Report no. 05/10

Two Variational Techniques for the Approximation of Curves of Maximal Slope

Christoph Ortner

This report generalizes several results on approximations of curves of maximal slope in the book by Ambrosio, Gigli and Savaré (*Gradient Flows*, 2005), allowing general approximations of the functional and of the metric. The conditions guaranteeing the convergence of the approximations are closely related to the conditions of Γ-convergence. Complete convergence results are obtained for λ-convex functionals together with general results indicating a possible procedure for even more general problems.

The theory presented provides convergence results as well as explicit a-priori error estimates for gradient flows of functionals which may be non-convex as well as non-differentiable.

Key words and phrases: curves of maximal slope, Gamma convergence, gradient flows, metric spaces

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD

June, 2005

Contents

1 Introduction **3**
1.1 Analysis in metric spaces 4
1.1.1 Metric derivative 5
1.1.2 Upper gradient and local slope 5

2 Approximation by Γ-Convergence **6**
2.1 The discrete scheme 6
2.2 Topological conditions 6
2.3 Auxiliary results 8
2.4 The abstract convergence result 9
2.5 The λ-convex case 14

3 Variational Inequalities **19**
3.1 Relaxed convexity and coercivity conditions 19
3.2 Evolutionary variational inequalities 22
3.3 Continuous case 25
3.4 Discrete case 30
3.5 Examples for (3.25) 35

1 Introduction

Let $(\mathscr{S}, d)$ be a metric space and $\phi : \mathscr{S} \to \mathbf{R}$ a functional defined on $\mathscr{S}$. The natural generalization of a gradient flow to the metric setting is a curve of maximal slope, i.e. an absolutely continuous curve $u : (a, b) \to \mathscr{S}$, satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) = -\frac{1}{2}|u'|^2(t) - \frac{1}{2}|\partial\phi|^2(u(t)), \tag{1.1}$$

where $|u'|$ is the metric derivative of u and $|\partial\phi|(u)$ the local slope of ϕ at u (see Sections 1.1.1 and 1.1.2 for the definitions). It can be easily seen that in the Hilbert space setting, (1.1) is equivalent to the classical formulation of a gradient flow if ϕ is differentiable. For, suppose that $\mathscr{S}$ is a Hilbert space, and that ϕ and the curve u are differentiable and satisfy $\dot{u} = -\phi'(u)$; then

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u) = (\phi'(u), \dot{u}) = -\|\dot{u}\|\|\phi'(u)\| = -\frac{1}{2}\|\dot{u}\|^2 - \frac{1}{2}\|\phi'(u)\|^2.$$

(Note that in this case $\|\dot{u}\| = |u'|$ and $\|\phi'(u)\| = |\partial\phi|(u)$.) On the other hand, if (1.1) is satisfied then

$$\begin{aligned} -\frac{1}{2}\|\dot{u}\|^2 - \frac{1}{2}\|\phi'(u)\|^2 &= \frac{\mathrm{d}}{\mathrm{d}t}\phi(u) \\ &= (\phi'(u), \dot{u}) \\ &\geq -\|\phi'(u)\|\|\dot{u}\| \\ &\geq -\frac{1}{2}\|\dot{u}\|^2 - \frac{1}{2}\|\phi'(u)\|^2. \end{aligned}$$

Hence equalities hold in the Cauchy-Schwarz and the Cauchy inequalities which we have applied; this, together with the fact that ϕ is decreasing along u, shows that $\dot{u} = -\phi'(u)$.

The motivation for the study of approximations of curves of maximal slope was the wish to analyze gradient flows of the Mumford-Shah functional, with an eye to applications in computational fracture mechanics. Interesting results for a regularization of the Mumford-Shah functional were recently obtained by Feng and Prohl [FP04] but, except in the one-dimensional setting in a work by Gobbino [Gob98], they could not be extended to gradient flows of the Mumford-Shah functional itself. Even the theory developed here could so far not be applied to this problem, but it is sufficiently general that it should be useful in many other applications.

Chapters 1-3 in [AGS05] provide an existence theory for curves of maximal slope which is based on an implicit (backward Euler) time-discretization, which requires, at the jth time-step, to minimize the functional

$$U^j \mapsto \frac{d^2(U^j, U^{j-1})}{2\Delta t} + \phi(U^j). \tag{1.2}$$

Using a compactness argument and lower semicontinuity assumptions for ϕ and $|\partial\phi|$ it can be shown that certain interpolants of (U_j) converge to a curve of maximal slope in a suitable topology.

In this paper, a useful generalization of this approach is presented. We allow that the functional ϕ in (1.2) is replaced by an approximation ϕ_h. More precisely, for every $h \in \mathbf{N}$, we consider an approximate functional ϕ_h and a time-partition $\boldsymbol{\tau}_h = (t_h^0, t_h^1, \dots)$ and compute the minimizers U_h^j of the functionals

$$U_h^j \mapsto \frac{d^2(U_h^j, U_h^{j-1})}{2(t_h^j - t_h^{j-1})} + \phi_h(U_h^j).$$

In Section 2 it is shown that, under suitable conditions on the family (ϕ_h), certain interpolants of $(U_h^j)_{j=0,1,\dots}$ converge to a curve of maximal slope. For some classes of functionals, these conditions reduce to $\Gamma(d)$-convergence of ϕ_h to ϕ, cf. Section 2.5.

Chapter 4 in [AGS05] provides an alternative theory for curves of maximal slope in a more restrictive setting. Under relaxed convexity conditions on the functional ϕ and strong conditions on the metric (cf. Condition (3.2)), it is shown that the curve of maximal slope $u(t)$ with a given initial value is unique and that it satisfies the evolutionary variational inequality

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} d^2(u(t), v) + \lambda d^2(u(t), v) + \phi(u(t)) \le \phi(v) \quad \forall v \in \mathscr{S},$$

where λ is related to the relaxed convexity condition (3.2). Based on this inequality, we provide a second theory for the approximation of curves of maximal slope, which yields significantly stronger results, although it is restricted to a much smaller class of problems. In Section 3.3 a time-continuous convergence analysis is developed, whereas Section 3.4 presents the corresponding theory for fully discrete approximations. Section 3 can be read independently from Section 2.

In the remainder of the Introduction, we review some results from the theory of metric spaces, which is required for the subsequent analysis.

1.1 Analysis in metric spaces

Let $(\mathscr{S}, d)$ be a complete metric space, endowed with an additional Hausdorff topology σ. Throughout this paper, we denote convergence as follows. If $d(u_j, u) \to 0$ as $j \to \infty$, then we write $u_j \to u$. If u_j converges to u in the topology σ, then we write $u_j \xrightarrow{\sigma} u$. As σ is typically a weak topology with good compactness properties, we call the convergence $u_j \xrightarrow{\sigma} u$ weak convergence. We assume throughout this paper, that σ is compatible with the topology induced by d, in the sense that if $u_j \to u$ then $u_j \xrightarrow{\sigma} u$ as well.

By $B(u, r)$ we denote the closed ball in $\mathscr{S}$ with radius r and center u.

1.1.1 Metric derivative

Let $v : (a,b) \to \mathscr{S}$, where $(a,b) \subset \mathbf{R}$. We say that the curve v belongs to $AC^p(a,b;\mathscr{S})$, for $1 \le p \le \infty$, if there exists a function $A \in L^p(a,b)$ such that

$$d(v(s),v(t)) \le \int_s^t A(r)\,\mathrm{d}r \quad \forall s < t \in (a,b).$$

We identify $AC^1 = AC$. If $v \in AC^p(a,b;\mathscr{S})$, then there exists function $|v'| \in L^p(a,b)$, the metric derivative of v, such that

$$\lim_{s\to t} \frac{d(v(t),v(s))}{|t-s|} = |v'|(t), \quad \text{for a.e. } t \in (a,b). \tag{1.3}$$

If $v \in AC^p(a,b;\mathscr{S})$ then v is uniformly continuous as a mapping from $[a,b]$ to $\mathscr{S}$. In particular, if $a \in \mathbf{R}$ (or $b \in \mathbf{R}$) then $v(a+)$ (or $v(b-)$) exist. If $v \in AC(a,b;\mathscr{S})$ then $v(a+)$ and $v(b-)$ exist even in the case $(a,b) = \mathbf{R}$.

If $\mathscr{S}$ is a Hilbert space and d the metric induced by its norm, then for any curve $v \in AC(a,b;\mathscr{S})$, $\dot v(t)$ exists for a.e. $t \in (a,b)$ and the metric derivative $|v'|$ coincides with $\|\dot v\|_{\mathscr{S}}$.

For further detail and proofs, see Section 1.1 in [AGS05].

1.1.2 Upper gradient and local slope

Let $\phi : \mathscr{S} \to (-\infty,\infty]$ be a functional on $\mathscr{S}$. We denote its domain by $D(\phi)$, i.e.,

$$D(\phi) = \{v \in \mathscr{S} : \phi(v) < \infty\}.$$

We say that a function $g : \mathscr{S} \to [0,\infty]$ is a strong upper gradient for ϕ if for every curve $v \in AC(a,b;\mathscr{S})$, the function $g \circ v$ is Borel and

$$|\phi(v(t)) - \phi(v(s))| \le \int_s^t g(v(r))|v'|(r)\,\mathrm{d}r. \tag{1.4}$$

A typical candidate for an upper gradient, the local slope of ϕ at a point v is defined as

$$|\partial\phi|(v) = \limsup_{w\to v} \frac{(\phi(v)-\phi(w))^+}{d(v,w)}. \tag{1.5}$$

Here and below we assume $p \in (1,\infty)$ and set $p' = p/(p-1)$ so that $1/p + 1/p' = 1$. We say that a curve $u \in AC^p_{\mathrm{loc}}(a,b;\mathscr{S})$ is a p-curve of maximal slope for the functional ϕ with respect to its upper gradient g if $t \mapsto \phi(u(t))$ is $\mathscr{L}^1$-a.e. equal to a non-increasing map and

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) \le -\frac{1}{p}|u'|^p(t) - \frac{1}{p'}g(u(t))^{p'} \tag{1.6}$$

for $\mathscr{L}^1$-a.e. $t \in (a,b)$.

If $\mathscr{S}$ is a Hilbert space and ϕ Fréchet differentiable at u, then $|\partial\phi|(u) = \|\phi'(u)\|_{\mathscr{S}}$. For further information and proofs, see Sections 1.2-1.4 in [AGS05].

2 Approximation by Γ-Convergence

2.1 The discrete scheme

For every $h \in \mathbf{N}$, let $\boldsymbol{\tau}_h$ be a mesh on $[0,\infty)$, $\boldsymbol{\tau}_h = (0 = t_h^0, t_h^1, \dots)$. We denote $\tau_h^j = t_h^j - t_h^{j-1}$ and $|\boldsymbol{\tau}_h| = \sup_{j\geq 1} \tau_h^j$, and we assume that $|\boldsymbol{\tau}_h| \to 0$ as $h \to \infty$.

Let $(\phi_h)_{h\in\mathbf{N}}$ be a family of functionals and $(d_h)_{h\in\mathbf{N}}$ be a family of metrics on $\mathscr{S}$, which approximate ϕ and d respectively as $h \to \infty$ in a sense that is made precise in Section 2.2. The discrete slopes are defined as

$$|\partial_h \phi_h|(v) = \limsup_{d_h(w,v)\to 0} \frac{(\phi_h(v) - \phi_h(w))^+}{d_h(v,w)}.$$

Let $(U_h^0)_{h\in\mathbf{N}}$ be a family of approximate initial values. We say that $(U_h^j)_{j\geq 0}$ is a *discrete solution* (of the gradient flow) if for every $j \geq 1$, U_h^j minimizes the functional

$$v \mapsto \Phi_h(U_h^{j-1}, \tau_h^j; v) := \frac{d_h^p(U_h^{j-1}, v)}{p(\tau_h^j)^{p-1}} + \phi_h(v) \tag{2.1}$$

over $\mathscr{S}$, and we define $J_{h,\tau}[u]$ to be the set of minimizers of $\Phi_h(u,\tau;\cdot)$.

Suppose that $(U_h^j)_{j\geq 0}$ is a discrete solution. We define the interpolants

$$\begin{aligned} \bar{U}_h(t) &= \begin{cases} U_h^0, & t = 0 \\ U_h^j, & t_h^{j-1} < t \leq t_h^j, j \geq 1 \end{cases} \\ |U_h'|(t) &= \frac{d_h(U_h^j, U_h^{j-1})}{\tau_h^j}, \quad \text{if } t_h^{j-1} < t < t_h^j. \end{aligned}$$

By $\tilde{U}_h(t)$ (De Giorgi's variational interpolant) we denote any interpolant of (U_h^j) such that $\tilde{U}_h(t) \in J_{h,t-t_h^{j-1}}[U_h^{j-1}]$ whenever $t_h^{j-1} < t \leq t_h^j$. Finally, we denote

$$G_h(t) = \frac{d_h(\tilde{U}_h(t), U_h^{j-1})}{t - t_h^{j-1}}, \quad \text{if } t_h^{j-1} < t \leq t_h^j.$$

It follows from Lemma 1 (b), that

$$|\partial_h \phi_h|^{p'}(\tilde{U}_h(t)) \leq G_h^p(t) \quad \forall t \in (0,\infty), \forall h \in \mathbf{N}, \tag{2.2}$$

thus providing a connection between G_h and the discrete slopes. The existence of an interpolation $\tilde{U}_h$ is guaranteed for sufficiently small $|\boldsymbol{\tau}_h|$ by condition (2.3) below.

2.2 Topological conditions

We start by a condition that guarantees the solubility of the discrete scheme for a sufficiently fine partition. We assume that

$$\exists \tau^* > 0 \text{ such that } \forall\, 0 < \tau \leq \tau^*, u \in \mathscr{S}, h \in \mathbf{N} \,:\, J_{h,\tau}[u] \neq \emptyset. \tag{2.3}$$

Detailed assumptions under which (2.3) is satisfied are standard. See for example [Dac89] and in particular Section 2.2 [AGS05].

