Remarks 2.8 (i) Regularity of minimizers, e.g., $u \in H^2(\Omega)$, can be proved if the boundary of Ω is C^2-regular and W is *strongly convex*, i.e., $D^2W(A)[B, B] \geq c|B|^2$ for all $A, B \in \mathbb{R}^{m\times d}$ and $c > 0$.
(ii) If the minimization problem involves a constraint, such as $G(u(x)) = 0$ for almost every $x \in \Omega$ with a continuously differentiable function $G : \mathbb{R}^m \to \mathbb{R}$, then one can formally consider a saddle point problem to derive the Euler–Lagrange equations, e.g., for $W(A) = |A|^2/2$, the problem

$$\inf_{u\in H^1_0(\Omega)} \sup_{\lambda\in L^q(\Omega)} \frac{1}{2}\int_\Omega |\nabla u|^2 \,\mathrm{d}x + \int_\Omega \lambda G(u)\,\mathrm{d}x.$$

The optimality conditions are with $g = DG$ given by

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_\Omega \lambda g(u) \cdot v \,\mathrm{d}x = 0, \quad \int_\Omega \mu G(u)\,\mathrm{d}x = 0$$

for all $v \in H^1_0(\Omega; \mathbb{R}^\ell)$ and all $\mu \in L^q(\Omega)$. The unknown variable λ is the *Lagrange multiplier* associated to the constraint $G(u(x)) = 0$ for almost every $x \in \Omega$.
(iii) On the part $\partial\Omega \setminus \Gamma_D$ where no *Dirichlet boundary conditions* $u|_{\Gamma_D} = u_D$ are imposed, the homogeneous *Neumann boundary conditions* $DW_0(\nabla u) \cdot n = 0$ are satisfied. Inhomogeneous Neumann conditions can be specified through a function $g \in L^q(\Gamma_N; \mathbb{R}^m)$ and a corresponding contribution to the energy functional, e.g.,

$$I(u) = \int_\Omega W_0(\nabla u(x))\,\mathrm{d}x - \int_{\Gamma_N} gu\,\mathrm{d}s.$$

(iv) The Euler–Lagrange equations define an operator $L : W^{1,p}(\Omega) \to W^{1,p}(\Omega)'$ and we look for $u \in W^{1,p}(\Omega)$ with $L(u) = b$ for a given right-hand side $b \in W^{1,p}(\Omega)'$. Under certain monotonicity conditions on L, the existence of solutions for this equation can be established with the help of discretizations and fixed-point theorems. This is of importance when the partial differential equation is not related to a minimization problem.

2.3 Gradient Flows

The direct method in the calculus of variations provides existence results for global minimizers of functionals but its proof is nonconstructive. In practice, the most robust methods to find stationary points are steepest descent methods. These can often be regarded as discretizations of time-dependent problems. To understand the stability and convergence properties of descent methods, it is important and insightful to analyze the corresponding continuous problems. In finite-dimensional situations we may think of a function $V : \mathbb{R}^n \to \mathbb{R}$ and the ordinary differential equation

$$y' = -\nabla V(y), \quad y(0) = y_0.$$

If $V \in C^2(\mathbb{R}^n)$, then the Picard–Lindelöf theorem guarantees the existence of a unique local solution $y : (-\widehat{T}, \widehat{T}) \to \mathbb{R}^n$. Taking the inner product of the differential equation with y' and using the chain rule to verify $\nabla V(y(t)) \cdot y'(t) = (V \circ y)'$ we have after integration over $(0, T')$

$$\int_0^{T'} |y'|^2 \, \mathrm{d}t + V(y(T')) = V(y(0)).$$

This is called an *energy law* and shows that the function $t \mapsto V(y(t))$ is decreasing. Since the evolution becomes stationary if $\nabla V(y(t)) = 0$, this allows us to find critical points of V with small energy. It is the aim of this section to justify gradient flows for functionals on infinite-dimensional spaces. For more details on this subject, we refer the reader to the textbooks [2, 6–8].

2.3.1 Differentiation in Banach Spaces

We consider a Banach space X and a functional $I : X \to \mathbb{R}$.

Definition 2.3 (a) We say that I is *Gâteaux-differentiable* at $v_0 \in X$ if for all $h \in X$ the limit

$$\delta I(v_0, h) = \lim_{s \to 0} \frac{I(v_0 + sh) - I(v_0)}{s}$$

exists and the mapping $DI(v_0) : X \to \mathbb{R}$, $h \mapsto \delta I(v_0, h)$ is linear and bounded.
(b) We say that I is *Fréchet-differentiable* at $v_0 \in X$ if there exist a bounded linear operator $A : X \to \mathbb{R}$ and a function $\varphi : \mathbb{R} \to \mathbb{R}$ with $\lim_{s \to 0} \varphi(s)/s = 0$ such that

$$I(v_0 + h) - I(v_0) = Ah + \varphi(\|h\|_X).$$

In this case we define $DI(v_0) = A$.

Remark 2.9 If I is Gâteaux-differentiable at every point in a neighborhood of v_0 and DI is continuous at v_0, then I is Fréchet-differentiable at v_0.

The gradient of a functional is the Riesz representative of the Fréchet derivative with respect to a given scalar product.

Definition 2.4 Let H be a Hilbert space such that X is continuously embedded in H. If I is Fréchet-differentiable at $v_0 \in X$ with $DI(v_0) \in H'$, then the *H-gradient* $\nabla_H I(v_0) \in H$ is defined by

$$(\nabla_H I(v_0), v)_H = DI(v_0)[v]$$

for all $v \in H$.

Example 2.2 For $X = H_0^1(\Omega)$ and $I(u) = (1/2)\int_\Omega |\nabla u|^2 \, dx$, we have

$$\begin{aligned} I(u + sv) - I(u) &= \frac{1}{2}\int_\Omega |\nabla u|^2 + 2s\nabla u \cdot \nabla v + s^2|\nabla v|^2 \, dx - \frac{1}{2}\int_\Omega |\nabla u|^2 \, dx \\ &= s\int_\Omega \nabla u \cdot \nabla v \, dx + \frac{s^2}{2}\int_\Omega |\nabla v|^2 \, dx \end{aligned}$$

and I is Fréchet differentiable with $DI(u)[v] = \int_\Omega \nabla u \cdot \nabla v \, dx$. For $H = X = H_0^1(\Omega)$ with scalar product $(v, w)_{H_0^1} = \int_\Omega \nabla v \cdot \nabla w \, dx$, we thus have $\nabla_{H_0^1} I(u) = u$. If $u \in H^2(\Omega) \cap H_0^1(\Omega)$, then Green's formula shows that

$$DI(u)[v] = \int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega (-\Delta u)v \, dx,$$

so that $DI(u)$ is a bounded linear functional on $L^2(\Omega)$. For $H = L^2(\Omega)$ with scalar product $(v, w) = \int_\Omega vw \, dx$, we therefore have $\nabla_{L^2} I(u) = -\Delta u$.

Remark 2.10 The Euler–Lagrange equations for $I(u) = \int_\Omega W(x, u, \nabla u) \, dx$ in the strong form corresponds to a vanishing L^2-gradient of I, i.e., $\nabla_{L^2} I(u) = 0$.

2.3.2 Bochner–Sobolev Spaces

For evolutionary partial differential equations we will consider functions $u : [0, T] \to X$ for a time interval $[0, T] \subset \mathbb{R}$. We assume that the Banach space X is separable and say that $u : [0, T] \to X$ is *weakly measurable* if, for all $\varphi \in X'$, the function $t \mapsto \langle \varphi, u(t) \rangle$ is Lebesgue measurable. In this case the *Bochner integral*

$$\int_0^T u(t) \, dt$$

is well defined with $\big\| \int_0^T u(t) \, dt \big\|_X \le \int_0^T \|u(t)\|_X \, dt$. The duality pairing between X and X' will be denoted by $\langle \cdot, \cdot \rangle$.

Definition 2.5 For $1 \le p \le \infty$, the *Bochner space* $L^p([0, T]; X)$ consists of all weakly measurable functions $u : [0, T] \to X$ with $\|u\|_{L^p([0,T];X)} < \infty$, where

$$\|u\|_{L^p([0,T];X)} = \begin{cases} \operatorname{esssup}_{t\in[0,T]} \|u(t)\|_X & \text{if } p=\infty, \\ \Big(\int\limits_{[0,T]} \|u(t)\|^p \,\mathrm{d}t\Big)^{1/p} & \text{if } 1 \le p < \infty. \end{cases}$$

Remark 2.11 The space $L^p([0,T];X)$ is a Banach space when equipped with the norm $\|\cdot\|_{L^p([0,T];X)}$.

