\mu \delta} \right) \|\nabla_h u_0\|_{h,h}^2 + \Gamma^{1/2}(u_0 - U_0)\|_h^2 \right] e^{6(\mu+\delta)t}.$$

In particular, if $u_0 = U_0$, then, for all $t \in [0, T]$ and all $\delta > 0$, there exists a $C_ > 0$ depending on μ , δ , T , and $\|u_0\|_{\infty}$ (details in (2.51)), such that*

$$\|u(t) - U(t)\|_h^2 \leq C_* h^2 \|\nabla_h u_0\|_{h,h}^2 e^{6(\mu+\delta)t}.$$

Proof Let $h \in \mathfrak{H}$. In most lines of the proof, to simplify the notation we will not write the explicit dependence on $t \in [0, T]$.

Similarly as in the proof of Proposition 2.3, we shall calculate $(v, v)_h = \langle v, \Gamma v \rangle_h = \|\Gamma^{1/2}v\|_h^2$ with $v := u - U$. Using equations (2.43) for u and (2.44) for U , together with the self-adjointness of Γ and the fact that Γ and $\frac{d}{dt}$ commute, we obtain that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\Gamma^{1/2}v\|_h^2 &= \langle v, \Gamma \dot{v} \rangle_h \\ &= \underbrace{\langle v, \Gamma \Delta_h u - \Delta_h U \rangle_h}_{=: I_1} - \underbrace{\langle v, \Gamma F \circ u - F \circ U \rangle_h}_{=: I_2} + \underbrace{\mu \langle v, \Gamma u - U \rangle_h}_{=: I_3}. \end{aligned} \quad (2.47)$$

We proceed by computing I_1 , using (2.27):

$$\begin{aligned} I_1 &= \langle v, \Delta_h v \rangle_h + \langle v, \Gamma \Delta_h u - \Delta_h u \rangle_h \\ &= -\langle \nabla_h v, \nabla_h v \rangle_{h,h} - \langle \nabla_h v, \nabla_h (\Gamma u - u) \rangle_{h,h} \end{aligned}$$

since $\Delta_h \Gamma = \Gamma \Delta_h$. Applying Young's inequality $ab \leq (a^2 + b^2)/2$, we obtain

$$\begin{aligned} I_1 &\leq -\|\nabla_h v\|_{h,h}^2 + \frac{1}{2} \|\nabla_h v\|_{h,h}^2 + \frac{1}{2} \|\nabla_h (\Gamma u - u)\|_{h,h}^2 \\ &= -\frac{1}{2} \|\nabla_h v\|_{h,h}^2 + \frac{1}{2} \|\nabla_h (u - \Gamma u)\|_{h,h}^2. \end{aligned}$$

We next calculate I_2 . Since F is nondecreasing, we see that

$$\begin{aligned} I_2 &= -\langle v, F \circ u - F \circ U \rangle_h - \langle v, \Gamma F \circ u - F \circ u \rangle_h \leq -\langle v, \Gamma F \circ u - F \circ u \rangle_h \\ &\leq \|v\|_h \|\Gamma F \circ u - F \circ u\|_h, \end{aligned} \quad (2.48)$$

where we used the Cauchy–Schwarz inequality to obtain the last inequality. Because

$$(\Gamma F \circ u)_k = \frac{2}{3} F(u_k) + \frac{1}{6} F(u_{k-1}) + \frac{1}{6} F(u_{k+1}),$$

a similar observation as the one for $u - \Gamma u$ in (2.45) shows that

$$F \circ u - \Gamma F \circ u = \frac{1}{6} \left[(F \circ u - F \circ \tau_+ u) + (F \circ u - F \circ \tau_- u) \right].$$

By Proposition 2.2 (ii), we know $\|u\|_\infty(t) \leq \|u_0\|_\infty e^{\mu t}$. Thus, pointwise on V_h ,

$$\begin{aligned} |F \circ u(t) - F \circ \tau_\pm u(t)| &\leq \left| \int_0^1 F'(\theta u(t) + (1-\theta)\tau_\pm u(t)) d\theta \right| |u(t) - \tau_\pm u(t)| \\ &\leq N |u(t) - \tau_\pm u(t)| \end{aligned}$$

with

$$N := \sup \{ F'(s) \mid |s| \leq \|u_0\|_\infty e^{\mu T} \}. \quad (2.49)$$

This estimate, combined with (2.48) and the triangle inequality, yields

$$\begin{aligned} I_2 &\leq \frac{1}{6} \|v\|_h (\|F \circ u - F \circ \tau_+ u\|_h + \|F \circ u - F \circ \tau_- u\|_h) \\ &\leq \frac{N}{6} \|v\|_h (\|u - \tau_+ u\|_h + \|u - \tau_- u\|_h). \end{aligned}$$

Similarly, using (2.45) and the Cauchy–Schwarz and triangle inequalities again, we estimate

$$\begin{aligned} I_3 &= \mu \|v\|_h^2 + \mu \langle v, \Gamma u - u \rangle_h \\ &\leq \mu \|v\|_h^2 + \frac{\mu}{6} \|v\|_h (\|u - \tau_+ u\|_h + \|u - \tau_- u\|_h). \end{aligned}$$

Let $\delta > 0$. Applying Young's inequality, $ab \leq \frac{\delta}{2} a^2 + \frac{1}{2\delta} b^2$, to the estimates for I_2 and I_3 above, we find

$$\begin{aligned} I_2 &\leq \frac{\delta}{2} \|v\|_h^2 + \frac{1}{2\delta} \left(\frac{N}{6} \right)^2 (\|u - \tau_+ u\|_h + \|u - \tau_- u\|_h)^2, \\ I_3 &\leq \mu \|v\|_h^2 + \frac{\delta}{2} \|v\|_h^2 + \frac{1}{2\delta} \left(\frac{\mu}{6} \right)^2 (\|u - \tau_+ u\|_h + \|u - \tau_- u\|_h)^2. \end{aligned}$$

Combining our estimates for I_1 , I_2 , and I_3 with (2.47) we find

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} \|\Gamma^{1/2} v\|_h^2(t) + \frac{1}{2} \|\nabla_h v\|_{h,h}^2 \\ &\leq \frac{1}{2} \|\nabla_h(u - \Gamma u)\|_{h,h}^2 + (\mu + \delta) \|v\|_h^2 + \frac{C_N}{2\delta} (\|u - \tau_+ u\|_h^2 + \|u - \tau_- u\|_h^2), \end{aligned}$$

with

$$C_N := \frac{N^2 + \mu^2}{36}. \quad (2.50)$$

Integrating over $(0, t)$ and multiplying by 2, we get, by rearranging terms,

$$\begin{aligned} &\|\Gamma^{1/2} v(t)\|_h^2 + \int_0^t \|\nabla_h v(s)\|_{h,h}^2 ds \leq \int_0^t \|\nabla_h(u - \Gamma u)(s)\|_{h,h}^2 ds \\ &+ 2(\mu + \delta) \int_0^t \|v(s)\|_h^2 ds \\ &+ \frac{C_N}{\delta} \left(\int_0^t \|u(s) - \tau_+ u(s)\|_h^2 ds + \int_0^t \|u(s) - \tau_- u(s)\|_h^2 ds \right) + \|\Gamma^{1/2}(u_0 - U_0)\|_h^2. \end{aligned}$$

By Proposition 2.3 we have

$$\begin{aligned} \int_0^t \|\nabla_h(u(s) - \Gamma u(s))\|_{h,h}^2 ds &\leq e^{2\mu T} \int_0^t e^{-2\mu s} \|\nabla_h(u(s) - \Gamma u(s))\|_{h,h}^2 ds \\ &\leq \frac{h^2}{18} e^{2\mu T} \|\nabla_h u_0\|_{h,h}^2. \end{aligned}$$

From (2.46) in the proof of Proposition 2.3, we see that

$$e^{-2\mu t} \|u(t) - \tau_{\pm} u(t)\|_h^2 + 2 \int_0^t e^{-2\mu s} \|\nabla_h(u - \tau_{\pm} u)(s)\|_{h,h}^2 ds \leq h^2 \|\nabla_h u_0\|_{h,h}^2.$$

This yields

$$\begin{aligned} \int_0^t \|u(s) - \tau_{\pm} u(s)\|_h^2 ds &\leq h^2 \|\nabla_h u_0\|_{h,h}^2 \int_0^t e^{2\mu s} ds = \frac{h^2}{2\mu} \|\nabla_h u_0\|_{h,h}^2 (e^{2\mu t} - 1) \\ &\leq \frac{h^2}{2\mu} e^{2\mu T} \|\nabla_h u_0\|_{h,h}^2. \end{aligned}$$

Applying these estimates, we deduce that

$$\begin{aligned} \|\Gamma^{1/2} v(t)\|_h^2 + \int_0^t \|\nabla_h v(s)\|_{h,h}^2 ds \\ \leq h^2 e^{2\mu T} \left(\frac{1}{18} + \frac{C_N}{\mu\delta} \right) \|\nabla_h u_0\|_{h,h}^2 + \|\Gamma^{1/2}(u_0 - U_0)\|_h^2 \\ + 2(\mu + \delta) \int_0^t \|v(s)\|_h^2 ds. \end{aligned}$$

After applying $\|v\|_h^2 \leq 3\|\Gamma^{1/2} v\|_h^2$ from (2.35) to the integrand on the right-hand side, Grönwall's inequality [37, Corollary I.6.6] yields, for all $t \in [0, T]$,

$$\|\Gamma^{1/2} v(t)\|_h^2 \leq \left[h^2 e^{2\mu T} \left(\frac{1}{18} + \frac{C_N}{\mu\delta} \right) \|\nabla_h u_0\|_{h,h}^2 + \|\Gamma^{1/2}(u_0 - U_0)\|_h^2 \right] e^{6(\mu+\delta)t}.$$

This proves the first result.

In the special case where $u_0 = U_0$, we find

$$\|v(t)\|_h^2 \leq C_* h^2 \|\nabla_h u_0\|_{h,h}^2 e^{6(\mu+\delta)t},$$

with

$$C_* := e^{2\mu T} \left(\frac{1}{6} + \frac{3C_N}{\mu\delta} \right). \quad (2.51)$$

This proves the second desired estimate. \square

We can now prove Theorem 2.8.

Proof (of Theorem 2.8) Let $h \in \mathfrak{H}$, $\delta > 0$, and $t \in [0, T]$. By Lemma 2.4 and Theorem 2.9 (with $u = u^h$, $U = U^h$, $\mu > \alpha$, and $F(x) = W'(x) + \mu x$) we have

$$\|I_h U^h(t) - I_h u^h(t)\|_{L^2(\mathbb{T})} = \|\Gamma^{1/2}(U^h(t) - u^h(t))\|_h \leq C_*^{1/2} h \|\nabla_h u_0^h\|_{h,h} e^{3(\mu+\delta)t}.$$

By assumption we also have that $\|u_0^h\|_\infty$ is bounded uniformly in h , for h small enough. Consequently so is $C_*^{1/2}$ (which depends on h only through $\|u_0^h\|_\infty$).

First taking the supremum over $t \in [0, T]$ in the inequality above and subsequently the limit for $h \rightarrow 0$ yields Theorem 2.8, since, by assumption, $h \|\nabla_h u_0^h\|_{h,h} \rightarrow 0$ as $h \rightarrow 0$. \square

2.5.7 Projections

Let $h \in \mathfrak{H}$. We define H_h to be the space of real-valued continuous piecewise-linear functions on \mathbb{T} whose derivatives can only have jumps at the points $kh \in \mathbb{T}$ with $k \in \{0, \dots, 1/h - 1\}$. (As an aside we observe that, if $f \in H_{h,i}$ —see (2.66)—then, keeping all coordinates but x_i fixed, the function $x_i \mapsto f(x)$ is in H_h .) We note that $I_h \bar{L}_h^2 = H_h$. Moreover, H_h is a linear subspace of $L^2(\mathbb{T})$ and, having finite dimension, it is also closed.

To be able to use Theorem 2.2 to prove Theorem 2.7, we require a projection P_h corresponding to the embedding $I_h : \bar{L}_h^2 \rightarrow L^2(\mathbb{T})$, as explained in Sect. 2.2. We will define this projection using a few intermediary functions, which will prove useful in their own right.

We define $\bar{P}_h : L^2(\mathbb{T}) \rightarrow H_h$ to be the orthogonal projection from $L^2(\mathbb{T})$ to H_h . This is well-defined, since H_h is a finite-dimensional, and thus closed, subspace of the Hilbert space $L^2(\mathbb{T})$. As one of the standard properties of orthogonal projections, it follows immediately that \bar{P}_h satisfies, for all $w \in L^2(\mathbb{T})$,

$$\|\bar{P}_h w\|_{L^2(\mathbb{T})} \leq \|w\|_{L^2(\mathbb{T})}. \quad (2.52)$$

Moreover, for all $w \in L^2(\mathbb{T})$, there exists a unique $w^\perp \in H_h^\perp$ (where H_h^\perp denotes the orthogonal complement of H_h in $L^2(\mathbb{T})$), such that $w = \bar{P}_h w + w^\perp$. Since $1 \in H_h$, we have

$$\int_{\mathbb{T}} \bar{P}_h w(x) dx = \int_{\mathbb{T}} w(x) \cdot 1 dx - \int_{\mathbb{T}} w^\perp(x) \cdot 1 dx = \int_{\mathbb{T}} w(x) dx. \quad (2.53)$$

With $C(\mathbb{T})$ being the set of continuous functions on \mathbb{T} , in this one-dimensional case we define the projection $\bar{\pi}_h : C(\mathbb{T}) \rightarrow \bar{L}_h^2$ by, for all $w \in C(\mathbb{T})$,

$$(\bar{\pi}_h w)(kh) = w(kh), \quad \text{for } k \in \left[\frac{1}{h} - 1 \right]_0.$$

Now we define $P_h : L^2(\mathbb{T}) \rightarrow \overline{L}_h^2$ as $P_h := \overline{\pi}_h \circ \overline{P}_h$. Since the restricted operator $\overline{P}_h|_{H_h}$ is the identity map Id_{H_h} on H_h , we have that $P_h \circ I_h = \overline{\pi}_h \circ \overline{P}_h \circ I_h = \overline{\pi}_h \circ \text{Id}_{H_h} \circ I_h = \text{Id}_{\overline{L}_h^2}$, the identity map on \overline{L}_h^2 , as required by our discussion in Sect. 2.2.

Furthermore, we note that $I_h \circ \overline{\pi}_h|_{H_h} = \text{Id}_{H_h}$. Hence $I_h \circ P_h = I_h \circ \overline{\pi}_h \circ \overline{P}_h = \overline{P}_h$, and thus for all $w \in L^2(\mathbb{T})$ we have, by (2.52),

$$\|P_h w\|_{\overline{h}} = \|I_h \circ P_h w\|_{L^2(\mathbb{T})} = \|\overline{P}_h w\|_{L^2(\mathbb{T})} \leq \|w\|_{L^2(\mathbb{T})}.$$

Hence the non-strict contraction requirement from (2.4) is also satisfied. Moreover, using again that \overline{P}_h is an orthogonal projection, we have, for all $v \in \overline{L}_h^2$ and for all $w \in L^2(\mathbb{T})$,

$$\begin{aligned} \|v - P_h w\|_{\overline{h}}^2 + \|I_h P_h w - w\|_{L^2(\mathbb{T})}^2 &= \|I_h v - I_h P_h w\|_{L^2(\mathbb{T})}^2 + \|I_h P_h w - w\|_{L^2(\mathbb{T})}^2 \\ &= \|I_h v - \overline{P}_h w\|_{L^2(\mathbb{T})}^2 + \|\overline{P}_h w - w\|_{L^2(\mathbb{T})}^2 \\ &= \|I_h v - w\|_{L^2(\mathbb{T})}^2. \end{aligned}$$

For the final equality, we used that $I_h v \in H_h$. This establishes that P_h satisfies condition (2.7). We postpone our check of condition (2.9) to the proof of Theorem 2.7 in Sect. 2.5.8.

Our goal in the remainder of Sect. 2.5.7 is to establish some other useful properties of \overline{P}_h , which we will need in our proof of Theorem 2.7. We collect these results in Theorem 2.10: in particular, Theorem 2.10 parts (i) and (iv) are used in the proof of Theorem 2.7, while parts (i) and (ii) are used in the proof of part (iii), which in turn is used to prove part (iv). In numerical analysis, the properties in parts (ii) and (iii) combined sometimes go by the name of stability in $H^1(\mathbb{T})$.

Theorem 2.10

(i) For all $w \in H^1(\mathbb{T})$ the estimate

$$\|(\overline{P}_h w)'\|_{L^2(\mathbb{T})} \leq C \|w'\|_{L^2(\mathbb{T})}$$

holds with $C := 4\sqrt{3}/\pi + 2$.

(ii) For all $w \in H^1(\mathbb{T})$, $\lim_{h \rightarrow 0} \|(\overline{P}_h w - w)'\|_{L^2(\mathbb{T})} = 0$.