Next, the initial conditions are required to satisfy

$$U_h^0 \overset{\sigma}{\rightharpoonup} u_0, \quad \phi_h(U_h^0) \to \phi(u_0), \quad \text{and } \sup_h d(U_h^0, u_0) < \infty. \tag{2.4}$$

A priori bounds on $\phi_h(u_h)$ are usually obtained by an equi-coercivity assumption on the familiy (ϕ_h). Here, in the time-dependent case, the following condition is sufficient. We assume that

$$\exists\, u^* \in \mathscr{S}, \tau^* > 0 \,:\, \inf_{h\in\mathbf{N}} \inf_{\mathscr{S}} \Phi_h(\tau^*, u^*; \cdot) = c_0 > -\infty. \tag{2.5}$$

We have used the time-step restriction τ^* both in (2.3) and (2.5). Of course, we may assume without loss of generality that they are the same, but more importantly, they are typically closely related, since it represents the threshold where $\Phi(\tau, u; \cdot)$ becomes bounded below.

To obtain compactness of the family of discrete solutions, we require that d_h- and ϕ_h-bounded sets are sequentially σ-compact:

$$\sup_h (d_h(U_h^0, u_h) + \phi_h(u_h)) < \infty \quad \Rightarrow \quad \exists (h_k)_{k\in\mathbf{N}} \uparrow \infty, u \in \mathscr{S} \text{ s.t. } u_{h_k} \overset{\sigma}{\rightharpoonup} u. \tag{2.6}$$

Here U_h^0 denotes the approximate initial value which should also satisfy (2.4).

To prove that limits of discrete solutions are curves of maximal slope three liminf conditions, similar to the one of Γ-convergence[1], are crucial. We will require for several results below that the following hold:

$$(u_h, v_h) \overset{\sigma}{\rightharpoonup} (u, v) \;\Rightarrow\; \liminf_{h\to\infty} d_h(u_h, v_h) \geq d(u, v) \text{ and} \tag{2.7}$$

$$u_h \overset{\sigma}{\rightharpoonup} u \quad \Rightarrow \quad \liminf_{h\to\infty} \phi_h(u_h) \geq \phi(u). \tag{2.8}$$

The third liminf condition is a liminf condition for the slopes. Let $g : \mathscr{S} \to [0, \infty]$ be defined as

$$g(u) = \big(\Gamma(\sigma)\text{–}\liminf_{h\to\infty} |\partial_h \phi_h|\big)(u) = \inf\{\liminf_{h\to\infty} |\partial_h \phi_h|(u_h), u_h \overset{\sigma}{\rightharpoonup} u\}. \tag{2.9}$$

The most general way to formulate the remaining liminf condition is to require that g is a strong upper gradient. In practice, it may be advantageous to show that $|\partial\phi|$ is an upper gradient and is bounded above by g, i.e., to require that

$$u_h \overset{\sigma}{\rightharpoonup} u \quad \Rightarrow \quad |\partial\phi|(u) \leq \liminf_{h\to\infty} |\partial_h \phi_h|(u_h). \tag{2.10}$$

A reader familiar with Γ-convergence will have noticed that no limsup condition was formulated. As it turns out, the limsup condition is implicitly hidden in the assumption that g is an upper gradient. This condition is probably the most difficult one to verify. In Section 2.5 we give an example of such a technique.

[1] For an introduction to Γ-convergence see for example [Bra02].

2.3 Auxiliary results

For every $\tau > 0, u, v \in \mathscr{S}$ let

$$\Phi(u, \tau; v) = \frac{d^p(u, v)}{p\tau^{p-1}} + \phi(v), \quad \text{and } J_\tau[u] = \operatorname{argmin} \Phi(u, \tau; \cdot)$$

We define the Moreau-Yosida approximation of ϕ to be

$$\phi_\tau(u) = \inf \Phi(\tau, u; \cdot).$$

If $J_\tau[u] \neq \emptyset$ then we also define

$$\mathrm{d}_\tau^+(u) = \sup_{u_\tau \in J_\tau[u]} d(u_\tau, u) \quad \text{and } \mathrm{d}_\tau^-(u) = \inf_{u_\tau \in J_\tau[u]} d(u_\tau, u).$$

The following Lemma is the first crucial step towards the convergence theorem. In particular, part (c) will provide a discrete version of (1.6).

Lemma 1 (Properties of the Moreau-Yosida Approximation) *Assume that ϕ is d-lower semicontinuous and that there exists $\tau^* > 0$ such that for $0 < \tau < \tau^*$, $J_\tau[u] \neq \emptyset$.*

(a) The mapping $(\tau, u) \mapsto \phi_\tau(u)$ is continuous in $(0, \tau^*) \times \mathscr{S}$ and if $0 < \tau_0 < \tau_1 < \tau^*$ and $u_{\tau_i} \in J_{\tau_i}[u]$, then

$$\begin{aligned} &\phi(u) \geq \phi_{\tau_0}(u) \geq \phi_{\tau_1}(u), \quad d(u_{\tau_0}, u) \leq d(u_{\tau_1}, u), \\ &\phi(u) \geq \phi(u_{\tau_0}) \geq \phi(u_{\tau_1}), \quad \mathrm{d}_{\tau_0}^+(u) \leq \mathrm{d}_{\tau_1}^-(u) \leq \mathrm{d}_{\tau_1}^+(u). \end{aligned} \tag{2.11}$$

In particular,

$$\lim_{\tau \downarrow 0} \phi_\tau(u) = \lim_{\tau \downarrow 0} \inf_{u_\tau \in J_\tau[u]} \phi(u_\tau) = \phi(u),$$

and if $u \in \overline{D(\phi)}$ then $\lim_{\tau \downarrow 0} \mathrm{d}_\tau^+(u) = 0$. Furthermore, there exists an at most countable set $\mathscr{N}_u \subset (0, \tau^*)$ such that

$$\mathrm{d}^-(u) = \mathrm{d}^+(u) \quad \forall \tau \in (0, \tau^*) \setminus \mathscr{N}_u.$$

(b) If $u_\tau \in J_\tau[u]$, then $u_\tau \in D(|\partial\phi|)$ and

$$|\partial\phi|^{p'}(u_\tau) \leq \frac{d^p(u_\tau, u)}{\tau^p}.$$

(c) For every $u \in \mathscr{S}$, the map $\tau \mapsto \phi_\tau(u)$ is locally Lipschitz continuous in $(0, \tau^*)$ and

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\phi_\tau(u) = -\frac{(\mathrm{d}_\tau^\pm(u))^p}{p'\tau^p} \quad \forall \tau \in (0, \tau^*) \setminus \mathscr{N}_u.$$

In particular, we have that, for every $\tau \in (0, \tau^*)$ and $u_\tau \in J_\tau[u]$,

$$\frac{d^p(u_\tau, u)}{p\tau^{p-1}} + \int_0^\tau \frac{(\mathrm{d}_r^\pm(u))^p}{p'r^p}\,\mathrm{d}r = \phi(u) - \phi(u_\tau). \tag{2.12}$$

For a proof of this lemma see Lemmas 3.1.2 to 3.1.5 as well as Remark 3.1.7 in [AGS05]. Note, however, that the completeness of the metric space is not required here. This allows a more general choice of the approximate metrics d_h.

We also state two compactness theorems which are used in the analysis of Section 2. For their proofs see Proposition 3.3.1 and Lemma 3.3.3 in [AGS05].

Lemma 2 (Ascoli-Arzelà Theorem) *Let $T > 0$, $K \subset \mathscr{S}$ be a sequentially compact set with respect to σ, and, for $h \in \mathbf{N}$, let $u_h : [0,T] \to \mathscr{S}$ be curves such that*

$$\begin{aligned} &\forall t \in [0,T], h \in \mathbf{N} : u_h(t) \in K, \text{ and} \\ &\forall s,t \in [0,T] : \limsup_{h\to\infty} d(u_h(s), u_h(t)) \le w(s,t) \end{aligned}$$

where w is a symmetric function $w : [0,T] \times [0,T] \to [0,\infty)$ such that

$$\forall r \in [0,T] \setminus \mathscr{N} : \lim_{(s,t)\to(r,r)} w(s,t) = 0,$$

where $\mathscr{N}$ may be at most countable. Then, there exists a subsequence (u_{h_k}) and a limit curve $u : [0,T] \to \mathscr{S}$ such that u is d-continuous in $[0,T] \setminus \mathscr{N}$ and

$$\forall t \in [0,T] : u_{n_k}(t) \xrightarrow{\sigma} u(t).$$

Lemma 3 (Helly's Theorem) *Suppose that, for $h \in \mathbf{N}$, $\varphi_h : [0,T] \to [-\infty,\infty]$ are non-increasing functions. Then, there exists a subsequence (φ_{h_k}) and a non-increasing map $\varphi : [0,T] \to [-\infty,\infty]$ such that $\lim_k \varphi_{h_k}(t) = \varphi(t)$ for every $t \in [0,T]$.*

2.4 The abstract convergence result

In this section we show all steps to prove the convergence of discrete solutions to a curve of maximal slope. Proposition 4 provides a discrete version of (1.6) and bounds on the energy and the distance, which are used in Proposition 5 to establish the weak sequential compactness of discrete solutions. Finally, in Proposition 6 we show that any accumulation point of the family of discrete solutions is a curve of maximal slope, using the liminf conditions of Section 2.

The following Lemma is a version of Lemma 3.2.2 in [AGS05], adapted to the present case, where it is crucial that all a-priori bounds are independent of h.

Proposition 4 (A-Priori Estimates) *Suppose that* (2.3), (2.4) *and* (2.5) *hold and suppose that for every $h \in \mathbf{N}$, $(U_h^j)_{j\in\mathbf{N}}$ is a discrete solution as described in Section 2.1.*

Then, for every $T > 0$, there exists a constant $C_0 \in \mathbf{R}$ such that

$$\forall j \in \mathbf{N}, \forall h \in \mathbf{N} \text{ such that } t_h^j \le T, \text{ we have } \phi_h(U_h^i) \ge C_0. \tag{2.13}$$

Furthermore, the following estimates hold for $t_h^{j-1} < t \le t_h^j \le T$*:*

$$\frac{1}{p}\int_{t_h^{j_1}}^{t_h^{j_2}} |U_h'|^p(t)\,\mathrm{d}t + \frac{1}{p'}\int_{t_h^{j_1}}^{t_h^{j_2}} |G_h(t)|^p\,\mathrm{d}t + \phi_h(U_h^{j_2}) = \phi_h(U_h^{j_1}), \quad (2.14)$$

$$\frac{1}{p} d_h^p(U_h^j, U_h^0) \le (t_h^j)^{p/p'}(\phi_h(U_h^0) - C_0), \quad (2.15)$$

$$\frac{1}{p} d_h^p(\tilde{U}_h(t), \bar{U}_h(t)) \le \frac{2^p|\boldsymbol{\tau}_h|^{p-1}}{p}\int_{t_h^{j-1}}^{t_h^j} |U_h'|^p(t)\,\mathrm{d}t \le 2^p|\boldsymbol{\tau}_h|^{p-1}(\phi_h(U_h^0) - C_0). \quad (2.16)$$

Proof Using equation (2.12) and the definitions of $|U_h'|$ and G_h gives

$$\frac{1}{p}\int_{t_h^{j-1}}^{t_h^j} |U_h'|^p(t)\,\mathrm{d}t + \frac{1}{p'}\int_{t_h^{j-1}}^{t_h^j} |G_h(t)|^p\,\mathrm{d}t + \phi_h(U_h^j) = \phi_h(U_h^{j-1}).$$

After summing over $j = j_1 + 1, \ldots, j_2$, we obtain (2.14).

To obtain a lower bound on the energy, consider

$$\begin{aligned} \phi_h(U_h^j) &= \Phi_h(\tau^*, u^*; U_h^j) - d^2(u^*, U_h^j) \\ &\ge c_0 - d^2(u^*, U_h^j), \end{aligned} \quad (2.17)$$

where we used assumption (2.5). It is therefore sufficient to compute an upper bound on $d(U_h^j, u^*)$. Following the proof in Lemma 3.2.2 [AGS05], we estimate

$$\begin{aligned} \frac{1}{2}d^2(U_h^j, u^*) - \frac{1}{2}d^2(U_h^0, u^*) &= \frac{1}{2}\sum_{i=1}^{j}\left(d^2(U_h^i, u^*) - d^2(U_h^{i-1}, u^*)\right) \\ &\le \sum_{i=1}^{j} d(U_h^i, U_h^{i-1}) d(U_h^i, u^*) \\ &\le \frac{\epsilon}{2}\int_0^{t_h^j} |U_h'|^2(t)\,\mathrm{d}t + \frac{1}{2\epsilon}\sum_{i=1}^{j}\tau_h^i d^2(U_h^j, u^*) \\ &\le \epsilon\phi_h(U_h^0) - \epsilon c_0 + \frac{\epsilon}{2\tau^*}d^2(U_h^j, u^*) + \frac{1}{2\epsilon}\sum_{i=1}^{j}\tau_h^i d^2(U_h^i, u^*). \end{aligned}$$

Choosing $\epsilon = \tau^*/2$, and using (2.4) and (2.5), we obtain

$$d^2(U_h^j, u^*) \le a + b\sum_{i=1}^{j}\tau_h^i d^2(U_h^i, u^*),$$

where a and b are constants which depend only on the bounds in (2.4) and (2.5), to which we can apply Gronwall's inequality, Lemma 31, to get a constant D, depending on T, but not on h such that

$$d^2(U_h^j, u^*) \le D \quad \forall h, j \in \mathbf{N} \text{ s.t. } t_h^j \le T. \quad (2.18)$$

This Together with (2.17), this gives (2.13).