Definition 2.6 For $u \in L^1([0,T];X)$ we say that $w \in L^1([0,T];X)$ is the *generalized derivative* of u, denoted by $u' = w$ if for all $\phi \in C_c^\infty((0,T))$, we have

$$\int_0^T \phi'(t)u(t)\,\mathrm{d}t = -\int_0^T \phi(t)w(t)\,\mathrm{d}t.$$

Remark 2.12 Since X is separable one can show that the generalized derivative u' coincides with the weak derivative $\partial_t u$ defined by

$$\int_0^T \langle u, \partial_t\phi\rangle\,\mathrm{d}t = -\int_0^T \langle \partial_t u, \phi\rangle\,\mathrm{d}t$$

for all $\phi \in C^1([0,T];X')$ with $\phi(0) = \phi(T) = 0$.

Definition 2.7 The *Sobolev–Bochner space* $W^{1,p}([0,T];X)$ consists of all functions $u \in L^p([0,T];X)$ with $u' \in L^p([0,T];X)$ and is equipped with the norm

$$\|u\|_{W^{1,p}([0,T];X)} = \begin{cases} \operatorname{esssup}_{t\in[0,T]} (\|u(t)\|_X + \|u'(t)\|_X) & \text{if } p=\infty, \\ \Big(\int_0^T \|u(t)\|_X^p + \|u'(t)\|_X^p \,\mathrm{d}t\Big)^{1/p} & \text{if } 1 \le p < \infty. \end{cases}$$

We write $H^1([0,T];X)$ for $W^{1,2}([0,T];X)$.

Remarks 2.13 (i) We have that $W^{1,p}([0,T];X)$ is a Banach space for $1 \le p \le \infty$. If X is a Hilbert space and $p = 2$, then $W^{1,2}([0,T];X)$ is a Hilbert space denoted by $H^1([0,T];X)$.
(ii) For $1 \le p \le \infty$ and $u \in W^{1,p}([0,T];X)$, we have that $u \in C([0,T];X)$ with $\max_{t\in[0,T]} \|u(t)\| \le c\|u\|_{W^{1,p}([0,T];X)}$.

Definition 2.8 If H is a separable Hilbert space that is identified with its dual H' and such that the inclusion $X \subset H$ is dense and continuous, then (X, H, X') is called a *Gelfand* or an *evolution triple*.

Remark 2.14 For a Gelfand triple (X, H, X') the duality pairing $\langle\varphi, v\rangle$ for $\varphi \in X'$ and $v \in X$ is regarded as a continuous extension of the scalar product on H, i.e., if $\varphi \in X' \cap H'$, then

$$\langle\varphi, v\rangle = (\varphi, v)_H.$$

Below, we always consider an evolution triple (X, H, X'). The Sobolev–Bochner spaces then have the following important properties.

Remarks 2.15 (i) If $u \in L^p([0,T]; X)$ with $u' \in L^{p'}([0,T]; X')$, then $u \in C([0,T]; H)$ with $\max_{t\in[0,T]} \|u(t)\|_H \le c(\|u\|_{L^p([0,T];X)} + \|u'\|_{L^{p'}([0,T];X')})$ and the *integration-by-parts formula*

$$(u(t_2), v(t_2))_H - (u(t_1), v(t_1))_H = \int_{t_1}^{t_2} \langle u'(t), v(t)\rangle + \langle v'(t), u(t)\rangle \, \mathrm{d}t$$

holds for all $v \in L^p([0,T]; X)$ with $v' \in L^{p'}([0,T]; X)$ and $t_1, t_2 \in [0,T]$. In particular, we have

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|_H^2 = \langle u'(t), u(t)\rangle$$

for almost every $t \in [0,T]$.
(ii) If X is compactly embedded in H, $1 < p < \infty$, and $1 < q \le \infty$, then according to the *Aubin–Lions lemma* the inclusion $L^p([0,T]; X) \cap W^{1,q}([0,T]; X') \subset L^p([0,T]; H)$ is compact.
(iii) For $1 \le p < \infty$ the space $L^p([0,T]; X)$ is separable. In particular, if $(f_n)_{n\in\mathbb{N}} \subset L^p(I)$ and $(v_n)_{n\in\mathbb{N}} \subset X$ are dense subsets, then $\mathrm{span}\{f_n v_m : n, m \in \mathbb{N}\}$ is dense in $L^p(I; X)$.
(iv) If $g \in L^{p'}([0,T]; X')$, then the mapping $f \mapsto \int_0^T \langle f(t), g(t)\rangle \, \mathrm{d}t$, defined for every $f \in L^p([0,T]; X)$, belongs to $(L^p([0,T]; X))'$ for $1 \le p \le \infty$. If $1 < p < \infty$, we have that $L^p([0,T]; X)$ is reflexive provided that X is reflexive. In particular, for $1 \le p < \infty$ we have $(L^p([0,T]; X))' = L^{p'}([0,T]; X')$.
(v) We have that $L^2(I; H)$ is a Hilbert space.

2.3.3 Existence Theory for Gradient Flows

We consider a Fréchet-differentiable functional $I : X \to \mathbb{R}$ with $DI : X \to X'$ and we want to derive conditions that guarantee existence of solutions for the H-gradient flow of I formally defined by

$$\partial_t u = -\nabla_H I(u), \quad u(0) = u_0.$$

We always let (X, H, X') be an evolution triple and assume that an abstract Poincaré inequality holds, i.e., that for a seminorm $|\cdot|_X$ on X we have

$$\|u\|_X \le c_P(|u|_X + \|u\|_H).$$

for all $u \in X$.

Definition 2.9 Given $u_0 \in H$, we say that $u \in L^p([0,T];X)$ is a *solution of the H-gradient flow for* I if $u' \in L^{p'}([0,T];X')$ and for almost every $t \in [0,T]$ and every $v \in X$, we have that

$$\langle u'(t), v\rangle + DI(u)[v] = 0$$

and $u(0) = u_0$.

Example 2.3 For $I(u) = (1/2)\int_\Omega |\nabla u|^2 \, dx$ defined on $H_0^1(\Omega)$, the $L^2(\Omega)$-gradient flow is the linear heat equation $\partial_t u - \Delta u = 0$.

We follow the *Rothe method* to construct solutions. This method consists of three steps: First, we consider an implicit time discretization that replaces the time derivative by difference quotients and establishes the existence of approximations. In the second step, a priori bounds that allow us to extract weakly convergent subsequences of the approximations as the time-step size tends to zero are proved. Finally, we pass to the limit and try to show that weak limits are solutions of the gradient flow.

Definition 2.10 The functional $I : X \to \mathbb{R}$ is called *semicoercive* if there exist $s > 0$, $c_1 > 0$, and $c_2 \in \mathbb{R}$ such that

$$I(v) \geq c_1 |v|_X^s - c_2 \|v\|_H^2$$

for all $v \in X$.

Proposition 2.1 (Implicit Euler scheme) *Assume that I is semicoercive and weakly lower semicontinuous. Then for every* $\tau > 0$ *with* $4\tau c_2 < 1$ *and* $k = 1, 2, \ldots, K$, $K = \lceil T/\tau \rceil$, *the functionals* $I^k : X \to \mathbb{R}$,

$$u \mapsto I^k(u) = \frac{1}{2\tau}\|u - u^{k-1}\|_H^2 + I(u),$$

with $u^0 = u_0$ *have minimizers that satisfy*

$$(d_t u^k, v)_H + DI(u^k)[v] = 0$$

for all $v \in X$ *with the* backward difference quotient $d_t u^k = (u^k - u^{k-1})/\tau$.

Proof Since I^k is coercive, bounded from below, and weakly lower semicontinuous, the direct method in the calculus of variations implies the existence of a minimum. Since I and $v \mapsto \|v\|_H^2$ are Fréchet-differentiable, the minimizers satisfy the asserted equations. □

Remarks 2.16 (i) More generally, one can consider a pseudomonotone operator $A : X \to X'$ and look for a solution u^k of the equation $(d_t u^k, v)_H + A(u^k)[v] = 0$ for all $v \in X$.
(ii) We have that u^k is uniquely defined if I^k is strictly convex. This is often satisfied for τ sufficiently small, i.e., if I is semiconvex.