(iii) For all $w \in L^2(\mathbb{T})$, $\lim_{h \rightarrow 0} \|\overline{P}_h w - w\|_{L^2(\mathbb{T})} = 0$.

(iv) For all $w \in H^1(\mathbb{T})$, $\overline{P}_h w \rightarrow w$ uniformly, as $h \rightarrow 0$.

In [32, Theorem 1.5], this is proved using the Ritz operator, i.e., the projection operator from $H^1(\mathbb{T})$ to H_h , and elliptic regularity in a general domain. A more general estimate in $L^1(\mathbb{T})$ and $W^{1,p}(\mathbb{T})$ has been established in [18, Theorem 2], where the two-dimensional case is also discussed.

Here we prove Theorem 2.10 by other means. To do so, we first require a different projection operator, π_h , and its properties which we present in Lemma 2.7. We present a proof of Theorem 2.10 after we have proven this lemma. We define $\pi_h : C(\mathbb{T}) \rightarrow H_h \subset C(\mathbb{T})$, by following $\bar{\pi}_h$ with linear interpolation on each interval $(kh, (k+1)h)$, i.e., $\pi_h := I_h \circ \bar{\pi}_h$.

In the following lemma we collect some basic properties of the error $e_h w := \pi_h w - w \in C(\mathbb{T})$. Recall that $H^2(\mathbb{T}) \subset H^1(\mathbb{T}) \subset C(\mathbb{T})$ by the Sobolev embedding [1] and so π_h is well-defined on $H^1(\mathbb{T})$ and $H^2(\mathbb{T})$. We also note that $H_h \subset H^1(\mathbb{T})$.

In the proof of the following lemma, we require two Poincaré–Wirtinger(-type) inequalities. In order not to interrupt the flow of the main argument, we defer the statement and proof of these inequalities to Lemma 2.14 in Appendix C2.9.

Lemma 2.7 *The following estimates are valid.*

- (i) $\|e_h w\|_{L^2(\mathbb{T})}^2 \leq (2h/\pi)^2 \|w'\|_{L^2(\mathbb{T})}^2$, for all $w \in H^1(\mathbb{T})$.
- (ii) $\|e_h w\|_{L^2(\mathbb{T})}^2 \leq (h/\pi)^4 \|w''\|_{L^2(\mathbb{T})}^2$, for all $w \in H^2(\mathbb{T})$.
- (iii) $\|(e_h w)'\|_{L^2(\mathbb{T})}^2 \leq \frac{1}{3} h^2 \|w''\|_{L^2(\mathbb{T})}^2$, for all $w \in H^2(\mathbb{T})$.
- (iv) $\|(\pi_h w)'\|_{L^2(\mathbb{T})} \leq \|w'\|_{L^2(\mathbb{T})}$, for all $w \in H^1(\mathbb{T})$.
- (v) $\|w'\|_{L^2(\mathbb{T})}^2 \leq (12/h^2) \|w\|_{L^2(\mathbb{T})}^2$, for all $w \in H_h$.

Proof We first prove (iv). Let $w \in H^1(\mathbb{T})$. By definition

$$\int_0^h |(\pi_h w)'(x)|^2 dx = \left| \frac{w(h) - w(0)}{h} \right|^2 h = \left| \frac{1}{h} \int_0^h w'(x) dx \right|^2 \cdot h.$$

By Hölder's inequality or Jensen's inequality [27, Appendix B Theorem 2], this is dominated by

$$\frac{1}{h} \int_0^h |w'(x)|^2 dx \cdot h = \int_0^h |w'|^2 dx.$$

Similarly,

$$\int_{kh}^{(k+1)h} |(\pi_h w)'(x)|^2 dx \leq \int_{kh}^{(k+1)h} |w'(x)|^2 dx.$$

Adding these inequalities from $k = 0$ to $k = 1/h - 1$ and taking square roots yields (iv).

We next prove (i). Let $w \in H^1(\mathbb{T})$. Since $e_h w(0) = e_h w(h) = 0$, by the Poincaré–Wirtinger inequality in Lemma 2.14 (i) we have

$$\int_0^h |e_h w(x)|^2 dx \leq (h/\pi)^2 \|(e_h w)'\|_{L^2((0,h))}^2.$$

By (iv) we have $\|(\pi_h w)'\|_{L^2(\mathbb{T})} \leq \|w'\|_{L^2(\mathbb{T})}$, so that $\|(e_h w)'\|_{L^2(\mathbb{T})} \leq 2\|w'\|_{L^2(\mathbb{T})}$. Hence

$$\int_0^h |e_h w(x)|^2 dx \leq (2h/\pi)^2 \|w'\|_{L^2((0,h))}^2.$$

Similar inequalities can be derived on the intervals $(kh, (k + 1)h)$. Adding these from $k = 0$ to $k = 1/h - 1$ yields (i).

For the proof of (ii), assume that $w \in H^2(\mathbb{T})$. Since $e_h w(0) = e_h w(h) = 0$, by the Poincaré–Wirtinger-type inequality in Lemma 2.14 (ii) we have

$$\int_0^h |e_h w(x)|^2 dx \leq (h/\pi)^4 \int_0^h |(e_h w)''(x)|^2 dx.$$

Since $(\pi_h w)'' = 0$ a.e. on $(0, h)$, this implies

$$\int_0^h |e_h w(x)|^2 dx \leq (h/\pi)^4 \int_0^h |w''(x)|^2 dx.$$

Similarly as before, this yields

$$\|e_h w\|_{L^2(\mathbb{T})}^2 \leq (h/\pi)^4 \|w''\|_{L^2(\mathbb{T})}^2.$$

To prove (iii), we again let $w \in H^2(\mathbb{T})$ and consider the integral

$$\int_0^h |(e_h w)'(x)|^2 dx = \int_0^h \left| w'(x) - \frac{w(h) - w(0)}{h} \right|^2 dx.$$

Setting $u = w'$, this integral becomes

$$\int_0^h \left| u(x) - \frac{1}{h} \int_0^h u(y) dy \right|^2 dx = \int_0^h \left| \frac{1}{h} \int_0^h (u(x) - u(y)) dy \right|^2 dx.$$

This is dominated by

$$\int_0^h \left(\frac{1}{h} \int_0^h |u(y) - u(x)|^2 dy \right) dx,$$

by the Hölder or Jensen inequality. Either of these inequalities also implies that

$$|u(y) - u(x)|^2 = \left| \int_x^y u'(s) ds \right|^2 \leq |x - y| \int_x^y |u'(s)|^2 ds \leq |x - y| \int_0^h |u'(s)|^2 ds,$$

where the final inequality holds for $x, y \in (0, h)$. Combining these observations, we have

$$\begin{aligned} \int_0^h |(e_h w)'(x)|^2 dx &\leq \frac{1}{h} \int_0^h \int_0^h |x - y| dx dy \int_0^h |w''(s)|^2 ds \\ &= \tilde{C} h^2 \int_0^h |w''(s)|^2 ds, \end{aligned}$$

with, as required,

$$\begin{aligned} \tilde{C} &= \int_0^1 \int_0^1 |x - y| dx dy = \int_0^1 \left[\int_0^y (y - x) dx + \int_y^1 (x - y) dx \right] dy \\ &= \int_0^1 \left[\left(y^2 - \frac{1}{2} y^2 \right) + \left(\frac{1}{2} - y - \frac{1}{2} y^2 + y^2 \right) \right] dy = \int_0^1 \left(y^2 - y + \frac{1}{2} \right) dy = \frac{1}{3}. \end{aligned}$$

We thus obtain (iii).

Finally, we shall prove the reverse Poincaré–Wirtinger-type estimate (v) for $w \in H_h$. It suffices to prove this inequality in $(0, h)$ for a linear function w , as the full result on \mathbb{T} will then follow by summing the contributions over all the intervals $(kh, (k+1)h)$ as in earlier parts of this proof. We may assume that $w(x) = ax - b$ for $a > 0, b \in \mathbb{R}$, by the symmetry provided by the square in the integrand. By overall multiplication of w by a constant, we may assume that $a = 1$. Then

$$\int_0^h (w'(x))^2 dx = h,$$

while

$$\begin{aligned} \int_0^h w^2(x) dx &= \int_0^h (x - b)^2 dx \\ &\geq \int_0^h (x - h/2)^2 dx = 2 \cdot \frac{1}{3} \left(\frac{h}{2} \right)^3 = \frac{1}{12} h^3. \end{aligned}$$

The inequality follows since $b \mapsto \int_0^h (x - b)^2 dx$ achieves its minimum value at $b = \frac{h}{2}$. Thus we have

$$\int_0^h (w'(x))^2 dx \leq (12/h^2) \int_0^h w^2(x) dx.$$

As is common practice in this proof by now, adding similar inequalities for each interval $(kh, (k+1)h)$ yields (iv). \square

Proof (of Theorem 2.10)

- (i) Let $w \in H^1(\mathbb{T})$. This implies that also $e_h w \in H^1(\mathbb{T})$. Since $\pi_h w \in H_h$ and \bar{P}_h is a linear operator, we observe that

$$((\bar{P}_h - I)w)' = -((\bar{P}_h(e_h w))' + (e_h w)').$$

From Lemma 2.7 (v) we see that

$$\left\| (\bar{P}_h(e_h w))' \right\|_{L^2(\mathbb{T})} \leq \frac{2\sqrt{3}}{h} \|\bar{P}_h(e_h w)\|_{L^2(\mathbb{T})} \leq \frac{2\sqrt{3}}{h} \|e_h w\|_{L^2(\mathbb{T})},$$

where for the last inequality we used (2.52). By Lemma 2.7 (iv), we have

$$\|(e_h w)'\|_{L^2(\mathbb{T})} \leq 2\|w'\|_{L^2(\mathbb{T})}.$$

Collecting these inequalities, we deduce that

$$\left\| ((\bar{P}_h - I)w)' \right\|_{L^2(\mathbb{T})} \leq \frac{2\sqrt{3}}{h} \|e_h w\|_{L^2(\mathbb{T})} + 2\|w'\|_{L^2(\mathbb{T})}.$$

Estimating the first term of the right-hand side by Lemma 2.7 (i) yields

$$\left\| ((\bar{P}_h - I)w)' \right\|_{L^2(\mathbb{T})} \leq \left(\frac{4\sqrt{3}}{\pi} + 2 \right) \|w'\|_{L^2(\mathbb{T})}.$$

- (ii) First we assume that $w \in H^2(\mathbb{T})$. As in (i) we have

$$\left\| ((\bar{P}_h - I)w)' \right\|_{L^2(\mathbb{T})} \leq \frac{2\sqrt{3}}{h} \|(e_h w)\|_{L^2(\mathbb{T})} + \|(e_h w)'\|_{L^2(\mathbb{T})}.$$

By Lemma 2.7 (ii) and (iii), this estimate yields

$$\left\| ((\bar{P}_h - I)w)' \right\|_{L^2(\mathbb{T})} \leq \frac{2\sqrt{3}}{h} (h/\pi)^2 \|w''\|_{L^2(\mathbb{T})} + \frac{h}{\sqrt{3}} \|w''\|_{L^2(\mathbb{T})}.$$

Thus, by part (i) of this theorem,

$$\left\| ((\bar{P}_h - I)w)' \right\|_{L^2(\mathbb{T})} \rightarrow 0, \quad \text{as } h \rightarrow 0.$$

Now assume $w \in H^1(\mathbb{T})$ instead and $\delta > 0$. By density, there exists a $w_\delta \in H^2(\mathbb{T})$ such that $\|(w - w_\delta)'\|_{L^2(\mathbb{T})} < \delta$. As observed above,

$$\left\| (\bar{P}_h w_\delta - w_\delta)' \right\|_{L^2(\mathbb{T})} \rightarrow 0, \quad \text{as } h \rightarrow 0.$$

By part (i) of this theorem and the triangle inequality, we have

$$\begin{aligned} \left\| (\overline{P}_h w - w)' \right\|_{L^2(\mathbb{T})} &\leq \left\| (\overline{P}_h w_\delta - w_\delta)' \right\|_{L^2(\mathbb{T})} + \left\| (\overline{P}_h(w - w_\delta))' \right\|_{L^2(\mathbb{T})} \\ &\quad + \left\| (w - w_\delta)' \right\|_{L^2(\mathbb{T})} \\ &\leq \left\| (\overline{P}_h w_\delta - w_\delta)' \right\|_{L^2(\mathbb{T})} + (C + 1) \left\| (w - w_\delta)' \right\|_{L^2(\mathbb{T})}. \end{aligned}$$

Sending h to zero yields

$$\overline{\lim}_{h \rightarrow 0} \left\| (\overline{P}_h w - w)' \right\|_{L^2(\mathbb{T})} \leq (C + 1) \left\| (w - w_\delta)' \right\|_{L^2(\mathbb{T})} < (C + 1)\delta.$$

Taking $\delta \rightarrow 0$ gives the desired convergence.

- (iii) First we assume that $w \in H^1(\mathbb{T})$. From (2.52) and part (i) of this theorem, we know that also $\overline{P}_h w \in H^1(\mathbb{T})$. By (2.53) we have that $\int_{\mathbb{T}} (\overline{P}_h w - w) dx = 0$, so by the Poincaré–Wirtinger inequality for functions with zero “mass” [44, Theorem 13.27 and Exercise 13.13], part (ii) of this theorem implies that $\overline{P}_h w \rightarrow w$ strongly in $L^2(\mathbb{T})$.

Now let $w \in L^2(\mathbb{T})$. By density of $H^1(\mathbb{T})$ in $L^2(\mathbb{T})$, for all $\delta > 0$ there exists a $w_\delta \in H^1(\mathbb{T})$ such that $\|w - w_\delta\|_{L^2(\mathbb{T})} < \delta$. From (2.52) and the triangle inequality it follows that

$$\begin{aligned} \left\| \overline{P}_h w - w \right\|_{L^2(\mathbb{T})} &\leq \left\| \overline{P}_h w_\delta - w_\delta \right\|_{L^2(\mathbb{T})} + \left\| \overline{P}_h w_\delta - \overline{P}_h w \right\|_{L^2(\mathbb{T})} \\ &\quad + \left\| w_\delta - w \right\|_{L^2(\mathbb{T})} \\ &\leq \left\| \overline{P}_h w_\delta - w_\delta \right\|_{L^2(\mathbb{T})} + 2 \left\| w_\delta - w \right\|_{L^2(\mathbb{T})} \\ &< \left\| \overline{P}_h w_\delta - w_\delta \right\|_{L^2(\mathbb{T})} + 2\delta. \end{aligned}$$

By the first part of the proof of (iii) we have that $\lim_{h \rightarrow 0} \left\| \overline{P}_h w_\delta - w_\delta \right\|_{L^2(\mathbb{T})} = 0$. Hence, if we first take the limit superior for $h \rightarrow 0$ in the inequality above and then the limit for $\delta \rightarrow 0$, the required result follows.

- (iv) Let (h_j) be a sequence of positive real numbers which converges to zero and let (h_{j_k}) be a subsequence. By Morrey’s inequality [44, Theorem 11.34], (2.52), and part (i) of this theorem, there exists some constants $\tilde{C}_1, \tilde{C}_2 > 0$ such that

$$\left\| \overline{P}_{h_{j_k}} w \right\|_{L^\infty(\mathbb{T})} \leq \tilde{C}_1 \left\| \overline{P}_{h_{j_k}} w \right\|_{H^1(\mathbb{T})} \leq \tilde{C}_2 \|w\|_{H^1(\mathbb{T})},$$

thus the sequence $(\overline{P}_{h_{j_k}} w)$ is uniformly bounded.

Let $x, y \in \mathbb{T}$ and assume $[x, y] \subset \mathbb{T}$ with appropriate periodicity taken into account. By Hölder’s inequality and part (i) of this theorem we have

$$\begin{aligned}
 |\overline{P}_{h_{j_k}} w(x) - \overline{P}_{h_{j_k}} w(y)| &= \left| \int_{[x,y]} \overline{P}_{h_{j_k}} w'(x) dx \right| \\
 &\leq \int_{[x,y]} |\overline{P}_{h_{j_k}} w'(x)| dx \leq \int_{\mathbb{T}} |\overline{P}_{h_{j_k}} w'(x)| dx \\
 &\leq \|1\|_{L^2(\mathbb{T})} \|\overline{P}_{h_{j_k}} w'\|_{L^2(\mathbb{T})} = \|\overline{P}_{h_{j_k}} w'\|_{L^2(\mathbb{T})} \\
 &\leq \|w'\|_{L^2(\mathbb{T})}.
 \end{aligned}$$

Thus the sequence $(\overline{P}_{h_{j_k}} w)$ is equicontinuous. By the Arzelà–Ascoli theorem, it follows that there exists a $\overline{w} \in C(\mathbb{T})$ and a subsubsequence $(\overline{P}_{h_{j_{k_l}}} w)$ such that $\overline{P}_{h_{j_{k_l}}} w \rightarrow \overline{w}$ uniformly as $l \rightarrow \infty$. Hence, since \mathbb{T} has finite measure, we also have $\overline{P}_{h_{j_{k_l}}} w \rightarrow \overline{w}$ in $L^2(\mathbb{T})$ and thus, by part (iii), $\overline{w} = w$. Since every subsequence of $(\overline{P}_{h_j} w)$ has a further subsubsequence which converges uniformly to w , it follows that $(\overline{P}_{h_j} w)$ itself also converges to w uniformly. From this we conclude the desired result. □

2.5.8 Proof of Theorem 2.7

We shall prove Theorem 2.7 as an application of Theorem 2.2.