Successive application of the triangle inequality gives the bound

$$d_h(U_h^j, U_h^0) \le \sum_{i=1}^{j} \tau_h^i \frac{d_h(U_h^i, U_h^{i-1})}{\tau_h^i} = \int_0^{t_h^j} |U_h'|(t)\,\mathrm{d}t$$

to which we apply Hölder's inequality to obtain

$$d_h(U_h^j, U_h^0) \le (t_h^j)^{1/p'} \left(\int_0^{t_h^j} |U_h'|^p(t)\,\mathrm{d}t \right)^{1/p}.$$

Equation (2.15) is obtained by taking the p-th power and applying (2.14) as well as (2.13). Suppose that $t_h^{j-1} < t < t_h^j$. From (2.11), we see that

$$d_h(\tilde{U}_h(t), \bar{U}_h(t)) \le d_h(\tilde{U}_h(t), U_h^{j-1}) + d_h(U_h^{j-1}, U_h^j) \le 2d_h(U_h^j, U_h^{j-1}).$$

Equation (2.16) can be obtained by taking the p-th power and applying (2.14) again. ■

Proposition 5 (Compactness) *Suppose now, that for every $h \in \mathbf{N}$, $(U_h^j)_{j\in\mathbf{N}}$ is a discrete solution and that all conditions* (2.3)–(2.8) *are satisfied. Then, there exists a subsequence $(h_n)_{n\in\mathbf{N}} \uparrow \infty$, and a curve $u \in AC^p_{\mathrm{loc}}([0,\infty); \mathscr{S})$, a non-increasing function $\varphi : [0,\infty) \to [-\infty, \infty]$, and a function $A \in L^p_{\mathrm{loc}}[0,\infty)$ such that*

$$\forall t \ge 0 : \bar{U}_{h_n}(t) \overset{\sigma}{\rightharpoonup} u(t), \quad \tilde{U}_{h_n} \overset{\sigma}{\rightharpoonup} u(t) \quad \text{as } n \to \infty, \tag{2.19}$$

$$\forall t \ge 0 : \lim_n \phi_{h_n}(\bar{U}_{h_n}(t)) = \varphi(t) \ge \phi(u(t)), \text{ and} \tag{2.20}$$

$$|U_{h_n}'| \rightharpoonup A \text{ in } L^p_{\mathrm{loc}}[0,\infty), \text{ and } A \ge |u'| \quad \mathscr{L}^1\text{−a.e.} \tag{2.21}$$

Proof Equation (2.14) shows that $|U_h'|$ is uniformly bounded and hence weakly compact in $L^p(0,T)$ for each $T > 0$, which guarantees the existence of a subsequence (h_n) such that $|U_{h_n}'|$ converges weakly in $L^p_{\mathrm{loc}}[0,\infty)$ to some function $A \in L^p_{\mathrm{loc}}[0,\infty)$. By Lemma 3, possibly upon extracing a further subsequence, we obtain a limit function φ, such that $\phi_{h_n}(\bar{U}_{h_n}(t)) \to \varphi(t)$ for every $t \ge 0$, and $\varphi(0) = \phi(u_0)$.

Condition (2.14) and (2.15) show that $d_h(\bar{U}_h(t), U_h^0)$ is uniformly bounded in h and t, for $0 \le t \le T$. By assumption (2.6) the ranges of the curves $\bar{U}_h$ are contained in a σ-compact set.

Fix $0 \le s < t$ and define

$$s(h) = \max\{t_h^j < s\} \quad \text{and} \quad t(h) = \min\{t_h^j > t\},$$

so that $s(h) \le s \le t \le t(h)$ and $\lim_{|\boldsymbol{\tau}_h|\to 0} s(h) = s$ and $\lim_{|\boldsymbol{\tau}_h|\to 0} t(h) = t$. We have

$$d_h(\bar{U}_h(s), \bar{U}_h(t)) \le d_h(U_h^{s(h)}, U_h^{s(h)+1}) + d_h(U_h^{s(h)+1}, U_h^{t(h)}) = \int_{s(h)}^{t(h)} |U_h'|(r)\,\mathrm{d}r,$$

and therefore, using the weak convergence of $|U'_{h_n}|$,

$$\limsup_{n\to\infty} d_{h_n}(\bar U_{h_n}(s),\bar U_{h_n}(t)) \le \int_s^t A(r)\,\mathrm{d}r.$$

This allows us to use Lemma 2 on bounded intervals $[0,T]$, so that, upon extracting a further subsequence, we may assume that there exists a limit-curve $u(t)$ such that $\bar U_{h_n}(t) \overset{\sigma}{\rightharpoonup} u(t)$ for every $t \ge 0$.

The liminf condition for the functionals (2.8) gives (2.20) and the liminf condition for the metric (2.8) gives the limit inequality

$$d(u(s),u(t)) \le \liminf_{n\to\infty} d_{h_n}(\bar U_{h_n}(s),\bar U_{h_n}(t)) \le \int_s^t A(r)\,\mathrm{d}r,$$

which shows that $u \in AC^p_{\mathrm{loc}}(0,\infty;\mathscr{S})$, and that $|u'| \le A$ a.e. in $(0,\infty)$. ■

Proposition 5 put us into the position to prove the convergence of discrete solutions to a curve of maximal slope. All we require is some information about the limit of G_{h_n}. A trivial lower bound for $\liminf_n G_{h_n}(t)$ is the function $g(u(t))$, defined in (2.9).

Proposition 6 *Suppose that for every $h \in \mathbf{N}$, $(U^j_h)_{j\in\mathbf{N}}$ is a discrete solution, as described in Section 2.1, with time-partition $\boldsymbol{\tau}_h$ and u is a limit curve, satisfying* (2.19) - (2.21). *Suppose also that* (2.3) *and* (2.4), *as well as* (2.7) *and* (2.8) *are satisfied, and that the functional g, defined in* (2.9) *is a strong upper gradient for ϕ.*

Then, u is a curve of maximal slope for ϕ with respect to g, $\phi\circ u$ is an element of $AC_{\mathrm{loc}}([0,\infty);\mathscr{S})$, the energy identity

$$\phi(u(t)) = \phi(u(0)) - \frac{1}{p}\int_0^t |u'|^p(t)\,\mathrm{d}t - \frac{1}{p'}\int_0^t g(u(t))^{p'}\,\mathrm{d}t \tag{2.22}$$

holds for all $t > 0$, and we have the following convergence properties for the discrete solution:

$$\begin{aligned}
\lim_{n\to\infty} \phi_{h_n}(\bar U_{h_n}(t)) &= \phi(u(t)) \quad \forall t \in [0,\infty), \\
\lim_{n\to\infty} |\partial_{h_n}\phi_{h_n}|(\bar U_{h_n}) &= g\circ u \quad \text{in } L^p_{\mathrm{loc}}[0,\infty), \text{ and} \\
\lim_{n\to\infty} |U'_{h_n}| &= |u'| \quad \text{in } L^p_{\mathrm{loc}}[0,\infty).
\end{aligned} \tag{2.23}$$

Proof Using (2.21), (2.9), and (2.20), we obtain for any $0 < t$,

$$\begin{aligned}
& \frac{1}{p}\int_0^t |u'|^p(r)\,\mathrm{d}r + \frac{1}{p'}\int_0^t g(u(r))^{p'}\,\mathrm{d}r + \phi(u(t)) \qquad (2.24)\\
\leq\ & \frac{1}{p}\int_0^t |A(r)|^p\,\mathrm{d}r + \frac{1}{p'}\int_0^t \liminf_n |\partial_{h_n}\phi_{h_n}|^{p'}(\tilde{U}_{h_n}(r))\,\mathrm{d}r + \lim_n \phi_{h_n}(\bar{U}_{h_n}(t))\\
\leq\ & \liminf_n \frac{1}{p}\int_0^t |U'_{h_n}|^p(r)\,\mathrm{d}r + \frac{1}{p'}\int_0^t \liminf_n |G_{h_n}(r)|^p\,\mathrm{d}r + \lim_n \phi_{h_n}(\bar{U}_{h_n}(t))\\
\leq\ & \limsup_n \frac{1}{p}\int_0^t |U'_{h_n}|^p(r)\,\mathrm{d}r + \liminf_n \frac{1}{p'}\int_0^t |G_{h_n}(r)|^p\,\mathrm{d}r + \lim_n \phi_{h_n}(\bar{U}_{h_n}(t))\\
\leq\ & \limsup_n \left(\frac{1}{p}\int_0^t |U'_{h_n}|^p(r)\,\mathrm{d}r + \frac{1}{p'}\int_0^t |G_{h_n}(r)|^p\,\mathrm{d}r + \phi_{h_n}(\bar{U}_{h_n}(t))\right)\\
=\ & \limsup_n \phi_{h_n}(\bar{U}_{h_n}(0)) \ =\ \varphi(0) \ =\ \phi(u_0).
\end{aligned}$$

Since g is a strong upper gradient for ϕ, we also have

$$\begin{aligned}
\phi(u(0)) - \phi(u(t)) &\leq \int_0^t g(u(r))|u'|(r)\,\mathrm{d}r \qquad (2.25)\\
&\leq \int_0^t \frac{1}{p}|u'|^p(r) + \frac{1}{p'}|g(u(r))|^{p'}\,\mathrm{d}r,
\end{aligned}$$

and therefore all of the estimates stated in (2.24) must be equalities. First, we immediately obtain line one in (2.23). Second, we find that the norms $\| |U'_{h_n}| \|_{L^p(0,t)}$ converge and hence $|U'_{h_n}|$ converges strongly in $L^p_{\mathrm{loc}}[0,\infty)$ to A, and furthermore that $A = |u'|$. To see that $g \circ u = \lim |\partial_{h_n}\phi_{h_n}(\tilde{U}_{h_n})|$, note that due to (2.14), the family $|\partial_{h_n}\phi_{h_n}|(\tilde{U}_{h_n})$ is bounded in $L^{p'}(0,\infty)$. Hence there exists a subsequence (h_{n_k}) and a function $G \in L^{p'}(0,\infty)$, so that $|\partial_{h_{n_k}}\phi_{h_{n_k}}|(\tilde{U}_{h_{n_k}}) \rightharpoonup G$, weakly in $L^{p'}(0,\infty)$. As above, we may deduce that $g \circ u = G$. This shows that the limit does not depend on the subsequence chosen and hence the entire sequence $|\partial_{h_n}\phi_{h_n}|(\tilde{U}_{h_n})$ converges weakly to $g \circ u$. To obtain strong convergence in $L^{p'}(0,t)$ and equation (2.23), we may use a similar argument as for the convergence of $|U'_{h_n}|$.

The energy identity (2.22) as well as the fact that $\phi \circ u \in AC_{\mathrm{loc}}[0,\infty)$ follow from

$$\begin{aligned}
\phi(u(t)) - \phi(u(0)) &= \int_0^t g(u(r))|u'|(r)\,\mathrm{d}r\\
&= \frac{1}{p}\int_0^t |u'|^p(r)\,\mathrm{d}r + \frac{1}{p'}\int_0^t |g(u(r))|^{p'}\,\mathrm{d}r
\end{aligned}$$

from which we may also deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)) = -\frac{1}{p}|u'|^p(t) - \frac{1}{p'}g(u(t))^{p'} \text{ for } \mathscr{L}^1\text{−a.e. } t \in (0,\infty).$$

■

Remark 7 *The convergence result contained in Proposition 6 combined with Proposition 5 merely states that in the family of approximations there exists one subsequence converging to a solution. This might seem a very weak result at first glance. However, note first that if only one solution exists then every convergent subsequence must converge to this solution and it follows easily that the entire family of approximations converges.*

In practical computations, we do not expect this to be a problem at all. First, we may not even refine the approximation, and even if we did, we would typically use the information previously obtained and thus compute a new discrete solution which lies close to the one we had obtained previously. Unless two exact solutions lie close together, it is therefore unlikely that we would obtain two subsequences converging to different solutions.

2.5 The λ-convex case

Equation (2.10) can be verified in the case when ϕ is convex and weak convergence is equivalent to convergence in the metric. The following Proposition is intended only as an example, as for many applications, the conditions are too strong.

Proposition 8 *Assume that $\mathscr{S}$ is a Banach space and that d is the metric induced by its norm. Suppose that ϕ and ϕ_h are convex and respectively d and d_h-lower semicontinuous and that ϕ_h $\Gamma(\sigma)$-converges to ϕ, i.e., that in addition to* (2.8)*, we have*

$$\forall\, v \in \mathscr{S}\ \exists\, (v_h)_{h\in\mathbf{N}} \in \mathscr{S} \text{ s.t. } v_h \xrightarrow{\sigma} v \text{ and } \phi_h(v_h) \to \phi(v). \tag{2.26}$$

Furthermore, suppose that the continuity condition

$$u_h \xrightarrow{\sigma} u \text{ and } v_h \xrightarrow{\sigma} v \quad \Rightarrow \quad d_h(u_h, v_h) \to d(u, v). \tag{2.27}$$

holds. Then, the slopes satisfy the liminf condition (2.10) *and $|\partial\phi|$ is an upper gradient for ϕ.*

Proof Note first, that if $\mathscr{S}$ is a Banach space and ϕ is convex and lower semicontinous, $|\partial\phi|$ satisfies

$$|\partial\phi|(u) = \sup_{v\neq u} \frac{(\phi(u) - \phi(v))^+}{d(u, v)}. \tag{2.28}$$

and based on this equation, one can show that $|\partial\phi|$ is an upper gradient for ϕ. For a proof see Theorem 1.2.5 and Proposition 1.4.4 in [AGS05]. Let v_j be a sequence satisfying

$$|\partial\phi|(u) = \lim_{j\to\infty} \frac{(\phi(u) - \phi(v_j))^+}{d(u, v_j)},$$

and for every $j \in \mathbf{N}$, let $(w_{j,h})_{h\in\mathbf{N}}$ be the recovery sequence for v_j, described by (2.26). Let $u_h \xrightarrow{\sigma} u$. With these preparations, we can now estimate

$$\begin{aligned} |\partial\phi|(u) &= \lim_j \frac{(\phi(u)-\phi(v_j))^+}{d(u,v_j)} \\ &\le \lim_j \frac{(\liminf_h \phi_h(u_h) - \lim_h \phi(w_{j,h}))^+}{\lim_h d_h(u_h,w_{j,h})} \\ &= \lim_j \liminf_h \frac{(\phi_h(u_h)-\phi(w_{j,h}))^+}{d_h(u_h,w_{j,h})} \\ &\le \liminf_h \limsup_j \frac{(\phi_h(u_h)-\phi(w_{j,h}))^+}{d_h(u_h,w_{j,h})} \\ &\le \liminf_h |\partial_h\phi_h|(u_h). \end{aligned} \tag{2.29}$$

■

We move on to a generalization of the above result, treating functionals whose non-convexity is controllable along certain curves in the metric space.