For the proof of the a priori bounds two important ingredients are required. The first is based on the binomial formula $2(a-b)a = (a-b)^2 + (a^2 - b^2)$ and shows that

$$\frac{1}{\tau}(u^k - u^{k-1}, u^k)_H = \frac{1}{2\tau}\|u^k - u^{k-1}\|_H^2 + \frac{1}{2\tau}\big(\|u^k\|_H^2 - \|u^{k-1}\|_H^2\big)$$

for $k = 1, 2, \ldots, K$. Equivalently, we have

$$(d_t u^k, u^k)_H = \frac{\tau}{2}\|d_t u^k\|_H^2 + \frac{1}{2} d_t \|u^k\|_H^2$$

which is a discrete version of the identity $2\langle u', u\rangle = (d/dt)\|u\|_H^2$. The second ingredient is the following discrete Gronwall lemma.

Lemma 2.2 (Discrete Gronwall lemma) *Let $(y^\ell)_{\ell=0,1,\ldots,L}$ be a sequence of nonnegative real numbers such that for nonnegative real numbers $a_0, b_0, b_1, \ldots, b_{L-1}$ and $\ell = 0, 1, \ldots, L$, we have*

$$y^\ell \le a_0 + \sum_{k=0}^{\ell-1} b_k y^k.$$

Then we have $\max_{\ell=0,1,\ldots,L} y^\ell \le a_0 \exp\big(\sum_{k=0}^{\ell-1} b_k\big)$.

Proof The proof follows from an inductive argument. □

We also have to assume a coerciveness property for the mapping $DI : X \to X'$.

Definition 2.11 We say that $DI : X \to X'$ is *semicoercive and bounded* if there exist $p \in (1, \infty)$, $c_1' > 0$, $c_2' \in \mathbb{R}$, and $c_3' > 0$ such that

$$DI(v)[v] \ge c_1' |v|_X^p - c_2' \|v\|_H^2$$

and

$$\|DI(v)\|_{X'} \le c_3'(1 + \|v\|_X^{p-1})$$

for all $v \in X$.

Proposition 2.2 (A priori bounds) *Suppose that $DI : X \to X'$ is semicoercive and bounded. If $4\tau c_2' \le 1$, then we have*

$$\max_{\ell=0,1,\ldots,K} \|u^\ell\|_H + \tau \sum_{k=1}^K \|u^k\|_X^p + \tau \sum_{k=1}^K \|d_t u^k\|_{X'}^{p'} + \tau \sum_{k=1}^K \|DI(u^k)\|_{X'}^{p'} \le C_0$$

with a constant $C_0 > 0$ that depends on p, T, u_0, c_P, c_1', c_2', and c_3'.

Proof Since $(d_t u^k, v)_H + DI(u^k)[v] = 0$ for all $v \in X$, we obtain by choosing $v = u^k$ that

$$\frac{\tau}{2}\|d_t u^k\|_H^2 + \frac{1}{2}d_t\|u^k\|_H^2 + c_1'|u^k|_X^p - c_2'\|u^k\|_H^2 \le (d_t u^k, u^k)_H + DI(u^k)[u^k] = 0.$$

Multiplication by τ and summation over $k = 1, 2, \dots, \ell$ for $1 \le \ell \le K$ lead to

$$\frac{1}{2}\|u^\ell\|_H^2 + \tau\sum_{k=1}^{\ell}\frac{\tau}{2}\|d_t u^k\|_H^2 + c_1'\tau\sum_{k=1}^{\ell}|u^k|_X^p \le \frac{1}{2}\|u^0\|_H^2 + c_2'\tau\sum_{k=1}^{\ell}\|u^k\|_H^2,$$

where we used the telescope effect $\tau\sum_{k=1}^{\ell} d_t\|u^k\|_H^2 = \|u^\ell\|_H^2 - \|u^0\|_H^2$. Since the second term on the left-hand side is nonnegative and since $4c_2'\tau \le 1$ so that we can absorb $c_2'\tau\|u^\ell\|_H^2$ on the left-hand side, we find that

$$\frac{1}{4}\|u^\ell\|_H^2 + c_1'\tau\sum_{k=1}^{\ell}|u^k|_X^p \le \frac{1}{2}\|u^0\|_H^2 + c_2'\tau\sum_{k=1}^{\ell-1}\|u^k\|_H^2.$$

For $\ell = 0, 1, \dots, K$, we set

$$y^\ell = \frac{1}{4}\|u^\ell\|_H^2 + c_1'\tau\sum_{k=1}^{\ell}|u^k|_X^p,$$

$a_0 = (1/2)\|u^0\|_H^2$ and $b = 4c_2'\tau$ so that

$$y^\ell \le a_0 + \sum_{k=1}^{\ell-1} b y^k$$

and we are in the situation to apply the discrete Gronwall lemma. This shows that

$$\max_{\ell=0,\dots,K}\frac{1}{4}\|u^\ell\|_H^2 + c_1'\tau\sum_{\ell=1}^{K}|u^k|_X^p \le \frac{1}{2}\|u^0\|_H^2\exp\big(4c_2'\sum_{k=1}^{K-1}\tau\big) \le \frac{1}{2}\|u^0\|_H^2\exp\big(4c_2'T\big)$$

and proves the bound for the first term on the left-hand side. The second bound follows with the abstract Poincaré inequality $\|u\|_X \le c_P(|u|_X + \|u\|_H)$ and

$$\|u^k\|_X^p \le c_P^p(|u^k|_X + \|u^k\|_H)^p \le c_P^p 2^{p-1}\big(|u^k|_X^p + c_P^p\|u^k\|_H^p\big).$$

For the third and fourth term on the left-hand side of the bound, we note that with the boundedness of DI and $(p-1)p' = p$, it follows that

$$\tau \sum_{k=1}^{L} \|DI(u^k)\|_{X'}^{p'} \le \tau (c_3')^{p'} \sum_{k=1}^{K} (1 + \|u^k\|_X^{p-1})^{p'} \le \tau (c_3')^{p'} 2^{p'-1} \sum_{k=1}^{K} (1 + \|u^k\|_X^p),$$

and the right-hand side is bounded according to the previous bounds. This proves the fourth estimate and the third bound follows immediately since $\|d_t u^k\|_{X'} = \|DI(u^k)\|_{X'}$ due to the identity $(d_t u^k, v) = -DI(u^k)[v]$ for all $v \in X$. □

To identify the limits of the approximations we define interpolants of the approximations $(u^k)_{k=0,\dots,K}$.

Definition 2.12 Given a time-step size $\tau > 0$ and a sequence $(u^k)_{k=0,\dots,K} \subset H$ for $K = \lceil T/\tau \rceil$, we set $t_k = k\tau$ for $k = 0, 1, \dots, K$ and define the piecewise constant and piecewise affine interpolants $u_\tau{}^-, u_\tau^+, \widehat{u}_\tau : [0, T] \to H$ for $t \in (t_{k-1}, t_k)$ by

$$u_\tau^-(t) = u^{k-1}, \quad u_\tau^+(t) = u^k, \quad \widehat{u}_\tau(t) = \frac{t - t_{k-1}}{\tau} u^k + \frac{t_k - t}{\tau} u^{k-1}.$$

The construction of the interpolants is illustrated in Fig. 2.14.

Remarks 2.17 (i) We have $\widehat{u}_\tau \in W^{1,\infty}([0, T]; H)$ with $\widehat{u}_\tau' = d_t u^k$ on (t_{k-1}, t_k) for $k = 1, 2, \dots, K$. Moreover, $u_\tau^+, u_\tau^- \in L^\infty([0, T]; H)$ and, e.g.,

$$\|u_\tau^+\|_{L^p([0,T];X)}^p \le \tau \sum_{k=1}^{K} \|u^k\|_X^p$$

with equality if $K\tau = T$.
(ii) We have $u_\tau^+(t) - u_\tau^-(t) = \tau \widehat{u}_\tau'(t)$ and $\widehat{u}_\tau(t) = u_\tau^-(t) + (t - t_{k-1})\widehat{u}_\tau'(t) = u_\tau^+(t) - (t_k - t)\widehat{u}_\tau'(t)$ for almost every $t \in (t_{k-1}, t_k)$.
(iii) If $\|\widehat{u}_\tau'\|_{L^1([0,T];X')} \le c$ for all $\tau > 0$, then it follows that $u_\tau^+ - u_\tau^- \to 0$ and $\widehat{u}_\tau - u_\tau^\pm \to 0$ in $L^1([0, T]; X')$ as $\tau \to 0$. In particular, all interpolants have the same limit if these exists.