Proof (of Theorem 2.7) Let $T > 0$. In the notation of Theorem 2.2, we set $M_h = \overline{L}_h^2$, $M = L^2(\mathbb{T})$, $i_h = I_h$, $p_h = P_h$, $\Phi_h = \varphi_{AC}^h$, and $\Phi = \varphi_{AC}$. Since \overline{L}_h^2 is a Hilbert space, it is closed and thus $\overline{\mathcal{D}(\varphi_{AC}^h)} = \overline{L}_h^2$. As derived at the end of Sect. 2.5.2, we also have $\overline{\mathcal{D}(\varphi_{AC})} = L^2(\mathbb{T})$. By our discussion in Sect. 2.5.1 we know that φ_{AC}^h and φ_{AC} satisfy condition $(-\alpha\text{-AGS})$. In Sect. 2.5.7 we established that condition (2.7) is satisfied. It remains to show that conditions (2.9), (2.10), and (2.11) are satisfied.

First we shall prove that (2.9) holds. Let $w \in L^2(\mathbb{T})$. Without loss of generality we may assume that $\varphi_{AC}(w) < +\infty$ and thus $w \in H^1(\mathbb{T})$. By (2.29) and Theorem 2.10 (i), we see that

$$\lim_{h \rightarrow 0} \varphi_D^h(P_h w) = \lim_{h \rightarrow 0} \varphi_D(I_h P_h w) = \lim_{h \rightarrow 0} \varphi_D(\overline{P}_h w) = \varphi_D(w).$$

By Theorem 2.10 (iv) we have that $P_h w \rightarrow w$ uniformly in \mathbb{T} , thus

$$\lim_{h \rightarrow 0} \varphi_W(\overline{P}_h w) = \varphi_W(w).$$

We claim that

$$\lim_{h \rightarrow 0} |\varphi_W^h(P_h w) - \varphi_W(\overline{P}_h w)| = 0. \tag{2.54}$$

Combining this with the limit above gives

$$\lim_{h \rightarrow 0} \varphi_W^h(P_h w) = \varphi_W(w),$$

which together with the convergence of $\varphi_D^h(P_h w)$ above establishes (2.9). To prove

the claim in (2.54), we recall that $\varphi_W^h(P_h w) = \varphi_W^h(\bar{\pi}_h \bar{P}_h w) = \sum_{k=0}^{\frac{1}{h}-1} h W(\bar{P}_h w(kh))$

and thus

$$\begin{aligned} |\varphi_W^h(P_h w) - \varphi_W(\bar{P}_h w)| &\leq \sum_{k=0}^{\frac{1}{h}-1} \left| h W(\bar{P}_h w(kh)) - \int_{kh}^{(k+1)h} W(\bar{P}_h w(x)) dx \right| \\ &= \sum_{k=0}^{\frac{1}{h}-1} \left| \int_{kh}^{(k+1)h} [W(\bar{P}_h w(kh)) - W(\bar{P}_h w(x))] dx \right| \\ &\leq \sum_{k=0}^{\frac{1}{h}-1} \int_{kh}^{(k+1)h} |W(\bar{P}_h w(kh)) - W(\bar{P}_h w(x))| dx. \end{aligned} \tag{2.55}$$

For the integrand on the right-hand side we estimate

$$\begin{aligned} |W(\bar{P}_h w(kh)) - W(\bar{P}_h w(x))| &\leq |W(\bar{P}_h w(kh)) - W(w(kh))| \\ &\quad + |W(w(kh)) - W(w(x))| \\ &\quad + |W(w(x)) - W(\bar{P}_h w(x))|, \end{aligned}$$

where we remember that in the integral $x \in [kh, (k+1)h]$ and thus $|x - kh| \leq h$. Let $\eta > 0$. Since $w \in H^1(\mathbb{T})$, we know by the Sobolev embedding that $w \in C(\mathbb{T})$ and thus, since W is continuous, $W \circ w \in C(\mathbb{T})$. Since \mathbb{T} is compact this means that $W \circ w$ is uniformly continuous. Hence there exists a $h_1 > 0$ such that, if $0 < h < h_1$, then $|W(w(kh)) - W(w(x))| < \eta/3$. Moreover, since $\bar{P}_h w \rightarrow w$ uniformly and since W is continuous, we have that $W(\bar{P}_h w) \rightarrow W(w)$ uniformly. Hence, there exists a $h_2 > 0$ such that, if $0 < h < h_2$, then $|W(w(kh)) - W(w(x))| < \eta/3$ and $|W(\bar{P}_h w(kh)) - W(\bar{P}_h w(x))| < \eta/3$. Defining $h^* := \min(h_1, h_2)$ we find that, if $0 < h < h^*$, then $|W(\bar{P}_h w(kh)) - W(\bar{P}_h w(x))| < \eta$. Substituting this into (2.55), we find, for $h \in (0, h^*)$,

$$|\varphi_W^h(P_h w) - \varphi_W(\bar{P}_h w)| < \sum_{k=0}^{\frac{1}{h}-1} \int_{kh}^{(k+1)h} \eta dx = \frac{1}{h} \cdot h \cdot \eta = \eta,$$

which proves the claim in (2.54) and thereby concludes the proof of (2.9).

Next we turn our attention to establishing the existence of a $\delta > 0$ and a nonnegative function $\Psi : L^2(\mathbb{T}) \rightarrow \mathbb{R} \cup \{+\infty\}$, such that $\Psi(u(\cdot)) \in L^1(0, T)$ and (2.10) is satisfied for all $h \in (0, \delta) \cap \mathfrak{H}$. In fact, we will show that for our choice of Ψ , (2.10) is satisfied for all $h \in \mathfrak{H}$, so that any choice of $\delta > 0$ is valid. Hence, let $h \in \mathfrak{H}$.

By Proposition 2.1 we know that for all $t > 0$, $u(t) \in \mathcal{D}(\varphi_{AC}) \subset H^1(\mathbb{T})$. Therefore, we define $\Psi := +\infty$ on $L^2(\mathbb{T}) \setminus H^1(\mathbb{T})$. Now let $w \in H^1(\mathbb{T})$. By (2.29) and Theorem 2.10 (i), we have

$$\varphi_D^h(P_h w) = \varphi_D(\overline{P}_h w) \leq \frac{C^2}{2} \|w'\|_{L^2(\mathbb{T})}^2.$$

Recalling that $W(x) = \alpha(x^2 - 1)^2/4$, we observe that

$$\varphi_W(\overline{P}_h w) \leq \frac{\alpha}{4} \left(\int_{\mathbb{T}} |\overline{P}_h w|^4 dx + 1 \right).$$

By Theorem 2.10 (i), the one-dimensional Sobolev (Morrey) inequality, and (2.52), there exists a constant $c_0 > 0$ such that

$$\begin{aligned} \|\overline{P}_h w\|_{L^\infty(\mathbb{T})} &\leq c_0 \left(\|(\overline{P}_h w)'\|_{L^2(\mathbb{T})} + \|\overline{P}_h w\|_{L^2(\mathbb{T})} \right) \leq c_0 \left(C \|w'\|_{L^2(\mathbb{T})} + \|w\|_{L^2(\mathbb{T})} \right) \\ &\leq c_1 \|w\|_{H^1(\mathbb{T})}, \end{aligned}$$

with $c_1 := c_0 C$.

By the interpolation inequality for L^p spaces [1, Theorem 2.11] (a special case of the Gagliardo–Nirenberg interpolation inequality), (2.52), and the inequality above, we have

$$\|\overline{P}_h w\|_{L^4(\mathbb{T})} \leq \|\overline{P}_h w\|_{L^2(\mathbb{T})}^{1/2} \|\overline{P}_h w\|_{L^\infty(\mathbb{T})}^{1/2} \leq C_0 \|w\|_{L^2(\mathbb{T})}^{1/2} \|w\|_{H^1(\mathbb{T})}^{1/2}, \tag{2.56}$$

with $C_0 := c_1^{1/2}$.

Thus

$$\varphi_W(\overline{P}_h w) \leq C_1 \|w\|_{L^2(\mathbb{T})}^2 \|w\|_{H^1(\mathbb{T})}^2 + \frac{\alpha}{4},$$

where $C_1 := \alpha C_0^4/4$. We note that C_1 is independent of h and w .

Using the explicit form of W , Lemma 2.4 (or, equivalently, (2.31) with $m = 2$), (2.32) with $m = 4$, and (2.56), we find

$$\begin{aligned} \varphi_W^h(P_h w) - \varphi_W(\overline{P}_h w) &= \varphi_W^h(P_h w) - \varphi_W(I_h P_h w) \\ &= \frac{\alpha}{4} \left(\|P_h w\|_{L_h^4}^4 - \|I_h P_h w\|_{L^4(\mathbb{T})}^4 - 2\|P_h w\|_h^2 + 2\|I_h P_h w\|_{L^2(\mathbb{T})}^2 \right) \\ &\leq \frac{\alpha}{2} \left(\|P_h w\|_{L_h^4}^4 - \|I_h P_h w\|_{L^4(\mathbb{T})}^4 \right) \\ &\leq \alpha \|I_h P_h w\|_{L^4(\mathbb{T})}^4 = \alpha \|\overline{P}_h w\|_{L^4(\mathbb{T})}^4 \\ &\leq C_2 \|w\|_{L^2(\mathbb{T})}^2 \|w\|_{H^1(\mathbb{T})}^2, \end{aligned}$$

with $C_2 := \alpha C_0^4$. It thus follows that

$$\varphi_W^h(P_h w) \leq \varphi_W(\overline{P}_h w) + C_2 \|w\|_{L^2(\mathbb{T})}^2 \|w\|_{H^1(\mathbb{T})}^2 \leq C_3 \|w\|_{L^2(\mathbb{T})}^2 \|w\|_{H^1(\mathbb{T})}^2 + \frac{\alpha}{4},$$

where $C_3 := C_1 + C_2 = \frac{5}{4}\alpha C_0^4$.

We define, for all $w \in L^2(\mathbb{T})$,

$$\Psi(w) := \begin{cases} \frac{C_3}{2} \|w'\|_{L^2(\mathbb{T})}^2 + C_3 \|w\|_{L^2(\mathbb{T})}^2 \|w\|_{H^1(\mathbb{T})}^2 + \frac{\alpha}{4}, & \text{if } w \in H^1(\mathbb{T}), \\ +\infty, & \text{otherwise.} \end{cases}$$

From the inequalities above, it follows that for all $w \in L^2(\mathbb{T})$ and all $h \in \mathfrak{H}$ the inequality in (2.10) is satisfied. Moreover, since $\alpha > 0$, Ψ is nonnegative.

As we observed earlier, by Proposition 2.1 we have that for all $t > 0$, $u(t) \in H^1(\mathbb{T})$, thus the functions $t \mapsto \|u(t)\|_{L^2(\mathbb{T})}$ and $t \mapsto \|\nabla u(t)\|_{L^2(\mathbb{T})}$ are well-defined from $(0, \infty)$ to $[0, \infty)$. By the same proposition we know that $t \mapsto \|u(t)\|_{L^2(\mathbb{T})}$ is continuous on $[0, \infty)$ (and hence measurable). It is also bounded on $[0, T]$ since, by the generalised contraction principle in (2.3) and the geodesic $(-\alpha)$ -convexity of φ_{AC} with α , we have

$$\|u(t) - 0\|_{L^2(\mathbb{T})} \leq e^{\alpha t} \|u_0 - 0\|_{L^2(\mathbb{T})} \leq e^{\alpha T} \|u_0 - 0\|_{L^2(\mathbb{T})}. \quad (2.57)$$

By Lemma 2.2 we have that $t \mapsto \varphi_{AC}(u(t))$ is a non-increasing function on $[0, \infty)$ and thus $\|\nabla u(t)\|_{L^2(\mathbb{T})}$ is also bounded on $[0, T]$. Moreover, by monotonicity, $t \mapsto \varphi_{AC}(u(t))$ is measurable on $[0, \infty)$ and, since $u \in C([0, T], L^2(\mathbb{T}))$ by Proposition 2.1, if we view u as real-valued function on $\mathbb{T} \times [0, T]$, it is also measurable.¹⁶ Thus so is $W \circ u$ and hence, by Fubini's theorem [28, Sect. I.4, Theorem 1], $t \mapsto \int_{\mathbb{T}} W \circ (u(t)) dx$ is measurable on $[0, T]$. Hence,

$$t \mapsto \frac{1}{2} \|\nabla u(t)\|_{L^2(\mathbb{T})}^2 = \varphi_{AC}(u(t)) - \int_{\mathbb{T}} W \circ (u(t)) dx$$

is measurable on $[0, T]$ as well.

¹⁶ To prove this, partition $[0, T]$ into intervals $[a_0, a_1], [a_1, a_2], \dots, [a_{m-1}, a_m]$ with $m \in \mathbb{N}$ and $a_{i+1} - a_i = T/m$. Set

$$u_m(t) := u(a_j) \quad \text{for } a_j \leq t < a_{j+1}, \quad j = 0, \dots, m-1.$$

Viewed as function on $\mathbb{T} \times [0, T]$, u_m is measurable and it is in $L^2(\mathbb{T} \times (0, T))$. Since $u \in C([0, T], L^2(\mathbb{T}))$ and $[0, T]$ is compact, the function u is uniformly continuous on $[0, T]$ with values in $L^2(\mathbb{T})$. Because, for all $t \in [0, T]$, there is an a_j such that $\|u(t) - u_m(t)\|_{L^2(\mathbb{T})} = \|u(t) - u(a_j)\|_{L^2(\mathbb{T})}$, uniform continuity of u implies convergence in $C([0, T], L^2(\mathbb{T}))$ of the sequence (u_m) to u . In particular, (u_m) is a Cauchy sequence in $L^2(\mathbb{T} \times (0, T))$. As $L^2(\mathbb{T} \times (0, T))$ is complete by the Fisher-Riesz theorem [10, Theorem 4.8], [64, Sect. I.9, Proposition 2], we have $u \in L^2(\mathbb{T} \times (0, T))$ and thus u is measurable.

It follows that $t \mapsto \Psi(u(t))$ is well-defined on $[0, T]$ and bounded on the same interval, so that

$$\int_0^T e^{-2\alpha t} \Psi(u(t)) dt < \infty.$$

It remains to prove that there exists a function $\varepsilon : (0, 1) \rightarrow (0, \infty)$ such that $\varepsilon(h) \rightarrow 0$ as $h \rightarrow 0$ and (2.11) is satisfied, i.e., for all $t \in [0, T]$,

$$\varphi_{AC}(I_h U^h(t)) \leq \varphi_{AC}^h(U^h(t)) + \varepsilon(h). \tag{2.58}$$

Let $t \in [0, T]$. By (2.29) we have $\varphi_D(I_h U^h(t)) = \varphi_D^h(U^h(t))$, and (2.31) with $m = 4$ shows that $\|I_h U^h(t)\|_{L^4(\mathbb{T})}^4 \leq \|U^h(t)\|_{L_h^4}^4$. Combined with the explicit form of W , which we recalled earlier in this proof, this implies that

$$\begin{aligned} & \varphi_{AC}(I_h U^h(t)) - \varphi_{AC}^h(U^h(t)) \\ &= \frac{\alpha}{4} \left(2 \|U^h(t)\|_h^2 - 2 \|I^h U^h(t)\|_{L^2(\mathbb{T})}^2 - \|U^h(t)\|_{L_h^4}^4 + \|I^h U^h(t)\|_{L^4(\mathbb{T})}^4 \right) \\ &\leq \frac{\alpha}{2} \left(\|U^h(t)\|_h^2 - \|I^h U^h(t)\|_{L^2(\mathbb{T})}^2 \right) \\ &= \frac{\alpha h^2}{12} \|\nabla_h U^h(t)\|_{h,h}^2. \end{aligned}$$

For the last inequality we invoked (2.40). To estimate $\|\nabla_h U^h(t)\|_{h,h}^2$, we recall equation (2.42) in the form (2.44), with $\mu > \alpha$ arbitrary. We apply ∇_h to both sides of (2.44) and take the inner product $\langle \cdot, \cdot \rangle_h$ with $-\Delta_h U^h$. By (2.27), we obtain

$$\begin{aligned} \langle \nabla_h U^h, \nabla_h \Gamma \dot{U}^h \rangle_{h,h} &= \langle -\Delta_h U^h, \Gamma \dot{U}^h \rangle_h \\ &= -\langle \Delta_h U^h, \Delta_h U^h \rangle_h - \mu \langle \Delta_h U^h, U^h \rangle_h + \langle \Delta_h U^h, F \circ U^h \rangle_h \\ &= -\|\Delta_h U^h\|_h^2 + \mu \|\nabla_h U^h\|_{h,h}^2 - \langle \nabla_h U^h, \nabla_h (F \circ U^h) \rangle_{h,h}. \end{aligned}$$

We note that, since $(F \circ U^h)_k = F(U_k^h)$,

$$\begin{aligned} (\nabla_h (F \circ U^h))_{k,k\pm 1} &= \frac{1}{h} \left\{ (F \circ U^h)_{k\pm 1} - (F \circ U^h)_k \right\} \\ &= (\nabla_h U^h)_{k,k\pm 1} \int_0^1 F'(\theta U_{k\pm 1}^h + (1-\theta)U_k^h) d\theta, \end{aligned}$$

so the monotonicity of F implies that

$$\langle \nabla_h U^h, \nabla_h (F \circ U^h) \rangle_{h,h} \geq 0.$$

Thus, we obtain

$$\langle \nabla_h U^h, \nabla_h \Gamma \dot{U}^h \rangle_{h,h} \leq - \|\Delta_h U^h\|_h^2 + \mu \|\nabla_h U^h\|_{h,h}^2 \leq \mu \|\nabla_h U^h\|_{h,h}^2.$$

Since $\mu > \alpha$ can be taken arbitrary, we may replace μ by α in the right-hand side.