Assumption 9 *For a given $\lambda \in \mathbf{R}$, $p \in (1,\infty)$, we assume that for any $v_0, v_1 \in D(\phi)$ there exists a curve $(v_t)_{t\in(0,1)}$ such that, for all $0 < \tau < 1/\lambda^-$, $\lambda^- = \max(0,-\lambda)$,*

$$\Phi(\tau, v_0; v_t) \le (1-t)\Phi(\tau, v_0; v_0) + t\Phi(\tau, v_0, v_1) - \frac{1}{p}\Big(\lambda + \frac{1}{\tau^{p-1}}\Big)t(1-t^{p-1})d^p(v_0,v_1), \tag{2.30}$$

where $\Phi(\tau,u;v) = d^p(u,v)/(p\tau^{p-1}) + \phi(v)$.

Remark 10 *Note that the curve v_t above is chosen independently of τ. This is used in the proof of Lemma 12 below.*

Remark 11 *An archetypical case for Assumption 9 is when the metric space $\mathscr{S}$ is geodesically complete, i.e., for any two points $v_0, v_1 \in \mathscr{S}$ there exists a curve v_t (a constant-speed geodesic) such that*

$$d(v_0, v_t) = td(v_0, v_1).$$

In this case, if a function ϕ satisfies

$$\phi(v_t) \le (1-t)\phi(v_0) + t\phi(v_1) - \frac{1}{p}\lambda t(1-t^{p-1})d^p(v_0,v_1) \tag{2.31}$$

along every geodesic $t \mapsto v_t$, then (2.30) is satisfied. The surprising result in [AGS05] was that much more general curves may be used if (2.30) is assumed directly. This procedure has, for example, applications to gradient flows with respect to the Wasserstein metric of probability measures (c.f. Part II in [AGS05]).

Lemma 12 *If Assumption 9 is satisfied, then the local slope admits the representation*

$$|\partial\phi|(v) = \sup_{w\neq v}\left(\frac{\phi(v)-\phi(w)}{d(v,w)} + \frac{1}{p}\lambda d^{p-1}(v,w)\right)^+. \tag{2.32}$$

If, in addition, ϕ is d-lower semicontinuous then $|\partial\phi|$ is a strong upper gradient for ϕ and is d-lower semicontinuous.

Proof The proof for $p = 2$ is given in Theorem 2.4.9 and Corollary 2.4.10 [AGS05]. The proof we give here is a simple extension to the case $p \neq 2$.

Note first that we can rewrite (2.30) as

$$\begin{aligned}
&\frac{d^p(v_0,v_t)}{p\tau^{p-1}} + \phi(v_t)\\
&\quad\leq\quad (1-t)\phi(v_0) + t\phi(v_1) + \frac{td^p(v_0,v_1)}{p\tau^{p-1}} - \frac{1}{p}\big(\tau^{1-p}+\lambda\big)t(1-t^{p-1})d^p(v_0,v_1)\\
&\quad=\quad (1-t)\phi(v_0) + t\phi(v_1) + \frac{td^p(v_0,v_1)}{p\tau^{p-1}}\big(t^{p-1} - \lambda\tau^{p-1}(1-t^{p-1})\big).
\end{aligned} \tag{2.33}$$

Multiplying (2.33) by τ^{p-1} passing to the limit as $\tau \to 0+$ (note that the v_t are independent of τ), we obtain

$$d(v_0,v_t) \leq td(v_0,v_1). \tag{2.34}$$

Neglecting the first term on the right-hand side of (2.33), we get

$$\phi(v_0) - \phi(v_t) \geq t\Big\{\phi(v_0) - \phi(v_1) + \frac{d^p(v_0,v_1)}{p\tau^{p-1}}\big(\lambda\tau^{p-1}(1-t^{p-1}) - t^{p-1}\big)\Big\}. \tag{2.35}$$

We divide (2.35) by $d(v,v_t)$ and use (2.34) to obtain

$$\begin{aligned}
\frac{\phi(v_0)-\phi(v_t)}{d(v_0,v_t)} &\geq \frac{td(v_0,v_1)}{d(v_0,v_t)}\Big\{\frac{\phi(v_0)-\phi(v_1)}{d(v_0,v_1)} + \frac{d^{p-1}(v_0,v_1)}{p\tau^{p-1}}\big(\lambda\tau^{p-1}(1-t^{p-1}) - t^{p-1}\big)\Big\}\\
&\geq \frac{\phi(v_0)-\phi(v_1)}{d(v_0,v_1)} + \frac{d^{p-1}(v_0,v_1)}{p\tau^{p-1}}\big(\lambda\tau^{p-1}(1-t^{p-1}) - t^{p-1}\big).
\end{aligned} \tag{2.36}$$

Fix $v, w \in \mathscr{S}$ and choose $v_0 = v$ and $v_1 = w$ in (2.36). Using (2.34), we have $v_t \to v$ as $t \to 0$ and hence, using the fact that $s \mapsto (s)^+$ is monotone and continuous,

$$\begin{aligned}
|\partial\phi(v)| &\geq \left(\limsup_{t\to 0}\frac{\phi(v)-\phi(v_t)}{d(v,v_t)}\right)^+\\
&\geq \left(\frac{\phi(v)-\phi(w)}{d(v,w)} + \frac{\lambda}{p}d^{p-1}(v,w)\right)^+,
\end{aligned}$$

which shows that the left-hand side is greater or equal the right-hand side in (2.32). The reverse inequality is trivial; hence we have shown (2.32).

To conclude, we simply note that the proof of Corollary 2.4.10 in [AGS05] does not depend on the specific formula for $|\partial\phi|$ but only on the structure. It can be repeated

word for word to show that $|\partial\phi|$ is lower semicontinuous and that it is a strong upper gradient for ϕ. ■

As we have obtained a formula for $|\partial\phi|$ which is quite similar to (2.28), it is natural to consider an analogous argument to the proof of Proposition 8 for λ-convex functionals.

Proposition 13 *Let $p \in (1,\infty)$. Suppose that Assumption 9 is satisfied for Φ as well as all Φ_h with the same λ. Suppose also that* (2.27) *holds and that ϕ_h $\Gamma(d)$-converges to ϕ (i.e.* (2.26) *and* (2.8) *are satisfied). Then, the liminf condition for the slope* (2.10) *holds.*

Proof We argue as in the proof of Proposition 8. Let v_j be a sequence satisfying

$$|\partial\phi|(u) = \lim_{j\to\infty} \left(\frac{\phi(u)-\phi(v_j)}{d(u,v_j)} + \frac{\lambda}{p} d^{p-1}(u,v_j) \right)^+ ,$$

and for every $j \in \mathbf{N}$, let $(w_{j,h})_{h\in\mathbf{N}}$ be the recovery sequences for v_j described in (2.26). Using the liminf condition (2.8) and Lemma 12 it is again easy to deduce that for $u_h \xrightarrow{\sigma} u$,

$$\begin{aligned}
|\partial\phi|(u) &= \lim_{j\to\infty} \left(\frac{\phi(u)-\phi(v_j)}{d(u,v_j)} + \frac{\lambda}{p} d^{p-1}(u,v_j) \right)^+ \\
&\le \lim_{j\to\infty} \left(\frac{\liminf_h \phi_h(u_h) - \lim_h \phi_h(w_{j,h})}{\lim_h d_h^{p-1}(u_h,w_{j,h})} + \frac{\lambda}{p} \lim_h d_h(u_h,w_{j,h}) \right)^+ \\
&= \lim_{j\to\infty} \liminf_{h\to\infty} \left(\frac{\phi_h(u_h)-\phi_h(w_{j,h})}{d_h(u_h,w_{j,h})} + \frac{\lambda}{p} d_h^{p-1}(u_h,w_{j,h}) \right)^+ \\
&\le \liminf_{h\to\infty} \sup_{j\in\mathbf{N}} \left(\frac{\phi_h(u_h)-\phi_h(w_{j,h})}{d_h(u_h,w_{j,h})} + \frac{\lambda}{p} d_h^{p-1}(u_h,w_{j,h}) \right)^+ \\
&\le \liminf_{h\to\infty} |\partial_h\phi_h|(u_h).
\end{aligned}$$

■

Remark 14 *1. A possible avenue for obtaining the liminf condition for the slopes if the functionals are not uniformly λ-convex, is to construct a more specific recovery sequence. If it were possible to strengthen the limsup condition* (2.26)*, i.e. to find $w_{h,j}$ satisfying also $\lim_{j\to\infty} d_h(w_{h,j},v_h) = 0$, then the proof of Proposition 13 applies even there is no representation for the slopes as the one obtained in Lemma 12.*

2. The case where the approach taken in this section fails completely is when σ is much weaker than d. This situation arises, for example, if the compactness in Condition (2.6) *is induced by the metric rather than the functional.*

Theorem 15 *Suppose that Assumption 9 holds for ϕ and all ϕ_h, and with the same λ, that equation* (2.27) *is satisfied and that ϕ_h $\Gamma(\sigma)$-converges to ϕ. Suppose*

also that (2.3) – (2.6) *are satisfied. If, for every* $h \in \mathbf{N}$, $(U_h^j)_{j\in\mathbf{N}}$ *is a discrete solution as described in Section 2.1, then there exists a subsequence* $(h_k)_{k\in\mathbf{N}}$ *of* $\mathbf{N}$ *such that* $(U_{h_k}^j)_{j\in\mathbf{N}}$ *converges in the sense of Proposition 5 to a curve of maximal slope for* ϕ *with respect to the strong upper gradient* $|\partial\phi|$*, with initial value* u_0*.*

Proof Simply combine Lemma 12 and Propositions 13 with Propositions 5 and 6. ■

3 Variational Inequalities

The condition of λ-convexity which we used in the previous section to obtain concrete convergence results can be tightened and used favourably to obtain much stronger convergence results based on variational inequalities. In the setting of this section we can only consider 2-curves of maximal slope. We shall call them simply curves of maximal slope for the remainder of the paper.

3.1 Relaxed convexity and coercivity conditions

We review the definition of λ-convexity. For $u, v \in \mathscr{S}, \tau > 0$, let

$$\Phi(\tau, u; v) = \frac{d^2(u, v)}{2\tau} + \phi(v).$$

Compare this with equation (1.2) to see the relationship to the time-discretization. Suppose for the moment that $(\mathscr{S}, d)$ is a Hilbert space. We say that ϕ is λ-convex, for some $\lambda \in \mathbf{R}$, if

$$\phi((1-t)v_0 + tv_1) \leq (1-t)\phi(v_0) + t\phi(v_1) - \frac{\lambda}{2}t(1-t)\|v_0 - v_1\|^2 \quad \forall t \in [0,1]. \tag{3.1}$$

In this case, it is easy to show that for $0 < \tau < 1/\lambda^-$, where we define

$$\lambda^- = \max(0, -\lambda),$$

we have

$$\Phi(\tau, u; (1-t)v_0 + tv_1) \leq (1-t)\Phi(\tau, u; v_0) + t\Phi(\tau, u; v_1) - \frac{1}{2}\big(\tau^{-1} + \lambda\big)t(1-t)\|v_0 - v_1\|^2.$$

Note that for sufficiently small τ, the 'time-step' (1.2) becomes a uniformly convex minimization problem, which is considerably easier to solve both theoretically and numerically.

In the general metric case, the following condition is required. We say that $\Phi(\tau, u; \cdot)$ is $(\tau^{-1} + \lambda)$-convex if for any two points $v_0, v_1 \in \mathscr{S}$, and $t \in (0,1)$, there exist $v_t \in \mathscr{S}$ such that, for all $0 < \tau < 1/\lambda^-$, we have

$$\Phi(\tau, u; v_t) \leq (1-t)\Phi(\tau, u; v_0) + t\Phi(\tau, u; v_1) - \frac{1}{2}\big(\tau^{-1} + \lambda\big)t(1-t)d^2(v_0, v_1), \tag{3.2}$$

for all $t \in [0,1]$. Condition (2.30) is considerably more general than Condition (3.2), as there, the convexity condition needed to be satisfied only for the base point v_0.

To obtain a better feel for the meaning of λ-convexity, consider the following simple proposition.

Proposition 16 *Let $\mathscr{S}$ be a closed, convex subset of a Hilbert space $\mathscr{H}$ with norm $\|\cdot\|$ and $\phi : \mathscr{H} \to (-\infty, +\infty]$. Then ϕ is λ-convex in $\mathscr{S}$ if and only if $u \mapsto \phi(u) - \frac{\lambda}{2}\|u\|^2$ is convex in $\mathscr{S}$.*

In particular, if ϕ is differentiable in $\mathscr{S}$ then ϕ is λ-convex if, and only if,

$$\big(\phi'(v_1) - \phi'(v_0), v_1 - v_0\big) \geq \lambda\|v_1 - v_0\|^2 \quad (\forall\, v_1, v_0 \in \mathscr{S}) \tag{3.3}$$

holds. If ϕ is twice differentiable in $\mathscr{S}$ and

$$\phi''(u; v, v) \geq \lambda\|v\|^2 \quad (\forall v \in \mathscr{H}) \tag{3.4}$$

holds for all points u which are not extremal in $\mathscr{S}$ then ϕ is λ-convex.

If $\phi = \phi_1 + \phi_2$, where $\phi_i : \mathscr{S} \to (-\infty, +\infty]$, ϕ_1 is convex and ϕ_2 is λ-convex, then ϕ is λ-convex.

Proof Throughout the proof, let $v_0, v_1 \in \mathscr{S}$, $v_t = (1-t)v_0 + tv_1$ and $F(v) = \phi(v) - \lambda/2\|v\|^2$.