Lemma 2.3 (Discrete evolution equation) *With the interpolants of the approximations $(u^k)_{k=0,\dots,K}$, we have*

$$(\widehat{u}_\tau'(t), v)_H + DI(u_\tau^+(t))[v] = 0$$

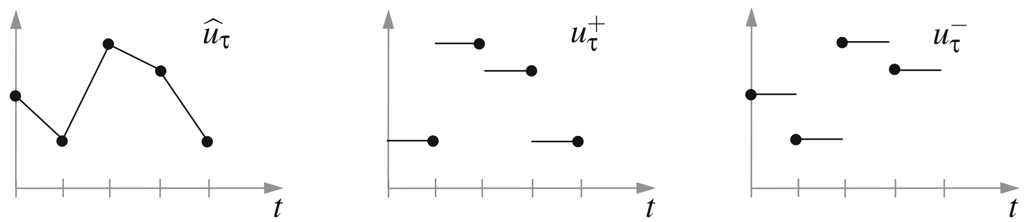

Fig. 2.14 Continuous interpolant $\widehat{u}_\tau$ (*left*) and piecewise constant interpolants u_τ^+ (*middle*) and u_τ^- (*right*)

for all $v \in X$ and almost every $t \in [0, T]$. Moreover, we have for all $\tau > 0$

$$\|u_\tau^+\|_{L^\infty([0,T];H)} + \|u_\tau^+\|_{L^p([0,T];X)} + \|\widehat{u}_\tau\|_{W^{1,p'}([0,T];X')} + \|DI(u_\tau^+)\|_{L^{p'}([0,T];X')} \le C_0.$$

Proof The identity follows directly from $(d_t u^k, v)_H + DI(u^k)[v] = 0$ for $k = 1, 2, \dots, K$ and all $v \in X$ with the definitions of the interpolants $\widehat{u}_\tau$ and u_τ^+. With the triangle inequality and $|t - t_k| \le \tau$ for $t \in (t_{k-1}, t_k)$, we observe that

$$\|\widehat{u}_\tau\|_{L^{p'}([0,T];X')} \le c_P T^{1/p'} \|u^+\|_{L^\infty([0,T];H)} + \tau \|\widehat{u}_\tau'\|_{L^{p'}([0,T];X')}.$$

The a priori bounds of Proposition 2.2 together with, e.g.,

$$\tau \sum_{k=1}^{K} \|u^k\|_X^p \ge \int_0^T \|u_\tau^+\|_X^p \, \mathrm{d}t,$$

where we used $K\tau \ge T$, imply the a priori bounds. □

The bounds for the interpolants allow us to select accumulation points.

Proposition 2.3 (Selection of a limit) *Assume that X is compactly embedded in H. Then there exist $u \in L^p([0, T]; X) \cap W^{1,p'}([0, T]; X')$ and $\xi \in L^{p'}([0, T]; X')$ such that for a sequence $(\tau_n)_{n\in\mathbb{N}}$ of positive numbers with $\tau_n \to 0$ as $n \to \infty$, we have*

$$\begin{aligned} \widehat{u}_{\tau_n}, u_{\tau_n} &\rightharpoonup^* u \quad \text{in } L^\infty([0, T]; H), \\ \widehat{u}_{\tau_n}, u_{\tau_n} &\rightharpoonup u \quad \text{in } L^p([0, T]; X), \\ \widehat{u}_{\tau_n} &\rightharpoonup u \quad \text{in } W^{1,p'}([0, T]; X'), \\ DI(u_{\tau_n}^+) &\rightharpoonup \xi \quad \text{in } L^{p'}([0, T]; X'). \end{aligned}$$

We have $u \in C([0, T]; H)$ with $u(0) = u_0$ and

$$\langle u'(t), v\rangle + \langle \xi(t), v\rangle = 0$$

for almost every $t \in [0, T]$ and all $v \in X$. In particular, if $\xi = DI(u)$, then u is a solution of the H-gradient flow for I.

Proof For a sequence $(\tau_n)_{n\in\mathbb{N}}$ of positive numbers with $\tau_n \to 0$ as $n \to \infty$, the a priori bounds yield the existence of weak limits for an appropriate subsequence which is not relabeled. Due to the bound for $\widehat{u}_\tau'$, the weak limits coincide. Multiplying the discrete evolution equation of Lemma 2.3 by $\phi \in C([0, T])$ and integrating the resulting identity over $[0, T]$ we find that

$$\int_0^T \langle \widehat{u}'_{\tau_n}, \phi v\rangle + DI(u^+_{\tau_n})[\phi v]\,\mathrm{d}t = 0$$

for every $v \in X$. Since $\phi v \in L^p([0,T];X)$ we can pass to the limit $n \to \infty$ in the equation and obtain

$$\int_0^T \langle u', \phi v\rangle + \langle \xi, \phi v\rangle\,\mathrm{d}t = 0.$$

Since this holds for every $\phi \in C([0,T])$ we deduce the asserted equation. The mapping $v \mapsto v(0)$ defines a bounded linear operator $L^p([0,T];X) \cap W^{1,p'}([0,T];X') \to H$ which is weakly continuous. Since $\widehat{u}_{\tau_n}(0) = u_0$ for all $n \in \mathbb{N}$, we deduce that $u(0) = u_0$. By continuous embeddings we also have $u \in C([0,T];H)$ which implies the continuous attainment of the initial data. □

Remark 2.18 The assumed identity $\xi = DI(u)$ in $L^{p'}([0,T];X')$, i.e., the convergence $DI(u^+_{\tau_n}) \rightharpoonup DI(u)$ can in general only be established under additional conditions on DI and requires special techniques from nonlinear functional analysis, e.g., based on concepts of pseudomonotonicity.

Example 2.4 For $F \in C^1(\mathbb{R})$ with $0 \le F(s) \le c_F(1+|s|^2)$ and $f(s) = F'(s)$ such that $|f(s)| \le c'_F(1+|s|)$, we consider

$$I(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\,\mathrm{d}x + \int_\Omega F(u)\,\mathrm{d}x.$$

Then, for $X = H^1_0(\Omega)$ and $H = L^2(\Omega)$, the conditions of the previous propositions are satisfied with $p = 2$ and

$$DI(u)[v] = \int_\Omega \nabla u \cdot \nabla v\,\mathrm{d}x + \int_\Omega f(u)v\,\mathrm{d}x.$$

We have $u^+_{\tau_n} \rightharpoonup u \in L^2([0,T];H^1_0(\Omega))$ so that $\nabla u^+_{\tau_n} \rightharpoonup \nabla u$ in $L^2([0,T];L^2(\Omega;\mathbb{R}^d))$ and thus

$$\int_0^T\int_\Omega \nabla u^+_{\tau_n} \cdot \nabla w\,\mathrm{d}x\,\mathrm{d}t \to \int_0^T\int_\Omega \nabla u \cdot \nabla w\,\mathrm{d}x\,\mathrm{d}t$$

for all $w \in L^2([0,T];H^1_0(\Omega))$. The compactness of the embedding

$$L^2([0,T];H^1_0(\Omega)) \cap W^{1,2}([0,T];H^1_0(\Omega)') \to L^2([0,T];L^2(\Omega)) = L^2([0,T]\times\Omega)$$

in combination with the generalized dominated convergence theorem shows that

$$\int_0^T \int_\Omega f(u_{\tau_n}^+) w \, dx \, dt \to \int_0^T \int_\Omega f(u) w \, dx \, dt.$$

Altogether this proves that

$$\int_0^T DI(u_{\tau_n}^+)[w] \, dt \to \int_0^T \int_\Omega \nabla u \cdot \nabla w \, dx \, dt + \int_0^T \int_\Omega f(u) w \, dx \, dt,$$

i.e., $\xi = DI(u)$.