Since Γ commutes with ∇_h and Γ is symmetric, the left-hand side equals

$$\frac{1}{2} \frac{d}{dt} \langle \nabla_h U^h, \Gamma \nabla_h U^h \rangle_{h,h}$$

which equals

$$\frac{1}{6} \frac{d}{dt} \|\nabla_h U^h\|_{h,h}^2$$

by (2.37). We thus obtain

$$\frac{1}{6} \frac{d}{dt} \|\nabla_h U^h\|_{h,h}^2 \leq \alpha \|\nabla_h U^h\|_{h,h}^2,$$

which yields by the differential form of Grönwall's inequality:

$$\|\nabla_h U^h(t)\|_{h,h}^2 \leq e^{6\alpha t} \|\nabla_h u_0^h\|_{h,h}^2.$$

Combining these observations, we conclude that

$$\varphi_{AC}(I_h U^h(t)) - \varphi_{AC}^h(U^h(t)) \leq \frac{\alpha h^2}{12} e^{6\alpha t} \|\nabla_h u_0^h\|_{h,h}^2.$$

If we define

$$\varepsilon(h) := \frac{\alpha h^2}{12} e^{6\alpha T} \|\nabla_h u_0^h\|_{h,h}^2,$$

then (2.58) follows by the assumption that $h \|\nabla_h u_0^h\|_h \rightarrow 0$ as $h \rightarrow 0$.

The result now follows by Theorem 2.2 and the additional comments about the “finite T ” case in Remark 2.6. \square

2.6 Conclusions and Future Work

In this paper we proved discrete-to-continuum convergence for total variation flow on a discretised n -dimensional torus and for Allen–Cahn flow on a discretised one-dimensional torus. Two potential generalisations immediately suggest themselves: Allen–Cahn flow convergence on a discrete torus of higher dimension and, for both flows, convergence of the flows on other graphs.

Our proof of convergence for the Allen–Cahn flow is not trivially extended to higher dimensions, since it makes extensive use of the properties of the operator Γ , which depend on the use of piecewise-linear interpolation to associate discrete functions with continuum functions. If piecewise- n -linear interpolation (bilinear, trilinear, etc.) were to be used instead, a different operator Γ (and corresponding embedding and projection operators) would need to be constructed. In higher dimensions, the space H^1 also loses a lot of its regularity.

Other graphs, besides the discrete torus, that are being considered by Samuel Mercer and the second author, are random geometric graphs, obtained by sampling points from Euclidean space (or possibly from a differentiable manifold) according to some probability density in order to construct the vertex set and subsequently connecting nodes with (potentially weighted) edges depending on their pairwise distances. The identification of discrete functions with continuum functions will then require a method different from the piecewise-constant and piecewise-linear interpolation methods that we used in this paper, such as the method based on optimal transport techniques that was introduced in [33].

Other flows can also be considered. For example, in the field of machine learning there is much current interest in proving consistency of methods through continuum limits of discrete dynamics.

Another possible direction for future research is the relationship between the continuum limit and the singular limit, for those dynamics that have an interesting singular limit. For example, if the potential term in the Allen–Cahn equation of (2.24) is rescaled by a factor ε^{-2} , then in the sharp interface limit $\varepsilon \downarrow 0$ solutions evolve according to flow by mean curvature [2, 14, 19, 20, 29, 30, 59, 60]. It cannot be expected that the continuum limit and singular limit commute, since in the limit $\varepsilon \downarrow 0$ the graph-based Allen–Cahn dynamics decouples per vertex and simply evolves into the well of W that is nearest the initial condition at the node [62]. A simultaneous limit, taking both $\varepsilon \downarrow 0$ and $h \downarrow 0$ (or, in a more general setting, the number of nodes $|V| \rightarrow \infty$), could potentially lead to nontrivial dynamics if ε scales in the correct way with h . A related example of such a simultaneous limit is the convergence of graph Merriman–Bence–Osher dynamics to mean curvature flow that has been established recently in [43].

2.7 Appendix A: A Proof of (2.21)

This section culminates with the proof of (2.21) in Lemma 2.12. Even though (2.21) seems to be a standard result, the authors did not find a proof in the literature. Along the way we prove some other useful results.

Lemma 2.8 *Let $p \in [1, \infty)$. Assume $g, \tilde{g} \in L^p(\mathbb{T}^n)$ and let $(g^j)_{j \in \mathbb{N}}$ be a sequence in $L^p(\mathbb{T}^n)$. If $g^j \rightarrow g$ in $L^p(\mathbb{T}^n)$ and $g^j \rightharpoonup^* \tilde{g}$ in $L^\infty(\mathbb{T}^n)$ as $j \rightarrow \infty$, then $g = \tilde{g}$ a.e. on \mathbb{T}^n .*

Consequently, if $w \in L^p(\mathbb{T}^n)$ and $(w^j)_{j \in \mathbb{N}}$ is a sequence in $L^p(\mathbb{T}^n)$ which is uniformly bounded in $L^\infty(\mathbb{T}^n)$ and such that $w^j \rightarrow w$ in $L^p(\mathbb{T}^n)$, then $w^j \rightharpoonup^* w$ in $L^\infty(\mathbb{T}^n)$.

Proof Let $U \subset \mathbb{T}^n$ be such that $g - \tilde{g} \geq 0$ on U and $g - \tilde{g} < 0$ on $U^c := \mathbb{T}^n \setminus U$. We note that as preimages of the measurable sets $[0, \infty)$ and $(-\infty, 0)$, respectively, U and U^c are also measurable. Define $\chi_U = 1$ on U and $\chi_U = 0$ on U^c . Then $2\chi_U - 1 \in L^1(\mathbb{T}^n) \cap L^{p^*}(\mathbb{T}^n)$, where p^* is defined via $1/p + 1/p^* = 1$. We compute

$$\begin{aligned} \int_{\mathbb{T}^n} |g(x) - \tilde{g}(x)| dx &= \int_{\mathbb{T}^n} (2\chi_U(x) - 1) (g(x) - \tilde{g}(x)) dx \\ &= \int_{\mathbb{T}^n} (2\chi_U(x) - 1) (g(x) - g^j(x)) dx \\ &\quad + \int_{\mathbb{T}^n} (2\chi_U(x) - 1) (g^j(x) - \tilde{g}(x)) dx. \end{aligned}$$

The first term on the right-hand side converges to zero as $j \rightarrow \infty$, by Hölder's inequality and $g^j \rightarrow g$ in $L^p(\mathbb{T}^n)$. The second term converges to zero, since $g^j \rightharpoonup^* \tilde{g}$ in $L^\infty(\mathbb{T}^n)$. This concludes the proof of the first part.

For the proof of the second part, we note that by the Banach–Alaoglu theorem each bounded sequence in $L^\infty(\mathbb{T}^n)$ has a weakly*-convergent subsequence. In particular each subsequence of (w^j) has a further subsubsequence which converges weakly*. By the first part of this result, each such subsubsequence has the same limit, w . Hence (w^j) converges weakly* in $L^\infty(\mathbb{T}^n)$ to w . \square

For $1 \leq p \leq \infty$ we denote by $L^p(\mathbb{T}; \mathbb{R}^n)$ the space of integrable functions $u : \mathbb{T}^n \rightarrow \mathbb{R}^n$ for which $\int_{\mathbb{T}^n} |u(x)|^p dx < \infty$ (if $p < \infty$), where $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^n , or for which, for almost every $x \in \mathbb{T}^n$, $|g(x)|_\infty < \infty$ (if $p = \infty$).

Lemma 2.9 *Let $g \in L^\infty(\mathbb{T}^n; \mathbb{R}^n)$ be such that, for all $i \in [n]$, $\frac{\partial g_i}{\partial x_i} \in L^\infty(\mathbb{T}^n)$ as weak partial derivatives and, for almost every $x \in \mathbb{T}^n$, $|g(x)|_\infty \leq 1$. Then there exists a sequence $(g^j)_{j \in \mathbb{N}}$ in $C^\infty(\mathbb{T}^n; \mathbb{R}^n)$ such that, as $j \rightarrow \infty$, $g^j \rightharpoonup^* g$ in $L^\infty(\mathbb{T}^n; \mathbb{R}^n)$ and, for all $i \in [n]$, $\frac{\partial g_i^j}{\partial x_i} \rightharpoonup^* \frac{\partial g_i}{\partial x_i}$. Moreover, for all $j \in \mathbb{N}$ and for almost all $x \in \mathbb{T}^n$, $|g^j(x)|_\infty \leq 1$.*

Proof This proof is a variant of the proofs in [25, Proposition 3.3], [15, Lemma A.3].

Let $\eta \in C^\infty(\mathbb{T}^n)$ be the standard mollifier (see, for example, [27, Appendix C.4]). In particular, $\eta \geq 0$ and $\int_{\mathbb{T}^n} \eta(x) dx = 1$. Define, for all $j \in \mathbb{N}$, $\eta^j(x) := j^n \eta(jx)$ and $g^j := g * \eta^j$. We note that $\eta^j \geq 0$ and $\int_{\mathbb{T}^n} \eta^j(x) dx = 1$. Moreover, since the convolution preserves the periodicity, g^j is well-defined on \mathbb{T}^n and by a standard result [27], $g^j \in C^\infty(\mathbb{T}^n)$.

Let $p \in [1, \infty)$. Since $g \in L^\infty(\mathbb{T}^n; \mathbb{R}^n)$ and $|\mathbb{T}^n| < \infty$, also $g \in L^p(\mathbb{T}^n; \mathbb{R}^n)$. Again by a standard result [27], $g^j \rightarrow g$ in $L^p(\mathbb{T}^n; \mathbb{R}^n)$. Moreover, since

$$\frac{\partial g_i^j}{\partial x_i} = \frac{\partial g_i}{\partial x_i} * \eta^j$$

and $\frac{\partial g_i}{\partial x_i} \in L^\infty(\mathbb{T}^n) \subset L^p(\mathbb{T}^n)$, by the same standard result as above, $\frac{\partial g_i^j}{\partial x_i} \rightarrow \frac{\partial g_i}{\partial x_i}$ in $L^p(\mathbb{T}^n)$.

By Lemma 2.8 it now suffices to prove uniform-in- j L^∞ bounds on g^j and $\frac{\partial g_i^j}{\partial x_i}$. For the former we use $|g(x)|_\infty \leq 1$ to compute

$$|g_i^j(x)| \leq j^n \int_{\mathbb{T}^n} \eta(jy) dy = 1.$$

For the latter,

$$\begin{aligned} \left| \frac{\partial g_i^j}{\partial x_i}(x) \right| &= \left| \int_{\mathbb{T}^n} \frac{\partial g_i}{\partial x_i}(y) \eta^j(x-y) dy \right| \leq \int_{\mathbb{T}^n} \eta^j(x-y) dy \left\| \frac{\partial g_i}{\partial x_i} \right\|_{L^\infty(\mathbb{T}^n)} \\ &= \left\| \frac{\partial g_i}{\partial x_i} \right\|_{L^\infty(\mathbb{T}^n)}. \end{aligned}$$

□

Using Lemma 2.9 we deduce (similar to [15, Corollary 3]) that the regularity conditions on the admissible vector fields g in the definition of anisotropic total variation can be relaxed:

$$\int_{\mathbb{T}^n} |Du|_{l^1} = \sup \left\{ \int_{\mathbb{T}^n} u(x) \operatorname{div} g(x) dx \mid g \in L^\infty(\mathbb{T}^n; \mathbb{R}^n), \right. \\ \left. \forall i \in [n] \frac{\partial g_i}{\partial x_i} \in L^\infty(\mathbb{T}^n), \text{ and for a.e. } x \in \mathbb{T}^n |g(x)|_\infty \leq 1 \right\}. \tag{2.59}$$

Moreover, since Lemma 2.9 tells us that each of the weak partial derivatives $\frac{\partial g_i}{\partial x_i}$ converges separately, the summation formula (2.20) also holds when the suprema in (2.20) are taken over all $g = (g_1, \dots, g_n) \in L^\infty(\mathbb{T}^n; \mathbb{R}^n)$ that satisfy the conditions in (2.59). It follows that identity (2.59) yields

$$\int_{\mathbb{T}^n} |Du|_{\ell^1} = \sum_{i=1}^n \int_{\mathbb{T}^n} |D_{x_i} u|, \tag{2.60}$$

if we define

$$\int_{\overline{U}} |D_{x_i} u| := \sup \left\{ \int_U u(x) \frac{\partial f}{\partial x_i}(x) dx \mid f \in L^\infty(U), \frac{\partial f}{\partial x_i} \in L^\infty(U), \text{ and} \right. \\ \left. \text{for a.e. } x \in U |f(x)| \leq 1 \right\}, \tag{2.61}$$

for subsets $U \subset \mathbb{T}^n$ that are open (in the Euclidean topology on \mathbb{T}^n) and functions $u \in L^1(U)$. We note that we have chosen the notation $\int_{\bar{U}} |D_{x_i} u|$ (rather than $\int_U |D_{x_i} u|$) to emphasize that this is not the standard total variation, since its behavior at the boundary of U is different. For example, if U is an interval and proper subset of \mathbb{T} and u is constantly equal to $c \in \mathbb{R}$ on U , then $\int_{\bar{U}} |D_{x_1} u| = 2c$, whereas the total variation of u on U is $\int_U |Du|_{\ell^1} = 0$. In general, if U has smooth boundary, we can interpret $\int_{\bar{U}} |D_{x_i} u|$ as the total variation on \mathbb{T}^n “in the direction of x_i ” of the function \bar{u} that is defined by $\bar{u} := u$ on U and $\bar{u} := 0$ on $\mathbb{T}^n \setminus U$.

For an open subset $U \subset \mathbb{T}^n$, we say that the boundary of U is *parallel to the x_i -axis* if ∂U (i.e., the topological boundary of U as subset of \mathbb{T}^n) is a non-empty subset of the union of countably many $n - 1$ -dimensional hyperplanes in \mathbb{R}^n (where we interpret \mathbb{T}^n as a subset of \mathbb{R}^n via identification with the hypercube Ω_n ; see Sect. 2.3.1) whose normal vectors are perpendicular to the x_i -axis.

We say a function $u : U \rightarrow \mathbb{R}$ is *independent of x_j* if the distributional derivative $\frac{\partial u}{\partial x_j}$ equals zero.

Lemma 2.10 *Let $i \in [n]$ and $n \geq 2$.*

(i) *If the boundary of U is parallel to the x_i -axis, then, for all $u \in L^1(\mathbb{T}^n)$,*

$$\int_{\mathbb{T}^n} |D_{x_i} u| = \int_{\bar{U}} |D_{x_i} u| + \int_{\mathbb{T}^n \setminus U} |D_{x_i} u|. \quad (2.62)$$

(ii) *Let $j \in [n]$, with $j \neq i$. If $u \in L^1(U)$ is independent of x_j , then*

$$\int_{\bar{U}} |D_{x_i} u| = \sup \left\{ \int_U u(x) \frac{\partial f}{\partial x_i}(x) dx \mid f \in L^\infty(U), \frac{\partial f}{\partial x_i} \in L^\infty(U), \right. \\ \left. \text{for a.e. } x \in U \mid f(x) \leq 1, \text{ and } f \text{ is independent of } x_j \right\}.$$

Proof (i) By definition, the inequality \leq holds for every open set U . Assume that the boundary of U is parallel to the x_i -axis. For notational convenience, define $V := \mathbb{T}^n \setminus \bar{U}$. Let $f_U \in L^\infty(U)$ with $\frac{\partial f_U}{\partial x_i} \in L^\infty(U)$ and, for a.e. $x \in U$, $|f_U(x)| \leq 1$, and let $f_V \in L^\infty(V)$ with $\frac{\partial f_V}{\partial x_i} \in L^\infty(V)$ and, for a.e. $x \in V$, $|f_V(x)| \leq 1$. Let $f : \mathbb{T}^n \rightarrow \mathbb{R}$ be such that $f|_U = f_U$ and $f|_V = f_V$. Then $f \in L^\infty(\mathbb{T}^n)$ and, since the boundary of U is parallel to the x_i -axis, $\frac{\partial f_U}{\partial x_i} \in L^\infty(U)$ and $\frac{\partial f_V}{\partial x_i} \in L^\infty(V)$. Moreover, for a.e. $x \in \mathbb{T}^n$, $|f(x)| \leq 1$. Thus, if f_U and f_V are admissible functions in the suprema on the right-hand side of (2.62), then f is admissible in the supremum on the left-hand side. This yields the reverse inequality of the identity.¹⁷

(ii) The inequality \geq follows immediately, since the admissible set of functions in the supremum on the right-hand side is a subset of the admissible set in (2.61).