The crucial observation is that $u \mapsto \frac{1}{2}\|u\|^2$ is 1-convex, in fact, we even have

$$\frac{1}{2}\|v_t\|^2 = (1-t)\frac{1}{2}\|v_0\|^2 + t\frac{1}{2}\|v_1\|^2 - \frac{1}{2}t(1-t)\|v_0 - v_1\|^2. \tag{3.5}$$

Suppose now that ϕ is λ-convex. Using (3.5), we have

$$\begin{aligned} F(v_t) &\leq (1-t)\phi(v_0) + t\phi(v_1) - \frac{\lambda}{2}t(1-t)\|v_0 - v_1\|^2 \\ &\quad -(1-t)\frac{\lambda}{2}\|v_0\|^2 - t\frac{\lambda}{2}\|v_1\|^2 + \frac{\lambda}{2}t(1-t)\|v_0 - v_1\|^2 \\ &= (1-t)F(v_0) + tF(v_1). \end{aligned}$$

On the other hand, if F is convex we may traverse the above inequality in the opposite direction, to obtain that ϕ is λ-convex.

The derivative of $\|v\|^2/2$ is $(v, \cdot)$, its second derivative is $(\cdot, \cdot)$. If ϕ satisfies (3.3) then

$$\begin{aligned} \big(F'(v_1) - F'(v_0), v_1 - v_0\big) &= \big(\phi'(v_1) - \phi'(v_0), v_1 - v_0\big) - \lambda\|v_1 - v_0\|^2 \\ &\geq 0, \end{aligned}$$

which is equivalent to F being convex. If ϕ is twice differentiable and satisfies (3.4) then

$$F''(u; v, v) = \phi''(u; v, v) - \lambda\|v\|^2 \geq 0,$$

for all points u which are not extremal and hence F is convex. The last statement of the proposition is trivial. ■

If Φ satisfies (3.2) then typical coercivity conditions can be generalized as well. Equation (2.4.10) in [AGS05] is a very general condition, if constant speed geodesics are available. Here, we formulate a condition which reduces the technicality of

some proofs but seems to be sufficiently general for most applications. We assume that

$$\exists \tau^* > 0, u^* \in D(\phi) \text{ s.t. } \inf_{\mathscr{S}} \Phi(\tau^*, u^*; \cdot) \geq m^* > -\infty. \tag{3.6}$$

Of course, $\inf_{\mathscr{S}} \phi > -\infty$ would be sufficient, but this condition may be too restrictive for some applications. If ϕ satisfies (3.6) we say that ϕ is coercive. It can be easily checked that (3.6) implies equation (2.4.10) in [AGS05].

A family $(\phi_i)_{i\in I}$ is called equi-coercive if (3.6) is simultaneously satisfied for all ϕ_i, with the same choice of τ^*, u^* and m^*. A typical case of equi-coercivity is covered in the following proposition.

Proposition 17 *Let I be an index set and for every $i \in I$ let $\phi_i : \mathscr{S} \to (-\infty, \infty]$. Suppose that for every $i \in I$, ϕ_i satisfies (3.2) for the same $\lambda \in \mathbf{R}$. Suppose also that there exists $u^* \in \mathscr{S}$ and $R > 0$ such that, for every $i \in I$, there exists $u_i^* \in B(u^*, R/2)$ so that*

$$\inf_{i\in I} \inf_{B(u^*,R)} \phi_i = c_0 > -\infty \quad \text{and} \quad \sup_{i\in I} \phi_i(u_i^*) = c_1 < \infty.$$

Then the family $(\phi_i)_{i\in I}$ is equicoercive.

Proof We assume without loss of generality that $\lambda < 0$.

For $v \in B(u^*, R)$, we trivially have $\Phi_i(\tau, u_i^*; v) \geq c_0$ for all $i \in I$ and all τ.

For $v \in \mathscr{S} \setminus B(u^*, R)$, let $\tau = 1/(2\lambda^-)$, $u = u_i^*$, $v_0 = u_i^*$ and $v_1 = v$ in (3.2) to obtain

$$\Phi_i(\tau, u_i^*; v_t) \leq (1-t)\Phi_i(\tau, u_i^*; u_i^*) + t\Phi_i(\tau, u_i^*; v) + \frac{\lambda}{2} t(1-t) d^2(u_i^*, v).$$

Set $t = R/(4d(u_i^*, v))$ and note that in the proof of 12 we had obtained that $d(v_0, v_t) \leq td(v_0, v_1)$, so that $v_t \in B(u^*, R)$. Hence, we get

$$\begin{aligned}
\Phi_i(\tau, u_i^*; v) &\geq \frac{4d(u_i^*, v)}{R} - \Phi_i(\tau, u_i^*; v_t) - \frac{4d(u_i^*, v) - R}{R}\phi_i(u_i^*) \\
&\quad -\frac{\lambda}{2}\left(1 - \frac{R}{4d(u_i^*, v)}\right) d^2(u_i^*, v) \\
&\geq -\frac{\lambda}{4} d^2(u_i^*, v) + \frac{c_0 - c_1}{R} d(u_i^*, v) + c_0
\end{aligned}$$

The last term in the above estimate is a quadratic polynomial in $d(u_i^*, v)$ with a positive leading order term and coefficients independent of i, and is hence bounded below by a constant independent of i, C_0, say. Using

$$\frac{d^2(u_i^*, v)}{2\tau} \leq \frac{d^2(u^*, v)}{\tau} + \frac{d^2(u_i^*, u^*)}{\tau} \leq \frac{d^2(u^*, v)}{\tau} + \frac{R^2}{4\tau},$$

we also have

$$\Phi_i(\tau, u^*; v) \geq \frac{1}{2}\Phi_i(\tau, u_i^*; v) - \frac{R^2}{2\tau} \geq \frac{1}{2} C_0 - \frac{R^2}{2\tau},$$

which shows that the family $(\phi_i)_{i\in I}$ is equicoercive. ∎

3.2 Evolutionary variational inequalities

Here, we summarize several results from [AGS05] which exploit the convexity conditions of the previous section. In particular, we present evolutionary variational inequalities that are satisfied by curves of maximal slope and their time-discretizations, as well as some approximation results.

Theorem 18 (Existence and uniqueness) *Let $(\mathscr{S}, d)$ be a complete metric space, ϕ a coercive and lower semicontinous functional on $\mathscr{S}$, satisfying* (3.2).

(a) For each $u_0 \in \overline{D(\phi)}$, there exists a unique curve of maximal slope u for ϕ.

(b) The curve u is locally Lipschitz continuous in $(0,\infty)$ with $u(t) \in D(|\partial\phi|) \subset D(\phi)$ for all $t > 0$ and it is the unique solution of the evolutionary variational inequality

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u,v) + \frac{\lambda}{2}d^2(u,v) + \phi(u) \le \phi(v) \quad \forall\, v \in \mathscr{S},\ \text{a.e. } t \in (0,\infty) \tag{3.7}$$

among all locally absolutely continuous curves v satisfying $v(0+) = u_0$.

Theorem 19 (Time-discretization) *Let $(\mathscr{S}, d)$ be a complete metric space, ϕ a coercive, lower semicontinous functional on $\mathscr{S}$, satisfying* (3.2).

(a) Let $\boldsymbol{\tau} = (0 = t^0, t^1, t^2, \dots)$, $\tau^j = t^j - t^{j-1}$, be a partition of $[0,\infty)$ with $|\boldsymbol{\tau}| = \sup_j \tau^j \le 1/\lambda^-$. Then, for every $U^0_{\boldsymbol{\tau}} \in \overline{D(\phi)}$, there exists a unique solution of the time-discrete scheme

$$\text{For } j = 1, 2, \dots \text{ find } U^j_{\boldsymbol{\tau}} \in \operatorname{argmin} \Phi(\tau^j, U^{j-1}_{\boldsymbol{\tau}}; \cdot). \tag{3.8}$$

(b) $(U^j_{\boldsymbol{\tau}})_{j=0,1,\dots}$ is the unique solution of the discrete evolutionary variational inequality

$$\frac{1}{2\tau^j}\Big(d^2(U^j_{\boldsymbol{\tau}}, v) - d^2(U^{j-1}_{\boldsymbol{\tau}}, v)\Big) + \frac{\lambda}{2}d^2(U^j_{\boldsymbol{\tau}}, v) + \phi(U^j_{\boldsymbol{\tau}}) \le \phi(v) \qquad \forall\, v \in \mathscr{S}, j = 0, 1, \dots \quad . \tag{3.9}$$

(c) Let $\bar{U}_{\boldsymbol{\tau}}(t) = U^j_{\boldsymbol{\tau}}$ if $t^{j-1} < t \le t^j$. If $\phi(U^0_{\boldsymbol{\tau}})$ is bounded and $U^0_{\boldsymbol{\tau}} \to u_0$ as $|\boldsymbol{\tau}| \to 0$ then $\bar{U}_{\boldsymbol{\tau}} \to u$ in $L^\infty_{\mathrm{loc}}([0,\infty); \mathscr{S})$ as $|\boldsymbol{\tau}| \to 0$.

For explicit error estimates for the time-discretization see especially Theorems 4.0.7, 4.0.9 and 4.0.10 in [AGS05]. To see, at least heuristically, where (3.7) comes from, see Proposition 20 below.

It turns out that the convergence analysis of Sections 3.3 and 3.4 is not restricted to gradient flows of functionals which are constant in time, but may be

applied to gradient flows of time-dependent functionals, i.e., to equations of the form

$$\dot{u}(t) = -\phi'(t, u(t)), \tag{3.10}$$

where ϕ' denotes the Fréchet derivative with respect to u. By setting $v = (v_1, v_2) \in \mathbf{R} \times \mathscr{S}$ and $A(v) = (1, -\phi'(v_1, v_2))$, (3.10) can be transformed into the autonomous nonlinear evolution equation $\dot{v} = A(v)$, for which an extensive theory exists; see in particular the recent work of Nochetto and Savaré [NS]. Note that condition (IS_3) in [NS] corresponds exactly to equation (3.3) in the present work.

For gradient flows in the metric setting, no such theory exists to the author's knowledge. It is conceivable, however, that the techniques of Chapter 4 in [AGS05] can be modified to give an existence and approximation theory for time-dependent gradient flows, if $\phi(t, u)$ is sufficiently regular in t. In particular, note that the time-dependent version of definition (1.6), for $p = 2$, is

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(t, u(t)) = -\frac{1}{2}|u'|^2(t) - \frac{1}{2}|\partial\phi|(t, u(t)) + \partial_t\phi(t, u(t)), \tag{3.11}$$

where

$$|\partial\phi|(t, u) = |\partial\phi(t, \cdot)|(u) \quad \text{and} \quad \partial_t\phi(t, u) = \lim_{s \to t} \frac{\phi(s, u) - \phi(t, u)}{s - t}.$$

We postpone this analysis to a later time and assume for now that such a solution does exist. The next two propositions show that the basis of our analysis, the evolutionary variational inequalities, do not change if we consider time-dependent functionals.

Proposition 20 *(a) Suppose that $\mathscr{H}$ is a Hilbert space, $\phi : (a, b) \times \mathscr{H} \to \mathbf{R}$ is differentiable in u, and that the curve $u \in AC(a, b; \mathscr{H})$ satisfies* (3.10)*. If $\phi(t, \cdot)$ is λ-convex in $\mathscr{H}$ for every $t \in (a, b)$, then u satisfies*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t) - v\|^2 + \frac{\lambda}{2}\|u(t) - v\|^2 + \phi(t, u(t)) \le \phi(t, v) \quad \forall\, v \in \mathscr{H},$$

for a.e. $t \in (a, b)$.

(b) Let $(\mathscr{S}, d)$ be a metric space, $\phi : (a, b) \times \mathscr{S} \to (-\infty, \infty]$ satisfy (3.2)*, and let $u_0 \in \mathscr{S}$. Then there exists at most one absolutely continuous curve u with $u(0) = u_0$ satisfying*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t), v) + \frac{\lambda}{2}d^2(u(t), v) + \phi(t, u(t)) \le \phi(t, v) \quad \forall\, v \in \mathscr{S}, \text{ a.e. } t \in (0, T). \tag{3.12}$$

Proof (a) Fix $v \in \mathscr{H}$. Take the inner product of $\dot{u} + \phi'(t,u)$ with $u - v$, to obtain

$$(\dot{u}, u - v) + (\phi'(t,u), u - v) = 0. \tag{3.13}$$

For $s \in [0,1]$, set $v_s = (1-s)u + sv = u + s(v-u)$. By a simple computation, we have

$$\phi(t,u) - \phi(t,v) - \int_0^1 (\phi'(v_s), u - v)\,\mathrm{d}s = 0. \tag{3.14}$$

Adding (3.14) to (3.13) gives

$$(\dot{u}, u - v) + \int_0^1 (\phi'(u) - \phi'(v_s), u - v)\,\mathrm{d}s + \phi(t,u) = \phi(t,v). \tag{3.15}$$

First note that $\frac{\mathrm{d}}{\mathrm{d}t}\|u - v\|^2 = 2(\dot{u}, u - v)$, and second, that $s(u - v) = u - v_s$. We use these facts, together with (3.3) to obtain from (3.15),

$$\begin{aligned}
\phi(t,v) &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u - v\|^2 + \int_0^1 \frac{1}{s}(\phi'(u) - \phi'(v_s), u - v_s)\,\mathrm{d}s + \phi(t,u) \\
&\geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u - v\|^2 + \lambda \int_0^1 \frac{1}{s}\|u - v_s\|^2\,\mathrm{d}s + \phi(t,u) \\
&= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u - v\|^2 + \lambda\|u - v\|^2 \int_0^1 s\,\mathrm{d}s + \phi(t,u) \\
&= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u - v\|^2 + \frac{\lambda}{2}\|u - v\|^2 + \phi(t,u),
\end{aligned}$$

which shows part (a).