Remarks 2.19 (i) For the semilinear heat equation $\partial_t u = \Delta u - f(u)$ of Example 2.4, one can establish the existence of a solution under more general conditions on f. Moreover, one can prove stronger a priori bounds and the energy law

$$I(u(T')) + \int_0^{T'} \|u'(t)\|_{L^2(\Omega)}^2 \, dt \le I(u_0)$$

for almost every $T' \in [0, T]$ provided $u_0 \in H_0^1(\Omega)$. The key ingredient is the convexity of I in the highest-order term.
(ii) An alternative method to establish the existence of solutions for gradient flows is the Galerkin method which is based on a discretization in space. This leads to a sequence of ordinary differential equations on finite-dimensional spaces and with appropriate a priori bounds, one can then show under appropriate conditions that the approximate solutions converge to a solution as the dimension tends to infinity.

2.3.4 Subdifferential Flows

The estimates for discretized gradient flows can be significantly improved if the functional I is convex, since then we would have

$$I(u^k) + DI(u^k)[u^{k-1} - u^k] \le I(u^{k-1}).$$

In particular, choosing $v = d_t u^k$ in the identity $(d_t u^k, v)_H + DI(u^k)[v] = 0$ gives

$$\tau \|d_t u^k\|_H^2 + I(u^k) \le I(u^{k-1})$$

and a summation over k yields the a priori bound $I(u^L) + \tau \sum_{k=1}^L \|d_t u^k\|_H^2 \le I(u^0)$. With these observations it is possible to establish a theory for convex functionals that are not differentiable. We always consider a Hilbert space H that is identified with its dual.

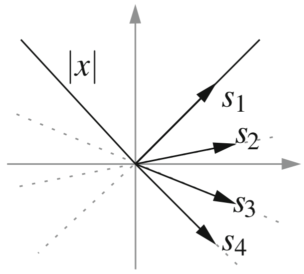

Fig. 2.15 Subdifferential of the function $x \mapsto |x|$ at $x = 0$; the arrows s_1, s_2, s_3, s_4 indicate subgradients at 0 which are the slopes of supporting hyperplanes at 0

Definition 2.13 We say that a functional $I : H \to \mathbb{R} \cup \{+\infty\}$ belongs to the class $\Gamma(H)$ if it is convex, lower semicontinuous, i.e., $I(u) \leq \liminf_{n\to\infty} I(u_n)$ whenever $u_n \to u$ in H as $n \to \infty$, and *proper*, i.e., there exists $u \in H$ with $I(u) \in \mathbb{R}$.

We assume that $I \in \Gamma(H)$ below.

Definition 2.14 The *subdifferential* $\partial I : H \to 2^H$ of I associates to every $u \in H$ the set

$$\partial I(u) = \{v \in H : I(w) \geq I(u) + (v, w - u)_H \text{ for all } w \in H\}.$$

The elements in $\partial I(u)$ are called *subgradients* of I at u.

Example 2.5 For $F(x) = |x|$, $x \in \mathbb{R}$, we have $\partial F(0) = [-1, 1]$, cf. Fig. 2.15.

Remarks 2.20 (i) The subdifferential $\partial I(u)$ consists of all slopes of affine functions that are below I and that intersect the graph of I at u.
(ii) For all $u_1, u_2 \in H$ and $v_1 \in \partial I(u_1)$, $v_2 \in \partial I(u_2)$ we have the *monotonicity estimate*

$$(v_1 - v_2, u_1 - u_2)_H \geq 0.$$

(iii) We have $0 \in \partial I(u)$ if and only if $u \in H$ is a global minimum for I.
(iv) We have $\partial I(u) = \{s\}$ for $s \in H$ if and only if I is Gâteaux-differentiable at u.
(v) For $I, J \in \Gamma(H)$ we have $\partial(I + J) \subset \partial I + \partial J$, and if there exists a point at which I and J are finite and I or J is continuous, we have equality.

Theorem 2.7 (Resolvent operator) *Let* $I \in \Gamma(H)$. *For every* $w \in H$ *and* $\lambda > 0$ *there exists a unique* $u \in H$ *with*

$$u + \lambda \partial I(u) \ni w.$$

This defines the resolvent operator $u = R_\lambda(w) = (\mathrm{Id} + \lambda \partial I)^{-1}(w)$.

Proof For a short proof we make the simplifying but nonrestrictive assumption that

$$I(v) \geq -c_1 - c_2 \|v\|_H.$$

For $\lambda > 0$ and $w \in H$ we consider the minimization problem defined through the functional

$$I_{\lambda,w}(u) = \frac{1}{2\lambda}\|u - w\|_H^2 + I(u) = \frac{1}{2\lambda}\|u\|_H^2 - \frac{1}{\lambda}(u, w)_H + \frac{1}{2\lambda}\|w\|_H^2 + I(u).$$

The identity $2(a^2 + b^2) = (a + b)^2 + (a - b)^2$ and the convexity of I show that for $u_1, u_2 \in H$ we have

$$\begin{aligned}\frac{1}{2}I_{\lambda,w}(u_1) + \frac{1}{2}I_{\lambda,w}(u_2) - I_{\lambda,w}\big(\frac{u_1 + u_2}{2}\big)\\ = \frac{1}{8\lambda}\|u_1 - u_2\|_H^2 + \frac{1}{2}I(u_1) + \frac{1}{2}I(u_2) - I\big(\frac{u_1 + u_2}{2}\big) \geq \frac{1}{8\lambda}\big\|u_1 - u_2\big\|_H^2,\end{aligned}$$

i.e., $I_{\lambda,w}$ is *strictly convex*. Thus, if $I_{\lambda,w}$ has a minimizer, then it is unique. Moreover, $u \in H$ minimizes $I_{\lambda,w}$ if and only if $0 \in \partial I_{\lambda,w}(u) = (1/\lambda)(u - w) + \partial I(u)$. It remains to show that there exists a minimizer. Since $I_{\lambda,w}$ is convex and lower semicontinuous it follows that I is weakly lower semicontinuous. We also have that $I_{\lambda,w}$ is coercive since two applications of Young's inequality lead to

$$\begin{aligned}I_{\lambda,w}(v) &\geq \frac{1}{2\lambda}\|v\|_H^2 - \frac{1}{\lambda}(v, w)_H + \frac{1}{2\lambda}\|w\|_H^2 - c_1 - c_2\|v\|_H\\ &\geq \frac{1}{4\lambda}\|v\|_H^2 - \frac{4}{\lambda}\|w\|_H^2 - c_1 - 4\lambda c_2^2.\end{aligned}$$

This estimate also proves the boundedness from below. The direct method in the calculus of variations thus implies the existence of a minimizer. □

Definition 2.15 The *Yosida regularization* $A_\lambda : H \to H$ is for $w \in H$ defined by $A_\lambda(w) = (1/\lambda)(w - R_\lambda(w))$.

Remark 2.21 The resolvent operator satisfies $\lim_{\lambda\to 0} R_\lambda(w) = w$. We have that A_λ is Lipschitz continuous with Lipschitz constant $2/\lambda$ and approximates ∂I in the sense that $A_\lambda(w) \in \partial I(R_\lambda w)$.

The theorem about the resolvent operator implies that for a time-step size $\tau > 0$ and an initial $u^0 \in H$, there exists a unique sequence $(u^k)_{k=0,\ldots,L} \subset H$ with

$$d_t u^k \in -\partial I(u^k)$$

since this is equivalent to $u^k = R_\tau(u^{k-1})$. We expect that as $\tau \to 0$ the approximations converge to a solution of the *subdifferential flow*

$$u' \in -\partial I(u), \quad u(0) = u_0.$$

Related a priori bounds that permit a corresponding passage to a limit will be discussed in Chap. 4.

Theorem 2.8 (Subdifferential flow, [2]) *For every* $u_0 \in H$ *such that* $\partial I(u_0) \neq \emptyset$ *and every* $T > 0$*, there exists a unique function* $u \in C([0,T]; H)$ *with* $u' \in L^\infty([0,T]; H)$ *such that* $u(0) = u_0$*,* $\partial I(u(t)) \neq \emptyset$ *for every* $t \in [0,T]$*, and*

$$u'(t) \in -\partial I(u(t))$$

for almost every $t \in [0,T]$.