¹⁷ We note that this may not hold in general if the boundary of U is not parallel to the x_i -axis.

To prove the inequality \leq , we show that we can replace any admissible function f by another admissible function that is independent of x_j , without changing the value of $\int_U u(x) \frac{\partial f}{\partial x_i}(x) dx$.

Without loss of generality, we take $j = 1$ and $i > 1$. Let f be admissible in the supremum in (2.61). We define the x_1 -average of f —which is independent of x_1 —for $(x_1, \tilde{x}) \in U$ as

$$\bar{f}(x_1, \tilde{x}) := \frac{1}{|U_{\tilde{x}}|} \int_{U_{\tilde{x}}} f(y, \tilde{x}) dy,$$

where $U_{\tilde{x}} := \{x_1 \in \mathbb{T} \mid (x_1, \tilde{x}) \in U\}$ and $|U_{\tilde{x}}|$ is the one-dimensional Lebesgue measure of $U_{\tilde{x}}$. We note that, for all $\tilde{x} \in \mathbb{T}^{n-1}$, $U_{\tilde{x}}$ is open in \mathbb{T} , since U is open in \mathbb{T}^n . In particular, if $(x_1, \tilde{x}) \in U$, then $U_{\tilde{x}} \neq \emptyset$ and thus $|U_{\tilde{x}}| > 0$. Hence \bar{f} is well-defined. We define $X := \{\tilde{x} \in \mathbb{T}^{n-1} \mid U_{\tilde{x}} \neq \emptyset\}$.

By differentiating under the integral sign [39, Theorem 7.40 (or Sect. 11.12)],¹⁸ we see that, for a.e. $(x_1, \tilde{x}) \in U$,

$$\frac{\partial \bar{f}}{\partial x_i}(x_1, \tilde{x}) = \frac{\partial f}{\partial x_i}(x_1, \tilde{x}).$$

Hence $\frac{\partial \bar{f}}{\partial x_i} \in L^\infty(U)$. Moreover, we observe that the value of $\int_U u(x) \frac{\partial f}{\partial x_i}(x) dx$ remains unchanged if we replace f by \bar{f} :

$$\begin{aligned} \int_U u(x) \frac{\partial \bar{f}}{\partial x_i}(x) dx &= \int_X \int_{U_{\tilde{x}}} u(x_1, \tilde{x}) \frac{\partial \bar{f}}{\partial x_i}(x_1, \tilde{x}) dx_1 d\tilde{x} \\ &= \int_X \int_{U_{\tilde{x}}} u(x_1, \tilde{x}) \frac{1}{|U_{\tilde{x}}|} \int_{U_{\tilde{x}}} \frac{\partial f(y, \tilde{x})}{\partial x_i} dy dx_1 d\tilde{x} \\ &= \int_X \frac{1}{|U_{\tilde{x}}|} \int_{U_{\tilde{x}}} u(y, \tilde{x}) \frac{\partial f(y, \tilde{x})}{\partial x_i} dy \int_{U_{\tilde{x}}} dx_1 d\tilde{x} \\ &= \int_X \int_{U_{\tilde{x}}} u(y, \tilde{x}) \frac{\partial f(y, \tilde{x})}{\partial x_i} dy d\tilde{x} \\ &= \int_U u(x) \frac{\partial f}{\partial x_i}(x) dx, \end{aligned} \tag{2.63}$$

where we have used that, for $x_1, y \in U_{\tilde{x}}$ and for $\tilde{x} \in X$, $u(x_1, \tilde{x}) = u(y, \tilde{x})$. Furthermore, for a.e. $x \in U$ we have $|f(x)| \leq 1$ since, for a.e. $x \in U$, $|f(x)| \leq 1$. This completes the proof of (ii). □

¹⁸ For a more directly applicable result published in the Japanese language, we refer to [38, Theorem 19.4].

Remark 2.15 By iterating the result of Lemma 2.10 (ii), it follows that, if $u \in L^1(U)$ is independent of all elements of $\{x_j\}_{j \in J}$, for a $J \subset [n] \setminus \{i\}$, then

$$\int_{\bar{U}} |D_{x_i} u| = \sup \left\{ \int_U u(x) \frac{\partial f}{\partial x_i}(x) dx \mid f \in L^\infty(U), \frac{\partial f}{\partial x_i} \in L^\infty(U), \right. \\ \left. \text{for a.e. } x \in U \mid f(x) \leq 1, \text{ and } f \text{ is independent of } \{x_j\}_{j \in J} \right\}.$$

In particular, extending the averaging argument from the proof of Lemma 2.10, if we average f over all elements of $\{x_j\}_{j \in J}$, the value of the integral remains unchanged as in (2.63). To be precise, without loss of generality let $\{x_j\}_{j \in J} = \{x_1, \dots, x_{|J|}\}$ and $i > |J|$, and define the $\{x_j\}_{j \in J}$ -average of f as

$$\bar{f}(x_1, \dots, x_{|J|}, \tilde{x}) := \frac{1}{|U_{\tilde{x}}|} \int_{U_{\tilde{x}}} f(y_1, \dots, y_{|J|}, \tilde{x}) dy_1 \dots dy_{|J|}, \quad (2.64)$$

for $(x_1, \dots, x_{|J|}, \tilde{x}) \in U$, $U_{\tilde{x}} := \{(x_1, \dots, x_{|J|}) \in \mathbb{T}^{|J|} \mid (x_1, \dots, x_{|J|}, \tilde{x}) \in U\}$, and $|U_{\tilde{x}}|$ to be the $|J|$ -dimensional Lebesgue measure of $U_{\tilde{x}}$, then by a similar argument as in (2.63) we obtain

$$\int_U u(x) \frac{\partial \bar{f}}{\partial x_i}(x) dx = \int_U u(x) \frac{\partial f}{\partial x_i}(x) dx. \quad (2.65)$$

We conclude this remark by noting that in the argument above, the fact that U is open is only used to guarantee that $|U_{\tilde{x}}| > 0$. The conclusion from (2.65) thus holds for any subset $U \subset \mathbb{T}^n$ on which f and its weak derivative $\frac{\partial f}{\partial x_i}$ are well-defined as elements of (equivalence classes in) $L^\infty(U)$ and for which, for all $\tilde{x} \in \mathbb{T}^{n-|J|}$ for which $U_{\tilde{x}} \neq \emptyset$, $|U_{\tilde{x}}| > 0$. In particular, if $U = Q_z^h$ for an $h \in \mathfrak{H}$ and a $z \in V_h$, then any non-empty $U_{\tilde{x}}$ will be the Cartesian product of sets $\{I_{z_i}^h\}_{i \in J}$ and thus have positive $|J|$ -dimensional Lebesgue measure. This allows us to prove Corollary 2.2 below, which is used in turn in our proof of Lemma 2.3.

Corollary 2.2 *Let $h \in \mathfrak{H}$ and $i \in [n]$. If $u \in L_h^2$, then*

$$\int_{\mathbb{T}^n} |D_{x_i} i_h u| = \sup_{g \in H_{h,i}} \int_{\mathbb{T}^n} i_h u(x) \frac{\partial g}{\partial x_i}(x) dx,$$

where

$$H_{h,i} := \left\{ f : \mathbb{T}^n \rightarrow \mathbb{R} \mid f \in L^\infty(\mathbb{T}^n), \frac{\partial f}{\partial x_i} \in L^\infty(\mathbb{T}^n), \forall x \in \mathbb{T} \mid f(x) \leq 1, \text{ and} \right. \\ \left. \forall z \in V_h \frac{\partial f}{\partial x_i} \text{ is constant a.e. on } \text{int } Q_z^h \right\}. \quad (2.66)$$

Proof Without loss of generality, we assume that $i = n$.

Since $H_{h,n}$ is a subset of the set of admissible functions in the supremum in the definition of $\int_{\mathbb{T}^n} |D_{x_n} i_h u|$ in (2.61) (with $i = n$), the inequality \geq follows.

We prove the inequality \leq by showing that for every admissible f in (2.61), there is an $\tilde{f} \in H_{h,n}$ such that the equality

$$\int_{\mathbb{T}^n} u(x) \frac{\partial \tilde{f}}{\partial x_n}(x) dx = \int_{\mathbb{T}^n} u(x) \frac{\partial f}{\partial x_n}(x) dx \tag{2.67}$$

holds and thus

$$\int_{\mathbb{T}^n} u(x) \frac{\partial \tilde{f}}{\partial x_n}(x) dx \leq \sup_{g \in H_{h,n}} \int_{\mathbb{T}^n} i_h u(x) \frac{\partial g}{\partial x_n}(x) dx,$$

from which the required inequality follows by first taking the supremum over all admissible f in the definition in (2.61).

To prove the equality (2.67), let f be admissible in (2.61). If $z \in V_h$ and $x \in Q_z^h$, then $i_h u(x) = u(z)$, thus in particular the function $i_h u$ is independent of all the variables in $\{x_j\}_{j \in [n-1]}$ on Q_z^h . For all $z \in V_h$ and for all $x \in Q_z^h$, define $\hat{f}(x) := \overline{f}_z(x)$, where \overline{f}_z is the $\{x_j\}_{j \in [n-1]}$ -average of f over Q_z^h , i.e., \overline{f}_z is as \overline{f} in (2.64) with $J = [n - 1]$ and $U = Q_z^h$. This choice of U is allowed, by the argument at the end of Remark 2.15. Then, for all $z \in V_h$ (2.65) holds (with $i_h u$ instead of u and \overline{f}_z instead of \overline{f}), i.e.,

$$\int_{Q_z^h} i_h u(x) \frac{\partial \hat{f}}{\partial x_n}(x) dx = \int_{Q_z^h} i_h u(x) \frac{\partial f}{\partial x_n}(x) dx.$$

Because \hat{f} is independent of $\tilde{x} = (x_1, \dots, x_{n-1})$ on each Q_z^h , we have that for a.e. $\tilde{x} \in \mathbb{T}^{n-1}$, $x_n \mapsto \hat{f}(\tilde{x}, x_n)$ is a Lipschitz-continuous function¹⁹ on \mathbb{T} , and thus in particular pointwise evaluation of $\hat{f}(\cdot, \tilde{x})$ is well-defined for a.e. $\tilde{x} \in \mathbb{T}^{n-1}$. We recall the definitions of $I_{z_n}^h$ and $\tilde{Q}_{z,n}^h$ from (2.17) and (2.18), respectively. Then

¹⁹ We recall that, if $h \in L^\infty(\mathbb{T})$ and $h' \in L^\infty(\mathbb{T})$, then h is Lipschitz-continuous (or, more accurately, has a representative in $L^\infty(\mathbb{T})$ that is Lipschitz-continuous) [44, Exercise 11.50], [27, Sect. 5.8, Theorem 4].

$$\begin{aligned}
\int_{\mathbb{T}^n} i_h u(x) \frac{\partial f}{\partial x_n}(x) dx &= \sum_{z \in V_h} \int_{Q_z^h} i_h u(x) \frac{\partial f}{\partial x_n}(x) dx = \sum_{z \in V_h} \int_{Q_z^h} i_h u(x) \frac{\partial \hat{f}}{\partial x_n}(x) dx \\
&= \sum_{z \in V_h} u(z) \int_{\tilde{Q}_{z,n}^h} \int_{I_{z_n}^h} \frac{\partial \hat{f}}{\partial x_n}(\tilde{x}, x_n) dx_n d\tilde{x} \\
&= \sum_{z \in V_h} u(z) \int_{\tilde{Q}_{z,n}^h} \int_{I_{z_n}^h} \frac{\partial \hat{f}}{\partial x_n}(x) dx_n d\tilde{x} \\
&= \sum_{z \in V_h} u(z) \int_{\tilde{Q}_{z,n}^h} (\hat{f}(\tilde{x}, z_n + h/2) - \hat{f}(\tilde{x}, z_n - h/2)) d\tilde{x} \\
&= \sum_{z \in V_h} u(z) \int_{\tilde{Q}_{z,n}^h} \int_{I_{z_n}^h} \frac{\partial \tilde{f}}{\partial x_n}(\tilde{x}, x_n) dx_n d\tilde{x} \\
&= \sum_{z \in V_h} \int_{\tilde{Q}_z^h} u(x) \frac{\partial \tilde{f}}{\partial x_n}(x) dx \\
&= \int_{\mathbb{T}^n} u(x) \frac{\partial \tilde{f}}{\partial x_n}(x) dx.
\end{aligned}$$

Here $\tilde{f} : \mathbb{T}^n \rightarrow \mathbb{R}$ is such that, for all $z \in V_h$, for all $x_n \in I_{z_n}^h$, and for all $\tilde{x} \in \tilde{Q}_{z,n}^h$, $\tilde{f}(\tilde{x}, x_n) := k_{\tilde{z}}(x_n)$, where $k_{\tilde{z}} : \mathbb{T} \rightarrow \mathbb{R}$ is the unique continuous piecewise-affine solution to

$$k'_{\tilde{z}} = h^{-1} \left(\hat{f}(\tilde{x}, z_n + h/2) - \hat{f}(\tilde{x}, z_n - h/2) \right) \quad \text{on each } \text{int } I_{z_n}^h.$$

We note that the function $k_{\tilde{z}}$ depends on $\tilde{z} := (z_1, \dots, z_{n-1})$, but is independent of $\tilde{x} \in \tilde{Q}_{z,n}^h$, because \hat{f} is independent of \tilde{x} on each Q_z^h . Hence $\frac{\partial \tilde{f}}{\partial x_n}$ is constant on every Q_z^h , which immediately implies that $\frac{\partial \tilde{f}}{\partial x_n} \in L^\infty(\mathbb{T}^n)$.

Furthermore, on each Q_z^h , \tilde{f} is independent of \tilde{x} and $x_n \mapsto \tilde{f}(\tilde{x}, x_n)$ is affine with bounded derivative—after all, since $f \in L^\infty(\mathbb{T}^n)$ also $\hat{f} \in L^\infty(\mathbb{T}^n)$ —and thus $\tilde{f} \in L^\infty(\mathbb{T}^n)$. Finally, since for a.e. $x \in \mathbb{T}^n$, $|f(x)| \leq 1$, we also have, for a.e. $x \in \mathbb{T}^n$, $|\tilde{f}(x)| \leq 1$. Thus, $\tilde{f} \in H_{h,n}$, which concludes the proof. \square

It is well known²⁰ that in the one-dimensional case, for functions $u \in BV(\mathbb{T})$,

$$\begin{aligned}
\int_{\mathbb{T}} |Du|_1 &= \text{ess Var } u := \inf_w \text{Var } w, \quad \text{with} \\
\text{Var } w &:= \sup_P \sum_{i=1}^q |w(x_i) - w(x_{i-1})| + |w(x_0) - w(x_q)|,
\end{aligned} \tag{2.68}$$

²⁰ The results from [44, Theorem 7.9], [34, Sect. 1.30] should be adapted to the torus.

where the infimum in essVar is taken over all functions $w : \mathbb{T} \rightarrow \mathbb{R}$ that satisfy $w(x) = u(x)$ for Lebesgue almost every $x \in \mathbb{T}$, and the supremum in Var is taken over all partitions²¹ P of \mathbb{T} .

Before we prove (2.21) in Lemma 2.12, we require the following result in the one-dimensional setting.