(b) The uniqueness follows directly from the error estimate (3.26), by taking $\phi_h = \phi$ and $v_h = u$. ■

The time-dependent implicit time-discretization of (3.10) is simply obtained by replacing $\phi(u)$ by $\phi(t,u)$ in (3.2). Let $\Phi(\tau, u; t, v) = d^2(u,v)/(2\tau) + \phi(t,v)$. We say that $(U^j_{\boldsymbol{\tau}})_{j=0,1,\ldots}$ is a discrete solution of the gradient flow with respect to the partition $\boldsymbol{\tau}$, if

$$\forall j = 1, 2, \ldots : U^j_{\boldsymbol{\tau}} \text{ minimizes } \Phi(\tau^j, U^{j-1}_{\boldsymbol{\tau}}; t^j, \cdot). \tag{3.16}$$

Proposition 21 *Let $(\mathscr{S}, d)$ be a complete metric space, and for every $t > 0$, let $\phi(t,\cdot)$ be d-lower semicontinuous and satisfy* (3.2) *for the same $\lambda \in \mathbf{R}$. If $|\boldsymbol{\tau}| < 1/\lambda^-$, then for a given initial value $U^0_{\boldsymbol{\tau}}$,* (3.16) *has a unique solution $(U^j_{\boldsymbol{\tau}})$, which is also the unique solution of*

$$\frac{1}{2\tau^j}\Big(d^2(U^j_{\boldsymbol{\tau}}, v) - d^2(U^{j-1}_{\boldsymbol{\tau}}, v)\Big) + \frac{\lambda}{2}d^2(U^j_{\boldsymbol{\tau}}, v) + \phi(t^j_{\boldsymbol{\tau}}, U^j_{\boldsymbol{\tau}}) \leq \phi(t^j_{\boldsymbol{\tau}}, v) \qquad \forall\, v \in \mathscr{S}, j = 0, 1, \ldots \quad , \tag{3.17}$$

given the same initial value $U^0_{\boldsymbol{\tau}}$.

Proof For every time-step, the problem can be reduced to (3.8), hence Theorem 19 applies except (possibly) for the convergence result of part (c). ■

3.3 Continuous case

In this section, we outline a general approach for analysing approximations of curves of maximal slope for a time-dependent functional $\phi : [0,T] \times \mathscr{S} \to \mathbf{R}$, by replacing it with a suitable approximation ϕ_h. Based on the relaxed convexity and coercivity conditions of Section 3.1 and the variational inequalities of Section 3.2, we are able to treat non-convex and non-differentiable functionals. The functional approximations are sufficiently general to allow for a wide variety of applications. They are not, for example, restricted to Galerkin or similar discretizations.

We summarize the basic assumptions for the approximation analysis in the following hypothesis:

Assumption 22 *(a)* Completeness: *$(\mathscr{S}, d)$ is a complete metric space.*

(b) Lower semicontinuity: *For every $0 < t \le T$ and $h \in \mathbf{N}$, $\phi(t,\cdot)$ and $\phi_h(t,\cdot)$ are d-lower semicontinuous functionals on $\mathscr{S}$.*

(c) Uniform λ-convexity: *There exists $\lambda \in \mathbf{R}$ such that* (3.2) *holds for $\phi(t,\cdot)$ and for $\phi_h(t,\cdot)$, for all $0 < t \le T$.*

(d) Equi-coercivity: *The families $(\phi(t,\cdot))_{0<t\le T}$ and $(\phi_h(t,\cdot))_{h\in\mathbf{N},0<t\le T}$ are equicoercive in the sense of* (3.6).

(e) Time-regularity: *There exist $a \in L^1(0,T)$, $u^* \in \mathscr{S}$ such that for every $h \in \mathbf{N}$ and for every $v \in \mathscr{S}$ and a.e. $t \in (0,T)$, the time-derivatives $\partial_t\phi(t,v)$ and $\partial_t\phi_h(t,v)$ exist and satisfy*

$$|\partial_t \phi_{(h)}(t,v)| \le a(t)\left(1 + d^2(u^*,v)\right). \tag{3.18}$$

Remark 23 *1. Conditions that are more general than Assumption 22(d) may be formulated. Condition* (3.6) *allows relatively simple proofs even for energies which are not bounded below. It only fails if the decay of ϕ_h along some sequence (v_j) is asymptotically faster than the growth of $d^2(u^*,v_j)$, in which case Assumption 22(c) would typically not be satisfied either.*

2. Although the conditions of Assumption 22 may seem very general at first glance, quite strong properties are hidden in the uniform λ-convexity conditions. Many typical (non-Hilbert) Banach spaces for example are excluded.

For time-independent functionals, if Assumption 22 holds, then Theorem 18 applies to ϕ as well as its approximation(s) ϕ_h. Otherwise, we shall assume (based on Proposition 20) that there exist curves $u, u_h \in AC^2(0,T;\mathscr{S})$ satisfying, respectively, the variational inequalities

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),v) + \frac{\lambda}{2}d^2(u(t),v) + \phi(t,u(t)) \;\le\; \phi(t,v) \quad \forall\, v \in \mathscr{S}, \tag{3.19}$$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u_h(t),v_h) + \frac{\lambda}{2}d^2(u_h(t),v_h) + \phi_h(t,u_h(t)) \;\le\; \phi_h(t,v_h) \quad \forall\, v_h \in \mathscr{S}. \tag{3.20}$$

for a.e. $t \in (0,T)$. It is typical that $D(\phi_h) \subsetneq D(\phi)$. Hence, we test (3.19) with $v = u_h$ but need to use an approximation $v_h \in D(\phi_h)$, which we call a recovery sequence of u to test (3.20).[2] The required properties for v_h will become apparent shortly. Adding (3.19) with $v = u_h(t)$ and (3.20) with $v_h = v_h(t)$ gives

$$\begin{aligned} &\frac{1}{2}\Big[\frac{\mathrm{d}}{\mathrm{d}t}\big(d^2(u(t),u_h(s)) + d^2(u_h(t),v_h(s))\big)\Big]_{s=t} \\ &\quad + \frac{\lambda}{2}\Big(d^2(u,u_h) + d^2(u_h,v_h)\Big) + \phi(t,u) + \phi_h(t,u_h) \le \phi(t,u_h) + \phi_h(t,v_h). \end{aligned} \tag{3.21}$$

We add and subtract terms, in order to bring the inequality into a form suitable for application of a Gronwall inequality. Using

$$\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),u_h(t)) = \Big[\frac{\mathrm{d}}{\mathrm{d}t}\big(d^2(u(t),u_h(s)) + d^2(u_h(t),u(s))\big)\Big]_{s=t},$$

which follows from $u, u_h \in AC^2_{\mathrm{loc}}(0,T;\mathscr{S})$, it is straightforward to find the rearrangement

$$\begin{aligned} &\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),u_h(t)) + \lambda d^2(u,u_h) \\ &\le \Big(\phi_h(t,v_h) - \phi(t,u)\Big) + \Big(\phi(t,u_h) - \phi_h(t,u_h)\Big) \\ &\quad + \frac{\lambda}{2}\Big(d^2(u,u_h) - d^2(v_h,u_h)\Big) + \frac{1}{2}\Big[\frac{\mathrm{d}}{\mathrm{d}t}\big(d^2(u_h(t),u(s)) - d^2(u_h(t),v_h(s))\big)\Big]_{s=t} \\ &=: \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4. \end{aligned} \tag{3.22}$$

The error terms ϵ_1 and ϵ_2 will depend strongly on the given approximation(s) ϕ_h. Natural conditions, guaranteeing convergence are $\phi_h(t,v_h(t)) \to \phi(t,u(t))$ for a.e. t and $\limsup_h(\phi_h(t,u_h) - \phi(t,u_h)) \le 0$. They are made precise in Theorem 25 (b). The third term is easily estimated, using the inverse triangle inequality and Cauchy's inequality, giving

$$\begin{aligned} \epsilon_3 &= \frac{\lambda}{2}\Big(d(u_h,u) + d(u_h,v_h)\Big)\Big(d(u_h,u) - d(u_h,v_h)\Big) \\ &\le \frac{|\lambda|}{2}\Big(d(u_h,u) + d(u_h,v_h)\Big)d(u,v_h) \\ &\le \frac{|\lambda|}{2}\Big(2d(u_h,u) + d(u,v_h)\Big)d(u,v_h) \\ &\le \frac{|\lambda|}{2}d^2(u_h,u) + |\lambda|d^2(u,v_h). \end{aligned}$$

Together with (3.22), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u(t),u_h(t)) + (\lambda - \frac{|\lambda|}{2})\,d^2(u,u_h) \le \epsilon_1 + \epsilon_2 + \epsilon_4 + |\lambda|d^2(u,v_h). \tag{3.23}$$

The fourth error term, ϵ_4, is in general more difficult to handle. In the following lemma, we consider the Hilbert-space case only.

[2] It is no coincidence, that we use the nomenclature from Γ convergence, compare (3.29).

Lemma 24 *Let $\mathscr{S}$ be a Hilbert space and $u \in AC(a,b;\mathscr{S})$. Then, for any $v_0, v_1 \in \mathscr{S}$, we have that*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\|u-v_0\|^2 - \|u-v_1\|^2\Big) \le \|\dot{u}\|\|v_0-v_1\| \quad \text{for a.e. } t \in (a,b). \tag{3.24}$$

Proof Equation (3.24) follows immediately from the elementary identity

$$\|u-v_0\|^2 - \|u-v_1\|^2 = 2(u, v_1-v_0) + \|v_0\|^2 - \|v_1\|^2.$$

■

The natural generalization of (3.24) is

$$\begin{aligned} &u \in AC(a,b;\mathscr{S}) \;\Rightarrow\; \forall\, v_0, v_1 \in \mathscr{S} \quad : \\ &\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big(d^2(u(t),v_0) - d^2(u(t),v_1)\Big) \le c_1|u'|(t)d(v_0,v_1), \quad \text{for a.e. } t \in (a,b), \end{aligned} \tag{3.25}$$

where c_1 is a constant, universal to $(\mathscr{S}, d)$. In general metric spaces, or even Banach spaces, inequality (3.25) does not hold, as simple counterexamples, for example for $\ell^1(\mathbf{R}^2)$ in Proposition 35 show. Its validity seems to rely on the smoothness of $u \mapsto \|u\|^2$; but compare with Section 3.5.

Theorem 25 *Suppose that Assumption 22 is satisfied and that, in addition, $(\mathscr{S}, d)$ satisfies* (3.25)*. Let u and u_h be curves of maximal slope in $(0,T)$ for, respectively, ϕ and ϕ_h, in the sense of* (3.11)*, and assume that they, respectively, satisfy* (3.19) *and* (3.20) *with initial values u^0 and u_h^0.*

(a) For every $0 < s \le T$ we have

$$\begin{aligned} e^{2\alpha s}d^2(u(s),u_h(s)) \;\le\; & d^2(u^0,u_h^0) + 2\int_0^s e^{2\alpha t}\big(\phi(t,u_h) - \phi_h(t,u_h)\big)\,\mathrm{d}t \\ & +2\int_0^s e^{2\alpha t}\Big[\big(\phi_h(t,v_h) - \phi(t,u)\big) + |\lambda| d^2(v_h,u) + c_1|u_h'|d(v_h,u)\Big]\,\mathrm{d}t, \end{aligned} \tag{3.26}$$

where $\alpha = \lambda - |\lambda|/2$ and c_1 is the constant from equation (3.25)*.*

(b) If, in addition, the conditions

$$d(u(0),u_h(0)) \to 0, \sup_{h\in\mathbf{N}} \phi_h(0,u_h(0)) < +\infty, \tag{3.27}$$

$$(w_h) \text{ bounded in } \mathscr{S} \Rightarrow \begin{cases} \limsup_h \big(\phi(t,w_h) - \phi_h(t,w_h)\big) \le 0, \text{and} \\ \phi(t,w_h) \le c_2(1 + \phi_h(t,w_h) + d^2(w_h,\hat{u})) \\ \qquad \text{for a.e. } t \in (0,T), \end{cases} \tag{3.28}$$

$$\exists v_h : (0,T) \to \mathscr{S} \;:\; \begin{cases} \phi_h(t,v_h(t)) \to \phi(t,u(t)), \text{ for a.e. } t \in (0,T), \\ \phi_h(t,v_h(t)) \le c_3(1 + \phi(t,u(t)) + d^2(v_h(t),\hat{u})), \\ \text{and } v_h \to u \text{ in } L^2(0,T;X) \end{cases} \tag{3.29}$$

for constants $c_2, c_3 \in \mathbf{R}$, $\hat{u} \in \mathscr{S}$, independent of h, are satisfied, then $u_h \to u$ in $L^\infty(0,T;\mathscr{S})$.

Proof Part (a) is a straightforward application of Lemma 26 to

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d^2(u,u_h)+(\lambda-|\lambda|/2)d^2(u,u_h) &\leq \big(\phi(t,u_h)-\phi_h(t,u_h)\big)+\big(\phi_h(t,v_h)-\phi(t,u)\big)\\ &\quad +|\lambda|d^2(v_h,u)+c_1|u_h'|d(v_h,u),\end{aligned}$$

which follows from (3.23) and (3.25).