Proof The existence of a solution is established by considering for every $\lambda > 0$ the problem

$$\partial_t u_\lambda = -A_\lambda(u_\lambda), \quad u_\lambda(0) = u_0$$

and studying the limit $\lambda \to 0$. Uniqueness of solutions follows from the convexity of I, i.e., if u_1 and u_2 are solutions then the monotonicity property of I shows that

$$-(u_1'(t) - u_2'(t), u_1(t) - u_2(t))_H \geq 0$$

for almost every $t \in [0,T]$ and this implies that

$$\frac{1}{2}\frac{d}{dt}\|(u_1 - u_2)(t)\|_H^2 = (u_1'(t) - u_2'(t), u_1(t) - u_2(t))_H \leq 0.$$

Since $u_1(0) = u_2(0)$ we deduce that $u_1(t) = u_2(t)$ for every $t \in [0,T]$. □

Remarks 2.22 (i) Negative subgradients are in general no descent directions. For the subdifferential flow one can however show that $u'(t) = -\partial^0 I(u(t))$ for almost every $t \in [0,T]$, where $\partial^0 I(v)$ is the subgradient $s \in \partial I(v)$ with minimal norm, i.e., $\|\partial^0 I(v)\|_H = \min_{r \in \partial I(v)} \|r\|_H$.
(ii) If $\partial I(u_0) = \emptyset$, then there exists a unique solution $u \in C([0,T]; H)$ such that $t^{1/2} u' \in L^2([0,T]; H)$.

References

1. Attouch, H., Buttazzo, G., Michaille, G.: Variational Analysis in Sobolev and BV Spaces. MPS/SIAM Series on Optimization, vol. 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2006)
2. Brézis, H.R.: Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert. North-Holland Publishing Co., Amsterdam (1973). North-Holland Mathematics Studies, No. 5. Notas de Matemática (50)
3. Ciarlet, P.G.: Mathematical Elasticity. Vol. I: Three-Dimensional Elasticity. Studies in Mathematics and its Applications, vol. 20. North-Holland Publishing Co., Amsterdam (1988)
4. Dacorogna, B.: Direct Methods in the Calculus of Variations. Applied Mathematical Sciences, vol. 78, 2nd edn. Springer, New York (2008)
5. Eck, C., Garcke, H., Knabner, P.: Mathematische Modellierung, 2nd edn. Springer Lehrbuch. Springer, Berlin (2011)
6. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19, 2nd edn. American Mathematical Society, Providence (2010)

7. Roubíček, T.: Nonlinear Partial Differential Equations with Applications. International Series of Numerical Mathematics, vol. 153, 2nd edn. Birkhäuser/Springer Basel AG, Basel (2013)
8. Ružička, M.: Nichtlineare Funktional Analysis. Springer, Berlin-Heidelberg-New York (2004)
9. Salsa, S., Vegni, F.M.G., Zaretti, A., Zunino, P.: A Primer on PDEs. Unitext, vol. 65, italian edn. Springer, Milan (2013)
10. Struwe, M.: Variational Methods, 4th edn. Springer, Berlin (2008)
11. Temam, R., Miranville, A.: Mathematical Modeling in Continuum Mechanics, 2nd edn. Cambridge University Press, Cambridge (2005)

and with $\|\nabla u\| = \|\nabla v\|$, we deduce that

$$\sup_{v \in H_0^1(\Omega)} \frac{a(u,v)}{\|\nabla v\|} \geq c_H \|\nabla u\|.$$

The second condition of the proposition is a direct consequence of the requirement that ω^2 is not an eigenvalue of $-\Delta$.

Remark 4.5 Proposition 4.2 is important for the analysis of saddle-point problems; the seminal paper [7] provides conditions that imply the assumptions of the proposition.

4.2.3 Abstract Subdifferential Flow

The subdifferential flow of a convex and lower semicontinuous functional $I : H \to \mathbb{R} \cup \{+\infty\}$ arises as an evolutionary model in applications, and can be used as a basis for numerical schemes to minimize I. The corresponding differential equation seeks $u : [0, T] \to H$, such that $u(0) = u_0$ and

$$\partial_t u \in -\partial I(u),$$

i.e., $u(0) = u_0$ and

$$(-\partial_t u, v - u)_H + I(u) \leq I(v)$$

for almost every $t \in [0, T]$ and every $v \in H$. An implicit discretization of this nonlinear evolution equation is equivalent to a sequence of minimization problems involving a quadratic term. We recall that $d_t u^k = (u^k - u^{k-1})/\tau$ denotes the backward difference quotient.

Theorem 4.7 (Semidiscrete scheme [15, 17]) *Assume that $I \geq 0$ and for $u^0 \in H$ let $(u^k)_{k=1,\dots,K} \subset H$ be minimizers for*

$$I_\tau^k(w) = \frac{1}{2\tau}\|w - u^{k-1}\|_H^2 + I(w)$$

for $k = 1, 2, \dots, K$. For $L = 1, 2, \dots, K$, we have

$$I(u^L) + \tau \sum_{k=1}^{L} \|d_t u^k\|_H^2 \leq I(u^0).$$

With the computable quantities

$$\mathscr{E}_k = -\tau \|d_t u^k\|_H^2 - I(u^k) + I(u^{k-1})$$

and the affine interpolant $\widehat{u}_\tau : [0,T] \to H$ of the sequence $(u^k)_{k=0,\dots,K}$ we have the a posteriori error estimate

$$\max_{t\in[0,T]} \|u-\widehat{u}\|_H^2 \le \|u_0-u^0\|_H^2 + \tau \sum_{k=1}^{L} \mathscr{E}_k.$$

We have the a priori error estimate

$$\max_{k=0,\dots,K} \|u(t_k)-u^k\|_H^2 \le \|u_0-u^0\|_H^2 + \tau I(u^0),$$

and under the condition $\partial I(u^0) \neq \emptyset$, the improved variant

$$\max_{k=0,\dots,K} \|u(t_k)-u^k\|_H^2 \le \|u_0-u^0\|_H^2 + \tau^2 \|\partial^o I(u^0)\|_H^2,$$

where $\partial^o I(u^0) \in H$ denotes the element of minimal norm in $\partial I(u^0)$.

Proof The direct method in the calculus of variations yields that for $k = 1, 2, \dots, K$, there exists a unique minimizer $u^k \in H$ for I_τ^k, and we have $d_t u^k \in -\partial I(u^k)$, i.e.,

$$(-d_t u^k, v-u^k)_H + I(u^k) \le I(v)$$

for all $v \in H$; the choice of $v = u^{k-1}$ implies that

$$-\mathscr{E}_k = \tau\|d_t u^k\|_H^2 + I(u^k) - I(u^{k-1}) \le 0$$

with $0 \le \mathscr{E}_k \le -\tau d_t I(u^k)$. A summation over $k = 1, 2, \dots, L$ yields the asserted stability estimate. If $\widehat{u}_\tau$ is the piecewise affine interpolant of $(u^k)_{k=0,\dots,K}$ associated to the time steps $t_k = k\tau$, $k = 0, 1, \dots, K$, and u_τ^+ is such that $u_\tau^+|_{(t_{k-1},t_k)} = u^k$ for $k = 1, 2, \dots$ and $t_k = k\tau$, then we have

$$(-\partial_t \widehat{u}_\tau, v-u_\tau^+)_H + I(u_\tau^+) \le I(v)$$

for almost every $t \in [0,T]$ and all $v \in H$. In introducing

$$\mathscr{C}_\tau(t) = (-\partial_t \widehat{u}_\tau, u_\tau^+ - \widehat{u}_\tau)_H - I(u_\tau^+) + I(\widehat{u}_\tau)$$

we have

$$(-\partial_t \widehat{u}_\tau, v-\widehat{u}_\tau)_H + I(\widehat{u}_\tau) \le I(v) + \mathscr{C}_\tau(t).$$

The choice of $v = u$ in this inequality and $v = \widehat{u}_\tau$ in the continuous evolution equation yield

$$\frac{d}{dt}\frac{1}{2}\|u-\widehat{u}\|_H^2 = (-\partial_t[u-\widehat{u}_\tau], \widehat{u}_\tau - u)_H \le \mathscr{C}_\tau(t).$$