Lemma 2.11 *Let h be fixed, $q \in \mathbb{N}$, and let $\{I_i\}_{i=0}^q$ be a finite set of non-degenerate disjoint intervals in $[-h/2, 1 - h/2)$ such that $[-h/2, 1 - h/2) \setminus \bigcup_{i=0}^q I_i = \emptyset$ and, for all $i \in \{0, \dots, q - 1\}$, for all $x \in I_i$, and for all $y \in I_{i+1}$, $x < y$. Assume $u \in BV(\mathbb{T})$ is constant (up to a null set) on each interval I_i (where we identify \mathbb{T} with $[-h/2, 1 - h/2)$), taking value c_i on interval I_i (up to a null set). Then*

$$\text{essVar } u = \sum_{i=1}^q |c_i - c_{i-1}| + |c_0 - c_q|.$$

Proof Since each interval I_i has positive Lebesgue measure, for every function w that agrees almost everywhere with u and for all i , there are $x_i \in I_i$ such that $w(x_i) = c_i$. Hence

$$\text{Var } w \geq \sum_{i=1}^q |w(x_i) - w(x_{i-1})| + |w(x_0) - w(x_q)| = \sum_{i=1}^q |c_i - c_{i-1}| + |c_0 - c_q|$$

and thus

$$\text{essVar } u \geq \sum_{i=1}^q |c_i - c_{i-1}| + |c_0 - c_q|.$$

The complement of the union of finitely many disjoint intervals in $[-h/2, 1 - h/2)$ is itself a union of finitely many (possibly degenerate) disjoint intervals. Since $[-h/2, 1 - h/2) \setminus \bigcup_{i=0}^q I_i$ has zero Lebesgue measure, it must consist of finitely many isolated points \tilde{x}_j . Let \tilde{u} be equal to u on $\bigcup_{i=0}^q I_i$ and for each \tilde{x}_j define $\tilde{u}(\tilde{x}_j) = c_i$, where I_i is the closest interval to \tilde{x}_j (make an arbitrary choice between candidates in case of nonuniqueness). Then

$$\text{essVar } u \leq \text{Var } \tilde{u} = \sum_{i=1}^q |c_i - c_{i-1}| + |c_0 - c_q|.$$

□

Now we are ready to prove (2.21).

²¹ If h is fixed and we identify \mathbb{T} with $[-h/2, 1 - h/2)$ with periodic boundary conditions, then by a partition of \mathbb{T} we mean a subset $P = \{x_0, \dots, x_q\} \subset [-h/2, 1 - h/2)$, for some $q \in \mathbb{N}$, such that $x_0 < x_1 < \dots < x_q$.

Lemma 2.12 (Proof of (2.21)) *Let $h \in \mathcal{S}$. If $u \in L_h^2$, then*

$$\int_{\mathbb{T}^n} |Di_h u|_{l^1} = \varphi_{TV}^h(u)$$

and thus consequently $i_h u \in BV(\mathbb{T}^n)$.

Proof First we note that the second claim in the lemma follows directly from the first, since $\varphi_{TV}^h(u) < +\infty$ for $u \in L_h^2$.

For the proof of the first claim it is useful to introduce some additional notation. For all $i \in [n]$, we define $R_i^h(z) := \{\tilde{z} \in V_h : \forall j \in [n] \text{ with } j \neq i, \tilde{z}_j = z_j\}$. The set $R_i^h(z)$ contains all nodes in V_h that lie on the same “row” as z in the i th direction.

We decompose the graph total variation functional into functionals that only consider differences along the coordinate axes:

$$\varphi_{TV}^h(u) = \sum_{i=1}^n \varphi_{TV,i}^h(u), \quad \text{with} \quad \varphi_{TV,i}^h(u) := \frac{1}{2} \sum_{\substack{z, \tilde{z} \sim z \\ \tilde{z} \in R_i^h(z)}} h^{n-1} |u(z) - u(\tilde{z})|,$$

where we use a variant of the notation from Sect. 2.4.1:

$$\sum_{\substack{z, \tilde{z} \sim z \\ \tilde{z} \in R_i^h(z)}} := \sum_{z \in V_h} \sum_{\tilde{z} \in R_i^h(z)}.$$

By (2.60) it suffices to prove that, for all $i \in [n]$,

$$\int_{\mathbb{T}^n} |D_{x_i} i_h u| = \varphi_{TV,i}^h(u).$$

Let $i \in [n]$. By permutation of variables, we may assume that $i = n$. Given a $z \in V_h$, by $\tilde{x} = (x_1, \dots, x_{n-1})$ we denote the coordinates on $\tilde{Q}_{z,n}^h$ (we recall the definition of $\tilde{Q}_{z,n}^h$ from (2.15)). We define V_h^0 to be the set containing all nodes $z \in V_h$ with $z_n = 0$, i.e.,

$$V_h^0 := \{z \in V_h \mid z_n = 0\} = \{0, h, 2h, \dots, 1-h\}^{n-1} \times \{0\}.$$

The space \mathbb{T}^n is represented as a disjoint union

$$\mathbb{T}^n = \bigcup_{z \in V_h^0} (\tilde{Q}_{z,n}^h \times \mathbb{T}).$$

For $z \in V_h^0$, we set

$$U_{z,n}^h := \text{int } \tilde{Q}_{z,n}^h \times \mathbb{T}.$$

Since the boundary of $U_{z,n}^h$ is parallel to the x_n -axis, repeated application of Lemma 2.10 (i) leads to

$$\int_{\mathbb{T}^n} |D_{x_n} i_h u| = \sum_{z \in V_h^0} \int_{U_{z,n}^h} |D_{x_n} i_h u|.$$

Because $i_h u$ is constant on each Q_z^h and thus independent of \tilde{x} in each $\tilde{Q}_{z,n}^h$, it is independent of \tilde{x} on each $U_{z,n}^h$. Temporarily fix $z \in V_h^0$. Then by Lemma 2.10 (ii) and Remark 2.15 we have that

$$\int_{U_{z,n}^h} |D_{x_n} i_h u| = \sup \left\{ \int_{U_{z,n}^h} (i_h u)(x) f'(x_n) dx \mid f \in L^\infty(\mathbb{T}), f' \in L^\infty(\mathbb{T}), \text{ and } \forall x \in \mathbb{T} |f(x)| \leq 1 \right\}.$$

Moreover, writing $\overline{i_h u}^z(x_n) := (i_h u)(\tilde{x}, x_n)$ for $x = (\tilde{x}, x_n) \in U_{z,n}^h$, we get

$$\int_{U_{z,n}^h} (i_h u)(x) f'(x_n) dx = h^{n-1} \int_{\mathbb{T}} \overline{i_h u}^z(x_n) f'(x_n) dx_n,$$

and thus

$$\int_{U_{z,n}^h} |D_{x_n} i_h u| = h^{n-1} \int_{\mathbb{T}} \left| D \overline{i_h u}^z \right|_{\ell^1}.$$

The problem is now reduced to a one-dimensional setting. Since $\overline{i_h u}^z$ is piecewise constant on \mathbb{T} , applying the one-dimensional result from Lemma 2.11 together with (2.68), we deduce that

$$\int_{\mathbb{T}} \left| D \overline{i_h u}^z \right|_{\ell^1} = \frac{1}{2} \sum_{\hat{z} \in R_h^h(z)} \sum_{\bar{z} \in R_h^h(z): \bar{z} \sim \hat{z}} |u(\hat{z}, \hat{z}_n) - u(\bar{z}, \bar{z}_n)|,$$

where $\hat{z} = (\tilde{z}, \hat{z}_n) \in V_h$ and $\bar{z} = (\tilde{z}, \bar{z}_n) \in V_h$. Now we unfix z and observe that

$$\begin{aligned} \int_{\mathbb{T}^n} |D_{x_n} i_h u| &= \sum_{z \in V_h^0} \int_{U_{z,n}^h} |D_{x_n} i_h u| = h^{n-1} \sum_{z \in V_h^0} \int_{\mathbb{T}} \left| D \overline{i_h u}^z \right|_{\ell^1} \\ &= \frac{1}{2} h^{n-1} \sum_{z \in V_h^0} \sum_{\hat{z} \in R_h^h(z)} \sum_{\bar{z} \in R_h^h(z): \bar{z} \sim \hat{z}} |u(\hat{z}, \hat{z}_n) - u(\bar{z}, \bar{z}_n)| \\ &= \frac{1}{2} h^{n-1} \sum_{\substack{z: z \sim \tilde{z} \\ \tilde{z} \in R_h^h(z)}} |u(z) - u(\bar{z})| = \phi_{TV,n}^h(u). \end{aligned}$$

This proves the lemma, so the proof of (2.18) is also complete now. □

2.8 Appendix B: Properties of Γ

In this section we prove some properties of the operator Γ , which is defined in (2.31), and its matrix exponential.

It will be useful to identify Γ with a circulant $1/h$ by $1/h$ -matrix, which we will again denote by Γ . This matrix has entries, for all $i, j \in [\frac{1}{h}]$,

$$\Gamma_{ij} = \begin{cases} \frac{2}{3}, & \text{if } i = j, \\ \frac{1}{6}, & \text{if } i \equiv j - 1 \pmod{\frac{1}{h}} \text{ or } i \equiv j + 1 \pmod{\frac{1}{h}}, \\ 0, & \text{otherwise.} \end{cases}$$

The general form of the eigenvectors and eigenvalues of circulant matrices is known. In this case it can be checked via direct computation that the eigenvalues of Γ are, for $k \in [\frac{1}{h} - 1]_0$,

$$\gamma_k := \frac{1}{3} (2 + \cos(2\pi hk)).$$

The corresponding normalized eigenvectors of Γ (in $\mathbb{C}^{1/h}$) are

$$z^{(k)} := \sqrt{h} \begin{pmatrix} 1 \\ \omega^k \\ \omega^{2k} \\ \vdots \\ \omega^{(\frac{1}{h}-1)k} \end{pmatrix}, \quad \text{thus } z_i^{(k)} = \sqrt{h} \omega^{(i-1)k}, \quad \text{for all } i \in [\frac{1}{h}],$$

where $\omega := e^{2\pi h i}$ is the $\frac{1}{h}$ th root of unity and we recall that i denotes the imaginary unit. We note that

$$\gamma_{\frac{1}{h}-k} = \gamma_k. \quad (2.69)$$

An alternative way to describe Γ is as the sum of three matrices,

$$\Gamma = \frac{2}{3}I + \frac{1}{6}A + \frac{1}{6}A^T, \quad (2.70)$$

where I is the $\frac{1}{h}$ -by- $\frac{1}{h}$ identity matrix and A and A^T are the matrix representations of the operators τ_+ and τ_- from (2.33), respectively, i.e.,

$$A_{ij} := \delta_{i, j+1},$$

where we always interpret the indices modulo $\frac{1}{h}$, so $\delta_{i, \frac{1}{h}+1} = \delta_{i, 1}$, etc. Remember that the Kronecker delta is defined by

$$\delta_{i,j} := \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

We note that, for all $n \in \mathbb{N}$,

$$A_{ij}^n = \delta_{i,j+n}, \tag{2.71}$$

as can be proved by induction on n : Taking as base case $n = 0$ or $n = 1$, we see it is true by definition of I or A , respectively. Assuming that it is true for $n = k > 1$, we find

$$A_{ij}^{k+1} = \sum_{l=1}^{\frac{1}{h}} A_{il} A_{lj}^k = \sum_{l=1}^{\frac{1}{h}} \delta_{i,l+1} \delta_{l,j+k} = \delta_{i,j+k+1}.$$

As a consequence of (2.71), we have that

$$A^{\frac{1}{h}} = A. \tag{2.72}$$

Furthermore $A^T = A^{-1}$, since

$$(AA^T)_{ij} = \sum_{l=1}^{\frac{1}{h}} A_{il} A_{jl} = \sum_{l=1}^{\frac{1}{h}} \delta_{i,l+1} \delta_{j,l+1} = \delta_{i,j}.$$

Thus we also have

$$(A^{-1})^{\frac{1}{h}} = (A^T)^{\frac{1}{h}} = (A^{\frac{1}{h}})^T = A^T = A^{-1}. \tag{2.73}$$

It is worth noting that

$$(A^{-1})_{ij} = (A^T)_{ij} = A_{ji} = \delta_{j,i+1} = \delta_{i,j-1} \quad \text{and, for all } n \in \mathbb{N}, \quad (A^{-n})_{ij} = \delta_{i,j-n}.$$

These two characterizations of Γ , one through its spectrum and the other using the matrix A , lead to two expressions for the matrix exponential $e^{\Gamma x}$ in the following lemma.

Lemma 2.13 *Let $x \in \mathbb{R}$. For all $i, j \in [\frac{1}{h}]$ we have*

$$(e^{\Gamma x})_{ij} = h e^{\frac{2}{3}x} \sum_{l=0}^{\frac{1}{h}-1} e^{\frac{1}{3} \cos(2\pi hl)x} \cos(2\pi hl(i-j)) \tag{2.74}$$

and

$$(e^{\Gamma x})_{ij} = e^{\frac{2}{3}x} \sum_{l=0}^{\frac{1}{h}-1} B_l(x) B_{l+j-i}(x), \tag{2.75}$$

where

$$B_l(x) := \sum_{m=0}^{\infty} \frac{1}{\left(\frac{m}{h} + l\right)!} \left(\frac{x}{6}\right)^{\frac{m}{h} + l}$$

has infinite radius of convergence. The subscripts of B are interpreted modulo $\frac{1}{h}$.

Proof The proof of the first expression follows an ODE approach: Consider the equation

$$u'(x) = \Gamma u(x). \quad (2.76)$$

By standard ODE theory [37, Sect. III.4], this equation has $\frac{1}{h}$ linearly independent solutions $u^j : \mathbb{R} \rightarrow \mathbb{R}^{1/h}$, for $j \in [\frac{1}{h}]$. Our discussion above shows that Γ has $\frac{1}{h}$ linearly independent eigenvectors and thus, for all $j \in [\frac{1}{h}]$, $u^j(x) = e^{\gamma_{j-1}x} z^{(j-1)}$. We can now compute

$$e^{\Gamma x} = X(x)X(0)^{-1}, \quad (2.77)$$

where $X : \mathbb{R} \rightarrow \mathbb{R}^{1/h \times 1/h}$ can be any fundamental matrix solution of (2.76). We choose the matrix which has the solutions u^j as columns: for all $i, j \in [\frac{1}{h}]$,

$$X_{ij}(x) := u_i^j(x) = \sqrt{h} e^{\gamma_{j-1}x} \omega^{(i-1)(j-1)}.$$

Then

$$X_{ij}(0) = u_i^j(0) = z_i^{(j-1)} = \sqrt{h} \omega^{(i-1)(j-1)}.$$

We claim that the inverse $X(0)^{-1}$ is given by, for all $i, j \in [\frac{1}{h}]$,

$$Y_{ij} := \sqrt{h} \omega^{-(i-1)(j-1)}.$$

To prove this claim, we compute

$$\begin{aligned} (X(0)Y)_{ij} &= \sum_{l=1}^{\frac{1}{h}} X(0)_{il} Y_{lj} = \sqrt{h} \sum_{l=1}^{\frac{1}{h}} z_i^{(l-1)} \omega^{-(l-1)(j-1)} = h \sum_{l=1}^{\frac{1}{h}} \omega^{(i-1)(l-1)} \omega^{-(l-1)(j-1)} \\ &= h \sum_{l=1}^{\frac{1}{h}} \omega^{(l-1)(i-j)}. \end{aligned}$$

Taking $i = j$ shows that $(X(0)Y)_{ii} = 1$. If $i \neq j$, then we multiply by ω^{i-j} :

$$\omega^{i-j} (X(0)Y)_{ij} = h \sum_{l=1}^{\frac{1}{h}} \omega^{l(i-j)} = h \sum_{k=2}^{\frac{1}{h}+1} \omega^{(l-1)(i-j)} = h \sum_{l=1}^{\frac{1}{h}} \omega^{(l-1)(i-j)} = (X(0)Y)_{ij},$$

where the third equality follows from the fact that $\omega^{(1/h+1-1)(i-j)} = (\omega^{1/h})^{i-j} = 1^{i-j} = 1 = \omega^{0(i-j)}$ —recall that ω is the $\frac{1}{h}$ th root of unity, and thus $\omega^l = 1$ if and

only if $l \equiv 0 \pmod{\frac{1}{h}}$. Since i is not congruent to j modulo $\frac{1}{h}$ and thus $\omega^{i-j} \neq 1$, it follows that $(X(0)Y)_{ij} = 0$. Thus $Y = X(0)^{-1}$.