Using Lemma 27 below, we obtain a constant C, independent of h, such that

$$\frac{1}{t}d^2(u_h(t),u_h(0))\leq\int_0^T|u_h'|^2\,\mathrm{d}t\leq C.$$

Hence, for every t, $(u_h(t))_{h\in\mathbf{N}}$ and $(|\phi_h(t,u_h(t)|)_{h\in\mathbf{N}}$ are bounded. We may use (3.28), in particular $\phi(t,u_h)\leq c_2(1+\phi_h(t,u_h)+d^2(u_h,\hat{u}))\leq C$, and apply Fatou's Lemma to obtain

$$\begin{aligned}\limsup_{h\to\infty}\int_0^s\big[\phi(t,u_h)-\phi_h(t,u_h)\big]\,\mathrm{d}t &\leq \limsup_{h\to\infty}\int_0^s\max(\phi(t,u_h)-\phi_h(t,u_h),0)\,\mathrm{d}t\\ &\leq \int_0^s\limsup_{h\to\infty}\max(\phi(t,u_h)-\phi_h(t,u_h),0)\,\mathrm{d}t\\ &\leq 0. \qquad (3.30)\end{aligned}$$

Let $v_h=v_h(t)$ be the recovery sequence from equation (3.29). Similarly to (3.30), using Fatou's Lemma again, we have

$$\limsup_{h\to\infty}\int_0^T|\phi_h(t,v_h)-\phi(t,u)|\,\mathrm{d}t\leq 0. \qquad (3.31)$$

The boundedness of $|u_h'|$ in $L^2(0,T)$ and (3.29) shows that

$$\int_0^T c_1|u_h'|d(u,v_h)+|\lambda|d^2(u,v_h)\,\mathrm{d}t\to 0. \qquad (3.32)$$

Combining (3.30) to (3.32) with the error estimate (3.26) concludes the proof. ∎

Lemma 26 *Let $x\in AC(0,T;\mathbf{R})$, $a,b\in L^1(0,T)$ and $\alpha\in\mathbf{R}$. If x satisfies*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}x(t)+\lambda x(t)\leq a(t)$$

for $\mathscr{L}^1$-a.e. $t\in(0,T)$, then

$$e^{2\lambda t}x(t)\leq x(0)+2\int_0^t e^{2\lambda s}a(s)\,\mathrm{d}s.$$

Proof Multiply the differential inequality by $e^{2\lambda t}$ to obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big(e^{2\lambda t}x(t)\big)\leq e^{2\lambda t}a(t).$$

Integrated from 0 to t this yields the desired result. ∎

Lemma 27 *Under the conditions of Theorem 25, there exist constants C_1, C_2, depending on the parameters of Assumption 22(c) and (e), on T, and on $\sup_h \phi_h(0, u_h(0))$, but not on h, such that for every $0 < s \leq T$, we have that*

$$\begin{aligned}
\frac{1}{2T} d^2(u_h(T), u_h(0)) &\leq \frac{1}{2} \int_0^T |u_h'|^2 \,\mathrm{d}t \qquad (3.33)\\
&\leq \phi_h(0, u_h(0)) - \phi_h(T, u_h(T)) + \int_0^T \partial_t \phi(t, u_h(t))\\
&\leq C_1, \quad \text{and}\\
|\phi_h(T, u_h(T))| &\leq C_2. \qquad (3.34)
\end{aligned}$$

Proof From (3.11), we obtain

$$\frac{1}{2} \int_0^T |u_h'|^2 \,\mathrm{d}t \leq \phi_h(u_h(0)) - \phi_h(u_h(T)) + \int_0^T \partial_t \phi_h(t, u_h(t)) \,\mathrm{d}t, \qquad (3.35)$$

which proves the second inequality in (3.33). Proving the first inequality is an easy exercise:

$$\begin{aligned}
d^2(u_h(T), u_h(0)) &= \left(\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} d(u_h(t), u_h(0)) \,\mathrm{d}t \right)^2\\
&\leq \left(\int_0^T |u_h'| \,\mathrm{d}t \right)^2 \leq T \int_0^T |u_h'|^2(t) \,\mathrm{d}t.
\end{aligned}$$

It remains to bound the terms in (3.35).

From (3.27), we obtain the boundedness of $\phi_h(0, u_h(0))$. To estimate $\phi_h(T, u_h(T))$ from below, we proceed as follows. Note first that, using (3.6) with $\phi_h(T, \cdot)$,

$$\begin{aligned}
\phi_h(T, u_h(T)) &= \Phi_h(\tau^*, u^*; T, u_h(T)) - \frac{d^2(u^*, u_h(T))}{2\tau^*}\\
&\geq m^* - \frac{d^2(u^*, u_h(0))}{\tau^*} - \frac{d^2(u_h(0), u_h(T))}{\tau^*}.
\end{aligned}$$

Next, we obtain an estimate for $d^2(u_h(T), u_h(0))$, which is ameanable to a Gronwall inequality-based argument:

$$\begin{aligned}
\frac{1}{2} d^2(u_h(T), u_h(0)) &= \frac{1}{2} \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} d^2(u_h(t), u_h(0)) \,\mathrm{d}t\\
&\leq \int_0^T d(u_h(t), u_h(0)) |u_h'|(t) \,\mathrm{d}t\\
&\leq \frac{\epsilon}{2} \int_0^T |u_h'|^2 \,\mathrm{d}t + \frac{1}{2\epsilon} \int_0^T d^2(u_h(t), u_h(0)) \,\mathrm{d}t\\
&\leq \epsilon \phi_h(0, u_h(0)) - \epsilon m^* + \frac{\epsilon}{\tau^*} d^2(u^*, u_h(0)) + \frac{\epsilon}{\tau^*} d^2(u_h(T), u_h(0))\\
&\quad + \int_0^T \partial_t \phi(t, u_h(t)) \,\mathrm{d}t + \frac{1}{2\epsilon} \int_0^T d^2(u_h(t), u_h(0)) \,\mathrm{d}t.
\end{aligned}$$

Setting $\epsilon = \tau^*/2$, and using equation (3.18) we obtain

$$d^2(u_h(T), u_h(0)) \leq C_1 + C_2 \int_0^T d^2(u_h(t), u_h(0))\,\mathrm{d}t.$$

Integrating Gronwall's Lemma 26 with $\lambda = -C_2/2$ and $a = C_1/2$, and applying it to the above inequality, we obtain a bound on $d^2(u_h(T), u_h(0))$, which is uniform in h and may immediately be translated into the third inequality of (3.33), using also (3.18).

Thus we have obtained a lower bound on $\phi_h(T, u_h(T))$. For an upper bound consider

$$\begin{aligned}
\phi_h(T, u_h(T)) &= \phi_h(0, u_h(0)) + \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\phi_h(t, u_h(t))\,\mathrm{d}t \\
&= \phi_h(0, u_h(0)) - \frac{1}{2}\int_0^T |u_h'|^2(t) + |\partial\phi_h|^2(t, u_h(t))\,\mathrm{d}t + \int_0^T \partial_t\phi_h(t, u_h(t))\,\mathrm{d}t \\
&\leq \sup_{h\in\mathbf{N}} \phi_h(0, u_h(0)) + \int_0^T a(t)(1 + d^2(u^*, u_h(t)))\,\mathrm{d}t.
\end{aligned}$$

Using (3.33) to bound $d(u^*, u_h(t))$, we obtain (3.34). ■

3.4 Discrete case

In this Section, we develop a discrete variant of the analysis performed in Section 3.3, comparing the discrete solution for the exact functional ϕ and the discrete solution for the approximation ϕ_h. The analysis follows along the same lines as in the continuous case, and is therefore kept brief at several places.

For every $h \in \mathbf{N}$, let $\boldsymbol{\tau}_h = (t_h^0, t_h^1, \dots)$, $\tau_h^j = t_h^j - t_h^{j-1}$, be a partition of $[0, T]$ such that $1/\lambda^- > |\boldsymbol{\tau}_h| = (\sup_j \tau_h^j) \to 0$ as $h \to \infty$. Choose initial values U_h^0 and let $U_h = (U_h^j)_{j\in\mathbf{N}}$ be the time-discrete approximations, defined as follows. For $\tau > 0$, $u, v \in \mathscr{S}$, let

$$\Phi_h(\tau, u; t, v) = \frac{d^2(u, v)}{2\tau} + \phi_h(t, v),$$

and define U_h^j to be the (unique) minimizer of $\Phi_h(\tau_h^j, U_h^{j-1}; t_h^j, \cdot)$. For comparison, let $u_h = (u_h^j)_{j\in\mathbf{N}}$ be the solution of the discrete scheme for ϕ, i.e., $u_h^0 = u_0$ and u_h^j is the minimizer of $\Phi(\tau_h^j, u_h^{j-1}; t_h^j, \cdot)$, where Φ is defined analogously to Φ_h.

Following Theorem 21, $(u_h^j)_{j=0,1,\dots}$ and $(U_h^j)_{j=0,1,\dots}$ satisfy the variational inequalities

$$\frac{d^2(u_h^j, v) - d^2(u_h^{j-1}, v)}{2\tau_h^j} + \frac{\lambda}{2}d^2(u_h^j, v) + \phi(t_h^j, u_h^j) \leq \phi(t_h^j, v) \quad \forall\, v \in \mathscr{S}, \tag{3.36}$$

$$\frac{d^2(U_h^j, v) - d^2(U_h^{j-1}, v)}{2\tau_h^j} + \frac{\lambda}{2}d^2(U_h^j, v) + \phi_h(t_h^j, U_h^j) \leq \phi_h(t_h^j, v) \quad \forall\, v \in \mathscr{S}. \tag{3.37}$$

As before, we test (3.36) with U_h^j and (3.37) with a family V_h^j, which satisfies good approximation properties. Adding the two inequalities and arguing as in the continuous case, we obtain

$$\frac{d^2(u_h^j, U_h^j) - d^2(u_h^{j-1}, U_h^{j-1})}{2\tau_h^j} + (\lambda - |\lambda|/2)d^2(u_h^j, U_h^j) \leq \epsilon_1^j + \epsilon_2^j + \epsilon_3^j + \epsilon_4^j, \quad (3.38)$$

where

$$\begin{aligned}
\epsilon_1^j &= \phi(t_h^j, U_h^j) - \phi_h(t_h^j, U_h^j), \\
\epsilon_2^j &= \phi_h(t_h^j, V_h^j) - \phi(t_h^j, u_h^j), \\
\epsilon_3^j &= |\lambda| d^2(u_h^j, V_h^j), \text{ and} \\
\epsilon_4^j &= \frac{d^2(U_h^{j-1}, V_h^j) - d^2(U_h^j, V_h^j) + d^2(u_h^{j-1}, U_h^j) - d^2(u_h^{j-1}, U_h^{j-1})}{2\tau_h^j}.
\end{aligned}$$

Again, the error term ϵ_4 is the most difficult one to control. Lemma 28 is the discrete counterpart of Lemma 24.

Lemma 28 *Let $(\mathscr{S}, d)$ be a Hilbert space, then*

$$\epsilon_4^j = \left(\frac{U_h^j - U_h^{j-1}}{\tau_h^j}, V_h^j - u_h^{j-1} \right)$$

In general, we shall assume that the space $(\mathscr{S}, d)$ satisfies the inequality

$$\frac{d^2(u^0, v_0) - d^2(u^1, v_0) + d^2(u^0, v_1) - d^1(u^1, v_1)}{2\tau} \leq c_1 \frac{d(u^1, u^0)}{\tau} \, d(v_0, v_1) \quad (3.39)$$

for all points $u^0, u^1, v_0, v_1 \in \mathscr{S}$. To simplify notation, we define

$$|U_h'|(t) = \frac{d(U_h^{j-1}, U_h^j)}{\tau_h^j} \quad \text{for } t \in (t_h^{j-1}, t_h^j),$$

and $|u_h'|$ analogously. Furthermore, we denote the backward piecewise constant interpolation of (U_h^j) by $\bar{U}_h$, i.e.,

$$\bar{U}_h(t) = U_h^j, \quad \text{if } t_h^{j-1} < t \leq t_h^j.$$

Theorem 29 *Suppose that Assumption 22 is satisfied and that, in addition, $(\mathscr{S}, d)$ satisfies* (3.39)*. For each $h \in \mathbf{N}$, let $(U_h^j)_{j=0,1,\ldots}$ and $(u_h^j)_{j=0,1,\ldots}$ be the discrete solutions of the gradient flow in the sense of* (3.16)*, for, respectively, ϕ_h and ϕ and for the time-partition $\boldsymbol{\tau}_h$.*

(a) For every $j = 1, 2, \ldots$ we have

$$\begin{aligned} e^{\alpha t_h^j} d^2(u_h^j, U_h^j) \;\leq\; & d^2(u_h^0, U_h^0) + 2\sum_{i=1}^{k} \tau_h^i e^{\alpha t_h^i} \Big\{ \big(\phi(t_h^i, U_h^i) - \phi_h(t_h^i, U_h^i)\big) \\ & + \big(\phi_h(t_h^i, V_h^i) - \phi(t_h^i, u_h^i)\big) + |\lambda| d^2(V_h^i, u_h^i) \\ & + c_1 \frac{d(U_h^i, U_h^{i-1})}{\tau_h^i} d(V_h^i, u_h^{i-1}) \Big\}, \quad \forall (V_h^i)_{j=1,2,\ldots} \subset \mathscr{S}, \end{aligned} \tag{3.40}$$

where $\alpha = 2(\lambda - |\lambda|/2)/(1 - 2(\lambda - |\lambda|/2)|\boldsymbol{\tau}_h|)$ and c_1 is the constant from equation (3.39).

(b) Suppose that, in addition, ϕ and ϕ_h are independent of t, and that the $\boldsymbol{\tau}_h$ are partitions of $[0, \infty)$ now, so that Theorems 18 and 19 apply. Suppose also, that (3.27)*, holds, that* (3.28) *holds for* every t*, and that*

$$\begin{aligned} & \forall h \in \mathbf{N}, \exists (V_h^j)_{j=1,2,\ldots} \text{ such that} \\ & \qquad \begin{cases} \phi_h(\bar{V}_h(t)) \to \phi(u(t)) \quad \text{for a.e. } t \in (0,T), \\ \phi_h(V_h^j) \leq c_2(1 + \phi(u(t_h^j))) \quad \text{and} \\ \bar{V}_h \to u \quad \text{in } L^2(0,T;\mathscr{S}) \end{cases} \end{aligned} \tag{3.41}$$

is satisfied. Then, $\bar{U}_h \to u$ in $L^\infty(0,T;\mathscr{S})$ for all $T > 0$.

Proof The error estimate (a) is a simple application of Lemma 31 to (3.38).

To obtain (b), note that $d(u_h^j, u(t_h^j)) \to 0$, uniformly for bounded t_h^j (cf. Theorem 19).