Noting $\widehat{u}_\tau - u_\tau^+ = (t - t_k)\partial_t \widehat{u}_\tau$ for $t \in (t_{k-1}, t_k)$ and using the convexity of I, i.e.,

$$I(\widehat{u}_\tau) \leq \frac{t_k - t}{\tau} I(u^{k-1}) + \frac{t - t_{k-1}}{\tau} I(u^k),$$

we verify for $t \in (t_{k-1}, t_k)$ using $u_\tau^+ = u^k$ that

$$\mathscr{C}_\tau(t) \leq (t - t_k)\|\partial_t \widehat{u}_\tau\|_H^2 - I(u_\tau^+) + \frac{t_k - t}{\tau} I(u^{k-1}) + \frac{t - t_{k-1}}{\tau} I(u^k) = \frac{t_k - t}{\tau}\mathscr{E}_k.$$

With $\mathscr{E}_k \leq -\tau d_t I(u^k)$ and $I \geq 0$ we deduce that

$$\int_0^{t_L} \mathscr{C}_\tau(t)\,\mathrm{d}t \leq \tau \sum_{k=1}^{L} \mathscr{E}_k \leq -\tau^2 \sum_{k=1}^{L} d_t I(u^k) = -\tau\big(I(u^L) - I(u^0)\big) \leq \tau I(u^0),$$

which implies the a posteriori and the first a priori error estimate. Assume that $\partial I(u^0) \neq \emptyset$ and define $u^{-1} \in H$ so that $d_t u^0 = (u^0 - u^{-1})/\tau = -\partial^o I(u^0)$, i.e., the discrete evolution equation also holds for $k = 0$,

$$(-d_t u^0, v - u^0)_H + I(u^0) \leq I(v)$$

for all $v \in H$. Choosing $v = u^k$ in the equation for $d_t u^{k-1}$, $k = 1, 2, \dots, K$, we observe that

$$(-d_t u^{k-1}, u^k - u^{k-1})_H + I(u^{k-1}) \leq I(u^k),$$

i.e., $-\tau d_t I(u^k) \leq \tau (d_t u^k, d_t u^{k-1})_H$, and it follows that

$$\begin{aligned}\mathscr{E}_k &= -\tau(d_t u^k, d_t u^k)_H - \tau d_t I(u^k) \leq -\tau(d_t u^k, d_t u^k)_H + \tau(d_t u^{k-1}, d_t u^k)_H \\ &= -\tau^2(d_t^2 u^k, d_t u^k)_H = -\tau^2 \frac{d_t}{2}\|d_t u^k\|_H^2 - \frac{\tau^3}{2}\|d_t^2 u^k\|_H^2 \leq -\tau^2 \frac{d_t}{2}\|d_t u^k\|_H^2.\end{aligned}$$

This implies that

$$\int_0^{t_L} \mathscr{C}_\tau(t)\,\mathrm{d}t \leq \tau \sum_{k=1}^{L} \mathscr{E}_k \leq \frac{\tau^2}{2}\|d_t u^0\|_H^2 = \frac{\tau^2}{2}\|\partial^o I(u^0)\|_H^2,$$

which proves the improved a priori error estimate. □

Remarks 4.6 (i) The condition $\partial I(u^0) \neq \emptyset$ is restrictive in many applications.
(ii) Subdifferential flows $\partial_t u \in -\partial I(u)$, i.e., $Lu \ni 0$ for $Lu = \partial_t u + v$ with $v \in \partial I(u)$, and with a convex functional $I : H \to \mathbb{R} \cup \{+\infty\}$ define monotone

problems in the sense that

$$\big(Lu_1 - Lu_2, u_1 - u_2\big)_H = (\partial_t(u_1 - u_2) + (v_1 - v_2), u_1 - u_2\big)_H$$
$$\geq \big(\partial_t(u_1 - u_2), u_1 - u_2\big)_H = \frac{1}{2}\frac{d}{dt}\|u_1 - u_2\|_H^2$$

for u_1, u_2 and v_1, v_2 with $v_i \in \partial I(u_i)$, $i = 1, 2$.
(iii) If $I : H \to \mathbb{R} \cup \{+\infty\}$ is strongly monotone in the sense that $(u_1 - u_2, v_1 - v_2)_H \geq \alpha\|u_1 - u_2\|_H^2$ whenever $v_\ell \in \partial I(u_\ell)$, $\ell = 1, 2$, and if there exists a solution $\overline{u} \in H$ of the stationary inclusion $\overline{v} = 0 \in \partial I(\overline{u})$, then we have $u(t) \to \overline{u}$ as $t \to \infty$. A proof follows from the estimate

$$\frac{1}{2}\frac{d}{dt}\|u - \overline{u}\|_H^2 = -(v - \overline{v}, u - \overline{u})_H \leq -\alpha\|u - \overline{u}\|_H^2,$$

where $v = -\partial_t u \in \partial I(u)$, and an application of Gronwall's lemma.

4.2.4 Weak Continuity Methods

Let $(u_h)_{h>0} \subset X$ be a bounded sequence in the reflexive, separable Banach space X such that there exists a weak limit $u \in X$ of a subsequence that is not relabeled, i.e., we have $u_h \rightharpoonup u$ as $h \to 0$. For an operator $F : X \to X'$, we define the sequence $(\xi_h)_{h>0} \subset X'$ through $\xi_h = F(u_h)$, and if the sequence is bounded in X', then there exists $\xi \in X'$, such that for a further subsequence $(\xi_h)_{h>0}$ which again is not relabeled, we have $\xi_h \rightharpoonup^* \xi$. The important question is now whether we have weak continuity in the sense that

$$F(u) = \xi.$$

Notice that weak continuity is a strictly stronger notion of continuity than strong continuity. For partial differential equations, this property is called *weak precompactness* of the solution set of the homogeneous equation, i.e., if $(u_j)_{j\in\mathbb{N}}$ is a sequence with $F(u_j) = 0$ for all $j \in \mathbb{N}$ and $u_j \rightharpoonup u$ as $j \to \infty$ then we may deduce that $F(u) = 0$. Such implications may also be regarded as properties of weak stability since they imply that if $F(u_j) = r_j$ with $\|r_j\|_{X'} \leq \varepsilon_j$ and $\varepsilon_j \to 0$ as $j \to \infty$, then we have $F(u) = 0$ for every accumulation point of the sequence $(u_j)_{j\in\mathbb{N}}$.

Theorem 4.8 (Discrete compactness) *For every $h > 0$ let $u_h \in X_h$ solve $F_h(u_h) = 0$. Assume that $F_h(u_h) \in X'$ with $\|F(u_h)\|_{X'} \leq c$ for all $h > 0$ and F is weakly continuous on X, i.e., $F(u_j)[v] \to F(u)[v]$ for all $v \in X$ whenever $u_j \rightharpoonup u$ in X. Suppose that for every bounded sequence $(w_h)_{h>0} \subset X$ with $w_h \in X_h$ for all $h > 0$, we have*

$$\|F(w_h) - F_h(w_h)\|_{X_h'} \to 0$$

as $h \to 0$ *and* $(X_h)_{h>0}$ *is dense in* X *with respect to strong convergence. If* $(u_h)_{h>0} \subset X$ *is bounded, then there exists a subsequence* $(u_{h'})_{h'>0}$ *and* $u \in X$ *such that* $u_h \rightharpoonup u$ *in* X *and* $F(u) = 0$.

Proof After extraction of a subsequence, we may assume that $u_h \rightharpoonup u$ in X as $h \to 0$ for some $u \in X$. Fixing $v \in X$ and using $F_h(u_h)[v_h] = 0$ for every $v_h \in X_h$, we have

$$F(u_h)[v] = F(u_h)[v - v_h] + F(u_h)[v_h] - F_h(u_h)[v_h].$$

For a sequence $(v_h)_{h>0} \subset X$ with $v_h \in X_h$ for every $h > 0$ and $v_h \to v$ in X, we find that

$$|F(u_h)[v - v_h]| \le \|F(u_h)\|_{X'} \|v - v_h\|_X \to 0$$

as $h \to 0$. The sequences $(u_h)_{h>0}$ and $(v_h)_{h>0}$ are bounded in X and thus

$$|F(u_h)[v_h] - F_h(u_h)[v_h]| \le \|F(u_h) - F_h(u_h)\|_{X_h'} \|v_h\|_X \to 0$$

as $h \to 0$. Together with the weak continuity of F we find that

$$F(u)[v] = \lim_{h\to 0} F(u_h)[v] = 0.$$

Since $v \in X$ was arbitrary this proves the theorem. □

The crucial part in the theorem is the weak continuity of the operator F. We include an example of an operator related to a constrained nonlinear partial differential equation that fulfills this requirement.