Using (2.77) we compute

$$\begin{aligned} (e^{\Gamma x})_{ij} &= h \sum_{l=1}^{\frac{1}{h}} e^{\gamma_{l-1}x} \omega^{(i-1)(l-1)} \omega^{-(l-1)(j-1)} = h \sum_{l=0}^{\frac{1}{h}-1} e^{\gamma_l x} \omega^{(i-j)l} \\ &= h + h \sum_{l=1}^{\frac{1}{h}-1} e^{\gamma_l x} \omega^{(i-j)l}. \end{aligned}$$

By (2.69) it follows that, if $\frac{1}{h}$ is odd,

$$\begin{aligned} (e^{\Gamma x})_{ij} &= h + h \sum_{l=1}^{\frac{1}{2}(\frac{1}{h}-1)} e^{\gamma_l x} (\omega^{(i-j)l} + \omega^{(i-j)(1/h-l)}) \\ &= h + \frac{h}{2} \sum_{l=1}^{\frac{1}{h}-1} e^{\gamma_l x} (\omega^{(i-j)l} + \omega^{(i-j)(1/h-l)}), \end{aligned}$$

where for the second equality we also used that $\omega^{(i-j)l} + \omega^{(i-j)(1/h-l)}$ remains unchanged when l is replaced by $\frac{1}{h} - l$. Since

$$\begin{aligned} &\omega^{(i-j)l} + \omega^{(i-j)(1/h-l)} \\ &= \cos(2\pi h(i-j)l) + i_l \sin(2\pi h(i-j)l) + \cos(2\pi h(i-j)(1/h-l)) \\ &\quad + i_l \sin(2\pi h(i-j)(1/h-l)) \\ &= \cos(2\pi h(i-j)l) + \cos(2\pi(i-j) - 2\pi(i-j)hl) \\ &\quad + i_l [\sin(2\pi h(i-j)l) + \sin(2\pi(i-j) - 2\pi(i-j)hl)] \\ &= \cos(2\pi h(i-j)l) + \cos(-2\pi(i-j)hl) + i_l [\sin(2\pi h(i-j)l) + \sin(-2\pi(i-j)hl)] \\ &= 2 \cos(2\pi h(i-j)l), \end{aligned} \tag{2.78}$$

the desired expression (2.74), for $\frac{1}{h}$ odd, now follows from the definition of γ_l . If, on the other hand, $\frac{1}{h}$ is even, we have

$$\begin{aligned} (e^{\Gamma x})_{ij} &= h + h \sum_{l=1}^{\frac{1}{2h}-1} e^{\gamma_l x} (\omega^{(i-j)l} + \omega^{(i-j)(1/h-l)}) + h e^{\gamma_{1/2} x} \omega^{(i-j)/(2h)} \\ &= h + \frac{h}{2} \sum_{\substack{l=1 \\ l \neq \frac{1}{2h}}}^{\frac{1}{h}-1} e^{\gamma_l x} (\omega^{(i-j)l} + \omega^{(i-j)(1/h-l)}) + h e^{\gamma_{1/2} x} \omega^{(i-j)/(2h)}. \end{aligned}$$

For the middle terms on the right-hand side we use (2.78) again. For the last term, we note that $\omega^{1/(2h)} = -1$ and thus

$$\omega^{(i-j)/(2h)} = (-1)^{i-j} = 1 = \cos(\pi(i-j)) = \cos\left(\frac{2\pi h(i-j)}{2h}\right).$$

Combined with the definition of γ_l , this now also gives (2.74) when $\frac{1}{h}$ is even.

To prove (2.75), we use the description of Γ from (2.70). First note that, since

$$\sum_{m=0}^{\infty} \left| \frac{1}{\left(\frac{m}{h} + l\right)!} \left(\frac{x}{6}\right)^{\frac{m}{h} + l} \right| = \sum_{m=0}^{\infty} \frac{1}{\left(\frac{m}{h} + l\right)!} \left| \frac{x}{6} \right|^{\frac{m}{h} + l} \leq \sum_{m=0}^{\infty} \frac{1}{m!} |x|^m = e^x,$$

the series B_l has an infinite radius of convergence.

Because A and $A^T = A^{-1}$ commute, we have

$$e^{\Gamma x} = e^{(\frac{2}{3}I + \frac{1}{6}A + \frac{1}{6}A^T)x} = e^{\frac{2}{3}x} e^{\frac{1}{6}Ax} e^{\frac{1}{6}A^{-1}x}. \quad (2.79)$$

Next we employ the series definition of the matrix exponential in combination with (2.72):

$$\begin{aligned} e^{\frac{1}{6}Ax} &= \sum_{m=0}^{\infty} \frac{1}{m!} A^m \left(\frac{x}{6}\right)^m \\ &= \left[\frac{1}{(0/h)!} \left(\frac{x}{6}\right)^{0/h} + \frac{1}{(1/h)!} \left(\frac{x}{6}\right)^{1/h} + \frac{1}{(2/h)!} \left(\frac{x}{6}\right)^{2/h} + \frac{1}{(3/h)!} \left(\frac{x}{6}\right)^{3/h} + \dots \right] I \\ &\quad + \left[\frac{1}{(0/h+1)!} \left(\frac{x}{6}\right)^{0/h+1} + \frac{1}{(1/h+1)!} \left(\frac{x}{6}\right)^{1/h+1} + \frac{1}{(2/h+1)!} \left(\frac{x}{6}\right)^{2/h+1} \right. \\ &\quad \quad \quad \left. + \frac{1}{(3/h+1)!} \left(\frac{x}{6}\right)^{3/h+1} + \dots \right] A \\ &\quad + \left[\frac{1}{(0/h+2)!} \left(\frac{x}{6}\right)^{0/h+2} + \frac{1}{(1/h+2)!} \left(\frac{x}{6}\right)^{1/h+2} + \frac{1}{(2/h+2)!} \left(\frac{x}{6}\right)^{2/h+2} \right. \\ &\quad \quad \quad \left. + \frac{1}{(3/h+2)!} \left(\frac{x}{6}\right)^{3/h+2} + \dots \right] A^2 \\ &\quad + \dots \\ &\quad + \left[\frac{1}{(0/h+1/h-1)!} \left(\frac{x}{6}\right)^{0/h+1/h-1} + \frac{1}{(1/h+1/h-1)!} \left(\frac{x}{6}\right)^{1/h+1/h-1} \right. \\ &\quad \quad \quad \left. + \frac{1}{(2/h+1/h-1)!} \left(\frac{x}{6}\right)^{2/h+1/h-1} + \frac{1}{(3/h+1/h-1)!} \left(\frac{x}{6}\right)^{3/h+1/h-1} \right. \\ &\quad \quad \quad \left. + \dots \right] A^{1/h-1} \\ &= \sum_{l=0}^{\frac{1}{h}-1} B_l(x) A^l. \end{aligned}$$

Similarly, by (2.73),

$$e^{\frac{1}{6}A^{-1}x} = \sum_{l=0}^{\frac{1}{h}-1} B_l(x)A^{-l},$$

where we use the standard notation $A^{-l} := (A^{-1})^l$.

Substituting both these identities into (2.79) we find

$$e^{\Gamma x} = e^{\frac{2}{3}x} \sum_{l=0}^{\frac{1}{h}-1} \sum_{r=0}^{\frac{1}{h}-1} B_l(x)B_r(x)A^lA^{-r}.$$

Hence,

$$\begin{aligned} (e^{\Gamma x})_{ij} &= e^{\frac{2}{3}x} \sum_{l=0}^{\frac{1}{h}-1} \sum_{r=0}^{\frac{1}{h}-1} B_l(x)B_r(x) \sum_{k=1}^{\frac{1}{h}-1} \delta_{i,k+l} \delta_{k,j-r} = e^{\frac{2}{3}x} \sum_{l=0}^{\frac{1}{h}-1} \sum_{r=0}^{\frac{1}{h}-1} B_l(x)B_r(x) \delta_{i,j-r+l} \\ &= e^{\frac{2}{3}x} \sum_{l=0}^{\frac{1}{h}-1} B_l(x)B_{l+j-i}(x). \end{aligned}$$

This completes the proof of (2.75). □

Remark 2.16 In the proof above we have given the deductions by which the two expressions for $e^{\Gamma x}$ were derived. Alternatively, we can also directly check that each column of $e^{\Gamma x}$ as given in (2.74) and (2.75) satisfies (2.76) with initial condition $e^{\Gamma 0} = I$.

2.9 Appendix C: Poincaré–Wirtinger-Type Inequalities

In this section we prove the Poincaré–Wirtinger-type inequalities we used in the proof of Lemma 2.7. The first one is a classic result, and the second one a higher order extension which is proved using similar methods as the first.

Lemma 2.14 (i) *Let $u \in H^1((0, h))$ with $u(0) = u(h) = 0$. Then the Poincaré–Wirtinger inequality holds:*

$$\int_0^h |u(x)|^2 dx \leq (h/\pi)^2 \int_0^h |u'(x)|^2 dx.$$

The constant $(h/\pi)^2$ is optimal.

(ii) Let $u \in H^2((0, h))$ with $u(0) = u(h) = 0$. Then

$$\int_0^h |u(x)|^2 dx \leq (h/\pi)^4 \int_0^h |u''(x)|^2 dx.$$

The constant $(h/\pi)^4$ is optimal.

Proof (i) We note that the boundary conditions are well-defined, since $H^1((0, h))$ is compactly embedded in the Banach space $\{u \in C((0, h)) \mid u \text{ is bounded}\}^{22}$ by the Rellich–Kondrachov theorem [1, Theorem 6.3].

The Poincaré–Wirtinger inequality for symmetric elliptic operators is found in [27, Sect. 6.5, Theorem 2 and the following remark], from which it immediately follows that there exists a $C > 0$ such that, for all $u \in H^1((0, 1))$ with $u(0) = u(h) = 0$, the inequality

$$\int_0^h |u(x)|^2 dx \leq C \int_0^h |u'(x)|^2 dx$$

holds. It also tells us that the optimal constant is $C = \nu^{-1}$, where ν is obtained

from its Rayleigh quotient formulation $\nu = \min_{\substack{w \in H^1((0, h)) \\ w \neq 0}} \frac{\int_0^h |w'(x)|^2 dx}{\int_0^h |w(x)|^2 dx}$ (where

the minimum is achieved), or as the minimal²³ eigenvalue in the eigenvalue problem²⁴

²² Equipped with the supremum norm $\sup_{x \in (0, h)} |u(x)|$.

²³ We emphasize that $w = 0$ is not allowed as eigenfunction.

²⁴ Additionally, the set of unit eigenfunctions of the differential operator $T(u) := u''$ is a complete orthonormal basis of the Hilbert space $L^2((0, h))$. This can be shown as follows. We claim that the inverse T^{-1} of the densely defined unbounded differential operator T on $L^2((0, h))$ exists and is compact. If this claim is true, then by the Riesz–Schauder theory [64, Sect. X.5, Theorem 2] all nonzero elements of the spectrum of T^{-1} are eigenvalues of T^{-1} and thus their reciprocals are eigenvalues of T with the same eigenfunctions. Moreover, since T^{-1} is also self-adjoint, its eigenfunctions form a complete orthogonal basis for $L^2((0, h))$ [64, Sect. XI.6, Theorem 1, and Sect. XI.9, Theorem 1 and Corollary 2]. Hence, the same is true for the eigenfunctions of T .

To prove that T^{-1} exists and is compact, we observe that, for $f \in L^2((0, h))$, $T^{-1}f = u$, where u is the unique solution to $u'' = f$ on $(0, h)$ and $u(0) = u(h) = 0$. Existence follows from an explicit computation that shows that

$$(T^{-1}f)(x) = \frac{1}{h} \int_0^x \int_0^h \left[\int_0^w f(y) dy - \int_0^z f(y) dy \right] dz dw.$$

If $\|f\|_{L^2((0, h))} \leq 1$ and $x, y \in (0, h)$ with $y > x$, then, by the Cauchy–Schwarz inequality,

$$\begin{aligned} |(T^{-1}f)(x)| &\leq \frac{2}{h} \int_0^h \int_0^h \int_0^h |f(y)| dy dz dw = 2h \int_0^h |f(y)| dy \\ &\leq 2h \|1\|_{L^2((0, h))} \|f\|_{L^2((0, h))} \leq 2h^{3/2}, \end{aligned}$$

$$\begin{cases} w'' + \tilde{\nu}w = 0, & \text{on } (0, h), \\ w(0) = w(h) = 0. \end{cases}$$

(Strict positivity of ν is given by [27, Sect.6.5, Theorem 1].) This ordinary differential equation can be solved by standard ODE techniques [37, Theorem III.4.1]; its general solution (without imposing boundary conditions) is given by

$$w(x) = c_1 \cos(\sqrt{\tilde{\nu}}x) + c_2 \sin(\sqrt{\tilde{\nu}}x),$$

for constants $c_1, c_2 \in \mathbb{R}$. From $w(0) = 0$, it follows that $c_1 = 0$. Since $w(h) = 0$, we have $c_2 = 0$ or $\sin(\sqrt{\tilde{\nu}}h) = 0$. Because $w = 0$ is not allowed as eigenfunction, the option $c_2 = 0$ is excluded, hence ν is the smallest $\tilde{\nu} > 0$ such that $\sin(\sqrt{\tilde{\nu}}h) = 0$. This proves that $\nu = \pi^2/h^2$, as required.

We note that [23, Sect. 1.7(2)] provides an alternative proof for continuously differentiable functions u based on Fourier series.

- (ii) Let $u \in \mathcal{A}_h := \{u \in H^2((0, h)) \mid u(0) = u(h) = 0\}$. We note that by the Rellich–Kondrachov theorem [1, Theorem 6.3] $H^2((0, h))$ is compactly embedded in the Banach space $\{u \in C^1((0, h)) \mid u, u' \text{ are bounded}\}^{25}$ and thus the boundary conditions in the definition of \mathcal{A}_h are well-defined.

We define, for this proof,

$$\nu := \inf_{\substack{w \in \mathcal{A}_h \\ w \neq 0}} \frac{\int_0^h |w''(x)|^2 dx}{\int_0^h |w(x)|^2 dx} = \inf_{w \in \mathcal{A}_h} \left\{ \|w''\|_{L^2((0,h))}^2 \mid \|w\|_{L^2((0,h))}^2 = 1 \right\}, \quad (2.22)$$

where the second expression follows by rescaling u by $\left(\int_0^h |w(x)|^2 dx\right)^{-1/2}$. Since the quotient is nonnegative, $\nu \geq 0$ exists. If, as we will prove, $\nu > 0$, then it follows that

$$\int_0^h |u(x)|^2 dx \leq \nu^{-1} \int_0^h |u''(x)|^2 dx.$$

and

$$\begin{aligned} |(T^{-1}f)(x_2) - (T^{-1}f)(x_1)| &\leq \frac{1}{h} \int_{x_1}^{x_2} \int_0^h \left| \int_0^w f(y) dy - \int_0^z f(y) dy \right| dz dw \\ &\leq 2|x_2 - x_1| \int_0^h |f(y)| dz \\ &\leq 2h^{1/2} \|f\|_{L^2((0,h))} |x_2 - x_1|. \end{aligned}$$

Hence by the Arzelà–Ascoli theorem [1, Theorem 1.33], [44, Theorem 5.37], T^{-1} is a compact operator.

²⁵ Equipped with the supremum norm $\max(\sup_{x \in (0,h)} |u(x)|, \sup_{x \in (0,h)} |u'(x)|)$.

Moreover, a minimizer of the minimization problem in the definition of ν exists by the direct method of the calculus of variations, as the following argument shows. We use the second characterization of ν in (2.22). Consider a minimizing sequence $(w_m)_m$ in $\left\{u \in \mathcal{A}_h \mid \|u\|_{L^2((0,h))}^2 = 1\right\}$. By definition $(\|w_m\|_{L^2((0,h))}^2)_m$ and $(\|w_m''\|_{L^2((0,h))}^2)_m$ are bounded. By part (i) of this lemma with $u = w'_m$, $(\|w'_m\|_{L^2((0,h))}^2)_m$ is also bounded, hence $(w_m)_m$ is bounded in $H^2((0, h))$ and thus an $H^2((0, h))$ -weakly converging subsequence exists with limit $w_* \in H^2((0, h))$. By the compact embedding of $H^2((0, h))$ into a space of continuous functions that was established above, $\|w_*\|_{L^2((0,h))}^2 = 1$. Moreover, since $u \mapsto \|u''\|_{L^2((0,h))}^2$ is lower semicontinuous with respect to weak- $H^2((0, h))$ convergence [10, Proposition 3.5], w_* minimizes $w \mapsto \|w''\|_{L^2((0,h))}^2$ over the admissible set.

From the first expression for ν in (2.22), we see that the minimizer w_* has to satisfy

$$\frac{d}{d\varepsilon} \frac{\int_0^h ((w_* + \varepsilon f)''(x))^2 dx}{\int_0^h (w_* + \varepsilon f)^2(x) dx} \Big|_{\varepsilon=0} = 0,$$

for all $f \in C^\infty([0, h])$ with $f(0) = f(h) = 0$. Computing this derivative and using that $\nu = \frac{\|w_*''\|_{L^2((0,h))}^2}{\|w_*\|_{L^2((0,h))}^2}$, we find

$$\frac{2}{\|w_*\|_{L^2((0,h))}^2} \int_0^h [w_*''(x)f''(x) - \nu w_*(x)f(x)] dx = 0.$$

Integrating by parts, we see

$$\begin{aligned} \int_0^h w_*''(x)f''(x) dx &= - \int_0^h w_*'''(x)f'(x) dx + w_*''f' \Big|_0^h \\ &= \int_0^h w_*''''(x)f(x) dx - w_*''f \Big|_0^h + w_*''f' \Big|_0^h. \end{aligned}$$

Since $f = 0$ on $\{0, h\}$, the term $w_*''f \Big|_0^h$ vanishes. Thus the weak form implies that

$$\int_0^h [w_*''''(x) - \nu(x)w_*(x)]f(x) dx + w_*''f' \Big|_0^h = 0,$$

for all $f \in C^\infty([0, h])$ with $f(0) = f(h) = 0$. Hence w_* solves

$$w_*'''' - \nu w_* = 0 \quad \text{on } (0, h).$$

Since f' can be chosen to have any two values at $\{0, h\}$, we get natural boundary conditions. Thus w_* solves

$$w''(0) = w''(h) = 0.$$

Since w_* satisfies $w_*|_{\{0,h\}} = 0$ and $w_* \neq 0$, we deduce that ν must be an eigenvalue of

$$\begin{cases} w'''' - \tilde{\nu}w = 0, & \text{on } (0, h), \\ w''(0) = w''(h) = 0, \\ w(0) = w(h) = 0. \end{cases} \tag{2.23}$$

We note that, if $\tilde{\nu} = 0$, then the ODE above implies that w is a polynomial of degree at most three and thus, due to the boundary conditions $w = 0$. Hence $\tilde{\nu} = 0$ is not an eigenvalue (since w is an eigenfunction and hence $w \neq 0$) and thus $\nu \neq 0$. Already we knew that $\nu \geq 0$, hence $\nu > 0$.