Fix $0 \leq s \leq T$ and let j_h be the smallest index such that $t_h^{j_h} \geq T$. To estimate the first error term in (3.40), we argue similarly as in the proof of Theorem 25, using (2.13) to obtain a bound which allows the application of Fatou's Lemma,

$$\begin{aligned} \limsup_{h\to\infty} \max_{j=1,\ldots,j_h} \sum_{i=1}^{j} \tau_h^i \big[\phi(U_h^i) - \phi_h(U_h^i)\big] \;&\leq\; \limsup_{h\to\infty} \int_0^{t_h^{j_h}} \big[\phi(\bar{U}_h) - \phi_h(\bar{U}_h)\big]^+ \,\mathrm{d}t \\ &\leq\; \limsup_{h\to\infty} \int_0^{T} \big[\phi(\bar{U}_h) - \phi_h(\bar{U}_h)\big]^+ \,\mathrm{d}t + \delta_1 \\ &\leq\; \int_0^{T} \limsup_{h\to\infty} \big[\phi(\bar{U}_h) - \phi_h(\bar{U}_h)\big] \,\mathrm{d}t + \delta_1 \\ &\leq\; 0 + \delta_1, \end{aligned}$$

where δ_1, the error made by replacing $\int_0^{t_h^{j_h}}$ with $\int_0^T$, can be estimated via

$$\begin{aligned}\delta_1 &= \limsup_h \int_T^{t_h^{j_h}} \big[\phi(\bar{U}_h) - \phi_h(\bar{U}_h)\big]\,\mathrm{d}t \\ &\le \limsup_h \int_T^{t_h^{j_h}} \big[c_1 + (c_1-1)\phi_h(\bar{U}_h)\big]\,\mathrm{d}t \\ &\le \limsup_h \int_T^{t_h^{j_h}} \big[c_1 + (c_1-1)\phi_h(U_h^0)\big]\,\mathrm{d}t \\ &= 0.\end{aligned}$$

The second error term can be treated similarly. Let w_h be the piecewise constant interpolant of $(u(t_h^j))_{j=1,2,\dots}$ (cf. Lemma 32). Let V_h^j be the recovery sequence from equation (3.41), then

$$\begin{aligned}\max_{j=1,\dots,j_h} \sum_{i=1}^{t_h^j} \tau_h^i \big[\phi_h(V_h^i) - \phi(u_h^i)\big] &\le \int_0^{t_h^{j_h}} |\phi_h(\bar{V}_h) - \phi(\bar{u}_h)|\,\mathrm{d}t \\ &\le \int_0^T |\phi_h(\bar{V}_h) - \phi(\bar{u}_h)|\,\mathrm{d}t + \delta_2\end{aligned}$$

The error term δ_2 can be estimated with the help of Lemma 32, showing that $\lim_{h\to\infty}\delta_2 = 0$. The bound $|\phi_h(\bar{V}_h) - \phi(\bar{u}_h)| \le C_1 + C_2\phi(u(0))$ shows that we can apply the dominated convergence theorem. Using

$$|\phi_h(\bar{V}_h) - \phi(\bar{u}_h)| \le |\phi_h(\bar{V}_h) - \phi(u)| + |\phi(u) - \phi(w_h)| + |\phi(w_h) - \phi(\bar{u}_h)| \to 0,$$

almost everywhere in $(0,\infty)$, we obtain that the second error term in (3.40) tends uniformly to zero as $h \to \infty$.

Finally, the third and fourth error terms can be estimated by using $\bar{V}_h \to u$ in L^2, together with the interpolation error estimate of Lemma 32. ■

Remark 30 *1. The result of part (b) in Theorem 29 is restricted to time-independent functionals only because, for the time being, we do not have a convergence theory for the time-discretization, like Theorem 19 (c). If this gap were closed, which can be expected, then Theorem 29 (b) will hold for time-dependent functionals as well.*

2. Estimates of the type (3.41) *should be easy to obtain with the help of Lemma 32.*

Lemma 31 *Let* $(x_j)_{j\in\mathbf{N}}, (a_j)_{j\in\mathbf{N}}, (\tau_j)_{j\in\mathbf{N}} \subset \mathbf{R}$, $x_j, \tau_j \ge 0$ *and* $\lambda \in \mathbf{R}$ *be such that*

$$\frac{x_j - x_{j-1}}{2\tau_j} + \lambda x_j \le a_j, \quad j = 1, 2, \dots.$$

If $|\boldsymbol{\tau}| = \sup_j \tau_j \le 1/(2\lambda^-)$ *then* (x_j) *satisfies*

$$e^{\alpha t_j} x_j \le x_0 + 2\sum_{i=1}^{j} \tau_i e^{\alpha t_{i-1}} a_i, \tag{3.42}$$

where $t_i = \sum_{k=1}^{i} \tau_k$ *and* $\alpha = 2\lambda$ *if* $\lambda \le 0$ *and* $\alpha = \lambda/(1 + 2\lambda \sup_j \tau_j)$, *if* $\lambda > 0$.

Proof Using $e^{x/(1+x)} \le (1+x)$ for $x > -1$, we have, for every $j = 1, 2, \ldots,$

$$e^{\alpha\tau_j} x_j \le e^{\frac{2\lambda\tau_j}{1+2\lambda\tau_j}} \le (1 + 2\lambda\tau_j) x_j \le x_{j-1} + 2\tau_j a_j, \tag{3.43}$$

which also proves (3.42) for $j = 1$. We proceed by induction. Assume that (3.42) holds for some j; then, using (3.43), we obtain

$$\begin{aligned} e^{\alpha t_{j+1}} x_{j+1} &\le e^{\alpha t_j}(1 + 2\lambda\tau_{j+1}) x_{j+1} \\ &\le e^{\alpha t_j} x_j + 2\tau_{j+1} e^{\alpha t_j} a_{j+1} \\ &\le x_0 + 2\sum_{i=1}^{j+1} \tau_i e^{\alpha t_{i-1}} a_i. \end{aligned}$$

■

Lemma 32 (Interpolation Error) *Suppose that Assumption 22 holds for a time-independent functional* ϕ. *Let* $u \in AC^2_{\mathrm{loc}}([0,\infty); \mathscr{S})$ *be a curve of maximal slope for* ϕ *and let* w *be its piecewise constant interpolant with respect to the mesh* $\boldsymbol{\tau}$, *i.e.,* $w(t) = u(t^j)$, *if* $t^{j-1} < t \le t^j$. *Then, the following two error estimates hold for all* $j > 0$*:*

$$\int_0^{t^j} d^2(u,w)\,\mathrm{d}t \le \frac{|\boldsymbol{\tau}|}{1-|\boldsymbol{\tau}|} \int_0^{t^j} |u'|^2(t)\,\mathrm{d}t \le \frac{|\boldsymbol{\tau}|}{1-|\boldsymbol{\tau}|}\big(\phi(u(0)) - \phi(u(t^j))\big), \tag{3.44}$$

$$\int_0^{t^j} \big|\phi(u) - \phi(w)\big|\,\mathrm{d}t \le |\boldsymbol{\tau}| \int_0^{t^j} |\partial\phi|(u)|u'|\,\mathrm{d}t \le |\boldsymbol{\tau}|\big(\phi(u(0)) - \phi(u(t^j))\big). \tag{3.45}$$

Proof For every j, we have

$$\begin{aligned} \int_{t^{j-1}}^{t^j} d^2(u,w)\,\mathrm{d}t &= \int_{t^{j-1}}^{t^j} \int_{s=t}^{t^j} \frac{\mathrm{d}}{\mathrm{d}t} d^2(u_h, w_h)\,\mathrm{d}s\,\mathrm{d}t \\ &\le \int_{t^{j-1}}^{t^j} \int_{s=t}^{t^j} 2d(u_h,w_h)|u'|\,\mathrm{d}s\,\mathrm{d}t \\ &\le \tau^j \int_{t^{j-1}}^{t^j} \big[d^2(u_h,w_h) + |u'|^2(t)\big]\,\mathrm{d}t, \end{aligned}$$

which gives (3.44). Similary, we can estimate the error in the energy, using the fact that $|\partial\phi|$ is an upper gradient for ϕ,

$$\begin{aligned}\int_{t^{j-1}}^{t^j} \big|\phi(u)-\phi(w)\big|\,\mathrm{d}t &= \int_{t^{j-1}}^{t^j}\left|\int_{s=t}^{t^j}\frac{\mathrm{d}}{\mathrm{d}s}\phi(u(s))\,\mathrm{d}s\right|\,\mathrm{d}t\\ &\le \int_{t^{j-1}}^{t^j}\int_{t}^{t^j}|\partial\phi|(u)|u'|\,\mathrm{d}s\\ &\le \tau^j\int_{t^{j-1}}^{t^j}|\partial\phi|(u)|u'|\,\mathrm{d}t.\end{aligned}$$

Estimate (3.45) now follows directly from the definition of a curve of maximal slope. ■

Remark 33 *1. For mere convergence results the theory presented may be sufficient. Further results concerning the convergence of the sequences of slopes $|\partial\phi_h|$ and metric derivatives $|U_h'|$ may be obtained by combining this analysis with the results of Section 2. In the Hilbert space case, these may be even extended to the convergence of the gradients $\phi_h'(U_h)$ and the speeds $\dot{U}_h$.*

2. For applications in numerical analysis, especially for obtaining a-posteriori *error estimates, one may prefer to compare an interpolated version of the discrete variational inequality* (3.9) *directly to the continuous variational inequality* (3.7)*. This is beyond the scope of this paper though.*

3. Using a different version of the Gronwall Lemma one can obtain a lower exponential factor in Lemmas 26 and (3.42) *at the expense of a slightly larger error bound and a somewhat more complicated error analysis. See also Lemma 4.1.8 [AGS05].*

3.5 Examples for (3.25)

We conclude by showing the relationship between the smoothness of the metric and equation (3.25).

Proposition 34 *Suppose that $(\mathscr{S},\|\cdot\|)$ is a normed space such that the functional $F(v)=\|v\|^2$ is Fréchet differentiable and its derivative has Lipschitz constant c_1. Then, for any curve $u\in AC(a,b;\mathscr{S})$ and any two points $v_0,v_1\in\mathscr{S}$, we have that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\|u-v_0\|^2-\|u-v_1\|^2\big)\le c_1\|v_0-v_1\|\|\dot{u}\|.$$

Proof If F is differentiable and $u\in AC(a,b;\mathscr{S})$, then $\frac{\mathrm{d}}{\mathrm{d}t}F(u-v_i)=\langle F'(u-v_i),\dot{u}\rangle$, where $\langle\cdot,\cdot\rangle$ is the duality pairing, and hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\big(F(u-v_0)-F(u-v_1)\big) &= \langle F'(u-v_0)-F'(u-v_1),\dot{u}\rangle\\ &\le \|F'(u-v_0)-F'(u-v_1)\|_*\|\dot{u}\|\\ &\le c_1\|v_1-v_0\|\|\dot{u}\|,\end{aligned}$$

where $\|\cdot\|_*$ is the norm in the dual Banach space $\mathscr{S}^*$. ■
For the most common norms, the L^p norms ($p \neq 2$), Proposition 34 does not apply and in fact, (3.25) does not hold. As L^p spaces with $p \neq 2$ are not candidates for the theory of this section, we do not go into further details, but show only on a simpler example, that the smoothness of the norm is crucial.

Proposition 35 *Let $\mathscr{S} = \mathbf{R}^2$ be equipped with the norm $\|v\|_{\ell^1} = |v_1| + |v_2|$. Then, there exists a constant c_0 such that for every $\epsilon > 0$, there exists a Lipschitz continuous curve u and points v, w such that*

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} \Big[\|u - v\|_{\ell^1}^2 - \|u - w\|_{\ell^1}^2 \Big] \right| \geq c_0 \frac{1 - 2\epsilon}{\epsilon} \|\dot{u}\|_{\ell^1} \|v - w\|_{\ell^1}. \tag{3.46}$$

Proof Let $v = (-\epsilon, 0)$, $w = (0, \epsilon)$, and $u(t)$ such that $-\epsilon \leq u_1(t) \leq \epsilon$, then

$$\begin{aligned} \left| \frac{\mathrm{d}}{\mathrm{d}t} \big(d^2(u, v) - d^2(u, w) \big) \right| &= |2(u_1 - \epsilon)\dot{u}_1 - 2(u_1 + \epsilon)\dot{u}_1 + 4\dot{u}_1 u_2 + 4u_1 \dot{u}_1| \\ &\geq 4|\dot{u}_1|\,|u_2| - 8\epsilon(|\dot{u}_1| + |\dot{u}_2|) \\ &= 4|\dot{u}_1|\,|u_2| - 4\|\dot{u}\|_{\ell^1} \|v - w\|_{\ell^1}. \end{aligned}$$

If we choose u_1 and u_2 accordingly, then we obtain the desired result. For example, let $\dot{u}_1 \in \{-1, 1\}$ and $u_2 \equiv 1$, then (3.46) holds with $c_0 = 2$. ■

Interestingly, some metrics which are quite far from being Hilbert space metrics satisfy (3.25) and (3.39) as well. Consider, for example, the following metric defined on $BV(\Omega) \cap L^2(\Omega)$, which is related to strict convergence in $BV(\Omega)$;

$$d(v, w) = \Big[\|v - w\|_{L^2}^2 + \big| |Dv|(\Omega) - |Dw|(\Omega) \big|^2 \Big]^{1/2},$$

where $|Du|(\Omega)$ is the total variation of u in Ω. It can be quite easily checked that this metric satisfies (3.25) and (3.39). The reason is the one-dimensional structure of the second term in d. It would, however, be very difficult to find curves v_t which are candidates for (3.2).

Acknowledgements I would like to thank Johannes Zimmer for bringing to my attention the excellent book by Luigi Ambrosio, Nicola Gigli and Giuseppe Savaré on the theory of curves of maximal slope. I have profited from several helpful discussions with Endre Süli and Johannes Zimmer during the writing of this paper. Section 3.5 is based on some comments by Bernd Kirchheim. I am indebted to him as well.

References

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Space and in the Space of Probability Measures.* Birkhäuser Verlag, 2005.

[Bra02] A. Braides. Γ-*convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[Dac89] B. Dacorogna. *Direct Methods in the Calculus of Variations*. Springer Verlag, 1989.

[FP04] X. Feng and A. Prohl. Analysis of gradient flow of a regularized Mumford-Shah functional for image segmentation and image inpainting. *M2AN Math. Model. Numer. Anal.*, 38(2):291–320, 2004.

[Gob98] M. Gobbino. Gradient flow for the one-dimensional Mumford-Shah functional. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 27(1):145–193 (1999), 1998.

[NS] R. Nochetto and G. Savaré. Nonlinear evolution governed by accretive operators in Banach spaces: Error control and applications. (unpublished) `http://www.math.umd.edu/~rhn/publications.html`.