Example 4.16 (*Harmonic maps*) Let $(u_j)_{j\in\mathbb{N}} \subset H^1(\Omega;\mathbb{R}^3)$ be a bounded sequence such that $|u_j(x)| = 1$ for all $j \in \mathbb{N}$ and almost every $x \in \Omega$. Assume that for every $j \in \mathbb{N}$ and all $v \in H^1(\Omega;\mathbb{R}^3) \cap L^\infty(\Omega;\mathbb{R}^3)$, we have

$$F(u_j)[v] = \int_\Omega \nabla u_j \cdot \nabla v \,\mathrm{d}x - \int_\Omega |\nabla u_j|^2 u_j \cdot v \,\mathrm{d}x = 0.$$

The choice of $v = u_j \times w$ shows that we have

$$\widetilde{F}(u_j)[w] = \int_\Omega \nabla u_j \cdot \nabla(u_j \times w)\,\mathrm{d}x = 0$$

for all $w \in H^1(\Omega;\mathbb{R}^3) \cap L^\infty(\Omega;\mathbb{R}^3)$. Using $\partial_k u_j \cdot \partial_k(u_j \times w) = \partial_k u_j \cdot (u_j \times \partial_k w)$ for $k = 1, 2, \ldots, d$, we find that

$$\widetilde{F}(u_j)[w] = \sum_{k=1}^d \int_\Omega \partial_k u_j \cdot (u_j \times \partial_k w)\,\mathrm{d}x = 0.$$

If $u_j \rightharpoonup u$ in $H^1_D(\Omega; \mathbb{R}^3)$, then $u_j \to u$ in $L^2(\Omega; \mathbb{R}^3)$ and thus, for every fixed $w \in C^\infty(\overline{\Omega}; \mathbb{R}^3)$, we can pass to the limit and find that

$$\widetilde{F}(u)[w] = 0.$$

Since up to a subsequence we have $u_j(x) \to u(x)$ for almost every $x \in \Omega$, we verify that $|u(x)| = 1$ for almost every $x \in \Omega$. A density result shows that this holds for all $w \in H^1(\Omega; \mathbb{R}^3) \cap L^\infty(\Omega; \mathbb{R}^3)$. Reversing the above argument by choosing $w = u \times v$ and employing the identity $a \times (b \times c) = (b \cdot a)c - (c \cdot a)b$ shows that $F(u)[v] = 0$ for all $v \in H^1(\Omega; \mathbb{R}^3) \cap L^\infty(\Omega; \mathbb{R}^3)$.

A general concept for weak continuity is based on the notion of pseudomonotonicity.

Example 4.17 (*Pseudomonotone operators*) The operator $F : X \to X'$ is a *pseudomonotone* operator if it is bounded, i.e., $\|F(u)\|_{X'} \le c(1 + \|u\|_X^s)$ for some $s \ge 0$, and whenever $u_j \rightharpoonup u$ in X, we have the implication that

$$\limsup_{j\to\infty} F(u_j)[u_j - u] \le 0 \quad \Longrightarrow \quad F(u)[u - v] \le \liminf_{j\to\infty} F(u_j)[u_j - v].$$

For such an operator we have that if $F(u_h)[v_h] = \ell(v_h)$ for all $v_h \in X_h$ with a strongly dense family of subspaces $(X_h)_{h>0}$ and $u_h \rightharpoonup u$ as $h \to 0$, then $F(u) = \ell$. To verify this, let $v \in X$ and $(v_h)_{h>0}$ with $v_h \in X_h$ such that $v_h \to u$ and note that

$$\begin{aligned} \limsup_{h\to 0} F(u_h)[u_h - u] &= \limsup_{h\to 0} F(u_h)[u_h - v_h] + F(u_h)[v_h - u] \\ &= \limsup_{h\to 0} \ell(u_h - v_h) + F(u_h)[v_h - u] = 0. \end{aligned}$$

Pseudomonotonicity yields for every $v_{h'} \in \cup_{h>0} X_h$ that

$$F(u)[u - v_{h'}] \le \liminf_{h\to 0} F(u_h)[u_h - v_{h'}] = \lim_{h\to 0} \ell(u_h - v_{h'}) = \ell(u - v_{h'}).$$

With the density of $(X_h)_{h>0}$ in X, we conclude that $F(u)[u - v] \le \ell(u - v)$ for all $v \in X$ and with $v = u \pm w$, we find that $F(u)[w] = \ell(w)$ for all $w \in X$.

Remarks 4.7 (i) Radially continuous bounded operators are pseudomonotone. Here, radial continuity means that $t \mapsto F(u + tv)[v]$ is continuous for $t \in \mathbb{R}$ and all $u, v \in X$. These operators allow us to apply Minty's trick to deduce from the inequality $\ell(u - v) - F(v)[u - v] \ge 0$ for all $v \in X$ that $F(u) = \ell$. To prove this implication, note that with $v = u + \varepsilon w$, we find that $\ell(w) - F(u + \varepsilon w)[w] \le 0$ and by radial continuity for $\varepsilon \to 0$, it follows that $\ell(w) - F(u)[w] \le 0$ and hence $F(u) = \ell$.
(ii) Pseudomonotone operators are often of the form $F = F_1 + F_2$ with a monotone operator F_1 and a weakly continuous operator F_2, e.g., a lower-order term described by F_2.

Example 4.18 (*Quasilinear diffusion*) The concept of pseudomonotonicity applies to the quasilinear elliptic equation

$$-\operatorname{div}\big(|\nabla u|^{p-2}\nabla u\big) + g(u) = f \text{ in } \Omega, \quad u|_{\partial\Omega} = 0 \text{ on } \partial\Omega,$$

with $g \in C(\mathbb{R})$ such that $|g(s)| \le c(1 + |s|^{r-1})$ and $1 < p < d$, $r < dp/(d-p)$.

4.3 Solution of Discrete Problems

We discuss in this section the practical solution of discretized minimization problems of the form

$$\text{Minimize } I_h(u_h) = \int_\Omega W(\nabla u_h) + g(u_h)\,\mathrm{d}x \text{ among } u_h \in \mathscr{A}_h.$$

In particular, we investigate four model situations with smooth and nonsmooth integrands and smooth and nonsmooth constraints included in $\mathscr{A}$. The iterative algorithms are based on an approximate solution of the discrete Euler–Lagrange equations. More general results can be found in the textbooks [4, 12].

4.3.1 Smooth, Unconstrained Minimization

Suppose that

$$\mathscr{A}_h = \{u_h \in \mathscr{S}^1(\mathscr{T}_h)^m : u_h|_{\Gamma_\mathrm{D}} = u_{\mathrm{D},h}\}$$

and I_h is defined as above with functions $W \in C^1(\mathbb{R}^{m\times d})$ and $g \in C^1(\mathbb{R}^m)$. The case $\Gamma_\mathrm{D} = \emptyset$ is not generally excluded in the following. A necessary condition for a minimizer $u_h \in \mathscr{A}_h$ is that for all $v_h \in \mathscr{S}^1_\mathrm{D}(\mathscr{T}_h)^m$, we have

$$F_h(u_h)[v_h] = \int_\Omega DW(\nabla u_h)\cdot\nabla v_h + Dg(u_h)\cdot v_h\,\mathrm{d}x = 0.$$

Steepest descent methods successively lower the energy by minimizing in descent directions defined through an appropriate gradient.

Algorithm 4.1 (*Descent method*) *Let* $(\cdot,\cdot)_H$ *be a scalar product on* $\mathscr{S}^1_\mathrm{D}(\mathscr{T}_h)^m$ *and* $\mu \in (0, 1/2)$. *Given* $u_h^0 \in \mathscr{A}_h$, *compute the sequence* $(u_h^j)_{j=0,1,\ldots}$ *via* $u_h^{j+1} = u_h^j + \alpha_j d_h^j$ *with* $d_h^j \in \mathscr{S}^1_\mathrm{D}(\mathscr{T}_h)^m$ *such that*

$$(d_h^j, v_h)_H = -F_h(u_h^j)[v_h]$$