In fact, substituting the (admissible) eigenfunction $w(x) = \sin(\pi x/h)$ in the minimization problem in (2.22), we find that $\nu \geq (\pi/h)^4$.

Finally we prove that every eigenvalue $\tilde{\nu}$ in the eigenvalue problem in (2.23) is larger than or equal to $(\pi/h)^4$, which establishes that $\nu = (\pi/h)^4$, as required.

Let $\tilde{\nu} > 0$ be an eigenvalue of (2.23) and let $\sigma > 0$ be the unique strictly positive real number such that $\tilde{\nu} = \sigma^4$. By standard ODE methods [37, Theorem III.4.1], the general solution to the ODE $w'''' - \tilde{\nu}w = 0$ is found to be

$$w(x) = c_1 e^{\sigma x} + c_2 e^{-\sigma x} + c_3 e^{i\sigma x} + c_4 e^{-i\sigma x},$$

for constants $c_1, c_2, c_3, c_4 \in \mathbb{R}$. We recall that i is the imaginary unit. Since $w(0) = 0$ and $w''(0) = 0$ imply $c_1 + c_2 + c_3 + c_4 = 0$ and $\sigma^2(c_1 + c_2) - \sigma^2(c_3 + c_4) = 0$, we deduce that $c_1 + c_2 = c_3 + c_4 = 0$. Therefore

$$w(x) = c_1(e^{\sigma x} - e^{-\sigma x}) + c_3(e^{i\sigma x} - e^{-i\sigma x}).$$

Thus

$$\begin{aligned} w(h) &= c_1(e^{\sigma h} - e^{-\sigma h}) + c_3(e^{i\sigma h} - e^{-i\sigma h}) = 0, \\ \sigma^{-2}w''(h) &= c_1(e^{\sigma h} - e^{-\sigma h}) - c_3(e^{i\sigma h} - e^{-i\sigma h}) = 0, \end{aligned}$$

which in matrix form reads

$$\begin{pmatrix} e^{\sigma h} - e^{-\sigma h} & e^{i\sigma h} - e^{-i\sigma h} \\ e^{\sigma h} - e^{-\sigma h} & -(e^{i\sigma h} - e^{-i\sigma h}) \end{pmatrix} \begin{pmatrix} c_1 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This equation has a nontrivial solution if and only if the determinant of the above matrix is zero: $-2(e^{\sigma h} - e^{-\sigma h})(e^{i\sigma h} - e^{-i\sigma h}) = 0$. Since $e^{\sigma h} \neq e^{-\sigma h}$, we obtain that $e^{i\sigma h} - e^{-i\sigma h} = 0$. In other words, $e^{2i\sigma h} = 1$. The minimum strictly positive σ for which this condition holds is π/h . Thus $\nu = (\pi/h)^4$. This concludes the proof.²⁶

□

Acknowledgements The first author has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 777826. The second author was partly supported by the Japan Society for the Promotion of Science through grants No. 19H00689 (Kiban A), No. 18H05323 (Kaitaku), and by Arithmer Inc. and Daikin Industries Ltd. through a collaborative grant. The work of the third author was done when he was a graduate student at the University of Tokyo and a Junior Research Associate at RIKEN. His work was partly supported by the Program for Leading Graduate Schools “Frontier of Mathematical Sciences and Physics”, MEXT, Japan.

We thank the two anonymous referees whose feedback on an earlier manuscript of this work have improved the final version. We also thank Joshua Willems for his comments on a previous draft, as well as the audiences at talks given by the authors on the topic of this work whose questions and comments have refined the presentation of some of our results.

References

1. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Pure and Applied Mathematics, vol. 140. Elsevier, Oxford, Amsterdam (2003)
2. Alfaro, M., Hilhorst, D., Matano, H.: The singular limit of the Allen-Cahn equation and the FitzHugh-Nagumo system. *J. Differ. Equ.* **245**(2), 505–565 (2008)
3. Allen, S.M., Cahn, J.W.: A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. *Acta Metall.* **27**, 1085–1095 (1979)
4. Ambrosio, L., Brué, E., Semola, D.: Lectures on Optimal Transport, vol. 130 of UNITEXT, 1st edn. Springer, Cham (2021). *La Matematica per il 3+2 (UNITEXTMAT)*
5. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows: in Metric Spaces and in the Space of Probability Measures, 2nd edn. Springer Science & Business Media (2008)
6. Andreu-Vaillo, F., Caselles, V., Mazón, J.M.: Parabolic quasilinear equations minimizing linear growth functionals. In: *Progress in Mathematics*, vol. 223. Birkhäuser Verlag, Basel (2004)
7. Bertozzi, A.L., Flenner, A.: Diffuse interface models on graphs for classification of high dimensional data. *Multiscale Model. Simul.* **10**(3), 1090–1118 (2012)
8. Braides, A.: Local Minimization, Variational Evolution and Γ -Convergence. *Lecture Notes in Mathematics*, vol. 2094. Springer International Publishing, Switzerland (2014)

²⁶ Additionally, we can conclude from our computations that the geometric multiplicity of the eigenvalue $\tilde{\nu} = (\pi/h)^4$ is equal to one, as any solution w to (2.23) with $\tilde{\nu} = (\pi/h)^4$ is of the form $C \sin \pi x$ for some constant $C \in \mathbb{R}$.

9. Brezis, H.: Opérateurs maximaux monotones et semi-groupes de contractions DANs les espaces de Hilbert, volume No. 5/ Notas de Matemática(50) of North-Holland Mathematics Studies. North-Holland Publishing Co./American Elsevier Publishing Co., Inc., Amsterdam-London/New York (1973)
10. Brezis, H.: Functional analysis. In: Sobolev Spaces and Partial Differential Equations. Universitext. Springer, New York (2011)
11. Budd, J., van Gennip, Y.: Graph Merriman-Bence-Osher as a semidiscrete implicit Euler scheme for graph Allen-Cahn flow. *SIAM J. Math. Anal.* **52**(5), 4101–4139 (2020)
12. Budd, J., van Gennip, Y.: Mass-conserving diffusion-based dynamics on graphs. *Eur. J. Appl. Math.* First View, 1–49 (2021)
13. Budd, J., van Gennip, Y., Latz, J.: Classification and image processing with a semi-discrete scheme for fidelity forced Allen-Cahn on graphs. *GAMM Mitteilungen Spec. Issue: Sci. Mach. Learn.-Part I* **44**, 1–43 (2021)
14. Chen, X.: Generation and propagation of interfaces in reaction-diffusion systems. *Trans. Am. Math. Soc.* **334**(2), 877–913 (1992)
15. Choksi, R., van Gennip, Y., Oberman, A.: Anisotropic total variation regularized L^1 approximation and denoising/deblurring of 2D bar codes. *Inverse Probl. Imaging* **5**(3), 591–617 (2011)
16. Chung, F.R.K.: Spectral Graph Theory, Volume 92 of CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC, by the American Mathematical Society, Providence, Rhode Island (1997)
17. Cohen, I., Berkov, T., Gilboa, G.: Total-variation—fast gradient flow and relations to Koopman theory (2021). [arXiv:2105.10044](https://arxiv.org/abs/2105.10044) [math.AP]
18. Crouzeix, M., Thomée, V.: The stability in l_p and w_p^1 of the l_2 -projection onto finite element function spaces. *Math. Comput.* **48**, 521–532 (1987)
19. de Mottoni, P., Schatzman, M.: Development of interfaces in \mathbb{R}^n . *Proc. R Soc. Edinburgh Sect. A* **116**(3–4), 207–220 (1990)
20. de Mottoni, P., Schatzman, M.: Geometrical evolution of developed interfaces. *Trans. Am. Math. Soc.* **347**(5), 1533–1589 (1995)
21. Degiovanni, M., Marino, A., Tosques, M.: Evolution equations with lack of convexity. *Nonlinear Anal. Theory Methods Appl.* **9**(12), 1401–1443 (1985)
22. Dondl, P., Frenzel, T., Mielke, A.: A gradient system with a wiggly energy and relaxed EDP-convergence. *ESAIM:COCV* **25**, 68 (2019)
23. Dym, H., McKean, H.P.: Fourier Series and Integrals, 1st edn. Probability and Mathematical Statistics. Academic Press, New York and London (1972)
24. Ekeland, I., Temam, R.: Convex analysis and variational problems. North-Holland Publishing Co., Amsterdam (1976). Translated from the French, Studies in Mathematics and its Applications, vol. 1
25. Esedoglu, S., Osher, S.J.: Decomposition of images by the anisotropic Rudin-Osher-Fatemi model. *Commnu. Pure Appl. Math.* **57**(12), 1609–1626 (2004)
26. Esposito, A., Patacchini, Schlichting, F.S., Slepčev, A., Slepcev, D.: Nonlocal-interaction equation on graphs: gradient flow structure and continuum limit. *Arch. Rational Mech. Anal.* **240**, 699–760 (2021)
27. Evans, L.C.: Partial Differential Equations, vol. 19 of Graduate Studies in Mathematics, 2nd edn. American Mathematical Society, Providence, RI (2010)
28. Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. Studies in Advanced Mathematics, 1st edn. CRC Press LLC, Boca Raton, Florida (1992)
29. Evans, L.C., Soner, H.M., Souganidis, P.E.: Phase transitions and generalized motion by mean curvature. *Commun. Pure Appl. Math.* **45**(9), 1097–1123 (1992)
30. Fischer, J., Laux, T., Simon, T.M.: Convergence rates of the Allen-Cahn equation to mean curvature flow: a short proof based on relative entropies. *SIAM J. Math. Anal.* **52**(6), 6222–6233 (2020)
31. Fleißner, F.: Γ -convergence and relaxations for gradient flows in metric spaces: a minimizing movement approach. *ESAIM:COCV* **25**, 28 (2019)

32. Fujita, H., Saito, N., Suzuki, T.: *Operator Theory and Numerical Methods. Studies in Mathematics and its Applications*, vol. 30. Elsevier, Amsterdam (2001)
33. García Trillos, N., Slepčev, D.: Continuum limit of total variation on point clouds. *Arch. Rat. Mech. Anal.* **220**, 193–241 (2016)
34. Giusti, E.: *Minimal Surfaces and Functions of Bounded Variation*, vol. 80 of *Monographs in Mathematics*, 1st edn. Birkhäuser, Boston (1984)
35. Hafiene, Y., Fadili, J., Chesneau, C., Elmoataz, A.: Continuum limit of the nonlocal p -Laplacian evolution problem on random inhomogeneous graphs. *ESAIM: Math. Model. Numer. Anal.* **54**(2), 565–589 (2020)
36. Hafiene, Y., Fadili, J., Elmoataz, A.: Nonlocal p -Laplacian evolution problems on graphs. *SIAM J. Numer. Anal.* **56**(2), 1064–1090 (2018)
37. Hale, J.K.: *Ordinary Differential Equations*, 2nd edn. Dover Publications Inc., Mineola, New York (2009)
38. Itô, S.: *Introduction to Lebesgue Integration, (Lebesgue sekibun nyu-mon)*. Shokabo (1965). in Japanese; reprint in 2017
39. Jones, D.S.: *The Theory of Generalised Functions*, 2nd edn. Cambridge University Press, Cambridge-New York (1982)
40. Keetch, B., van Gennip, Y.: A Max-Cut approximation using a graph based MBO scheme. *Discrete Contin. Dyn. Syst. Ser. B* **24**(11), 6091–6139 (2019)
41. Kōmura, Y.: Nonlinear semi-groups in Hilbert space. *J. Math. Soc. Jpn.* **19**, 493–507 (1967)
42. Michal, Ł., Moll, S., Mucha, P.B.: Total variation denoising in ℓ^1 anisotropy. *SIAM J. Imaging Sci.* **10**, 1619–1723 (2017)
43. Laux, T., Lelmi, J.: Large data limit of the MBO scheme for data clustering: convergence of the dynamics (2022). [arXiv:2209.05837](https://arxiv.org/abs/2209.05837) [math.AP]
44. Leoni, G.: *A First Course in Sobolev Spaces. Graduate Studies in Mathematics*, vol. 105. American Mathematical Society, Providence, Rhode Island (2009)
45. Loeb, P.A.: *Real Analysis*. Birkhäuser/Springer, [Cham] (2016)
46. Mielke, A.: On evolutionary Γ -convergence for gradient systems. In: Muntean, A., Rademacher, J., Zagaris, A. (eds.) *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, vol. 3 of *Lecture Notes in Applied Mathematics and Mechanics*, pp. 187–249. Springer, Cham (2016)
47. Mielke, A., Montefusco, A., Peletier, M.A.: Exploring families of energy-dissipation landscapes via tilting: three types of EDP convergence. *Continuum Mech. Thermodyn.* **33**, 611–637 (2021)
48. Mielke, A., Peletier, M.A., Stephan, A.: EDP-convergence for nonlinear fast-slow reaction systems with detailed balance. *Nonlinearity* **34**, 5762–5798 (2021)
49. Mielke, A., Rossi, R., Savaré, G.: Variational convergence of gradient flows and rate-independent evolutions in metric space. *Milan J. Math.* **80**, 381–410 (2012)
50. Mielke, A., Stephan, A.: Coarse-graining via EDP-convergence for linear fast-slow reaction systems. *Math. Model. Methods Appl. Sci.* **30**(9), 1765–1807 (2020)
51. Modica, L., Mortola, S.: Un esempio di Γ^- -convergenza. *Boll. Un. Mat. Ital. B* **14**(1), 285–299 (1977)
52. Muratori, M., Savaré, G.: Gradient flows and evolution variational inequalities in metric spaces. I: structural properties. *J. Funct. Anal.* **278**(4), 108347 (2020)
53. Ortner, C.: Two variational techniques for the approximation of curves of maximal slope. Technical Report Technical Report NA-05/10, University of Oxford Numerical Analysis Group Research Reports, Oxford University Computing Laboratory, Numerical Analysis Group, Wolfson Building, Parks Road, Oxford, England OX1 3QD (2005)
54. Krishnan, P.: *Topology*, vol. 134 of *UNITEXT*, 1st edn. Springer, Singapore (2022). *La Matematica per il 3+2 (UNITEXTMAT)*
55. Roman, S.: *Advanced Linear Algebra*, vol. 135 of *Graduate Texts in Mathematics*, 3rd edn. Springer, New York, NY (2008)
56. Sandier, E., Serfaty, S.: Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Commun. Pure Appl. Math.* **57**, 1627–1672 (2004)

57. Serfaty, S.: *Gamma-Convergence of Gradient Flows and Applications to Ginzburg–Landau Vortex Dynamics*, vol. 2 of *Lecture Notes of the Unione Matematica Italiana*, pp. 267–292. Springer, Berlin (2006)
58. Serfaty, S.: Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.* **31**, 1427–1451 (2011)
59. Soner, H.M.: Ginzburg–Landau equation and motion by mean curvature. I. Convergence. *J. Geom. Anal.* **7**(3), 437–475 (1997)
60. Soner, H.M.: Ginzburg–Landau equation and motion by mean curvature. II. Development of the initial interface. *J. Geom. Anal.* **7**(3), 477–491 (1997)
61. van Gennip, Y., Bertozzi, A.L.: Γ -convergence of graph Ginzburg–Landau functionals. *Adv. Differ. Equ.* **17**(11–12), 1115–1180 (2012)
62. van Gennip, Y., Guillen, N., Osting, B., Bertozzi, A.L.: Mean curvature, threshold dynamics, and phase field theory on finite graphs. *Milan J. Math.* **82**(1), 3–65 (2014)
63. von Luxburg, U.: A tutorial on spectral clustering. *Stat. Comput.* **17**(4), 395–416 (2007)
64. Yosida, K.: *Functional analysis*. In: *Classics in Mathematics*. Springer, Berlin (1995). Reprint of the sixth edition (1980)